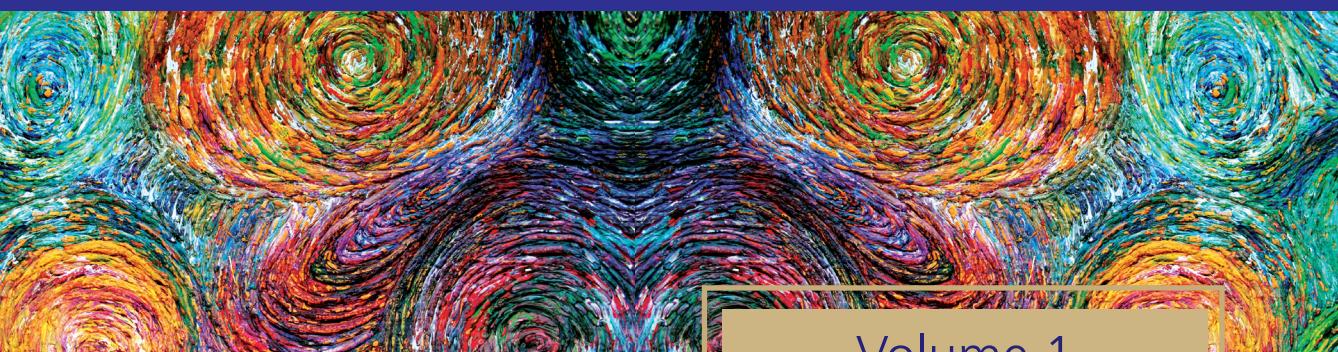


Handbook of Theories of Social Psychology



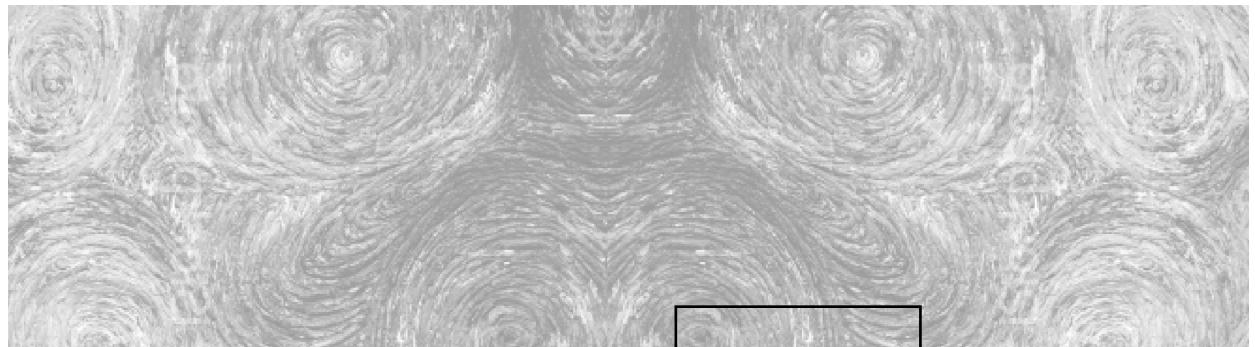
Volume 1

Edited by
Paul A. M. Van Lange
Arie W. Kruglanski
and E. Tory Higgins



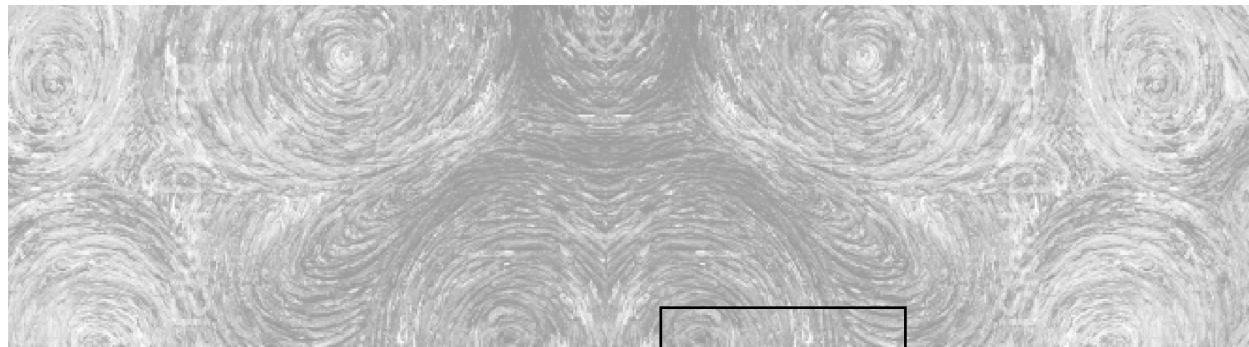
Handbook of

Theories of Social Psychology



Volume 1

Handbook of Theories of Social Psychology



Volume 1

Edited by
Paul A. M. Van Lange,
Arie W. Kruglanski,
and E. Tory Higgins



Los Angeles | London | New Delhi
Singapore | Washington DC

Introduction and Editorial arrangement © Paul A.M. Van Lange, Arie W. Kruglanski, and E. Tory Higgins 2012

Chapter 1 © Douglas T. Kenrick 2012	Chapter 14 © Norbert Schwarz 2012
Chapter 2 © Shelley E. Taylor 2012	Chapter 15 © Gün R. Semin 2012
Chapter 3 © John T. Cacioppo, Gary G. Berntson, Catherine J. Norris, and Jackie K. Gollan 2012	Chapter 16 © Robin R. Vallacher and Daniel M. Wegner 2012
Chapter 4 © E. Tory Higgins 2012	Chapter 17 © Albert Bandura 2012
Chapter 5 © Fritz Strack and Roland Deutsch 2012	Chapter 18 © Joel Cooper 2012
Chapter 6 © Yaacov Trope and Nira Liberman 2012	Chapter 19 © Jeffrey Greenberg and Jamie Arndt 2012
Chapter 7 © Bernard Weiner 2012	Chapter 20 © Edward L. Deci and Richard M. Ryan 2012
Chapter 8 © Robert S. Wyer, Jr. 2012	Chapter 21 © Icek Ajzen 2012
Chapter 9 © Chester A. Insko 2012	Chapter 22 © Jerry Suls and Ladd Wheeler 2012
Chapter 10 © Arie W. Kruglanski 2012	Chapter 23 © E. Tory Higgins 2012
Chapter 11 © Richard E. Petty and Pablo Briñol 2012	Chapter 24 © Charles S. Carver and Michael F. Scheier 2012
Chapter 12 © Shelly Chaiken and Alison Ledgerwood 2012	Chapter 25 © Peter M. Gollwitzer 2012
Chapter 13 © Susan T. Fiske 2012	

First published 2012

Apart from any fair dealing for the purposes of research or private study, or criticism or review, as permitted under the Copyright, Designs and Patents Act, 1988, this publication may be reproduced, stored or transmitted in any form, or by any means, only with the prior permission in writing of the publishers, or in the case of reprographic reproduction, in accordance with the terms of licences issued by the Copyright Licensing Agency. Enquiries concerning reproduction outside those terms should be sent to the publishers.

SAGE Publications Ltd
1 Oliver's Yard
55 City Road
London EC1Y 1SP

SAGE Publications Inc.
2455 Teller Road
Thousand Oaks, California 91320

SAGE Publications India Pvt Ltd
B 1/I 1 Mohan Cooperative Industrial Area
Mathura Road
New Delhi 110 044

SAGE Publications Asia-Pacific Pte Ltd
33 Pekin Street #02-01
Far East Square
Singapore 048763

Library of Congress Control Number: 2010936852

British Library Cataloguing in Publication data

A catalogue record for this book is available from the British Library

ISBN 978-0-85702-960-7

Typeset by Glyph International
Printed in India at Rephika Pvt Ltd
Printed on paper from sustainable resources

Contents

VOLUME 1

Notes on Editors and Contributors	ix
Preface	xix

Theories of Social Psychology: An Introduction <i>Paul A.M. Van Lange, Arie W. Kruglanski, and E. Tory Higgins</i>	1
---	---

PART I BIOLOGICAL/EVOLUTIONARY LEVEL OF ANALYSIS 9

1 Evolutionary Theory and Human Social Behavior <i>Douglas T. Kenrick</i>	11
2 Tend and Befriend Theory <i>Shelley E. Taylor</i>	32
3 The Evaluative Space Model <i>John T. Cacioppo, Gary G. Berntson, Catherine J. Norris, and Jackie K. Gollan</i>	50

PART II COGNITIVE LEVEL OF ANALYSIS 73

4 Accessibility theory <i>E. Tory Higgins</i>	75
5 A Theory of Impulse and Reflection <i>Fritz Strack and Roland Deutsch</i>	97
6 Construal Level Theory <i>Yaacov Trope and Nira Liberman</i>	118
7 An Attribution Theory of Motivation <i>Bernard Weiner</i>	135
8 A Theory of Social Information Processing <i>Robert S. Wyer, Jr.</i>	156
9 Balance-Logic Theory <i>Chester A. Insko</i>	178

10	Lay Epistemic Theory <i>Arie W. Kruglanski</i>	201
11	The Elaboration Likelihood Model <i>Richard E. Petty and Pablo Briñol</i>	224
12	A Theory of Heuristic and Systematic Information Processing <i>Shelly Chaiken and Alison Ledgerwood</i>	246
13	The Continuum Model and the Stereotype Content Model <i>Susan T. Fiske</i>	267
14	Feelings-as-Information Theory <i>Norbert Schwarz</i>	289
15	The Linguistic Category Model <i>Gün R. Semin</i>	309
16	Action Identification Theory <i>Robin R. Vallacher and Daniel M. Wegner</i>	327
17	Social Cognitive Theory <i>Albert Bandura</i>	349
PART III MOTIVATIONAL/AFFECTIVE LEVEL OF ANALYSIS		375
18	Cognitive Dissonance Theory <i>Joel Cooper</i>	377
19	Terror Management Theory <i>Jeffrey Greenberg and Jamie Arndt</i>	398
20	Self-Determination Theory <i>Edward L. Deci and Richard M. Ryan</i>	416
21	The Theory of Planned Behavior <i>Icek Ajzen</i>	438
22	Social Comparison Theory <i>Jerry Suls and Ladd Wheeler</i>	460
23	Regulatory Focus Theory <i>E. Tory Higgins</i>	483
24	A Model of Behavioral Self-Regulation <i>Charles S. Carver and Michael F. Scheier</i>	505
25	Mindset Theory of Action Phases <i>Peter M. Gollwitzer</i>	526
	Name Index	547
	Subject Index	553

VOLUME 2

Notes on Editors and Contributors	ix
Preface	xi
26 Self-Control Theory <i>Walter Mischel</i>	1
27 Self-Verification Theory <i>William B. Swann, Jr.</i>	23
28 Implicit Theories <i>Carol S. Dweck</i>	43
29 Uncertainty-Identity Theory <i>Michael A. Hogg</i>	62
30 Optimal Distinctiveness Theory <i>Marilynn B. Brewer</i>	81
31 A Cognitive-Neoassociation Theory of Aggression <i>Leonard Berkowitz</i>	99
PART IV INTERPERSONAL LEVEL OF ANALYSIS	119
32 Need-to-Belong Theory <i>Roy F. Baumeister</i>	121
33 Sociometer Theory <i>Mark R. Leary</i>	141
34 Attachment Theory <i>Phillip R. Shaver and Mario Mikulincer</i>	160
35 Shared-Reality Theory <i>Gerald Echterhoff</i>	180
36 Equity Theory in Close Relationships <i>Elaine Hatfield and Richard L. Rapson</i>	200
37 The Investment Model of Commitment Processes <i>Caryl E. Rusbult, Christopher R. Agnew, and Ximena B. Arriaga</i>	218
38 A Theory of Communal (and Exchange) Relationships <i>Margaret S. Clark and Judson R. Mills</i>	232
39 Interdependence Theory <i>Paul A.M. Van Lange and Caryl E. Rusbult</i>	251

PART V GROUP AND CULTURAL LEVEL OF ANALYSIS	273
40 A Theory of Cooperation – Competition and Beyond <i>Morton Deutsch</i>	275
41 The Focus Theory of Normative Conduct <i>Robert B. Cialdini</i>	295
42 System Justification Theory <i>John T. Jost and Jojanneke van der Toorn</i>	313
43 Justice Theory <i>Tom R. Tyler</i>	344
44 Minority Influence Theory <i>Charlan Jeanne Nemeth</i>	362
45 Social Identity Theory <i>Naomi Ellemers and S. Alexander Haslam</i>	379
46 Self-Categorization Theory <i>John C. Turner and Katherine J. Reynolds</i>	399
47 Social Dominance Theory <i>Jim Sidanius and Felicia Pratto</i>	418
48 The Common Ingroup Identity Model <i>Samuel L. Gaertner and John F. Dovidio</i>	439
49 Social Role Theory <i>Alice H. Eagly and Wendy Wood</i>	458
50 Social Representation Theory <i>Patrick Rateau, Pascal Moliner, Christian Guimelli, and Jean-Claude Abric</i>	477
51 A Theory of Individualism and Collectivism <i>Harry C. Triandis and Michele J. Gelfand</i>	498
Name Index	521
Subject Index	528

Notes on Editors and Contributors

Icek Ajzen is professor of social psychology at the University of Massachusetts Amherst (USA). He received his PhD degree from the University of Illinois and has, in past years, been a visiting professor at Tel Aviv University and the Hebrew University in Jerusalem. Dr. Ajzen has conducted research on the attitude–behavior relation and is best known for his theory of planned behavior, a widely applied model of behavioral prediction. He has published several books and numerous scholarly articles in professional journals, and he has been recognized as a highly cited researcher in ISI Web of Knowledge. Among his books are *Attitudes, Personality, and Behavior* (Open University Press, 2005) and coauthored with Professor Martin Fishbein, *Predicting and Changing Behavior: The Reasoned Action Approach* (Psychology Press, 2010).

Jamie Arndt is professor of psychology and director of the Social/Personality Program at the University of Missouri. He received his PhD from University of Arizona. His research interests focus on the motivational and existential dynamics of the human condition and how this interfaces with various forms of social and health behavior. These interests have led him to study the self, psychological defense, and unconscious motivation, among other topics. His applications of these ideas to health-related behavior have been funded by the National Cancer Institute. He has published articles regularly in a variety of journals, including *Psychological Bulletin*, *Psychological Review*, *Psychological Science*, *Journal of Experimental Psychology: General*, *Journal of Personality and Social Psychology*, *Health Psychology*, *Personality and Social Psychology Bulletin*, and the *Journal of Experimental Social Psychology*.

Albert Bandura is professor of psychology at Stanford University. The major focus of his work centers on the mechanisms of human agency through which people exercise some measure of influence over personal and social change. Human agency is exercised individually over what is personally controllable, in proxy form by influencing others to act on one's behalf and collectively by working together (see attachment). His book, *Social Foundations of Thought and Action: A Social Cognitive Theory* (Prentice-Hall, 1986), provides the conceptual framework of his theory and analyzes the large body of knowledge bearing on it. His most recent book, *Self-Efficacy: The Exercise of Control* (Worth, 1997), presents belief in one's efficacy to produce effects by one's actions as an important vehicle of human agency. His diverse programs of research blend his theoretical interests with an abiding concern for the use of our knowledge for human enlightenment and betterment.

Gary G. Berntson is a professor of psychology at the Ohio State University. He studied biology and psychology at the University of Minnesota, and was awarded a PhD (psychobiology and life sciences) in 1971. He is currently the president of the Society for Psychophysiological Research and is on the editorial board for the *International Journal of Psychophysiology*. His interest is the behavioral neurosciences broadly, and the multilevel analysis of neurobiological substrates of stress and emotion in particular. In addition, his research focuses on the psychophysiology of autonomic control and its role in health and disease. He has published over 200 papers and several books, including the *Handbook of Neuroscience for the Behavioral Sciences* (Wiley, 2009) and the *Handbook of Psychophysiology* (Cambridge University Press, 2000).

Pablo Briñol is associate professor of social psychology at the Universidad Autónoma de Madrid (Spain), and a visiting scholar at Ohio State University. His research interests focus on the study of the psychological mechanisms underlying attitudes and persuasion. These fundamental processes range from the least thoughtful automatic processes (e.g., self-perception) to the most thoughtful metacognitive processes (e.g., thought validation). Dr. Briñol has published several books and book chapters in the domain of persuasion. His research has appeared in leading journals in the field, including *Psychological Bulletin*, *Psychological Science*, *Journal of Personality and Social Psychology*, *Advances in Experimental Social Psychology*, *Personality and Social Psychology Bulletin*, and the *Journal of Experimental Social Psychology*.

John T. Cacioppo is the Tiffany and Margaret Blake Distinguished Service Professor and Director of the Center for Cognitive and Social Neuroscience at the University of Chicago. He also served as the founding director of the Arete Initiative of the Office of the Vice President for Research and National Laboratories at the University of Chicago. Cacioppo takes a social neuroscientific approach in his investigations of affect and social behavior, where he and his colleagues utilize a variety of methods, including functional magnetic resonance imaging (fMRI), standard and high-density electroencephalography and event-related brain potentials, psychophysiological assessments, neuroendocrine and immune assays, and quantitative genetics. Among his books are the *Handbook of Neuroscience for the Behavioral Sciences* (Wiley, 2009), *Loneliness: Human Nature and the Need for Social Connection* (W.W. Norton, 2008), *Social Neuroscience: People Thinking About Thinking People* (MIT Press, 2006), *Foundations in Social Neuroscience* (MIT Press, 2002), *Emotional Contagion* (Cambridge University Press, 1994), *Communication and Persuasion: Central and Peripheral Routes to Attitude Change* (Springer, 1986), and *Attitudes and Persuasion: Classic and Contemporary Approaches* (Westview Press, 1981).

Charles S. Carver is distinguished professor of psychology at the University of Miami. He has worked on a wide range of topics in personality, social, health, and clinical psychology. These topics include basic concepts regarding behavioral self-regulation, optimism, and the origins and functions of affect. He has also worked on applied topics such as vulnerability factors in both depression and mania. He has received awards for outstanding professional contributions from Division 38 (Health Psychology) and Division 8 (Personality and Social Psychology) of the American Psychological Association. He has been an editor of the *Journal of Personality and Social Psychology*, and currently is associate editor of *Psychological Review*. He and his longtime collaborator Michael F. Scheier are authors of *On the Self-regulation of Behavior*

(Cambridge University Press, 2001) and seven editions of *Perspectives on Personality* (6th ed., Allyn & Bacon, 2007). His current work focuses on genetic correlates of personality and psychopathology.

Shelly Chaiken received her PhD in social psychology in 1978 from the University of Massachusetts at Amherst, and she has since held professorial appointments at New York University, University of Toronto, and Vanderbilt University. Her research has centered on attitudes, persuasion, and social cognition, and she was awarded the Society for Experimental Social Psychology's Scientific Impact Award in 2009 for her influential work on dual process theories in attitudes. She is now retired and lives in Berkeley.

Joel Cooper is professor of psychology and former chair of the department of psychology at Princeton University. His research has focused on attitudes and attitude change, with particular emphasis on cognitive dissonance. Cooper is the author of *Cognitive Dissonance: 50 Years of a Classic Theory* (Sage, 2007) and co-author of *Gender and Computers: Understanding the Digital Divide* (Lawrence Erlbaum, 2003). He is also co-editor of the *Sage Handbook of Social Psychology* (Sage, 2003; 2007). A former chair of the executive committee of the Society of Experimental Social Psychology, Cooper is currently the editor-in-chief of the *Journal of Experimental Social Psychology*.

Edward L. Deci is Helen F. and Fred H. Gowen Professor in the Social Sciences at the University of Rochester. He holds a psychology PhD from Carnegie-Mellon University, studied at the University of Pennsylvania, the University of London, and Hamilton College, and was a Postdoctoral fellow at Stanford. For more than 40 years he has done research on human motivation, much of it in collaboration with Richard M. Ryan. Deci has published ten books, including: *Intrinsic Motivation and Self-Determination in Human Behavior* (with Ryan; Plenum Press, 1985). A grantee of the National Institutes of Health, the National Science Foundation, and the Institute of Education Sciences, he has lectured and consulted for universities, organizations, and governmental agencies in 23 countries on five continents.

Roland Deutsch is professor of social psychology at the Technical University Dresden. His research is focused on social cognition and motivation. Current projects address processes of automatic evaluation, indirect attitude measures, effects of deprivation, and approach/avoidance motivation. He is associate editor of the journal *Social Psychology* and has been awarded the Theoretical Innovation Prize of the *Society of Personality and Social Psychology* (jointly with Fritz Strack).

Susan T. Fiske is Eugene Higgins Professor of Psychology, Princeton University. She received her PhD from Harvard University and honorary doctorates from Université Catholique de Louvain-la-Neuve, Belgium; Universiteit Leiden, Netherlands. Author of over 250 publications, she investigates social cognition, especially cognitive stereotypes and emotional prejudices, at cultural, interpersonal, and neural levels. She recently edited *Beyond Common Sense: Psychological Science in the Courtroom* (Wiley/Blackwell, 2007) and the *Handbook of Social Psychology* (5th edition, Wiley, 2010). She wrote *Social Beings: Core Motives in Social Psychology* (Wiley, 2003) and *Social Cognition: From Brains to Culture* (McGraw-Hill, 2007). Her forthcoming Russell-Sage-Foundation book is *Envy Up and Scorn Down: How Status*

Divides Us (2011). Recently, she won a Guggenheim, American Psychological Association's Distinguished Scientific Contribution Award, and Association for Psychological Science's William James Award. She is an American Academy of Arts and Sciences Fellow.

Jackie K. Gollan is assistant professor of psychiatry and behavioral Sciences at Northwestern University's Feinberg School of Medicine. She served as the Director of the Northwestern University Regional Center for the Depression Treatment Network (DTN), as part of a collaborative network of mood disorder researchers from 15 academic institutions. She directs a translational affective science laboratory, Stress and Depression Laboratory (SADLAB), and with the National Institute of Mental Health and foundation funding, conducts clinical research studies on substrates of disease onset and progression, biological and behavioral predictors of treatment response, new treatment development, and clinical efficacy testing. Her published work is on affective information processing, behavioral activation treatment, and predictors of treatment response.

Peter M. Gollwitzer received his BA and MA in psychology in Germany, and his PhD at the University of Texas at Austin. He headed the Research Unit "Intention and Action" at the Max Planck Institute for Psychological Research in Munich, Germany. Since 1993 he has been a Full Professor of Psychology at the University of Konstanz and since 1999 at New York University. Gollwitzer has developed various models of action control: the theory of symbolic self-completion (with Robert A. Wicklund), the mindset model of action phases (with Heinz Heckhausen), the auto-motive model of automatic goal striving (with John A. Bargh), and most recently the theory of intentional action control, which distinguishes implementation intentions from goal intentions, and describes how if-then planning (i.e., the formation of implementation intentions) automates action control. Gollwitzer is the editor of *Psychology of Action: Linking Cognition and Motivation to Behavior* (with John A. Bargh; Guilford Press, 1996) and the *Oxford Handbook of Human Action* (with Ezequiel Morsella and John A. Bargh; Oxford University Press, 2008).

Jeffrey Greenberg is professor of psychology at the University of Arizona. He has published over 200 articles and chapters, primarily focused on understanding self-esteem, prejudice, and intergroup conflict. In collaboration with Sheldon Solomon and Tom Pyszczynski, he developed terror management theory, a broad theoretical framework which explores the role of existential fears in diverse aspects of human behavior. He has received numerous grants for his research from the National Science Foundation and the National Institute of Aging and has received the International Society for Self and Identity Lifetime Career Award. He is coauthor of two books, including *In the Wake of 9/11: The Psychology of Terror* (American Psychological Association, 2003), and coeditor of two books, including the *Handbook of Experimental Existential Psychology* (Guilford Press, 2004).

E. Tory Higgins is the Stanley Schachter Professor of Psychology, Professor of Business, and Director of the Motivation Science Center at Columbia (where he also received his PhD in 1973). He has received a MERIT Award from the National Institute of Mental Health, the Thomas M. Ostrom Award in Social Cognition, the Donald T. Campbell Award for Outstanding Contributions to Social Psychology (Society of Personality and Social Psychology), and the Lifetime Contribution Award from the International Society for Self and Identity. He has also

received the Distinguished Scientist Award from the Society of Experimental Social Psychology, the William James Fellow Award for Distinguished Achievements in Psychological Science (from the American Psychological Society), and the American Psychological Association Award for Distinguished Scientific Contributions. He is a Fellow of the American Academy of Arts and Sciences. He is also a recipient of Columbia's Presidential Award for Outstanding Teaching.

Chester A. Insko is a professor of psychology at the University of North Carolina at Chapel Hill. He received his AB in philosophy from the University of California at Berkeley in 1957, his MA in psychology from Boston University in 1958 and his PhD in Psychology from the University of California at Berkeley in 1963. He has spent time as a visiting professor at the Department of Psychology at the University of California at Berkeley, and at the University of Tilburg in Holland. He is a past associate editor of the *Journal of Experimental Social Psychology* and past editor of the Interpersonal Relations and Group Processes section of the *Journal of Personality and Social Psychology*. For many years most of his research was on attitude change, influence, and interpersonal attraction. While some of that research continues, more recently his research has focused on interindividual-intergroup discontinuity—the tendency in some social contexts for relations between groups to be more conflict prone than relations between individuals.

Douglas T. Kenrick is professor of psychology at Arizona State University. His research attempts to integrate ideas from evolutionary biology, cognitive science, and dynamical systems theory. That work has been funded by the National Science Foundation and the National Institute of Mental Health and been reported in *Behavioral and Brain Sciences*, *Psychological Review*, *Perspectives on Psychological Science*, *Personality and Social Psychology Review*, and *Evolution and Human Behavior*. Kenrick has edited several books on evolutionary psychology, and contributed chapters to the *Handbook of Social Psychology* and *Handbook of Evolutionary Psychology*. He is author of *Sex, Murder, and the Meaning of Life: A Psychologist Investigates How Evolution, Cognition, and Complexity Illuminate Human Nature* (Basic Books, 2011) and (with Steven Neuberg and Robert Cialdini) of *Social Psychology: Goals in Interaction* (5th ed., Allyn & Bacon, 2010).

Arie W. Kruglanski is a distinguished university professor at the University of Maryland, College Park. He is recipient of the National Institute of Mental Health Research Scientist Award, the Senior Humboldt Award, the Donald Campbell Award for Outstanding Contributions to Social Psychology from the Society for Personality and Social Psychology, The University of Maryland Regents Award for Scholarship and Creativity, and the Distinguished Scientific Contribution Award from the Society of Experimental Social Psychology, and is recipient of the Regesz Chair at the University of Amsterdam. He was Fellow at the Center for Advanced Studies in the Behavioral Sciences, and is Fellow of the American Psychological Association and the American Psychological Society. He has served as editor of the *Journal of Personality and Social Psychology: Attitudes and Social Cognition*, editor of the *Personality and Social Psychology Bulletin*, and associate editor of the *American Psychologist*. His interests have been in the domains of human judgment and decision making, the motivation-cognition interface, group and intergroup processes, and the psychology of human goals. His work has been disseminated in over 200 articles,

chapters and books and has been continuously supported by grants from the National Science Foundation, the National Institute of Mental Health, Deutsche Forschungs Gemeineschaft, the Ford Foundation, and the Israeli Academy of Science. He has recently served as member of the National Academy of Science panels on counterterrorism, and educational paradigms in homeland security. Kruglanski is now a co-director of START (National Center for the Study of Terrorism and the Response to Terrorism), at the University of Maryland.

Paul A.M. Van Lange is professor of social psychology and chair of the department of social and organizational psychology at the VU University at Amsterdam, the Netherlands. Most of his research on human cooperation and trust is grounded in interdependence theory, through which he seeks to understand the functions of forgiveness, generosity, empathy, competition, and general beliefs of human nature in various situations. Van Lange has coauthored the *Atlas of Interpersonal Situations* (Cambridge University Press, 2003), edited *Bridging Social Psychology* (Lawrence Erlbaum, 2006), and served as an associate editor for various journals, including the *Journal of Personality and Social Psychology*. He has been a Director of the Kurt Lewin Institute and currently serves as Member and President of the Executive Committee of the Society of Experimental Social Psychology.

Alison Ledgerwood is an assistant professor in the psychology department at the University of California, Davis. She received her PhD in 2008 from New York University. Her research focuses on understanding when and why evaluations shift in response to the social context, and her work has been published in the *Journal of Personality and Social Psychology*, the *Journal of Experimental Social Psychology*, *Psychological Science*, *Social Cognition*, and *Advances in Experimental Social Psychology*.

Catherine J. Norris is an assistant professor in the department of psychological and brain sciences at Dartmouth College. She received her BA, MA, and PhD all from the University of Chicago. Her research uses neuroscience and psychophysiological measures to understand emotional and evaluative processes. She is interested in how individual differences in responses to emotional stimuli and events affect mental and physical health, as well as the role that the social context plays in these relationships. She has recently been published in the *Journal of Cognitive Neuroscience*, *Emotion*, *Psychophysiology*, and *Biological Psychology*.

Richard E. Petty is a distinguished university professor of psychology at Ohio State University. He received his BA from the University of Virginia and his PhD from Ohio State. Petty's research focuses on the situational and individual factors responsible for changes in attitudes and behaviors with a current emphasis on both implicit and metacognitive factors. He has published eight books and over 275 articles and chapters. Petty's honors include receipt of the Scientific Impact Award from the Society of Experimental Social Psychology and Distinguished Scientific Contribution Awards from the Societies for Personality and Social Psychology and Consumer Psychology. He is past editor of the *Personality and Social Psychology Bulletin* and former President of the Society for Personality and Social Psychology and the Midwestern Psychological Association.

Richard M. Ryan is professor of psychology, psychiatry, and education at the University of Rochester, and the director of clinical training. He is a widely published researcher and theorist

in the areas of human motivation, personality, and wellbeing, and codeveloper (with Edward L. Deci) of self-determination theory. Ryan, who has lectured in more than 60 universities around the globe, is a Fellow of the American Psychological Association, the American Educational Research Association, the Society for Personality and Social Psychology, and an Honorary Member of the German Psychological Society (DGP). He is editor-in-chief of the journal *Motivation & Emotion* and has been a James McKeen Cattell Fellow and a visiting scientist at the Max Planck Institute for Human Development.

Michael F. Scheier is professor and head of psychology at Carnegie Mellon University. His research falls at the intersection of personality, social, and health psychology. His current research focuses on the effects of dispositional optimism on psychological and physical well-being, and on the health benefits of goal adjustment when confronting adversity. He is a fellow in Divisions 8 and 38 of the American Psychological Association (APA), and in the Society of Behavioral Medicine. He has received awards for Outstanding Contributions to Health Psychology and the Donald T. Campbell Award for distinguished lifetime contributions to social psychology (offered by APA Divisions 38 and 8, respectively). He has served Division 38 (Health Psychology) in the past as chair of the Nominations and Election Committee, as associate editor of *Health Psychology*, and as president.

Nira Liberman is professor of psychology at Tel Aviv University. Her doctoral degree is from Tel Aviv University. As one of the authors of Construal Level Theory, much of her research focuses on psychological distance – how it affects and is being affected by mental construal, prediction, decision making, persuasion, performance, interpersonal relations, and more. She also made contributions to other areas of theory and research, all of which come under the general umbrella of the interface between motivation and cognition: An attributional theory of thought suppression, the question of how goals affect construct accessibility, how regulatory foci affect decision making. Her research has been funded by the Israeli Science Foundation and the US-Israel Binational Science Foundation. She has served as an Associate Editor for *Journal of Personality and Social Psychology*.

Norbert Schwarz is the Charles Horton Cooley Collegiate Professor of psychology at the University of Michigan, where he also holds appointments at the Ross School of Business and the Institute for Social Research. His research focuses on the socially situated and embodied nature of cognition, the interplay of feeling and thinking, and the implications of basic cognitive and communicative processes for public opinion, consumer behavior, and social science research.

Gün R. Semin is an academy professor with The Netherlands Royal Academy of Arts and Sciences at Utrecht University and also professor of psychology at Koç University, Istanbul. He has previously served as professor and chair of the department of social psychology at the Free University, where he also was the research director of the Faculty of Psychology and Education. He was also the founding scientific director of the Kurt Lewin Institute, the inter-university graduate school in social psychology and its applications and was chair of the Standing Committee on Social Sciences (SCSS) of the European Science Foundation. Currently, he is chair of Psychology Panel of the Portuguese National Science Foundation (FCT – Foundation for Science and Technology), Lisbon, Portugal, member of the Expert Committee of the Deutsche Forschungsgemeinschap and Science Council for the Excellence

initiative by the German Government (since 2007), and Member of the ERC Advanced Grant evaluation panel. He is also a fellow of Association of Psychological Science (APS) and the Society of Experimental Social Psychology (SESP), and chair of the Internationalization Committee of APS and secretary to the APS Board.

Fritz Strack is professor and chair in psychology at the University of Würzburg. He was editor of the *European Journal of Social Psychology* and President of the European Association of Social Psychology (EASP). His research focuses on issues in the domain of social cognition. In particular, he is interested in the interaction between subjective experience and cognition in their determination of judgments and behavior. He has been awarded the Wilhelm-Wundt Medal of the German Psychological Society (jointly with Norbert Schwarz) and the Theoretical Innovation Prize of the Society of Personality and Social Psychology (jointly with Roland Deutsch).

Jerry Suls is professor of psychology and collegiate fellow at the University of Iowa. His main interests have been social comparison, the role of psychological factors in physical wellbeing, and biases concerning social norms, such as the above-average effect and false distinctiveness. He has edited the *Handbook of Social Comparison* (with Ladd Wheeler; Springer, 2000) and several other volumes. He was a past editor of *Personality and Social Psychology Bulletin* and is the current editor-in-chief of *Social and Personality Psychology Compass*.

Shelley E. Taylor is a distinguished professor of psychology at UCLA. She studies social relationships and how they protect against stress. Her tend-and-befriend model builds on the fact that, in response to stress, people come together with others for joint protection of self and offspring. Professor Taylor also studies self-regulation, stress, and coping and explores the skills that people develop and use for anticipating stressful events and for minimizing their adverse effects when they do occur. Finally, Taylor studies how positive beliefs are protective of mental and physical health. She shows that optimism, self-enhancement, a perception of control, and social support can protect against threats or traumas, not only psychologically but also in terms of physical health. She publishes in both biological and psychological journals.

Yaacov Trope is a professor of psychology at New York University. He received his PhD from the University of Michigan in 1974. His general areas of interest are social cognition, motivation, and self-regulation. His current research emphasizes self control processes, social judgment, and the cognitive, motivational, and social processes that enable people to focus on the “here-and-now” and those that enable them to transcend the “here-and-now” and traverse psychological distance. He is editor of two books: *Dual-process Theories in Social Psychology* (Guilford Press, 1998) and *Self Control in Society, Mind, and Brain* (Oxford University Press, 2010).

Robin R. Vallacher is professor of psychology at Florida Atlantic University, and research associate in the Center for Complex Systems at University of Warsaw, and the Advanced Consortium on Cooperation, Conflict, and Complexity at Columbia University. His research interests include action identification, self-concept structure and dynamics, social coordination, conflict, and social change. In recent years, he and his colleagues have adapted principles and

methods of complexity science to reframe and investigate these and other topics. His books include *Implicit Psychology* (Oxford University Press, 1977) and *A Theory of Action Identification* (Psychology Press, 1985), both with Daniel Wegner, and *Dynamical Systems in Social Psychology* (Academic Press, 1994) and *Dynamical Social Psychology* (Guilford Press, 1998), both with Andrzej Nowak.

Daniel M. Wegner is professor of psychology at Harvard University. A PhD of Michigan State University (1974), he studies thought suppression (how we keep unwanted thoughts out of mind), transactive memory (how we remember things cooperatively with others), action identification (what we think we are doing), mind perception (how we gauge whether entities have minds), and apparent mental causation (what gives us the sense that we consciously will our actions). His books include *Implicit Psychology* (Oxford University Press, 1977) and *A Theory of Action Identification* (Psychology Press, 1985), both with R.R. Vallacher, *White Bears and Other Unwanted Thoughts* (Guilford Press, 1989), and *The Illusion of Conscious Will* (Bradford Books, 2002).

Bernard Weiner is currently a distinguished professor of psychology and has been at the University of California, Los Angeles since 1965. He received his undergraduate degree from the University of Chicago and his PhD from the University of Michigan in 1963. He has written, coauthored, or edited 16 books, including *Judgments of Responsibility* (Guilford Press, 1995), and *Social Motivation, Justice, and the Moral Emotions* (Laurence Erlbaum, 2006), as well as published more than 200 articles. He has been awarded the Donald Campbell Research Award and the Edward L. Thorndike Lifetime Achievement Award from the American Psychological Association, the Distinguished Scientist Award from the Society of Experimental Social Psychology, and the Palmer Johnson Publication Award from the American Educational Research Association. In addition to a Distinguished Teaching Award, he holds honorary degrees from the University of Bielefeld, Germany; the Turku University, Finland; and the University of Manitoba, Canada.

Ladd Wheeler is currently a professor at the Macquarie University in Sydney, Australia. He was a longstanding member of the faculty of the University of Rochester. His major substantive interests have been social comparison, behavioral contagion, physical attractiveness, and social interaction. He is a pioneer in the use of diary methods, such as the Rochester Interaction Record and the Social Comparison Record. He is a past president of the Society for Personality and Social Psychology and the founding editor of the *Review of Personality and Social Psychology*.

Robert S. Wyer, Jr. is a visiting professor at the Chinese University of Hong Kong and professor (Emeritus) at the University of Illinois, Urbana-Champaign. His research interests include social comprehension, memory and judgment, affect, consumer information processing. Dr. Wyer is the author or coauthor of four books, the most recent being *Social Comprehension and Judgment* (Laurence Erlbaum, 2004) and an editor of several others including the *Handbook of Social Cognition* (2nd ed., Psychology Press, 1994), the *Advances in Social Cognition* series (Psychology Press), and *Understanding Culture: Theory and Research and Application* (Psychology Press, 2009). Dr. Wyer is a former editor of the *Journal of Experimental Social Psychology* and, more recently, the *Journal of Consumer Psychology*.

He is a recipient of the Alexander von Humboldt Special Research Prize for Distinguished Scientists, the first annual Thomas M. Ostrom Award for Distinguished Contributions to Person Memory and Social Cognition, and the Distinguished Scientific Contribution Awards given by the Society for Experimental Social Psychology (2009) and the Society for Consumer Psychology (2011).

Preface

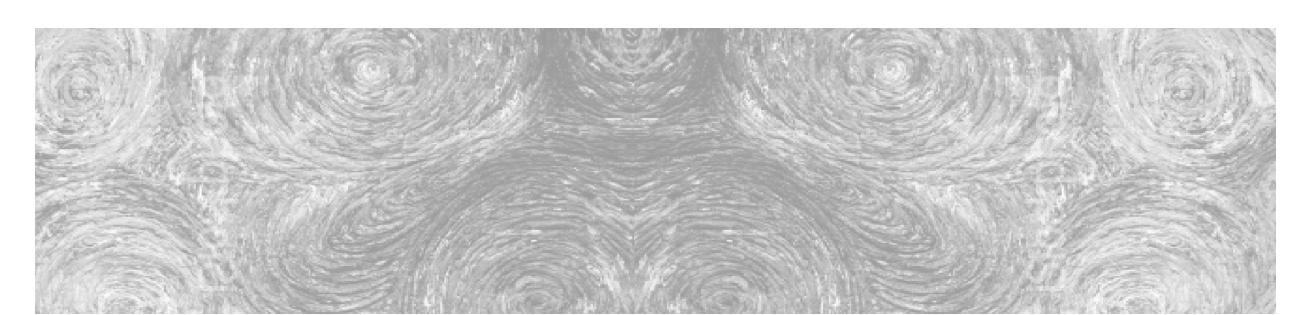
Ideas make the world go around – especially good ideas and especially in science. Indeed, science is all about ideas and their implementation in empirical research. This is true for the science of social psychology as well. Indisputably, the quintessential carriers of scientific ideas are theories. It is theories that get to the underlying essences of phenomena and trace their implications for myriads of concrete situations. It is theories that pull the strands of seemingly disparate occurrences and tie them into coherent systems guided by common principles. Good theories are not just practical, as Lewin noted; they are essential to the scientific enterprise. It is, therefore, hardly surprising that from its early beginnings social psychological research has been guided by theories of various kinds. Numerous theoretical frameworks have been added by creative thinkers in the course of time. By now, the field of social psychology is rich in theoretical contributions in its many domains of endeavor. Some social psychological theories have been around for a long time, others for little more than a decade. Some have been tested, revised, and extended, while others have remained in their original form and continued to inspire research on the force of their timeless insights. Some theories have intriguingly morphed into other theories, others remained pristinely faithful to their initial version. Some theories have been wonderfully elaborated and articulated. Others have been adumbrated in vague outline, representing work in progress or diamonds in the rough. In this volume, we are interested in all such theories not only because they provide a comprehensive overview of the theories in social psychology, but also because we felt it is important that authors share with the readers the process of theory construction, development, and nurturance that serves such an important function for science. Here is why.

The process of theorizing, and the skills of theory construction, have been shrouded in a cloak of mystery in our field. They are rarely taught in graduate programs in social psychology, nor do they constitute a recognized and trusted tool in the kit of young researchers. A major purpose of the present project was to demystify the process of theorizing and expose its hidden underbelly and intricate entrails. Indeed, chapters by our contributors reveal how serendipity born of personal circumstances often determines the course that one's theory construction would take; how theory development often requires tenacity, persistence, patience, and “blood, sweat and tears.” Another purpose of the book was to illustrate how the work of theory construction is indispensable to scientific development, and how important and gratifying it can be to those who manage to stay on the course of constructing and testing their theory.

Our own conviction, stemming from our earlier work, and presented in the introductory chapter, has been that theories should be guided by the regulatory ideas of truth, abstraction, progress, and applicability. This notion served as the basis of a research grant, “Social Psychology: Bridging Theory and Application in Society,” (NWO. grant, nr. 400-07-710), awarded by the Netherlands Organization for Scientific Research, which gave the first editor extra time to devote to this Handbook. Because of the immensity of the project and common

interest in theory, he invited the second and third editors to join in, and they enthusiastically agreed. After initial discussion, we concurred that this volume should carry a unique mission: illuminating theory construction from the inside out. Accordingly, the instructions we gave to our contributors were explicit and precise. We asked authors not only to give an overview of their theory or model, but also touch on three essential aspects: (1) a personalized history of the theory's beginnings and development over time as recounted by the theoretician; (2) the theory's place in the intellectual space in a given domain (i.e., the contribution it makes to the history of ideas on its topic); and (3) the theory's relevance to real-world concerns (i.e., its potential contribution to solving real-world problems). Inevitably, the various chapters in this volume differed in their primary focus, and in the emphasis accorded to each of these aspects. But overall, these three foci are amply represented across the chapters. Of greatest importance, they tell a fascinating tale documenting the challenges, adversities, and joys that theory construction brings its practitioners, and the rich conceptual endowment that it brings our discipline.

The Editors



Theories of Social Psychology: An Introduction

Paul A.M. Van Lange, Arie W. Kruglanski,
and E. Tory Higgins

The advancement of theory is a key goal of science. Ideally, theory helps us explain particular events and phenomena and find the underlying *truth, beyond surface appearances*. Theory helps us see the coherent structures in seemingly chaotic environments and make inroads into previously uncharted domains, thus affording progress in the way we understand the world around us. Because it elucidates the causal mechanisms that produce manifest effects, theory points to ways of intervening in phenomena and changing the course of events; hence, theory is of essential pragmatic value and constitutes an indispensable tool for application.

This book is about theories of social psychology, a field broadly defined by reciprocal influences between individuals and their social environments – that is, other people, whether present, imagined, or implied (Allport, 1954; cf. Deutsch and Krauss, 1965: 1; Jones and Gerard, 1967: 1; Shaver, 1977: 4; Van Lange, 2006: 13). Consistent with this definition, social psychology encompasses a wide range of domains addressing the manifold targets of influence (including individuals' thoughts, feelings, and actions), as well as different kinds of

influence (conscious and unconscious; implicit and explicit). Because of its subject matter, social psychology is critically relevant to a vast multitude of events in the social realm spanning the gamut from intraindividual judgments and decisions to problems in interpersonal relations, group and intergroup dynamics all the way to the impacts of culture and intercultural encounters. Indeed, social psychologists have been carrying out important conceptual and empirical work in all these domains of endeavor.

Given the scope and relevance of social psychology, and the importance of theory for a field of science, it is somewhat surprising that considerable time has passed since a volume was published specifically devoted to social psychological theories. The latest effort of this kind was the 1980 primer by West and Wicklund, which complemented the 1970 book by Shaw and Costanzo (revised in 1982), and the earlier classic work by Deutsch and Krauss (1965). Yet considerable theoretical work in social psychology has been accomplished in the three decades that followed. Numerous conceptual frameworks have been constructed at the various levels of social psychological analysis. Moreover,

theoretical debates and commentaries have appeared on approaches to social theory construction (e.g., Higgins, 2004; Kelley, 2000; Kruglanski and Higgins, 2004), the state of social/personality theory (Kruglanski, 2001; Mischel, 2004), and the theoretical bridging of social psychological analyses with other fields and scientific disciplines (Kruglanski, 2006; Van Lange, 2006, 2007). The *Handbook of Theories of Social Psychology* reflects these discussions and provides a comprehensive perspective on major theoretical developments since the time that social psychology was still in its infancy. As such, it depicts our discipline's journey from childhood to adulthood, spanning more than half a century of growth and conceptual progress in scientific understanding of the social world.

THEORY: REGULATORY IDEALS

It is not easy to precisely define what a theory is – and more importantly, what would qualify as theory and what would not. Like earlier writers, we suggest that a theory may minimally be defined as a set of interrelated propositions (or principles) concerning a phenomenon or set of phenomena (Mandler and Kessen, 1959: 159; Shaw and Costanzo, 1982: 4). Clearly, theories may differ in their generality, precision, and origins. Our approach in this volume has been inclusive rather than exclusive, and based on a minimalist definition of the concept of “theory” to guide our final decision. We reasoned that a conceptual framework that inspired significant empirical research is worthy of inclusion even if it is incomplete or otherwise imperfect from a “purist” metatheoretical perspective. Besides, social psychological theories tend to be “middle range” (Merton, 1949: 5) anyway and hence “intermediate to minor working hypotheses” rather than grand theoretical edifices. Social psychology is rich in such mid-range theories, each representing a “work in progress”

rather than being an etched-in-stone master scheme (see Pinker, 2002: 241; Van Lange, 2006: 8).

Whatever definition of a theory one opts for, it is important to ascertain what constitutes a good theory. Though numerous constructs have been advanced to outline various qualifications, standards, and criteria for theoretical “goodness,” there is a fair amount of consensus at least concerning these matters: theories are believed to be better if they have greater explanatory power; are more suitable to empirical tests and modeling; are more “logical,” in the sense of coherence and internal consistency; are capable of explaining more (phenomena) with less (by way of assumptions) reflecting the criterion of parsimony or Occam’s Razor; and, critically, inspire new research that yields empirical discoveries (see, for example, Fiske, 2004; Higgins, 2004). For the present purposes, we focus on a framework that is characterized by four broad regulatory ideals, namely: truth, abstraction, progress, and applicability (see also Kruglanski, 2006). We discuss each of these in turn.

Ideal 1: Truth

A theory should be dealing with the truth; it should separate fact from fiction; it should establish what’s real rather than what’s imaginary. Although an inaccurate, fictitious theory can serve important functions (such as serving an heuristic function to stimulate further research), it should be clear that theories seek to pursue “the truth and nothing but the truth.” This is what hypothesis testing is all about. The entire logic of experimental design is to eliminate (or prove invalid) possible alternative interpretations of empirical facts. Critical experiments are designed to set apart competing theories and decide which one appears more valid, and is better supported by the available evidence, than its competitors. And holding onto a theory known to be false simply because it had heuristic or communicative value (i.e., it was

easy to convey to others) is self-contradictory because subscribing to a theory is tantamount to believing it to be true.

Yet, as a regulatory ideal, truth can be striven for but never securely attained. No theory, however successful, is secure, for alternative accounts of the same data are always possible in the future even if they may not be apparent in the present. You can disprove a theory but never prove it. You can only find support for a theory based on what is currently known. Moreover, the empirical “facts” are far from absolute. As noted by Popper (1959: 111) the empirical basis of science is both conjectural and fallible:

The empirical basis of objective science has nothing “absolute” about it. Science does not rest upon rock bottom. The bold structure of its theories rises, as it were, above a swamp. It is like a building erected upon piles. The piles are driven down from above into the swamp; and when we cease our attempt to drive our piles into a deeper layer, it is not because we have reached firm ground. We simply stop when we are satisfied that they are firm enough to carry the structure, at least for the time being.

Thus, although a primary ideal of science, the pursuit of truth is a mission that can never be completely accomplished.

Ideal 2: Abstraction

A theory should be the result of abstraction, in that the particulars (e.g., phenomena, events) need to be described in terms of the general (concepts, assumptions, principles). While a particular phenomenon may be interesting in and of itself, one needs a theory to understand the psychological principles that underlie the phenomenon – the same principles that underlie other seemingly disparate phenomena as well. It is the higher level of aggregation that a theory should pursue, to transcend particular observations and link them at a deeper (i.e., more abstract) level to other observations. Thus, theories focus on the heart of the matter in terms of understanding and insight, as it deals with

essential causal mechanisms underlying observed effects.

Ideal 3: Progress

Any new theory is expected to make a contribution beyond what was previously known; it should improve or expand our explication of a given realm of phenomena representing the ideal of progress. It should replace myths by wisdom, and it should add truth to existing knowledge enlarging the scope of our understanding. Ideally then, newer theories relate to and build upon past theories, replacing inaccurate with accurate principles, or complementing a predecessor theory with new principles that had not previously been identified. Science is unlikely to progress if theories are not subject to refinement through a process of sharpening and empirical testing. In this sense, the principle of progress is closely linked with the principle of truth, as the theoretical refinements and modifications are in the service of ever-greater validity and precision. Because truth is an important regulatory (though largely unattainable) goal of a theory, a theory is often subject to refinement and precision; for example, by outlining the conditions in which the hypotheses derived from the theory should be supported. Also, a theory often inspires new ways of thinking, because it serves (implicitly, at least) as a tool for theorists and researchers to see connections and relationships that would not have been evident on the basis of data alone (cf. Shaw and Costanzo, 1982). Finally, a theory is often an inspiration for new research questions, along with new tools, methodologies, and paradigms (Fiedler, 2004; Fiske, 2004). As such, theories function both as bridges to the past (the past findings a theory accounts for) and the future (future research and findings inspired by the theory). Because the implications of a theory inspire new predictions that in turn inspire new research to test them, a theory is the driving force behind new empirical discoveries about what the

world is and how it works – progress from new discoveries (Higgins, 2004).

“Olympus of pure thought” and explore the possible contribution of their ideas to the solution of the multitude of real-world problems to which they pertain.

Ideal 4: Applicability

Ideally, a psychological theory should speak to many events and issues in everyday life. It should be applicable to real-world concerns and afford interventions aimed to alter the course of events in desirable ways. As Edward E. Jones (1986: 100) aptly remarked, “The future of social psychology is assured not only by the vital importance of its subject matter but also by its unique conceptual and methodological strengths that permit the identification of underlying processes in *everyday social life*.” [italics added]. Just as scientific progress is closely linked to the quest for truth, a theory’s applicability is closely linked to the precept of abstraction. In other words, the more abstract the theory, the greater its empirical content (Popper, 1959) and the broader the range of situations to which it applies. Of course, theoretical breadth in and of itself is not tantamount to application, and an appreciable measure of ingenuity is needed to translate a theory’s implications into specific procedures and interventions of practical value.

In fact, despite the intimate relation between theory and application, the two have been often juxtaposed with each other, and viewed as fundamentally disparate – theory versus practice. Theory has been often associated with logic, deduction, and knowledge (“knowing”), whereas application has been often associated with intuition, induction, and implementation (“doing”). Perhaps Kurt Lewin’s famous dictum, “Nothing is as practical as a good theory,” received so much attention because it was surprising in light of the general tendency to view application as the very antithesis of theorizing. Nonetheless, the notion of “translational research” highlights the intimate connection between theory and application and encourages theoreticians (often by means of funding opportunities) to descend from the

Theories need to be TAPAS-proof

We have conceptualized theory construction in terms of four regulatory ideals: namely, Truth, Abstraction, Progress, and Applicability. These may serve as Standards for critically evaluating a theory (TAPAS). Implicitly, these TAPAS are used pervasively for assessing psychological research. Truth claims are evaluated via appropriate experimental design, the carrying out of critical experiments, and by conceptual reviews and meta-analyses of the available evidence. Abstraction is often used to evaluate theoretical breadth and, at a more concrete level, in testing general psychological hypotheses via specific empirical observations. Progress is assessed through the ubiquitous demand that research be innovative and makes a significant new contribution to knowledge. Finally, applicability is often intimated by the plea for a body of psychological work to have “broader impact,” (e.g., Buunk, 2006; Fiedler, 2006; Van Lange, 2006) and it is deemed an important criterion for the funding of research grants at federal granting agencies in the US (like the National Science Foundation and the National Institutes of Health) because it is understood that scientific knowledge should have societal benefits. We highlight TAPAS explicitly in the hope that this will enhance their accessibility and use among social psychological theorists.

THEORY CONSTRUCTION AND DEVELOPMENT

Where do ideas come from? What are the sources of theorists’ inspirations? How do social psychologists turn tacit hunches and inchoate intuitions into overtly articulated

theoretical statements? And how does one facilitate theoretical progress? How “responsible” is the theorist for the future fate of her or his intellectual offspring? There is little public discussion of these matters in the social psychological literature though experience on these matters has likely accumulated in the personal memory stores of individual theorists. The latter assumption served as a departure point for a special issue published in the *Personality and Social Psychology Review* (2004, Vol. 8), entitled “Theory construction in social personality theory: Personal experiences and lessons learned.” Contributors to this special issue shared with the readers their innermost insights and metatheoretical self-reflections as well as “tricks of the trade” and personal strategies of theory construction and development. For example, there was a discussion about how detecting features that different phenomena share in common can lead to the development of overarching theories and the unveiling of deeper structures masked by striking surface differences (e.g., Kruglanski, 2004). Another paper discussed ways of carrying out theoretically inspired research programs and ensuring that a theory’s potential for contribution and impact is realized to its fullest (Higgins, 2004). Yet other papers discussed how to benefit from sabbatical leaves by developing a broader perspective on a body of empirical work (Zanna, 2004), and how collaboration with one’s colleagues may deepen and enrich one’s theoretical frameworks (Levine and Moreland, 2004). The chapters in the present volume illustrate such fascinating strategies and approaches via numerous theorists’ personal narratives of their intellectual journeys; journeys wherein their conceptual insights were explicitly articulated, and their theoretical structures were erected and developed.

IMPARTING THEORY CONSTRUCTION

It has been argued that over the course of its history, social psychology’s attention to

theory has declined, while the field has become increasingly data-driven and phenomenon-focused (e.g., Fiedler, 2004; Kruglanski, 2001; see also Jones, 1986). Part of the problem might have been the lack of systematic ways of teaching theory construction and incorporating courses and workshops on this topic in graduate curricula in social psychology. To a considerable extent, our graduate training focuses on issues of method, design, and data analysis. Theory construction is generally treated as unteachable and largely a matter of inspiration. Yet a great deal about theorizing can be articulated and imparted. The important ingredients of a successful theory, what we called TAPAS, can be defined, explained, and deliberately striven for. And the way those theoretical properties have been cultivated and developed by various theorists has also been explicitly described (Kruglanski and Higgins, 2004). In this vein, the present volume too purports to afford greater appreciation of the theorizing enterprise through personal stories of successful theoreticians, thereby teaching theoretical skills through examples.

Admittedly, passive exposure to historical accounts of theorizing strategies is unlikely to suffice. Concrete practice in translating the general principles of theory construction into actual attempts at conceptualization seems essential. One of us has briefly described a seminar course that aims to do this (Higgins, 2004). And regarding the teaching of applicability, a recent book by Buunk and Van Vugt (2007) provides a platform for practicing the skills of making theoretical concepts applicable to specific real-world problems. It does so by challenging the student to formalize the key properties of an important social issue (e.g., how to reduce vandalism in football stadiums; how to increase environmental concerns), analyze it in cause-and-effect terms using social psychological concepts and principles, and suggest possible policy measures based on such theoretical analysis of the problem.

ABOUT THIS HANDBOOK

Like the earlier books on social psychological theories, the present handbook reviews major conceptual developments in our field. Yet it also differs from prior reviews in significant respects. To begin with, the handbook covers the decades that passed since the last such volumes appeared in print (namely, West and Wicklund, 1980; Shaw and Costanzo, 1982) during which significant conceptual developments have taken place. Within that period social psychology has undergone exponential growth worldwide, yielding an explosion of theoretical frameworks by social psychologists at all levels of analysis – the biological system, the cognitive system, the motivational (and affective) system, the interpersonal system, and the group and cultural system (cf. Higgins and Kruglanski, 1996; Kruglanski and Higgins, 2007). At the same time, classic theories continued to inspire research in their different incarnations, such as Festinger's dissonance theory (1957) that received a "new look" in Cooper and Fazio's interpretation, and theories that have undergone a transition, such as Bandura's social learning theory that shifted into his social cognitive theory (for more information, see Chapters by Cooper and Bandura in this handbook).

In addition to offering a more up-to-date portrait of the theoretical landscape in social psychology, this handbook complements its predecessors in at least three important ways. First, each author has provided a personal, historical narrative of the theory's development, including the various inspirations, serendipitous events, critical junctures, and problem-solving efforts that affected theoretical choices and influenced the theory's evolution and impact. These personal narratives are unique to the present handbook and provide a richly textured background for understanding how theories are actually created, nurtured, and shaped over time.

Second, each author has placed her or his theory within the intellectual history of

the topic it addresses, and has commented on the theory's unique contribution to the field against this intellectual backdrop. This places each theory within a second kind of history – *the history of ideas* – that is also unique to this handbook. This aspect of the book provides a strong answer to the question, "Why should I care? What's the added value gained from the theory's contribution?" Third, each author has evaluated her or his theory in terms of its *applicability* for not only understanding but also for solving critical social issues and real-world problems. This aspect of the book provides a second strong answer to the question, "Why should I care? What's the added value?"

The overarching principle underlying these three foci was that theories are ultimately about *ideas*. Hence, it is important to learn: (a) *where* the ideas came from and how they developed; (b) *why* the ideas matter intellectually and historically; and (c) *what* difference the ideas make for dealing with current societal concerns.

Criteria for inclusion

When is a conceptual framework a theory, when is it a model, and when is it a hypothesis? Though presumably social psychologists would agree on the theoretical "status" of numerous conceptual frameworks in the field, they might disagree with respect to some. We do not regard this surprising because the field of social psychology has grown considerably – and in many directions – over the past three decades (since the 1982 book by Shaw and Costanzo). Further, social psychology lacks a widely received "grand theory" of the sort that seeks to account for all diversity and serves as the consensual platform for more specialized analyses (like evolutionary theory in biology or rational choice theory in economics).

In selecting theories for the present handbook, we used the following broad guidelines. First, we decided to include

theories with a bit of a “track record”; theories whose development could be traced back more than a decade. Indeed, one of our objectives was to provide insights into the ways that social psychologists design, develop, and “nurture” their theories. And this process, inevitably, takes time. It takes time to develop a theory’s validity (the *truth* criterion), a theory’s generality (the *abstraction* criterion), a theory’s generative power (the *progress* criterion), and a theory’s usefulness in solving real-world problems (the *applicability* criterion).

Second, we decided to include theories that have survived the passage of time, those that continue to guide research in the “here and now.” It is, of course, possible that some theories will be revitalized in some way in the future, in which case they may well be included in future handbooks on theory.

Finally, we decided to include theories developed in the tradition of social psychology, rather than ones developed outside of the field of social psychology. Admittedly, there are many influential theories and models on social psychological topics that could be included. Examples are theories in fields such as cognitive neuroscience, decision-making, economics, sociology, and political science. But to include such theories, we felt, would excessively broaden the scope of our enterprise and defocus our perspective from conceptual work carried out in social psychology proper.

We should note that not all theories that met these guidelines ended up in this book. In some cases, the author of a theory had passed away. In some cases, the author of a theory declined our invitation. We regret not having these theories represented. But we are very thankful for the theories that are represented, and we are extremely grateful to our authors for their contributions to this Handbook. It has been a wonderful experience for us as Editors to learn more about the history of their landmark theories.

REFERENCES

Allport, G.W. (1954) The historical background of modern social psychology. In G. Lindzey (ed.), *Handbook of Social Psychology*, 1, 3–56. Cambridge, MA: Addison-Wesley.

Buunk, A.P. (2006) Social psychology deserves better: Marketing the pivotal social science. In P.A.M. Van Lange (ed.), *Bridging Social Psychology: Benefits of Transdisciplinary Approaches*, pp. 83–89. Mahwah, NJ: Erlbaum.

Buunk, A.P. and Van Vugt, M. (2007) *Applying Social Psychology: From Problems to Solutions*. Thousand Oaks, CA: Sage.

Deutsch, M. and Krauss, R.M. (1965) *Theories in Social Psychology*. New York: Basic Books.

Einstein, A. (1934) *Mein Weltbild*. Amsterdam: Querido.

Festinger, L. (1957) *A Theory of Cognitive Dissonance*. Stanford, CA: Stanford University Press.

Fiedler, K. (2004) Tools, toys, truisms, and theories: Some thoughts on the creative cycle of theory formation. *Personality and Social Psychology Review*, 8, 123–131.

Fiedler, K. (2006) On theories and societal practice: Getting rid of a myth. In P.A.M. Van Lange (ed.), *Bridging Social Psychology: Benefits of Transdisciplinary Approaches*, pp. 65–70. Mahwah NJ: Erlbaum.

Fiske, S.T. (2004) Mind the gap: In praise of informal sources of formal theory. *Personality and Social Psychology Review*, 8, 138–145.

Higgins, E.T. (2004) Making a theory useful: Lessons handed down. *Personality and Social Psychology Review*, 8, 138–145.

Higgins, E.T. and Kruglanski, A.W. (eds) (1996) *Social Psychology: Handbook of Basic Principles*. New York: Guilford Press.

Jones, E.E. (1986) Major developments in social psychology during the past five decades. In G. Lindzey and E. Aronson (eds), *The Handbook of Social Psychology*, 3rd Edition, pp. 47–135. New York: McGraw-Hill.

Jones, E.E. (1988) Major developments in five decades of social psychology. In D.L. Gilbert, Susan T. Fiske and G. Lindzey (eds), *The Handbook of Social Psychology*, 1, 3–57. New York: McGraw-Hill.

Jones, E.E. and Gerard, H.B. (1967) *Foundations of Social Psychology*. New York: Wiley.

Kelley, H.H. (2000) The proper study of social psychology. *Social Psychology Quarterly*, 63, 3–15.

Kruglanski, A.W. (2001) That 'vision thing': The state of theory in social and personality psychology at the edge of the new millennium. *Journal of Personality and Social Psychology*, 80, 871–875.

Kruglanski, A.W. (2004) The quest for the gist: On challenges of going abstract in social and personality psychology. *Personality and Social Psychology Review*, 8, 156–163.

Kruglanski, A.W. (2006) Theories as bridges. In P.A.M. Van Lange (ed.), *Bridging Social Psychology: Benefits of Transdisciplinary Approaches*, pp. 21–34. Mahwah, NJ: Erlbaum.

Kruglanski, A.W. and Higgins, E.T. (2004) Theory construction in social and personality psychology: Personal experiences and lessons learned. *Personality and Social Psychology Review*, 8, 96–97.

Kruglanski, A.W. and Higgins, E.T. (eds) (2007) *Social Psychology: Handbook of Basic Principles*, 2nd Edition. New York: Guilford Press.

Levine, J.M. and Moreland, R.L. (2004) Collaboration: The social context of theory development. *Personality and Social Psychology Review*, 8, 164–172.

Mandler, G. and Kessen, W. (1959) *The Language of Psychology*. New York: Wiley.

Merton, R.K. (1949) *Social Theory and Social Structure*. Glencoe, IL: The Free Press.

Mischel, W. (2004) Toward an integrative science of the person. *Annual Review of Psychology*, 55, 1–22.

Pinker, S. (2002) *The Blank Slate: The Modern Denial of Human Nature*. New York: Viking.

Popper, K.R. (1959) *The Logic of Scientific Discovery*. New York: Harper. (Original work published as *Logik der Forschung*, 1935).

Shaver, K.G. (1977) *Principles of Social Psychology*. Cambridge, MA: Winthrop.

Shaw, M.E. and Costanzo, P.R. (1982) *Theories of Social Psychology*. New York: McGraw-Hill. (Originally published 1970).

Van Lange, P.A.M. (ed.) (2006) *Bridging Social Psychology: Benefits of Transdisciplinary Approaches*. Mahwah, NJ: Erlbaum.

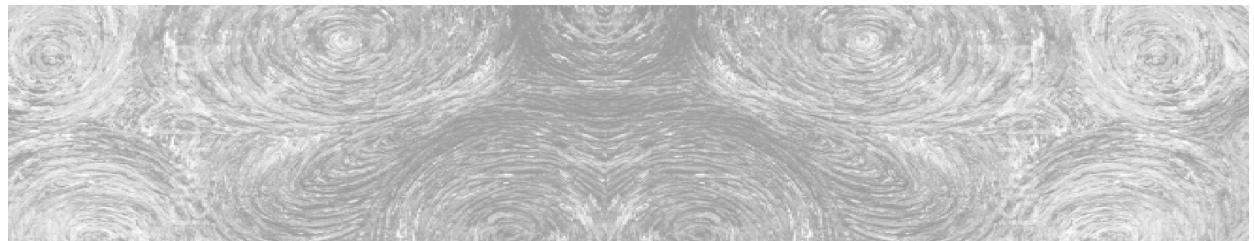
Van Lange, P.A.M. (2007) Benefits of bridging social psychology and economics. *Academy of Management Review*, 32, 671–674.

West, S.G. and Wicklund, R.A. (1980) *A Primer of Social Psychological Theories*. Monterey, CA: Brooks /Cole.

Zanna, M. (2004) A naïve epistemology of a working social psychologist (Or the working epistemology of a social psychologist): The value of taking "temporary givens" seriously. *Personality and Social Psychology Review*, 8, 210–218.

PART I

Biological/Evolutionary Level of Analysis



Evolutionary Theory and Human Social Behavior

Douglas T. Kenrick

ABSTRACT

From an evolutionary perspective, all the reactions people typically have to one another reflect the influence of psychological predispositions that helped our ancestors survive and reproduce. Any behavior can be understood at several levels of analysis, involving immediate triggers in the environment and the person's current biochemistry, developmental experiences, and evolutionary analyses of the adaptive function of certain choices over others. The evolutionary approach is thus not an alternative to other approaches discussed in this book. Instead, researchers adopting this perspective attempt to integrate research findings on ongoing social cognition and interpersonal relationships with theory and research from evolutionary biology, anthropology, and cognitive neuroscience. This chapter reviews research applying evolutionary models to: (1) biases in information processing (such as outgroup stereotypes), (2) the influence of affect and motivation on the decisions we make about other people, and (3) how simple, evolved decision-biases contribute to social dynamics and the emergence of culture. From this perspective, the mind is a coloring book, rather than either a blank slate or an unfolding blueprint. The approach has implications for important everyday behaviors and social problems, and the chapter concludes with a discussion of some of those

implications for research on prejudice and economic decision-making.

INTRODUCTION

Evolutionary social psychology begins with a simple assumption, namely that people's interactions with one another are influenced by mental and emotional mechanisms shaped by natural selection. From this perspective, the way individual people think about one another, the way families, friends, and enemies feel about one another, and even the societies human beings construct, can be better understood by considering local social interactions in the broader context of other societies, and by considering all human societies in the still-broader context of other animal species. In a classic example of this approach, Charles Darwin suggested in 1872 that emotional expressions served an adaptive function – communicating one person's motivations and intentions to others (e.g., overt expressions of anger reduce the odds of

costly physical conflict; smiles increase the odds of cooperation). Along with physical features such as upright stance and opposable thumbs, evolutionary theorists assume humans inherit brains equipped with mechanisms for managing our movement through the physical world (e.g., seeing in color, discriminating sugars from poisonous alkaloids) and the social world (e.g., speaking languages, bonding between mother and infant).

Although the central assumption at the heart of the evolutionary approach may not sound controversial, there has been debate about the extent to which evolved mechanisms are involved in human social behavior (see Alcock and Crawford, 2008; Kenrick, 2006; Tybur et al., 2008). Psychologists have wondered about questions such as: how can researchers sort out the influence of evolved mechanisms from effects of culture or learning? Consider the difficulties involved in trying to determine why women in North America are less likely to commit homicides than men. That difference could be a product of evolved mechanisms, or a product of the fact that American women grow up seeing movies and reading books in which men are depicted as more violent, or seeing boys, but not girls, encouraged to fight. If one assumes that culture, learning, and evolved mechanisms interact with one another, such questions become rather complex, requiring inputs from many different fields of research. In fact, modern evolutionary theory involves the integration of a broad network of ideas and evidence from different disciplines, including evolutionary biology, anthropology, and cognitive science.

Despite their complexity, the questions raised by an evolutionary perspective involve profoundly important issues about human nature and society. An evolutionary perspective has implications for every domain of human social behavior, from altruism, friendship, and love, to aggression, prejudice, and stereotyping. Evolutionary social psychology involves questions that are not only theoretically interesting, but also immensely

important in a practical sense, with implications for law, business, medicine, and political science.

ON BECOMING AN EVOLUTIONARY SOCIAL PSYCHOLOGIST

My initial interest in the evolutionary perspective could be attributed either to a broad-ranging curiosity about nature, or, less nobly, to an inability to decide what I wanted to be when I grew up. As an undergraduate I started off in biology, switched to psychology, and then considered switching to anthropology. I started graduate school in clinical psychology, but in my second year of graduate school I got to teach a lecture class in general psychology. I discovered that, if I were to become a research psychologist, I could study artistic creativity and mindless conformity, parental love and homicidal violence, sexual attraction and racial bigotry, or any other topic that involves thought, feeling, or behavior. This array of choices didn't require any premature foreclosures on my life options, so I decided to switch from clinical psychology to a research track in social psychology.

My conversion to an evolutionary perspective came two years later, as I was preparing for comprehensive examinations in social psychology. I should have been holed up in the library, reading all I could about experiments on dissonance theory, attributional processes, and objective self-awareness. But whenever I have a daunting amount of work to do, I develop intense interests in anything unrelated to the task at hand. In this spirit, I drifted into the campus bookstore and picked up a copy of *Primate Behavior and the Emergence of Human Culture*, by anthropologist Jane Lancaster (1975). This particular volume seemed comfortably outside the domain of experimental social psychology, so in my work-avoidance mode I felt compelled to buy it, bring it home, and read it. Lancaster's book indeed had little

to do with the questions my social psychology professors asked during my comprehensive examination. But I came to believe it raised many questions they should have asked.

Despite the fascinating range of research topics in social psychology, the scope of theory in the field was rather narrow at the time. Preparing for my comprehensive examination, I encountered a scattered disarray of unconnected small range theories, each independently advanced to explain a particular facet of social behavior. One mini-theory addressed frustration-induced aggression, another covered interpersonal attraction between people with similar attitudes, another addressed responses to one-sided versus two-sided arguments, and on and on. I missed the grand theories of personality and behavior I'd been exposed to during my clinical training, but when I mentioned this to the social faculty, they wagged their fingers and proudly informed me that social psychology was a "mini-theory" discipline.

Social psychologists at the time prided themselves not only on being theoretically constricted, but on being empirically narrow as well – studying anorexically thin slices of thought and behavior. Social psychologists in 1975 rejected the study of stable "traits" and focused mainly on people's thoughts and behavior in response to the "immediate situation," or at least those situations that could be captured within the half-hour duration of a typical psychology experiment. There were reasons for these strictures – experimental studies were designed to maximize control, and theoretical restraint was intended to cut down on rampant speculation about unobservable events inside the head or body – things scientists couldn't easily observe and count. But to a curious young student interested in the roots of human behavior, those constraints seemed excessive. I was not alone in being troubled by these limitations. Indeed, social psychology was undergoing an "identity crisis" at that time, with several of the field's leaders calling for an expansion of our theories and our research methods.

In this context, I took an almost guilty delight in glimpsing the very broad theoretical perspective suggested in Lancaster's book on primate social behavior. Instead of a narrow focus on one or another aspect of the social behavior of the members of our particular culture under specific laboratory conditions, Lancaster's evolutionary perspective offered a tantalizing suggestion – that we ought to erase the lines between psychology, biology, and anthropology, and consider how all these vast subjects fit together.

I began raving about Lancaster's book to anyone who would listen. Some of my fellow graduate students and faculty advisors just gave me an uncomfortable smile, as if I was earnestly explaining why I had just joined a cult. But a new assistant professor named Ed Sadalla had just picked up a copy of a recent book titled *Sociobiology* (Wilson, 1975), which considered common evolutionary principles underlying the social behavior of animals from ants to humans. Sadalla suggested that the evolutionary approach had great promise for generating testable hypotheses about human behavior, and he already had one he thought we should test. As part of a process Darwin called "sexual selection," females in many species choose males who have proven their dominance over other males, whereas males, with less to lose from an ill-chosen mating, tend to be less selective. Along with Beth Vershure, Sadalla and I began a series of studies suggesting analogous processes in humans. Although our results were clear and reliable across several experiments, it took us over a decade to publish them (Sadalla et al., 1987). As it turns out, we had unintentionally walked into an intellectual firestorm. There was an academic tumult surrounding sociobiology, which became the subject of a fascinating chapter in the history of science (Segerstrale, 2000).

The various controversies surrounding evolutionary social psychology were discouraging to many young researchers, but they have not been fruitless. Controversy often contributes to empirical and theoretical progress, as a theory's proponents search for

new findings to address critics' skepticism. The evolutionary approach has generated many new findings and ideas, and the field's top journals have since published hundreds of social psychological studies testing evolutionarily informed hypotheses about the whole range of social behaviors, from altruism to xenophobia (e.g., Griskevicius et al., 2007; Navarette et al., 2009; Schaller and Murray, 2008). Many of today's prominent social psychologists, including several contributors to this volume, have incorporated evolutionary perspectives into their research (e.g., Brewer and Caporael, 2006; Cialdini et al., 1997; Deci and Ryan, 2000; Fiske et al. 2007; Higgins and Pittman, 2008; Huang and Bargh, 2008; Leary and Baumeister, 2000; Mikulincer and Shaver, 2006; Sedikides and Skowronski, 1997; Taylor et al., 2000; Van Knippenberg and Van Baaren, 2006; Van Vugt and Van Lange, 2006). At the same time, the approach continues to generate new empirical questions, and new theoretical puzzles to solve.

WHAT IS EVOLUTIONARY SOCIAL PSYCHOLOGY?

From an evolutionary perspective, all recurrent human social behaviors reflect the influence of physical and psychological predispositions that helped our ancestors survive and reproduce. This does not mean every individual social behavior is successful in promoting survival and reproduction, and it does not mean people (or other social animals) consciously think about survival and reproduction all, most, or even much of the time. It does imply that any social animal's brain is composed, in part, of mechanisms that helped its ancestors succeed in interactions with other members of its species. Thus, humans' reactions to other humans are presumed to *reflect the influence* of mechanisms shaped to solve the kinds of problems and opportunities our ancestors regularly encountered. I emphasize "reflect the influence"

because an evolutionary approach does not imply that human behaviors are robotically determined by instinctive mechanisms over which we have no conscious control, or which are impervious to environmental inputs. People can and often do exercise control over powerful and fundamental emotional and motivational inclinations, including anger, fear, and sexual arousal. Furthermore, most mental mechanisms reflect the operation of flexible trade-offs, determined in interaction with current environmental conditions and past learning experiences (e.g., Ohman and Mineka, 2001). Although flexible, the influences of evolved predispositions (like hunger, thirst, sexual arousal, and anger) are nevertheless powerful vectors in our decision-making.

To understand the importance of evolved mental mechanisms, it helps to step back from our own species, and consider the interaction between other animals' bodily features and their environments. Killer whales, for example, though related to cows, would not do well with a cow's brain, since a killer whale's brain must control a body that tracks prey in the ocean rather than eating grass in a meadow. Likewise, bats, though also mammals, need brains designed to run tiny bodies that fly around catching insects at high speeds in the dark. Because all organisms' brains are composed of mechanisms evolved to deal with recurrent environmental threats and opportunities, evolutionary theorists ask: what are the implications of human evolutionary history (e.g., living in omnivorous and hierarchical primate groups populated by kin) for the design of the human mind? Evolutionary social psychologists focus on the subset of questions dealing with recurrent social conditions of human life, and their hypotheses reflect anthropological data about social interactions common in societies around the world (e.g., close relationships with family, dominance hierarchies, long-term bonds between parents, common dangers from other groups competing for resources and territory, etc.), as well as general principles derived from placing humans

in the context of other species confronting diverse adaptive problems (e.g., Allen-Arave et al., 2008; Lummaa, 2007).

WHAT ARE THE ROOTS OF EVOLUTIONARY SOCIAL PSYCHOLOGY?

Around the time I was studying for my comprehensive examinations in 1975, there was an explosion of interest in, and controversy surrounding, the evolutionary approach to human behavior, centered around E.O. Wilson's book *Sociobiology*. But the perspective had been developing for some time, and represented the convergence of several streams of influence. A major contribution came from the field of ethology – the study of the behavior of animals in their natural habitats. In 1973, three European ethologists received the Nobel Prize for work on the adaptive significance of animal behaviors. One of them, Niko Tinbergen, experimentally studied how stickleback fish respond with complex behavioral displays to other sticklebacks (e.g., males demonstrate an aggressive display on seeing a red underbelly on another male, and the mechanism can be tricked by using “supernormal stimuli” – fishlike shapes with bright red paint on the underside). The second, Karl Von Frisch, demonstrated that bees engage in complex communications – informing other colony members about the location of nectar-bearing flowers. The third Nobel Prize-winning ethologist, Konrad Lorenz, conducted research on imprinting, the process by which young geese became attached to their mothers. Each of these lines of research demonstrated an interaction between an innate mechanism and inputs from the social environment.

Did the ethological work on other animals have implications for human social behavior? Many behavioral scientists at the time believed the answer was yes. Paralleling Tinbergen's findings on innate communication in bees, human language seemed a

beautiful example of instinctive preparedness interacting flexibly with inputs from the environment (Pinker, 1994). For example, comparative linguists had uncovered evidence that languages the world over have a similar underlying grammatical structure, developmental psychologists found that children make similar linguistic mistakes regardless of the language they are learning, and the languages spoken by adults everywhere were found to be equally complex, regardless of the speaker's education level. Paralleling Lorenz's work on animal imprinting, developmental psychologists had found that young children go through predictable phases in their relationships with parents and strangers (such as a fear of strangers that peaks around nine months), and that infants engage in complex nonverbal mimicry of their mothers. Paralleling Tinbergen's research on fixed action patterns in sticklebacks, Irenaus Eibl-Eibesfeldt (1975) found a sequence of nonverbal flirtation gestures in societies around the world, and other research revealed that children born blind and deaf nevertheless demonstrated appropriate facial expressions in appropriate contexts (e.g., smiling when tickled) (Eibl-Eibesfeldt, 1973).

I had personally observed the power of innate mechanisms on animal social behavior as a child, when I raised tropical fish for a hobby. Following instructions in a tropical fish manual, I placed a male *Betta splendens* (or Siamese fighting fish) in a tank, fed him an abundant supply of live food, and raised the water's temperature and acidity (thereby simulating conditions during breeding season in Southeast Asia, from whence his ancestors came). He responded by building a nest of bubbles on the water's surface, exactly as my manual had predicted. When the nest reached the right size, I introduced a female *Betta* who had been exposed to similar conditions and developed a bulge in her belly, indicating a supply of eggs. Upon seeing the female, the male puffed out his colorful fins, and began swimming around her in a mating dance. She responded with movements indicating her readiness to mate, and he then wrapped his

body around hers, triggering her release of several eggs and his release of sperm. He scooped up the eggs in his mouth, placed them in the bubble nest, and repeated the sequence several times until her supply of eggs was depleted. He then chased her away and began a period of guarding the eggs and later the emerging tadpole-like fry. This sequence was complex, but none of it was learned. I know this because I watched one of his offspring perform the exact same ritual with another female, despite the fact that he had been raised in isolation (which is common with male Siamese fighting fish, who are notoriously pugnacious and likely to attack other fish).

The ethologists' research made it clear that much so-called instinctive behavior (such as imprinting) involved what Ernst Mayr called "open instincts" – innate proclivities requiring environmental inputs to be fully operative. Although a fixed sequence of movements is fine for mating rituals, some innate predispositions require flexibility to be useful. For example, animals need a rapid response system for avoiding other dangerous animals and poisonous foods, but exactly which other species in the local environs pose threats is often quite variable over the range into which they might be born (consider young English sparrows who, like humans, may be born in deserts, forests, farmlands, or cities on several continents). Hence, those avoidance systems need to be calibrated to reflect local threats (Ohman and Mineka, 2001).

Psychologists studying learning in the laboratory had presumed mammals possessed only a very few innate drives (such as hunger and thirst, often referred to as "primary drives"), and that experiences after birth led to the development of "secondary drives" (desires for other stimuli associated with satisfaction of hunger and thirst). The development of secondary drives was believed to depend on two simple forms of learning – classical and operant conditioning, the rules of which presumably applied similarly to many kinds of learning across many species (Skinner, 1953). This view had the advantage

of being parsimonious, meaning it could explain much with few assumptions. However, several behavioral psychologists began to uncover findings challenging this view. The rules of conditioning changed depending on what was being learned and which species did the learning (Seligman and Hager, 1972). Consider the rules involved in learning to avoid poisonous foods. Some learning requires instantaneous feedback (e.g., a jolt of pain immediately after touching a hot stove facilitates learning not to touch it again), but people and other animals learn to avoid foods that made them sick many hours after the food was eaten. Learning food aversion is also unlike many other types of learning in that it requires only one trial and is difficult to extinguish. Furthermore, the types of stimuli that get conditioned to nausea vary in ways consistent with the organism's evolutionary history and typical ecology. For example, rats, which have poor vision and rely on taste and smell to find food at night, easily condition aversions to novel tastes, but not to novel visual stimuli (Garcia and Koelling, 1966). Quail, on the other hand, have excellent vision and rely on visual cues in food choice, and show the opposite bias, conditioning easily to visual cues but not to taste (Wilcoxon et al., 1971).

Although not many social psychologists were radical behaviorists, most shared the assumption that social behaviors were products of individual learning experiences and not innate mechanisms; however, two other influences at the time began to challenge that simple empiricist worldview. Behavioral geneticists were uncovering evidence that complex human behaviors could be passed on genetically. For example, twins raised separately shared many adult personality traits, including rare forms of psychopathology, whereas adopted children shared surprisingly few such traits with the families in which they were raised (Rowe, 1994). At the same time, advances in cognitive science, including the study of artificial intelligence, were beginning to suggest that complex cognitive processing could be programmed

from very simple “on–off” components (Gardner, 1985).

At the time, social psychologists were, like ethologists, becoming more open to studying behavior in naturalistic contexts. In particular, researchers interested in relationships were arguing that processes unfolding over months and years of social interactions might not emerge in a half hour between strangers in laboratory experiments (e.g., Murstein, 1981).

MULTIPLE LEVELS OF CAUSAL ANALYSIS

Even in the 1970s, most behavioral scientists accepted natural selection as an explanation of animal morphology (why giraffes have long necks and porpoises have fins, for example). They also accepted that human morphology, including upright stance, opposable thumbs, and large brain size, were products of natural selection; likewise for the behavior of other animals, such as peacocks’ mating displays. But back then, it was less clear how an understanding of natural selection applied to human thought and behavior, and if so, how that understanding would translate into hypotheses of the sort typically considered by psychologists. My friend Jim Sherman, for example, had taken E.O. Wilson’s biology course when he was a young student at Harvard. Jim did not doubt the theory of natural selection, but when we discussed these issues in the 1980s, he questioned whether evolutionary theory could yield testable implications for social cognition. Jim suggested I consider switching my field of interest to anthropology or biology.

Other behavioral scientists at the time and since have questioned how one can ever study the evolution of behavioral inclinations. Unlike bones, behaviors generally do not leave fossil evidence. Although anthropologists are clever in deriving evidence about ancestral behavior, many interesting behaviors, such as flirtatious gestures or

gossip-filled conversations, leave no traces anyone has yet deciphered using archaeological or geological methods. But even if there were major advances in such inferential techniques, most evolutionary social psychologists would not buy a pith helmet and rush off to the Olduvai Gorge. Why? Because an evolutionary perspective on psychology does not typically involve a search for historical roots of behaviors – for when and how particular behavioral mechanisms evolved. This is also true of animal ethologists studying insects, birds, and fish; they are not typically concerned with the evolutionary history of the inclinations they study either. To understand why not, it is important to appreciate a distinction biologists frequently make between explanations at different levels of analysis.

Historical controversies have been fueled by failures to distinguish between different levels of causation within the field of biology. Consequently, biological theorists have stressed the importance of differentiating between causal explanations of behavior involving evolutionary history, adaptive function, ontogenetic development, and proximate determinants (e.g., Sherman, 1988). For example, consider the question of why human mothers nurse their offspring. This can be addressed at four different levels of analysis:

- 1 *Historical* explanations consider *ancestral roots* of behaviors. Researchers adopting a historical perspective might consider human nursing in the comparative context of other animal species. Unlike human language, nursing capacity does not pose much of a historical puzzle, since all primate females nurse their offspring, as do all mammals.
- 2 *Functional* explanations, on the other hand, are concerned with ultimate *adaptive purposes* of behaviors. A functional explanation might suggest mothers nurse offspring because it increases offspring survival rates.
- 3 *Developmental* explanations are concerned with *lifespan-specific inputs* that sensitize the organism to particular cues. A developmental explanation might suggest that mothers nurse

offspring because pregnancy and childbirth trigger puberty-dependent shifts in hormones and milk production in mammalian females.

4 *Proximate* explanations focus on *immediate triggers* for particular behaviors. A proximate explanation might suggest that nursing occurs because an infant has begun suckling on the female's nipple, which leads to immediate hormonal changes inside the mother, which stimulates milk release.

Sometimes there are obvious connections between the different levels of analysis. In the case of nursing, developmental changes in lactation capacity accompany other changes during pregnancy, and the infant, who receives obvious functional benefits from nursing, triggers the proximate release of milk. But connections between different levels of analysis are not always obvious. Consider why birds migrate each year. A proximate explanation is that birds migrate because days are getting shorter – the immediate cue triggering migration. The ultimate reason for such migration, however, is survival and reproduction – the distribution of desirable food and mating sites varies seasonally. There are two key implications here: (1) animals, including humans, need not be consciously aware of the ultimate functions of their behaviors; and (2) the connection between long-term goals and immediate goals is often indirect and not obvious.

A given researcher is usually concerned with one or two levels of analysis. However, an explanation at one level of analysis must be compatible with explanations at other levels. Positing a proximate or developmental mechanism that reliably leads people to make functionally maladaptive decisions (such as Freud's death instinct) is problematic. Evolutionary psychologists typically advance hypotheses about links between proximate mechanisms and adaptive function, not about the historical roots of the mechanism. In deriving those hypotheses, however, psychologists adopting an evolutionary perspective attempt to take into account pertinent findings from evolutionary biology and/or anthropology. Psychologists

can derive hypotheses about proximate causes and development without thinking in evolutionary terms, but disregarding evidence and theory derived from research on other cultures and other species can lead to hypotheses incompatible with other levels of analysis. For example, psychologists during the last century often assumed most sex differences in social behavior (such as differences in violent aggression) were products of "American culture," unaware that similar differences were found in other cultures, and even other species. Indeed, the study of sex differences in social behavior led to some of the first critical findings in evolutionary social psychology, as I will describe below.

CRITICAL TESTS VERSUS NOMOLOGICAL NETWORKS

Could an evolutionary perspective provide new and testable hypotheses about human social behavior? A natural place to begin looking was by considering sex differences in mating-related behaviors. Sex differences in mating behavior are found throughout the animal kingdom, and comparative biologists have made progress understanding not only why those differences exist, but also why they are sometimes relatively larger or smaller. Darwin himself wondered why there are such dramatic sex differences in some species (such as peacocks and peahens), and why some features (such as peacock's brilliant and attention-getting feathers) involve features likely to decrease survival. He coined the term *sexual selection* to refer to the evolution of features useful not for survival, but for either competing with members of the same sex or attracting members of the other sex. A century later, Trivers (1972) connected sexual selection to *differential parental investment*. Within any given species, there is often an initial discrepancy in minimum resources or effort one sex is required to invest in the offspring. In mammals, for example, females must carry

offspring inside their bodies, and later nurse them for weeks, months, or years. On the other hand, the minimum obligatory investment by mammalian males is a few moments of courtship and the energy required to copulate. Because mammalian females are limited to fewer offspring than a male could have (if he could attract and inseminate multiple females), females choose more carefully, picking males demonstrating either superiority to other males, or willingness to invest effort and resources in the female and their offspring.

The research Sadalla, Vershure, and I started in 1976 was designed to test a simple derivation from the literature on sexual selection and differential parental investment – that women, like females in other mammalian species, would be attracted to males giving behavioral indications of dominance relative to other males. We presumed that men, like males in other mammalian species, would be relatively less attentive to female dominance. In a series of experiments manipulating social dominance in various ways, we found support for that hypothesis. Early reviewers, however, were not convinced these findings clearly favored an evolutionary explanation over what they considered a “more straightforward” explanation based on learned social norms. According to this alternative, people are attracted to members of the other sex who conform to societal expectations, which differ for men and women. Reviewers argued that American men were expected to act dominant, and women to act submissive. We saw problems with this alternative. First, it did not explain why men in our studies were not attracted to submissive females (who were, after all, conforming to the presumed societal expectations of the time). Second, the alternative explanation tacitly assumed the cultural norms were arbitrary, without considering why it was more socially normative for men to act dominant in the first place. Third, although the reviewers expressed confidence that sex differences in dominance were products of American culture, they did

not present evidence of cultures in which females were more dominant than males, and in which men, but not women, preferred dominant partners. As later research indicated, women all around the world care more about social dominance and status in their mates than men do (Buss, 1989). Fourth, the alternative could not explain why the sex difference among North Americans was also found in many other animal species. Finally, it could not account for findings that the hormone testosterone, found in much larger quantities in male mammals, was closely linked to dominance behaviors and mating displays. Our findings did not, by themselves, rule out the possibility that North American society had constructed this sex difference arbitrarily, and that the apparent links with other species were mere coincidence. However, our findings fit into a broad *nomological network* of research findings from several disciplines. Indeed, to appreciate the strength of evolutionary explanations, it is often necessary to appreciate how a particular finding fits with many other findings.

Our findings on dominance and sexual attraction came from laboratory experiments involving North American college students, and could not, in themselves, rule in favor of an evolutionary explanation over a cultural or learning alternative. In later research, Rich Keefe and I adopted a different approach, examining archival data on sex differences from many different cultures. Keefe and I were interested in another sex difference – the common tendency for younger women to pair up with relatively older men (Kenrick and Keefe, 1992). Several social scientists had conducted analyses of singles’ advertisements, and observed that men typically advertised for a partner a couple of years younger, and women typically wanted a partner slightly older than themselves. Although this violated the general tendency to prefer partners as similar as possible, it was typically attributed to the influence of sex-discrepant norms in American society (e.g., Cameron et al., 1977).

Keefe and I suspected the sex difference in age preferences might be better explained in terms of ideas from a branch of evolutionary theory called *life history theory*. Life history theorists explore how each animal's life cycle – from conception to death – is shaped by natural selection to facilitate reproductive success (Stearns et al., 2008). A life history is a genetically organized developmental plan – a set of general strategies and specific tactics by which organisms allocate energy to survival, growth, and reproduction. Different species have very different life histories; some (such as salmon) reproduce in a single burst of effort, others (such as humans) reproduce many times over their life span. Some animals begin reproducing shortly after birth; others, such as humans and elephants, wait over a decade to begin reproducing. Within any species, there are often differences in life history, some linked to the differences in parental investment we discussed earlier. In the case of humans, females typically reach sexual maturity a year or two earlier than males, a difference common in species in which larger males compete more successfully for mates (Geary, 1998). Human females, unlike most other mammals, go through age-linked physiological changes terminating their reproductive capacities. Menopause may have evolved because ancestral women lived long lives, and typically invested heavily in grandchildren, balancing costs of direct reproduction against benefits of indirect kin care.

Unlike women, men do not directly invest bodily resources in offspring. They do, however, frequently invest indirect resources, providing food, care, and protection. Anthropological studies of hunter-gatherers suggest that men become increasingly skilled at providing resources as they age (Kaplan et al., 2000). Even in modern societies, men's social status and resources typically increase for several decades after reaching sexual maturity. Males do not go through menopause and are capable of fathering children into their 80s. All of these considerations led Keefe and I to guess that sex differences in

age preference were better explained in terms of human life history than American cultural norms. On the life history view, females choose older males who traditionally had more access to resources; males prefer females with features indicating peak fertility (between late teens and early thirties). If the evolutionary life-history account were true, then preferred age differences between partners ought not to be constant, but to change depending on each partner's age. In particular, preferences for younger females would not be pronounced among young men, but would get stronger as men aged (and women their age had fewer remaining years of fertility). Young men, although typically most sex-typed, would violate this presumed norm by expressing attraction to women slightly older than themselves. Those predictions were confirmed in several studies, raising difficulty for explanations based on sex-typed norms (Kenrick and Keefe, 1992; Kenrick et al., 1996a).

More critically, the life history perspective suggested that this sex difference would be found across human societies, and not be in any way associated with North American cultural norms. Indeed, we found the sex difference in age preferences across societies ranging from Germany to India (despite many other differences in marital and mating customs). We also found it across historical periods, on a small remote island in the Philippines in 1913 and in the Netherlands during the 1600s and 1700s. Consistently, the United Nations data on marriages indicated that young men around the world are several times less likely to marry than are young women, whereas older men are much more likely to marry than are older women (Kenrick and Keefe, 1992). The size of the sex difference varies somewhat across societies (Eagly and Wood, 1999; Kenrick and Keefe, 1992), but contrary to social norms explanations, the difference is actually least pronounced in North America and Europe in comparison to less Westernized societies. As I'll discuss later, the cultural variations are themselves likely to be

linked to biological factors (Kenrick et al., submitted).

Other researchers have found various universals in human social behavior. For example, romantic love, once believed to be a cultural invention of medieval European societies, is found all around the world (Jankowiak and Fisher, 1992). Sex differences in violent crime, again often attributed to American cultural norms, are found around the world and over historical periods (Daly and Wilson, 1988). Buss (1989) and Schmitt (2003) and their colleagues have also found numerous universals in human mating arrangements.

CROSS-FERTILIZATION VERSUS ARTIFICIAL INSEMINATION: BRIDGING SOCIAL PSYCHOLOGY AND EVOLUTIONARY THEORY

Much of the original debate involved attempts to uncover findings that fit better with evolutionary models than with traditional social science models, to support the premise that pursuing an evolutionary perspective was not a waste of time. So far, the story sounds more like artificial insemination than cross-fertilization, with evolutionary biology providing important ideas and psychology absorbing them. But social psychologists have many ideas, findings, and research tools that have contributed to a better understanding of the links between evolution and social behavior.

Consider gender differences in mating-relevant behaviors. Evolutionary theorists initially emphasized these differences, following considerations of life history and parental investment. Although sex differences are often large, however, they are sometimes small. Social psychological research tools and theory allowed us to address why. Recall that social psychologists in the 1970s had begun arguing that laboratory experiments might not be ideal for studying friendships and romantic relationships. Those researchers

demonstrated that what people find rewarding at a relationship's beginning is not necessarily what they find rewarding in long-term relationships or marriages (Murstein, 1981). This insight, when integrated with the theory of differential parental investment, was important to understanding variations in sex differences and similarities in mate choice. In the initial phases of romantic relationships, men have invested very little; yet if a woman were to become pregnant, she would invest a great deal in any resulting offspring. Over time, men invest more resources and effort in romantic relationship, as in a marriage. Across species, when males invest little in offspring, only females are choosy about mating; as males invest more, they become increasingly selective. Indeed, in some species of birds (such as phalaropes) and fish (such as seahorses), males invest more than females in the offspring, and everything changes (males become more choosy about mating; females become more competitive with one another to obtain males). Many species, including humans, have more than one mating option (Gangestad et al., 2007). As in most other mammals, it is possible for human males to make little investment, in which case, females ought to select males based on evidence of "good genes" relative to other males (using cues such as appearance and behavior, as peahens do in distinguishing between peacocks). When, alternatively, human mating involves high levels of male commitment, men ought to be more selective, as in other species with high male parental investment.

We tested the hypothesis that human mating preferences show contingent changes with level of commitment in several ways. In some studies, we simply asked men and women to specify their "minimum standards" for various characteristics in single dates, sexual partners, steady partners, or marriage partners. Men and women specified similarly high standards for marriage partners, but very different standards for sexual partners (Kenrick et al., 1990). Indeed, in considering one-night stands, men were

willing to consider partners below their normal standards for a date (Kenrick et al., 1993).

Although both sexes became more demanding for long-term partners, there were still differences in the particular traits they most valued. In particular, women placed much more emphasis on status and wealth in long-term partners; men placed somewhat more importance on physical attractiveness. Other evolutionarily oriented researchers have found evidence that physical attractiveness is linked to cues signaling a woman's fertility, whereas status and wealth are linked to men's ability to generate resources. In later research, my former student Norm Li, who had been trained in economics before switching to social psychology, devised a method for distinguishing between traits people regard as "luxuries" versus "necessities." Previous researchers had often simply asked people what they most preferred in mates, without considering what those people might actually be able to "afford." If you simply ask someone what they prefer in a mate, both men and women are attracted to someone who is as physically attractive as a movie star, as creative and funny as a comedian, as warm as a loving grandmother, and as wealthy as Bill Gates. Indeed, when Norm Li gave participants a high budget of "mating dollars," men and women differed only slightly in their criteria for mates. But what if they are given a more realistic budget, one like most mortal humans actually confront when choosing a mate? With a limited budget, when choosing a lot of x meant compromising on y , the sexes expressed very different preferences. Women regarded social status as a "necessity," and were willing to compromise on physical attractiveness. Men regarded good looks as a necessity, and were willing to disregard status and other characteristics that would otherwise be desirable (Li and Kenrick, 2006; Li et al., 2002).

In recent years, there have been a number of other findings especially difficult to explain except in light of an evolutionary perspective. For example, several recent

studies suggest that when females are ovulating (and most likely to become pregnant), they are more attracted to men showing indications of "good genes," including high levels of masculinity and symmetry (e.g., Penton-Voak et al., 2003). Ovulating women are also more interested in extra-pair sexual relations, particularly with men more attractive than their long-term partners (Pillsworth and Haselton, 2006). Most interestingly, ovulating women are more likely to express attraction toward the smell of highly masculine and symmetrical men, based merely on exposure to t-shirts worn by these men (Gangestad and Thornhill, 1998).

Why the greater emphasis on physical attractiveness for extra-pair sexual liaisons than for long-term relationships? Presumably, symmetry and high masculinity in men, like colorful and symmetrical displays in peacocks, indicates the possession of genetic traits well-suited to survival. If a female has a casual sexual liaison, it could result in the transmission of the attractive male's beneficial genes to offspring, even though the male will not be around to provide resources. Hence, females are presumed to face some degree of trade-off between choosing a male who will stay around and provide resources and one who is highly attractive to other females. Concealed extra-pair liaisons with highly attractive males during the period of maximal fertility are presumably a way of attempting to have it both ways. Presumably, such affairs are temporally limited and hidden because they result in potentially high cost (loss of the long-term partner) if discovered. None of this is consciously mediated, and the cyclic effects are not found for women on birth-control pills (which change normal hormonal patterns).

BIASES IN INFORMATION PROCESSING: THE OLD LOOK

The evolutionary perspective was emerging during the period when the social cognitive

perspective came to predominate in social psychology. When Jim Sherman and I debated the relevance of an evolutionary perspective during the early 1980s, we perhaps viewed the two perspectives as alternative paradigms for our field. Many researchers have since realized that cognitive and evolutionary perspectives are not only compatible, but that a full understanding of human nature requires an integration of the two perspectives (e.g., Cosmides and Tooby, 1992; Haselton and Nettle, 2006; Kenrick, 1994, 2001).

An appealing feature of cognitive methods is that they can help bypass social desirability biases. If a research participant fills out a questionnaire regarding mating behaviors or preferences, he or she may strategically choose a socially appropriate response. But if challenged to identify a particular type of face in a crowd as rapidly as possible, or to say whether he or she remembers seeing a face previously flashed on the screen, or to judge the frequency of faces in a particular category, the bias is to try to get the right answer. Along with several students and colleagues, I have been involved in a program of studies examining the implications of evolutionary ideas for processes of attention, encoding, and retrieval (e.g., Ackerman et al., 2009; Kenrick et al., 1994, 2007; Roese et al., 2006). We find that people selectively attend to other people with functionally relevant characteristics, and are better at encoding and remembering those individuals.

Harking back to Darwin's evolutionary view of emotional expressions, another person's anger is a very functionally relevant stimulus, foreboding potential harm. If that other person is a male, the dangers are substantially higher: males commit approximately 90 percent of homicides, and even when females do commit murder, it is often a self-protective response to male harassment (Daly and Wilson, 1988). Males are also more likely to do physical damage if they hit or strike someone (Archer, 2000). Women, on the other hand, are more likely to afford cooperative opportunities to others (e.g., Taylor et al., 2000). Consistent with

functionally derived hypotheses about emotion recognition, we found that participants not only identify whether a face is angry or happy in a fraction of a second, but are significantly quicker and more accurate at recognizing anger on a male than a female face, and happiness on a female face (Becker et al., 2007). Interestingly, the reverse is also true, people can identify a face as male or female in a fraction of a second, but are significantly faster at identifying a face as male if it is angry and as female if it is happy. To deal with the possible confound that men might be better at expressing anger and women better at expressing happiness, we showed participants computer-generated faces made to look like males or females, and found the same effects. If a face was completely androgynous, but made to look slightly angry, participants judged it to be a man. Initially, some reviewers were confident these effects were products of cultural stereotypes associated with masculinity and femininity. We were able to design an experiment to pit the alternative explanations against one another. When we created androgynous faces and gave them masculine cultural cues (a suit and tie), they were judged as highly masculine, but not angry. When the identical faces bore feminine cultural cues (a blouse, earrings, and necklace), they were judged as feminine, but not happy. One critical cue was the masculine eyebrow ridge (which is lowered in males and in anger). When the androgynous face was given a lower eyebrow ridge, it was judged as significantly more angry, but only slightly more masculine (not as much as when it wore a suit and tie).

In another series of studies linking attention, encoding, and memory, participants were shown crowds of faces containing attractive and average-looking people of both sexes, and later asked to remember which faces they saw. Using an eyetracker, we found that both sexes looked relatively more at attractive women. When asked to judge the frequency of attractive women in a crowd, both sexes overestimated the number if the

crowds were presented rapidly, suggesting that attractive women are more immediately cognitively accessible. Consistently, people of both sexes are good at remembering whether or not they have seen a particular attractive woman before (Maner et al., 2003). For attractive male targets, on the other hand, there is an interesting disjunction between different levels of processing. In several studies, we found that women (but not men) attend more to good-looking men, but then neither remember them well, nor overestimate their frequency in previously presented crowds. These findings make sense in terms of male and female mating strategies discussed earlier. For men, strange attractive women are mating opportunities. For women, who are typically more inclined toward long-term relationships with familiar men, and who typically do not make advances toward strangers, there is less inclination to devote processing time to strange attractive men.

Another study in this program used functional logic to reconsider a bias in memory linked to “outgroup homogeneity” and the failure to remember individual members of other racial groups. In functional terms, it makes sense to make finer discriminations between members of one’s own group (with whom one has more frequent and more cooperative interactions). However, there are also times when it makes functional sense to pay close attention to, and remember the identity of, members of other groups – when those individuals might pose threats to one’s welfare (members of outgroups often have less motivation to inhibit aggressive inclinations than members of one’s own group). Consistent with that logic, we found the standard outgroup homogeneity effect when white participants were asked to remember faces of emotionally neutral black men, but a tendency to remember black men especially well if they were angry (Ackerman et al., 2006). When cognitive resources were limited, in fact, the outgroup homogeneity effect was reversed for angry faces, with white subjects remembering angry black men better than white men.

AFFECT, COGNITION, AND THE MODULAR MIND

Evolutionarily informed research on learning and memory suggests that brains are composed of a number of specialized “domain-specific” mechanisms. For example, birds use different memory systems and different rules to remember species song, tastes of poisonous food, and locations of food caches (Sherry and Schacter, 1987). Many birds learn their species song during an early critical period, then reproduce it perfectly during the next breeding season, without ever having practiced it. On the other hand, birds learn the characteristics of poisonous foods in a single trial during any time of life. Following still different rules, food locations are learned, updated, and erased on a daily basis. Using the same decision rules for each of these problems would be highly inefficient, and different memory systems in birds are anatomically distinct. Likewise, humans inherit different memory systems to deal with different, sometimes conceptually incompatible tasks, including learning language, learning to avoid poisonous foods, and remembering other people’s faces.

The notion of domain specificity has profound implications for studying the links between motivation and social cognition (Kenrick and Shiota, 2008; Kenrick et al., 2003; Neuberg et al., 2010). This approach assumes cognitive processes differ qualitatively depending on which fundamental social motive is currently active (see Table 1.1). Thinking in these terms leads to a number of hypotheses about how attention, memory, and downstream social interactions change in functionally specific ways as a function of activating motives such as self-protection, mating, or affiliation.

For example, in line with our earlier discussion of sex differences in parental investment and sexual selection, activating mating goals leads men to erroneously perceive sexual arousal (but not other emotional states) in photographs of neutrally expressive

Table 1.1 Adaptive problems domains and examples of associated decision constraints

<i>Social problem domain</i>	<i>Example of decision bias</i>
Coalition formation	Cooperate with those who (a) share your genes, (b) have cooperated with you in past
Status seeking	Males will take more risks to gain and maintain status
Self-protection	Potential threats or costs trigger reciprocal aggressive behavior, particularly among non-kin
Mate choice	Males, compared to females, generally more inclined toward unrestricted sexual strategy
Mating relationship maintenance	Breaking a bond likely for: (a) males when a mate is sexually unfaithful or when physically attractive alternatives are available (b) females when a mate compromises resources or when high status alternatives are available
Parental care	Familial provision of resources and care will follow the order: (a) self > siblings (b) own offspring > stepchildren

attractive women. No such effect emerges among female perceivers (Maner et al., 2005). Activating self-protection goals, on the other hand, triggers particular outgroup stereotypes connoting threat or danger, but not equally negative, but threat-irrelevant stereotypes. For example, white North Americans who are in a dark room, or have just seen a fear-inducing movie, view black men, but not whites or black women, as dangerous or angry (e.g., Maner et al., 2005; Schaller et al., 2003).

Fundamental goals that seem, on the surface, specific to a narrow kind of behavior (e.g., mating) may have functional consequences spanning a broad range of behavior. For example, activating self-protective goals enhances conformity among both men and women; activating mating goals enhances conformity among women but leads to counter-conforming behavior among men (Griskevicius et al., 2006b). Another series of studies finds that activating short-term mating goals leads men, but not women, to be more creative (Griskevicius et al., 2006a). These findings all fit with predictions about differential consequences of different social behaviors depending on the functional goal accessible at the time.

Interesting questions arise from considering implications of this approach for other views of motivation (Kenrick, Griskevicius, et al., 2010). For example, regulatory focus

theory (Higgins, 1998) links prevention-based thoughts and behaviors to the goal of security and promotion-based thoughts and behaviors to nurturance. This would suggest that the domain of self-protection is more closely related to prevention goals, and coalition and parental care to promotion goals. At the same time, each of the domains described in Table 1.1 probably involves, to some degree, both promotion and prevention subgoals (finding new friends versus avoiding loss of existing friends, encouraging bonds with a mating partner versus avoiding infidelity, etc.). There are potentially productive empirical questions involving the extent to which similar or different types of cognitive biases are associated with different types of promotion and prevention in particular domains.

SOCIAL DYNAMICS AND THE EMERGENCE OF CULTURE: THE MIND AS A COLORING BOOK

Evolutionary social psychology is one component of a broader interdisciplinary movement toward integrative science (Kenrick, 2001; Wilson, 1998). Dynamical systems theory is another attempt to combine ideas from biology, mathematics, computer science, and cognitive psychology (Latané, 1996;

Nowak and Vallacher, 1998). Two fascinating conclusions have emerged from research on complex systems ranging from ant colonies to world economies. Dazzling fractal images represent one lesson of dynamical systems theory: awe-inspiring complexity can emerge from a few variables interacting according to a few simple rules. Such emergent processes can be intellectually overwhelming, but there is another side to the story. Complex systems at every level from molecules to genes to neurons to ecosystems often reveal *self-organization* – order emerging out of initial disorder.

Exactly which patterns of self-organization arise in complex systems depends on simple decision rules at the individual level. Combined with ideas from evolutionary modularity (e.g., biases in Table 1.1), this helps explain why different patterns of social organization arise in different domains of social life, such as the hierarchies associated with status, dyadic pairings associated with mating, and large aggregations associated with intergroup conflict (Kenrick et al., 2003). This work has important implications for another area of modern social psychology: the emergence of cultural norms (Harton and Bourgeois, 2004; Kenrick et al., in press, b). A dynamical evolutionary approach helps us understand how cultural norms encompass underlying universal biases as well as diverse cultural norms. For example, most human societies involve more marriages between older men and younger women than the converse, as discussed. There are, however, a few societies, such as Tiwi aborigines of Australia, in which older women marry younger men. On closer examination, those societies do not involve exceptions to underlying human nature, but a new dynamic emerging from multiple evolved biases. Men in Tiwi society are in fact attracted to younger women, and Tiwi women are attracted to older high status men. The Tiwi arrangement is an accommodation to another general feature of human social life – male competition for status and emergent patterns of patriarchy. In Tiwi society, older men have all the

power, and, since they are polygynous, they broth their daughters to other patriarchs, who reciprocate. This leaves the younger men out of luck. Another norm in Tiwi society, however, is that all women – of all ages – must be married. When an older woman becomes widowed, she must remarry, but the older patriarchs with younger wives do not wish to marry her. Instead, she marries a younger man, who thereby does a favor for her older male relatives, and puts himself into the game of acquiring younger wives in the future. These arrangements suggest that the mind–culture interaction is neither captured by the “blank slate” metaphor or by the metaphor of a predetermined blueprint. Instead, we can fruitfully conceptualize the interaction using the metaphor of a coloring book, which has some pre-drawn lines, but a great deal of flexibility within which local colors can be recombined in diverse ways.

CONCLUSION: APPLYING EVOLUTIONARY PRINCIPLES TO SOCIAL LIFE

There is now abundant evidence supporting the premise that human social behaviors reflect the influence of physical and psychological predispositions that helped our ancestors survive and reproduce. Social psychology is increasingly moving to the center of the interdisciplinary synthesis constituting evolutionary psychology. Researchers are realizing that understanding human nature also has implications for social problems at all levels – from family conflicts and neighborhood crime to overpopulation and international conflict (Crawford and Salmon, 2004; Penn, 2003). Social psychologists have begun applying functional analyses to a wide range of social issues and interpersonal problems, including prejudice and intergroup conflict (Kurzban et al., 2001; Navarette et al., 2009), sexual harassment (Haselton and Buss, 2000; Kenrick et al., 1996b), homicide (Daly and Wilson, 1988; Kenrick and Sheets,

1994), and political conflict (Sidanius and Pratto, 1999; Weeden et al., 2008). An evolutionary approach also has implications for understanding clinical disorder (Keller and Nesse, 2006) and positive psychology (Buss, 2000; Kenrick et al., submitted).

Consider for example, the topic of prejudice. Although social psychologists typically conceived of prejudice as “negative affect,” Cottrell and Neuberg (2005) fruitfully expanded this conception by applying notions of functional domain specificity. They presented evidence that different social groups (e.g., African American males, Asian females, gay men, religious fundamentalists) often elicited qualitatively different emotional reactions linked to the particular functional threats posed by those groups. Other research suggests stereotypical and prejudicial responses to particular groups vary in functional ways with the nature of the threat and the functional context. For example, Schaller et al. (2003) found that being in a dark room specifically enhanced perceived physical threats from black and Arab men, but did not enhance other stereotypical negative characteristics. And although prejudicial responses to these groups appear to be linked to fear, prejudicial responses to other groups (such as people from little-known foreign countries) are more associated with concerns about disease (Faulkner et al., 2004). As these authors point out, attempts to reduce prejudice and discrimination will be more successful to the extent that they target the particular functional motivations and the particular threats associated with particular groups.

An exciting recent development involves bridges between evolutionary psychology and economics (e.g., Ermer et al., 2007; Kenrick et al., 2008, in press, a; Wilson and Daly, 2003). As one example, a classic finding from behavioral economics involves “loss aversion” (Tversky and Kahneman, 1991). People are willing to pay more to ensure against the loss of \$100 they already have in their wallets than to gain an additional \$100. In one recent series of studies, we find that activating mating or status motives erases

loss aversion (so that participants now value a given-sized gain as much as, or more than, an equal loss). In line with findings reviewed earlier that mating strategy involves more competitive showing off for males than for females, this mating-based reversal of loss aversion was found only for males (Li and Kenrick, submitted).

Thus, the evolutionary perspective scores highly on two key criteria for powerful scientific theories: comprehensiveness and heuristic potential. I have only touched on a somewhat idiosyncratic sampling of the broad range of hypotheses generated by this perspective (see Crawford and Krebs, 2008; Dunbar and Barrett, 2007, and Schaller et al., 2006 for a diverse sampling of other perspectives). Studying social psychology from an evolutionary perspective has proven an ideal occupational choice for a curious young person who couldn’t decide what he wanted to be when he grew up. Indeed, reviewing the history of this perspective reignites my postadolescent awe at the number of fascinating questions about human social psychology yet to be asked.

REFERENCES

Ackerman, J.M., Shapiro, J.R., Neuberg, S.L., Kenrick, D.T., Becker, D.V., Griskevicius, V., Maner, J.K. and Schaller, M. (2006) They all look the same to me (unless they’re angry): From out-group homogeneity to out-group heterogeneity. *Psychological Science*, 17, 836–840.

Ackerman, J.M., Becker, D.V., Mortensen, C.R., Sasaki, T., Neuberg, S.L. and Kenrick, D.T. (2009) A pox on the mind: Disjunction of attention and memory in the processing of physical disfigurement. *Journal of Experimental Social Psychology* 45, 478–485.

Alcock, J. and Crawford, C. (2008) Evolutionary questions for evolutionary psychologists. In C. Crawford and D.L. Krebs (eds), *Foundations of Evolutionary Psychology*, pp. 25–45. Mahwah NJ: Lawrence Erlbaum Associates.

Allen-Arave, W., Gurven, M. and Hill, K. (2008) Reciprocal altruism, rather than kin selection, maintains nepotistic food transfers on an Ache reservation. *Evolution and Human Behavior*, 29, 305–318.

Archer, J. (2000) Sex differences in aggression between heterosexual partners: A meta-analytic review. *Psychological Bulletin*, 126, 651–680.

Becker, D.V., Kenrick, D.T., Guerin, S. and Maner, J.K. (2005) Concentrating on beauty: Sexual selection and sociospatial memory. *Personality and Social Psychology Bulletin*, 31, 1643–1652.

Becker, D.V., Kenrick, D.T., Neuberg, S.L., Blackwell, K.C. and Smith, D.M. (2007) The confounded nature of angry men and happy women. *Journal of Personality and Social Psychology*, 92, 179–190.

Brewer, M.B. and Capra, L.R. (2006) An evolutionary perspective on social identity: Revisiting groups. In M. Schaller, J. Simpson, and D.T. Kenrick (eds), *Evolution and Social Psychology*, pp. 143–161. New York: Psychology Press.

Buss, D.M. (1989) Sex differences in human mate preferences: Evolutionary hypotheses tested in 37 cultures. *Behavioral and Brain Sciences*, 12, 1–49.

Buss, D.M. (2000) The evolution of happiness. *American Psychologist*, 55, 15–23.

Cameron, C., Oskamp, S. and Sparks, W. (1977) Courtship American style – newspaper ads. *Family Coordinator*, 26, 27–30.

Cialdini, R.B., Brown, S.L., Lewis, B.P., Luce, C. and Neuberg, S.L. (1997) Reinterpreting the empathy–altruism relationship: When one into one equals oneness. *Journal of Personality and Social Psychology*, 73, 481–494.

Cosmides, L. and Tooby, J. (1992) Cognitive adaptations for social exchange. In J.H. Barkow, L. Cosmides, and J. Tooby (eds), *The Adapted Mind*, pp. 163–228. New York: Oxford University Press.

Cottrell, C.A. and Neuberg, S.L. (2005) Different emotional reactions to different groups: A sociofunctional threat-based approach to 'prejudice'. *Journal of Personality and Social Psychology*, 88, 770–789.

Crawford, C.B. and Krebs, D. (2008) *Foundations of Evolutionary Psychology*. New York: Erlbaum.

Crawford, C.B. and Salmon, C. (2004) *Evolutionary Psychology, Public Policy, and Personal Decisions*. Mahwah, NJ: Lawrence Erlbaum Associates.

Daly, M. and Wilson, M.I. (1983) *Sex, Evolution and Behavior: Adaptations for Reproduction* (2nd edition). Boston, MA: Willard Grant Press.

Daly, M. and Wilson, M.I. (1988) *Homicide*. New York: Aldine deGruyter.

Darwin, C. (1872) *The Expression of Emotions in Man and Animals*. London: Murray.

Deci, E.L. and Ryan, R.M. (2000) The 'what' and 'why' of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, 11, 227–268.

Dunbar, R.I.M. and Barrett, L. (2007) *The Oxford Handbook of Evolutionary Psychology*. Oxford: Oxford University Press.

Eagly, A.H. and Wood, W. (1999) The origins of sex differences in human behavior: Evolved dispositions versus social roles. *American Psychologist*, 54, 408–423.

Eibl-Eibesfeldt, I. (1973) The expressive behavior of the deaf-and-blind-born. In M. von Cranach and I. Vine (eds), *Social Communication and Movement*, pp. 163–194. New York: Academic Press.

Eibl-Eibesfeldt, I. (1975) *Ethology: The Biology of Behavior* (2nd edition). New York: Holt, Rinehart, and Winston.

Ermer, E., Cosmides, L. and Tooby, J. (2008) Relative status regulates risky decision making about resources in men: Evidence for the co-evolution of motivation and cognition. *Evolution and Human Behavior*, 29, 106–118.

Faulkner, J., Schaller, M., Park, J.H. and Duncan, L.A. (2004) Evolved disease-avoidance mechanisms and contemporary xenophobic attitudes. *Group Processes and Intergroup Relations*, 7, 333–353.

Fiske, S.T., Cuddy, A.J.C. and Glick, P. (2007) Universal dimensions of social perception: Warmth and competence. *Trends in Cognitive Science*, 11, 77–83.

Gangestad, S.W., Garver-Apgar, C.E., Simpson, J.A. and Cousins, A.J. (2007) Changes in women's mate preferences across the ovulatory cycle. *Journal of Personality and Social Psychology*, 92, 151–163.

Gangestad, S.W. and Thornhill, R. (1998) Menstrual cycle variation in women's preferences for the scent of symmetrical men. *Proceedings of the Royal Society of London B*, 265, 727–733.

Garcia, J. and Koelling, R.A. (1966) Relation of cue to consequence in avoidance learning. *Psychonomic Science*, 4, 123–124.

Gardner, H. (1985) *The Mind's New Science: A History of the Cognitive Revolution*. New York: Basic Books.

Geary, D.C. (1998) *Male, Female: The Evolution of Human Sex Differences*. Washington, DC: American Psychological Association.

Griskevicius, V., Cialdini, R.B. and Kenrick, D.T. (2006a) Peacocks, Picasso, and parental investment: The effects of romantic motives on creativity. *Journal of Personality and Social Psychology*, 91, 63–76.

Griskevicius, V., Goldstein, N., Mortensen, C., Cialdini, R.B. and Kenrick, D.T. (2006b) Going along versus going alone: When fundamental motives facilitate strategic (non)conformity. *Journal of Personality and Social Psychology*, 91, 281–294.

Griskevicius, V., Tybur, J.M., Sundie, J.M., Cialdini, R.B., Miller, G.F. and Kenrick, D.T. (2007) Blatant benevolence and conspicuous consumption: When romantic motives elicit strategic costly signals. *Journal of Personality and Social Psychology*, 93, 85–102.

Harton, H.C. and Bourgeois, M.J. (2004) Cultural elements emerge from dynamic social impact. In M. Schaller and C. Crandall (eds), *The Psychological Foundations of Culture*, pp. 41–75. Mahwah, NJ: Lawrence Erlbaum Associates.

Haselton, M.G. and Buss, D.M. (2000) Error management theory: A new perspective on biases in cross-sex mind reading. *Journal of Personality and Social Psychology*, 78, 81–91.

Haselton, M.G. and Nettle, D. (2006) The paranoid optimist: An integrative evolutionary model of cognitive bias. *Personality and Social Psychology Review*, 10, 47–66.

Higgins, E.T. (1998) Promotion and prevention: regulatory focus as a motivational principle. In M.P. Zanna (ed.), *Advances in Experimental Social Psychology*, 30, 1–46. New York: Academic Press.

Higgins, E.T. and Pittman, T.S. (2008) Motives of the human animal: Comprehending, managing, and sharing inner states. *Annual Review of Psychology*, 59, 361–385.

Huang, J.Y. and Bargh, J.A. (2008) Peak of desire: activating the mating goal changes life-stage preferences across living kinds. *Psychological Science*, 19, 573–578.

Jankowiak, W.R. and Fischer, E.F. (1992) A cross-cultural perspective on romantic love. *Ethnology*, 31, 149–155.

Kaplan, H.S., Hill, K., Lancaster, J.B. and Hurtado, A.M. (2000) A theory of human life history evolution: Diet, intelligence, and longevity. *Evolutionary Anthropology*, 9, 156–185.

Keller, M.C. and Nesse, R.M. (2006) The evolutionary significance of depressive symptoms: Different adverse situations lead to different depressive symptom patterns. *Journal of Personality and Social Psychology*, 91, 316–330.

Kenrick, D.T. (1994) Evolutionary social psychology: From sexual selection to social cognition. In M.P. Zanna (ed.), *Advances in Experimental Social Psychology*, 26, 75–121. San Diego, CA: Academic Press.

Kenrick, D.T. (2001) Evolution, cognitive science, and dynamical systems: An emerging integrative paradigm. *Current Directions in Psychological Science*, 10, 13–17.

Kenrick, D.T. (2006) Evolutionary psychology: Resistance is futile. *Psychological Inquiry*, 17, 102–108.

Kenrick, D.T., Delton, A.W., Robertson, T., Becker, D.V. and Neuberg, S.L. (2007) How the mind warps: A social evolutionary perspective on cognitive processing disjunctions. In J.P. Forgas, M.G. Haselton, and W. Von Hippel (eds), *The Evolution of the Social Mind: Evolution and Social Cognition*, pp. 49–68. New York: Psychology Press.

Kenrick, D.T., Gabrielidis, C., Keefe, R.C. and Cornelius, J. (1996a) Adolescents' age preferences for dating partners: Support for an evolutionary model of life-history strategies. *Child Development*, 67, 1499–1511.

Kenrick, D.T., Griskevicius, V., Neuberg, S.L. and Schaller, M. (2010) Renovating the pyramid of needs: Contemporary extensions built upon ancient foundations. *Perspectives on Psychological Science*, 5, 292–314.

Kenrick, D.T., Griskevicius, V., Sundie, J.M., Li, N.P., Li, Y.J. and Neuberg, S.L. (2009) Deep rationality: The evolutionary economics of decision-making. *Social Cognition*, 27, 764–785.

Kenrick, D.T., Groth, G.R., Trost, M.R. and Sadalla, E.K. (1993) Integrating evolutionary and social exchange perspectives on relationships: Effects of gender, self-appraisal, and involvement level on mate selection criteria. *Journal of Personality and Social Psychology*, 64, 951–969.

Kenrick, D.T. and Keefe, R.C. (1992) Age preferences in mates reflect sex differences in human reproductive strategies. *Behavioral and Brain Sciences*, 15, 75–133.

Kenrick, D.T., Li, N.P. and Butner, J. (2003) Dynamical evolutionary psychology: Individual decision-rules and emergent social norms. *Psychological Review*, 110, 3–28.

Kenrick, D.T., Neuberg, S.L., Zierk, K. and Krones, J. (1994) Evolution and social cognition: Contrast effects as a function of sex, dominance, and physical attractiveness. *Personality and Social Psychology Bulletin*, 20, 210–217.

Kenrick, D.T., Nieuweboer, S. and Buunk, A.P. (2010) Universal mechanisms and cultural diversity: Replacing the blank slate with a coloring book. In M. Schaller, S. Heine, A. Norenzayan, T. Yamagishi and T. Kameda (eds), *Evolution, Culture, and the Human Mind*, pp. 257–271. New York: Psychology Press.

Kenrick, D.T., Sadalla, E.K., Groth, G. and Trost, M.R. (1990) Evolution, traits, and the stages of human courtship: Qualifying the parental investment model. *Journal of Personality*, 58, 97–116.

Kenrick, D.T. and Sheets, V. (1994) Homicidal fantasies. *Ethology and Sociobiology*, 14, 231–246.

Kenrick, D.T. and Shioya, M.N. (2008) Approach and avoidance motivation(s): An evolutionary perspective. In A.J. Elliot (ed.), *Handbook of Approach and Avoidance Motivation*, pp. 273–288. New York: Psychology Press.

Kenrick, D.T., Sundie, J.M. and Kurzban, R. (2008) Cooperation and conflict between kith, kin, and strangers: Game theory by domains. In C. Crawford and D. Krebs (eds), *Foundations of Evolutionary Psychology*, pp. 353–370. New York: Lawrence Erlbaum Associates.

Kenrick, D.T., Trost, M.R. and Sheets, V.L. (1996b) The feminist advantages of an evolutionary perspective. In D.M. Buss and N. Malamuth (eds), *Sex, Power, Conflict: Feminist and Evolutionary Perspectives*, pp. 29–53. New York: Oxford University Press.

Kurzban, R., Tooby, J. and Cosmides, L. (2001) Can race be erased? Coalitional computation and social categorization. *Proceedings of the National Academy of Sciences*, 98, 15387–15392.

Lancaster, J.B. (1975) *Primate Behavior and the Emergence of Human Culture*. New York: Holt, Rinehart, and Winston.

Latané, B. (1996) Dynamic social impact: The creation of culture by communication. *Journal of Communication*, 46, 13–25.

Leary, M.R., and Baumeister, R.F. (2000) The nature and function of self-esteem: Sociometer theory. *Advances in Experimental Social Psychology*, 32, 1–62.

Li, N.P., Bailey, J.M., Kenrick, D.T. and Linsenmeier, J.A. (2002) The necessities and luxuries of mate preferences: Testing the trade-offs. *Journal of Personality and Social Psychology*, 82, 947–955.

Li, N.P. and Kenrick, D.T. (2006) Sex similarities and differences in preferences for short-term mates: What, whether, and why. *Journal of Personality and Social Psychology*, 90, 468–489.

Li, Y.J., Kenrick, D.T., Griskevicius, V. and Neuberg, S.L. (2010, June) *The Evolutionary Roots of Decision Biases: Erasing and Exacerbating Loss Aversion*. Paper presented at the conference for the Human Behavior and Evolution Society, Eugene, Oregon.

Lummaa, V. (2007) Life-history theory, reproduction and longevity in humans. In R.I.M. Dunbar and L. Barrett (eds), *Oxford Handbook of Evolutionary Psychology*, pp. 397–413. Oxford: Oxford University Press.

Maner, J.K., Kenrick, D.T., Becker, D.V., Delton, A.W., Hofer, B., Wilbur, C.J. and Neuberg, S.L. (2003) Sexually selective cognition: Beauty captures the mind of the beholder. *Journal of Personality and Social Psychology*, 6, 1107–1120.

Maner, J.K., Kenrick, D.T., Becker, D.V., Robertson, T.E., Hofer, B., Neuberg, S.L., Delton, A.W., Butner, J. and Schaller, M. (2005) Functional projection: How fundamental social motives can bias interpersonal perception. *Journal of Personality and Social Psychology*, 88, 63–78.

Mikulincer, M. and Shaver, P.R. (2006) The behavioral systems construct: A useful tool for building an integrative model of the social mind. In P.A.M. Van Lange (ed.), *Bridging Social Psychology*, pp. 279–284. Mahwah, NJ: Erlbaum.

Murstein, B.I. (1981) Process, filter, and stage theories of attraction. In M. Cook (ed.), *The Bases of Human Sexual Attraction*, pp. 179–211. San Diego: Academic Press.

Navarrete, C.D., Olsson, A., Ho, A.K., Mendes, W.B., Thomsen, L. and Sidanius, J. (2009) Fear extinction to an out-group face: The role of target gender. *Psychological Science*, 20, 155–158.

Neuberg, S.L., Kenrick, D.T. and Schaller, M. (2010) Evolutionary social psychology. pp. 761–796 in S.T. Fiske, D.T. Gilbert, and G. Lindzey (eds), *Handbook of Social Psychology* (5th edition). New York: John Wiley and Sons.

Nowak, A. and Vallacher, R.R. (1998) *Dynamical Social Psychology*. New York: Guilford.

Ohman, A. and Mineka, S. (2001) Fears, phobias, and preparedness: Toward an evolved module of fear and fear learning. *Psychological Review*, 108, 483–522.

Penn, D.J. (2003) The evolutionary roots of our environmental problems: Toward a Darwinian ecology. *Quarterly Review of Biology*, 78, 275–301.

Penton-Voak, I.S., Little, A.C., Jones, B.C., Burt, D.M., Tiddeman, B.P. and Perrett, D.J. (2003) Female condition influences preferences for sexual dimorphism in faces of male humans (*Homo sapiens*). *Journal of Comparative Psychology*, 117, 264–271.

Pillsbury, E.G. and Haselton, M.G. (2006) Male sexual attractiveness predicts differential ovulatory shifts in female extra-pair attraction and male mate retention. *Evolution and Human Behavior*, 27, 247–258.

Pinker, S. (1994) *The Language Instinct*. New York: Harper Collins.

Roese, N., Pennington, G.L., Coleman, J., Janicki, M., Li, N.P. and Kenrick, D.T. (2006) Sex differences in regret: All for love or some for lust? *Personality and Social Psychology Bulletin*, 32, 770–780.

Rowe, D.C. (1994) *The Limits of Family Experience: Genes, Experience, and Behavior*. New York: Guilford.

Sadalla, E.K., Kenrick, D.T. and Vershure, B. (1987) Dominance and heterosexual attraction. *Journal of Personality and Social Psychology*, 52, 730–738.

Schaller, M. and Murray, D.R. (2008) Pathogens, personality and culture: Disease prevalence predicts worldwide variability in sociosexuality, extraversion, and openness to experience. *Journal of Personality and Social Psychology*, 95, 212–221.

Schaller, M., Park, J.H. and Mueller, A. (2003) Fear of the dark: Interactive effects of beliefs about danger and ambient darkness on ethnic stereotypes. *Personality and Social Psychology Bulletin*, 29, 637–649.

Schaller, M., Simpson, J. and Kenrick, D.T. (eds) (2006) *Evolution and Social Psychology*. New York: Psychology Press.

Schmitt, D.P. and 118 members of International Sexuality Description Project. (2003) Universal sex differences in the desire for sexual variety: Tests from 52 nations, 6 continents, and 13 islands. *Journal of Personality and Social Psychology*, 85, 85–104.

Sedikides, C. and Skowronski, J.J. (1997) The symbolic self in evolutionary context. *Personality and Social Psychology Review*, 1, 80–102.

Segerstale, U. (2000) *Defenders of the Truth: The Battle for Science in the Sociology Debate and Beyond*. Oxford: Oxford University Press.

Seligman, M.E.P. and Hager, J.L. (1972) *Biological Boundaries of Learning*. New York: Appleton-Century-Crofts.

Sherman, P.W. (1988) The levels of analysis. *Animal Behavior*, 36, 616–619.

Sherry, D.F. and Schacter, D.L. (1987) The evolution of multiple memory systems. *Psychological Review*, 94, 439–454.

Sidanius, J. and Pratto, F. (1999) *Social Dominance: An Intergroup Theory of Social Hierarchy and Oppression*. New York: Cambridge University Press.

Skinner, B.F. (1953) *Science and Human Behavior*. New York: Free Press.

Stearns, S.C., Allal, N. and Mace, R. (2008) Life history theory and human development. In C. Crawford and D. Krebs (eds), *Foundations of Evolutionary Psychology*, pp. 47–70. New York: Lawrence Erlbaum Associates.

Taylor, S.E., Klein, L.C., Lewis, B.P., Gruenwald, T.L., Gurung, R.A.R. and Updegraff, J.A. (2000) Biobehavioral responses to stress in females: Tend-and-befriend, not fight-or-flight. *Psychological Review*, 107, 411–429.

Trivers, R.L. (1972) Parental investment and sexual selection. In B. Campbell (ed.), *Sexual Selection and the Descent of Man*, pp. 136–179. Chicago: Aldine-Atherton.

Tversky, A. and Kahneman, D. (1991) Loss aversion in riskless choice: A reference dependent model. *Quarterly Journal of Economics*, 106, 1039–1061.

Tybur, J.M., Miller, G.F. and Gangestad, S.W. (2008) Testing the controversy: An empirical examination of adaptationists' attitudes toward politics and science. *Human Nature: An Interdisciplinary Biosocial Perspective*, 18, 313–328.

Van Knippenberg, A. and Van Baaren, R. (2006) Baboons, brains, babies, and bonding: A multidisciplinary approach to mimicry. In P.A.M. Van Lange (ed.), *Bridging Social Psychology*, pp. 173–178. Mahwah, NJ: Erlbaum.

Van Vugt, M. and Van Lange, P.A.M. (2006) The altruism puzzle: Psychological adaptations for prosocial behavior. In M. Schaller, J.A. Simpson, and D.T. Kenrick (eds), *Evolution and Social Psychology*, pp. 237–261. Madison, CT: Psychology Press.

Weeden, J., Cohen, A.B. and Kenrick, D.T. (2008) Religious attendance as reproductive support. *Evolution and Human Behavior*, 29, 327–334.

Wilcoxon, H.C., Dragoin, W.B. and Kral, P.A. (1971) Illness-induced aversions in rat and quail: Relative salience of visual and gustatory cues. *Science*, 171, 826–828.

Wilson, M. and Daly, M. (2003) Do pretty women inspire men to discount the future? *Proceedings of the Royal Society London B*, 271, S177–S179.

Wilson, E.O. (1975) *Sociobiology*. Cambridge, MA: Harvard University Press.

Wilson, E.O. (1998) *Consilience: The Unity of Knowledge*. New York: Alfred A. Knopf.

Tend and Befriend Theory

Shelley E. Taylor

ABSTRACT

The “tend and befriend” theory builds on the observation that human beings affiliate in response to stress. Under conditions of threat, they tend to offspring to ensure their survival and affiliate with others for joint protection and comfort. These responses are underpinned by an affiliative neurocircuitry that appears to be based on oxytocin and endogenous opioid peptides. When close relationships are threatened or one is socially isolated, a rise in plasma oxytocin occurs, a biological marker that may signal a need for affiliation. Oxytocin prompts affiliative behavior in response to stress, in conjunction with the opioid system. Together with positive social contacts, oxytocin attenuates biological stress responses that would otherwise arise in response to social threats. These social responses to stress and their biological underpinnings appear to be more characteristic of women than men. This model helps to unravel puzzles not only in the research literature but also with respect to health and may shed light on why women tend to live longer than men.

Human beings’ social nature may have evolved in part because of the needs that we have for contact with others to ensure safety from threat. Whereas other species have thick skin, sharp teeth, quick reflexes, or camouflage to protect themselves, human beings have adopted group living as the primary solution to problems of survival and reproduction. Social isolation is both psychologically and physically toxic, and it is associated with a heightened risk of death and early mortality (Cacioppo and Hawkley, 2003; House et al., 1988). In addition to survival benefits, social relationships serve important regulatory functions. For example, contact with a nurturant caregiver early in life is essential for the development of biological stress regulatory systems (e.g., Repetti et al., 2002, 2007), and social support is protective across the lifespan (see Taylor, 2009, for a review).

INTRODUCTION

Social relationships are vital resources for managing the demands of the environment.

RESPONSES TO THREAT: FIGHT OR FLIGHT

Despite the centrality of the social group to human wellbeing and survival, research on stress and threat has, until recently, largely

ignored the importance of social mechanisms for addressing responses to threat. Instead, stress research has been guided heavily by the fight or flight metaphor, responses that may protect individuals but that are not social in nature. First articulated by Walter Cannon (1932), the fight or flight response has two components: a behavioral component and a biological component. In response to a threat, a person can become aggressive and mount an active or antagonistic response to the threatening circumstances, or the person can flee, either literally or metaphorically. Among the responses that contemporary stress researchers interpret as flight behavior are social withdrawal and substance use, especially alcohol and drug use.

The biological component of fight or flight depends on two interacting stress systems: the sympathetic nervous system (SNS) and the hypothalamic pituitary adrenocortical (HPA) axis. The actions of the sympathetic adrenomedullary (SAM) system are mediated primarily by the catecholamines norepinephrine and epinephrine, which produce, among other changes, increased heart rate and blood pressure, dilation of the airways, and enhanced availability of glucose and fatty acids for energy. Engagement of the HPA axis begins with the release of corticotropin-releasing hormone (CRH), which stimulates the secretion of adrenocorticotropin hormone (ACTH) by the anterior pituitary gland, resulting in the release of glucocorticoids such as cortisol. Cortisol helps to restore processes that prime homeostatic defense mechanisms, and as such, HPA axis reactivity modulates a wide range of functions, including energy release and immune activity.

These two systems are important because they account for the protective short-term effects of stress responses but also their long-term costs. In the short term, they shunt reserves of energy for fight or flight, and the subjective experience is often arousal and feelings of anger, fear, or anxiety. As such, these systems mobilize the body to meet the demands of pressing situations and then

engage homeostatic mechanisms that return the body to its previous functioning. With repeated or recurrent stress, however, these biological stress responses have long-term costs that have implications for health (McEwen, 1998). The theory of allostatic load (McEwen, 1998) maintains that repeated or chronic engagement of stress systems interacts with genetic vulnerabilities and acquired risks (such as poor health habits) to erode the resiliency of biological stress regulatory systems and increase the likelihood of disease. These include such chronic disorders as coronary heart disease, hypertension, type II diabetes, and some cancers. These long-term costs provide important conceptual underpinnings to the tend and befriend theory to be articulated, because, as will be seen, social contact during times of stress and social relationships more generally exert protective effects against these potential long-term costs of stress.

When stress researchers began to study stress in human beings, they borrowed from animal research in ways conducive to identifying fight or flight responses in humans (Taylor et al., 2000). Although fighting and fleeing are unquestionably part of the human repertoire for responding to threat, there are at least two reasons to suspect that they are unlikely to be the only or primary responses. First, fighting can leave vulnerable offspring at risk for predation. Likewise, fleeing with an infant or toddler might slow the caregiver down to be at enhanced risk for attack.

Human beings would not have survived as a species had they not developed stress responses that protected their offspring as well as self in times of danger. Humans evolved in small hunter-gatherer groups, and so coming together as a group instead of fighting or fleeing would provide greater defense as well as information about resource location for combating threats. In short, there are many reasons to believe that humans have used social relationships not only as a basic accommodation to the exigencies of life, but also as a primary resource for dealing with stressful circumstances.

RESPONSES TO THREAT: TEND AND BEFRIEND

To characterize these social responses to threat, we developed the theory “tend and befriend” (Taylor, 2002; Taylor et al., 2000). The theory maintains that under conditions of threat, tending to offspring and affiliating with others (befriending) are common responses in humans.

Development of the theory

The theory of tend and befriend was initiated in 1998 when the members of our laboratory group went to hear a prominent animal stress researcher present his work. He made the statement: “We shocked the animals and, of course, they all attacked each other.” This was an arresting statement, in large part because it is quite contrary to what one sees in human beings. When there is an immediate threat, human beings are more likely to affiliate with one another and to offer each other aid or solace than to attack one another. Accordingly, our lab group began a series of discussions to develop what would become the theory of tend and befriend as a characterization of human social responses to stress.

Our point of departure was evolutionary theory. A dilemma with any evolutionary-based theory of social behavior, however, is the potential for “just so” stories, a reference to Rudyard Kipling’s fables about how the leopard got his spots, the elephant his trunk, and the like. Such stories can seem plausible, but have no basis in fact. Accordingly, in building our theory, we imposed constraints on its development. Our method of constraint was to build parallel and mutually constraining biological and behavioral models. For every hypothesis we developed that generated a behavioral statement for the tend and befriend theory, we required evidence at the biological level. The reverse was also true. Any biological literature that yielded a potential insight for the theory was constrained by a requirement of evidence at the

behavioral level. Using these criteria, we jettisoned a number of otherwise promising hypotheses early on. Thus, the following characterization of the theory meets the criteria for evidence at both the biological and behavioral levels.

Overview of the theory

Drawing on animal and human studies, we maintain that there is an affiliative neurocircuitry that monitors the adequacy of social contact in light of the demands of the environment and prompts affiliation when necessary, regulating social approach behavior. We suggest that this system works in much the same way as occurs for other appetitive needs. That is, just as people have basic needs of hunger, thirst, and sexual drive, they also need to maintain an adequate level of protective and rewarding social relationships.

We hypothesize that there is a biological signaling system that comes into play when affiliations fall below an adequate level. Once signaled, the appetitive need is met through purposeful social behavior, such as affiliation. If social contacts are hostile and unsupportive, then psychological and biological stress responses are heightened and efforts toward social affiliation may be redoubled. If social contacts are supportive and comforting, stress responses decline. These positive contacts, then, lead to a decline in the need for affiliation and, in the context of stress, a decline in biological stress responses. The model is pictured in Figure 2.1.

On the biological level, the theory draws heavily on oxytocin and the opioid system. Oxytocin and endogenous opioid peptides are released in response to at least some stressors, especially those that trigger affiliative needs. Oxytocin prompts affiliative behavior in response to stress in conjunction with the opioid system, and oxytocin together with positive social contacts attenuates biological stress responses that otherwise arise in response to social threats. We suggest that

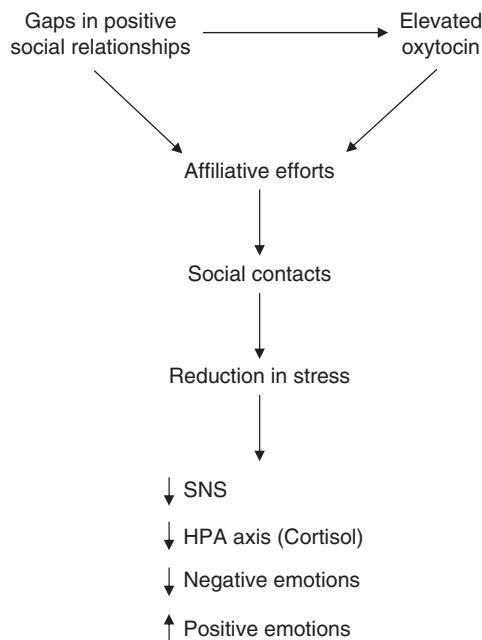


Figure 2.1 Tend and befriend

this oxytocin–opioid system is an appetitive system that regulates social approach behavior and recruits the neurocircuitry for reward in its enactment. Finally, we maintain that some of the health benefits associated with social support and social integration are mediated by this appetitive social approach system via attenuation of threat responses.

Tend and befriend responses to stress may be particularly characteristic of women. At the time when human stress responses evolved, generally thought to be the Pleistocene Era, tasks of daily living were largely sex-segregated, with men heavily responsible for protection and hunting and women primarily responsible for childcare and foraging. Consequently, women's responses to threat would have evolved so as to protect not only self but also immature offspring in their care. Affiliating with the social group for joint protection of self and offspring would have had substantial survival benefits, helping to ensure that offspring would reach an age when they could reproduce.

This is not to suggest that tend and befriend responses are exclusive to women. A large

literature indicates that under stress, both men and women turn to others for protection and solace (see Taylor, 2009, for a review). A gender difference exists such that women are somewhat more likely to seek and use social support in response to stress than men are (Tamres et al., 2002). However, this difference, although robust, is relatively modest in size, and thus, men's social responses to stress are well documented.

THEORETICAL PRINCIPLES AND EVIDENCE FOR TEND AND BEFRIEND

Biological signaling system: social pain/separation

Because affiliation is vital to the survival of human beings, there are likely to be biobehavioral mechanisms that are sensitive to social threats or loss of social contact, resulting in social distress and consequent efforts to establish or restore positive social contacts. One paradigm for such a system is separation distress, which has been studied primarily in young animals and human infants. When the young are separated from the caregiver, separation distress can result leading to distress vocalizations (e.g., crying in human infants) or active efforts to find the caregiver.

This system appears to depend in part on brain opioids. Evidence consistent with this pathway includes the fact that brain opioids reduce separation distress and opioid-based drugs, such as morphine, reduce distress vocalizations in animal offspring separated from the mother (Panksepp, 1998). Similarly, depriving animals of companionship can increase their consumption of exogenous opioids. Genetic evidence, likewise, suggests a role for opioids in the separation distress process. Mice that lack the μ -opioid receptor gene emit fewer distress vocalizations when separated from their mothers (Moles et al., 2004). Of interest, the opioid system is also vital to the experience and management of physical pain, and so researchers have

inferred that the experience of social pain and separation distress may build on the neurocircuitry for physical pain. Recent genetic and neuroimaging studies with humans have lent credibility to this idea (Eisenberger et al., 2003; Way and Taylor, 2011).

Oxytocin also appears to be implicated in distress due to social isolation. Adults as well as young children encounter gaps in their social relationships and may experience an analogue of separation distress. To address this point, in a laboratory study, we gave women measures of psychological and social functioning and related their responses to their levels of plasma oxytocin (Taylor et al., 2006a). The questionnaires assessed gaps in the women's relationships, including recent declines in contacts with significant others and how positive and negative their relationships were. Women who were experiencing gaps in their social relationships had high levels of oxytocin. They were more likely to report reduced contact with their mothers, with their best friend, with a pet, and with the social groups to which they belonged. Oxytocin levels were also related to the absence of positive relations with a partner. Specifically, women who reported that their husbands were not supportive, did not understand how they felt about things, and did not care for them, had higher levels of oxytocin. These women reported that they could not open up to their husbands if they needed to share their concerns. Poor quality of the marital relationship and infrequent displays of affection were also associated with high levels of oxytocin. These findings suggest that oxytocin is sensitive to the absence of positive significant social relationships. Similar results were reported by Turner et al. (1999) who found that elevated oxytocin was associated with anxiety over relationships, not being in a primary relationship, or having cold, intrusive relationships (see also Taylor et al., 2009).

It is of course possible that women with high levels of oxytocin are inclined to construe their social relationships as unsupportive. Some evidence argues against this

direction of causality: women who reported declines in contact with a pet and mother often had experienced their deaths, and women's oxytocin levels are unlikely to have caused these deaths. A recent animal study provides more definitive evidence regarding the direction of causality. Specifically, Grippo et al (2007) isolated female prairie voles and found that oxytocin levels increased in response to social isolation. In summary, social pain and separation are psychologically distressing, and the opioid and oxytocin systems are implicated in these responses.

Affiliation

The tend and befriend theory maintains that in response to either a psychological or biological impetus to affiliate or both, people seek contact with others. As an affiliative hormone, oxytocin may provide this impetus for social contact. Manifold evidence from animal studies shows that exogenous administration of oxytocin leads to affiliation in species as varied as rats, monkeys, and sheep. The injection of oxytocin leads to increases in maternal behavior, in grooming, and in other prosocial behavior (e.g., see Carter et al., 1999, for a review).

Opioid mechanisms also appear to be implicated in these processes. Administration of an opioid antagonist, for example, naloxone or naltrexone, results in less caregiving and protective behavior toward infants in rhesus monkeys (Martel et al., 1993), inhibits maternal behavior in sheep (Kendrick and Keverne, 1989), and diminishes the rewarding aspect of maternal cues in rats (Panksepp et al., 1999). Animals demonstrate a preference for other animals in whose presence they have previously experienced high levels of oxytocin and opioid activity (Panksepp, 1998). Administration of an opioid antagonist can suppress juvenile social behavior (Jalowiec et al., 1989), and opioid-blocking agents have been associated with reduced social activity and grooming in rhesus monkeys (Martel et al., 1993).

There is suggestive evidence that opioid-blocking agents may suppress human social behavior as well. Specifically, in one study (Jamner et al., 1998), administration of an opioid-blocking agent increased the amount of time that women chose to spend alone, reduced the amount of contact they had with friends, reduced the likelihood that they would contact their friends, and reduced the pleasantness of interactions with friends. Exogenous administration of oxytocin appears to enhance prosocial behavior and instill a sense of trust (Kosfeld et al., 2005; Zak et al., 2004). Thus, a broad array of affiliative behaviors appear to be subserved by oxytocin and opioid mechanisms in both animals and humans, although the animal data are more plentiful at this point in time.

nurturant, soothing immediate stimulation. On the short term, this contact reduces SAM and HPA axis responses to stress in the pups.

Over the long term, this maternal behavior results in a better regulated HPA axis response to stress and novelty, and better regulation of somatic growth and neural development, especially hippocampal synaptic development in the pup. These rat pups also show more open field exploration, which suggests lower levels of fear. This compelling animal model suggests that nurturant stimulation by the mother in response to stressful encounters modulates the responses of offspring to stress in ways that have permanent effects on the offspring's HPA axis responses to stress, on behavior suggestive of anxiety/fearfulness, and on cognitive function (see also Suomi, 1999).

TENDING AND RESPONSES TO STRESS

The original tend and befriend model maintained that maternal nurturance and tending under stressful conditions have specific advantages for offspring by protecting them from physical harm and increasing the likelihood that they will grow to adulthood and reproduce. Recent empirical developments indicate that the importance of tending to offspring in response to threat is far broader than the early theoretical statements would imply.

Animal studies

In an important animal model, Meaney and colleagues (Francis et al., 1999; Liu et al., 1997) explicitly linked early nurturant maternal contact following a stressful encounter to the development of stress responses in offspring. In their paradigm, infant rats are removed from the nest, stroked, and then returned to the nest. The response of the mother to this separation and reunification is intense licking and grooming and arched-back nursing, which provides the pup with

Human studies

Warm, nurturant, and supportive contact with a caregiver affects physiological and neuroendocrine stress responses in human infants and children just as in these animal studies. Early research on orphans reported high levels of emotional disturbance, especially depression, in infants who failed to receive nurturant stimulating contact from a caregiver (Spitz and Wolff, 1946). More recent findings from Eastern European abandoned infants confirm that, without the affectionate attentions of caregivers, infants may fail to thrive and many die (Carlson and Earls, 1997).

Building on observations such as these, attachment theory characterizes how vital early nurturant contact is to psychological and biological development (Bowlby, 1969/1982). Through the comforting behavior of attachment figures, typically parents, infants learn to understand and respond to the world. Bowlby regarded attachment behavior as regulated by an innate motivational system to ensure physical and psychological proximity to a caregiver. When a child experiences the attachment figure as responsive, effective self-regulation, exploratory behavior, and

normative biological responses to challenges may result. However, when there is a threat to this relationship, the child may react biologically and behaviorally with signs of stress and seek attention and comfort.

Attachment also moderates biological responses to stress. Studying 15-month-old children receiving well-baby examinations, Gunnar and her associates found that securely attached infants were less likely to show elevated cortisol responses to normal stressors such as inoculations than less securely attached infants (Gunnar et al., 1996; see also Nachmias et al., 1996). The protective effects of secure attachment were especially evident for socially fearful or inhibited children (see also Hart et al., 1996; Levine and Wiener, 1988; see Collins and Feeney, 2000, for a discussion of attachment in adult supportive relationships).

Early on, children depend critically on physical contact with primary caregivers, but over time they develop internal working models of these interactions that may buffer them when the primary caregiver is absent. If significant others are perceived to be warm, responsive, and available, a secure attachment will result. If significant others are cold, rejecting, unpredictable, or insensitive, however, an anxious or insecure attachment may result. Instead of being able to draw on an internal working model of caregivers for comfort and solace, the child may instead regulate his or her behavior by withdrawing from others or excessively demanding attention.

To a degree, the attachments laid down early in life provide a model for adult attachments (Fraley, 2002). That is, warm, nurturant contact with parents provides a model not only for stress responses and emotion regulation throughout childhood, but also for adult social relationships, and children who come from families in which they experienced warm, attentive behavior are more likely to develop the social skills that serve them well across their lifespan (Repetti et al., 2002). A broad array of evidence demonstrates that children from supportive families are more

likely than those from unsupportive families to form secure attachments, and to develop effective emotion regulation skills and social competencies that help them to regulate their responses to stress (Repetti et al., 2002).

In essence, then, the early family environment may provide the groundwork for emotion regulation skills and social competencies for managing stress across the lifespan. In families that are warm and nurturant, children develop secure attachments and learn to manage threat effectively with a lesser physiological/neuroendocrine toll. If they are raised in non-nurturant or conflict-ridden families, children instead experience threatening events more commonly and learn fewer socioemotional skills for managing stress.

Biological mechanisms

Families characterized by unsupportive relationships have damaging outcomes for the mental, physical, and social health of their offspring, not only in the short term, but across the lifespan. Overt family conflict, manifested in recurrent episodes of anger and aggression, deficient nurturing, and family relationships that are cold, unsupportive, and/or neglectful have been associated with a broad array of adverse mental and physical health outcomes long into adulthood (Repetti et al., 2002, 2007). The chronic stress of unsupportive families produces repeated or chronic SNS activation in children, which in turn may lead to wear and tear on the cardiovascular system.

Epigenetic factors appear to be involved in these pathways as well. That is, maternal nurturance can induce long-lasting changes in the function of genes, which is an additional mechanism by which experiences of early nurturance can induce long-term behavioral alterations in emotional and social functioning. For example, Suomi (1987) reported that highly reactive monkeys cross-fostered to nurturant mothers develop good socioemotional skills and achieve high status in the

dominance hierarchy, whereas monkeys with reactive temperaments who are peer-raised develop poor socioemotional skills and end up at the bottom of the dominance hierarchy.

Such long-term effects of maternal care appear to be a result of epigenetic structural alterations (methylation) to the glucocorticoid receptor gene that affect its expression throughout the lifespan (Meaney and Szyf, 2005). Mothers showing high levels of nurturant behavior exhibit greater increases in oxytocin receptors during pregnancy, which is thought to trigger maternal responsivity (Meaney, 2001), and have higher levels of dopamine release when caring for their pups (Champagne et al., 2004). This especially nurturant mothering triggers greater increases in serotonin turnover in the pup, which initiates the cascade leading to the altered glucocorticoid receptor expression that affects adult reactivity to stress (Meaney and Szyf, 2005).

Related evidence has been uncovered with humans. For example, the harshness or nurturance of the early family environment is implicated in the expression of the serotonin transporter gene (5-HTTLPR). People with two copies of the short allele (short/short) of this gene and who have experienced childhood maltreatment are more likely to be diagnosed with major depressive disorder than individuals with one or two copies of the long allele who have experienced similar environments (Caspi et al., 2003; Kaufman et al., 2004). However, a study from our laboratory (Taylor et al., 2006b) qualifies this conclusion: the short allele may function not only as a risk allele for depression in conjunction with an adverse early environment, but as an allele reflecting general sensitivity to the environment, providing protection from symptoms of depression when the environment is nurturant. Using a nonclinical sample of 118 adult men and women, we assessed nurturance of the early family environment, depressive symptomatology, and 5-HTTLPR genotype. As expected, a stressful early family environment by itself was significantly related to depressive symptomatology. However, a significant

gene-by-environment interaction between the 5-HTTLPR and the nurturance of the early family environment qualified this risk. Specifically, individuals with two copies of the short allele had greater depressive symptomatology if they had experienced early familial adversity compared to participants with the short/long or long/long genotypes, but significantly less depressive symptomatology if they reported a nurturant early environment.

Thus, long-term, often permanent effects of early nurturance are evident not only at the behavioral level, but also at the biological level and can include the functioning of relevant genes. Tending to offspring in times of stress, then, offers not only immediate protection, but also long-term protection in the form of biological and behavioral responses to stress.

BEFRIENDING AND RESPONSES TO STRESS

Just as offspring are benefited through tending, so befriending confers stress regulatory benefits.

Animal studies

Animal studies with rats, sheep, prairie voles, and other species show that exogenous administration of oxytocin or stimulation of oxytocin secretion via stroking decreases sympathetic reactivity, blood pressure, pain sensitivity, and glucocorticoid levels, among other findings suggestive of reduced biological stress responses (Carter, 1998; Insel, 1997; Petersson et al., 1996; Uvnäs-Moberg, 1997; Uvnäs-Moberg et al., 1994). Oxytocin also reduces psychological distress, having anxiolytic properties (McCarthy, 1995). For example, exogenous administration of oxytocin enhances sedation and relaxation, is tied to signs of reduced fearfulness in rodent studies, and is tied to enhanced exploratory

behavior (Uvnäs-Moberg et al., 1994; Mantella et al., 2003; McCarthy, 1995).

Human studies

Similar findings have been identified in human studies. For example, high levels of oxytocin or exogenous administration of oxytocin in humans produce decreases in sympathetic activity (e.g., Light et al., 2000; Uvnäs-Moberg, 1997) and inhibit secretion of ACTH and cortisol (Altemus et al., 1995; Chiodera and Legros, 1981). Heinrichs and colleagues (Heinrichs et al., 2003) found that exogenous administration of oxytocin produced lower anxiety and lower cortisol levels during a laboratory stress challenge; the reduced cortisol response was especially pronounced in men who also experienced social support from a friend (see also Kosfeld et al., 2005; Zak et al., 2004). Breastfeeding mothers in whom oxytocin levels are high have lower anxiety, depression, and stress following breastfeeding as compared with bottle-feeding (Modahl and Newton, 1979). Oxytocin increases the sensitivity of brain opioid systems and so, when oxytocin injection is accompanied by an opioid-blocking agent, cortisol levels do not change. Thus, some of the anti-stress pathways properties of oxytocin are probably mediated via an opioid pathway.

Social support

When people affiliate in response to stress, they commonly experience social support. Social support is defined as the perception or experience that one is loved and cared for by others, esteemed and valued, and part of a social network of mutual assistance and obligations (Wills, 1991). Social support may come from a partner, relatives, friends, coworkers, social and community ties, strangers, and even a devoted pet (Allen et al., 2002).

Taxonomies of social support typically classify it into several specific forms.

Informational support occurs when one person helps another to understand a stressful event better by providing information about the event. Instrumental support involves the provision of tangible assistance, such as services, financial assistance, and other specific aid or goods. Emotional support involves providing warmth and assistance to another person and reassuring that person that he or she is a valuable person for whom others care. Social support may involve the reality of using the social network for benefits such as these, but it can also involve simply the perception that such resources are available should they be needed. That is, just knowing that one is cared for and that one could obtain support from others is often comforting in its own right.

Social contacts and social support are psychologically beneficial. Social support reduces psychological distress such as depression or anxiety during times of stress (e.g., Fleming et al., 1982; Sarason et al., 1997; Lin et al., 1999). It promotes psychological adjustment to chronically stressful conditions, such as coronary artery disease (Holahan et al., 1997), diabetes, HIV (Turner-Cobb et al., 2002), and cancer (Penninx et al., 1998; Stone et al., 1999), among many other health-related disorders. Social support protects against cognitive decline in older adults (Seeman et al., 2001), heart disease among the recently widowed (Sorkin et al., 2002), and psychological distress in response to traumatic events, such as 9/11 (Simeon et al., 2005), among other psychological benefits.

Health benefits of social support

Social support has been tied to a variety of specific health benefits among individuals sustaining health risks. These include fewer complications during pregnancy and childbirth (Collins et al., 1993), less susceptibility to herpes attacks among infected individuals (VanderPlate et al., 1988), lower rates of myocardial infarction among individuals with diagnosed disease, a reduced likelihood of mortality from myocardial infarction

(Kulik and Mahler, 1993; Wiklund et al., 1988), faster recovery from coronary artery disease surgery (King et al., 1993; Kulik and Mahler, 1993), better diabetes control (Marteau et al., 1987), longer survival in patients with end-stage renal disease (Cohen et al., 2007), and less pain among arthritis patients (Brown et al., 2003b).

Social support also contributes to survival. In a classic study of 7,000 community residents in Alameda County, CA, epidemiologists Lisa Berkman and Leonard Syme (1979) found that people who lacked social and community ties over the previous nine years were more likely to die of all causes during the follow-up period than those who cultivated or maintained their social relationships. Having social contacts predicted an average 2.8 years increased longevity among women and 2.3 years among men, and these differences persisted after controlling for socioeconomic status, health status at the beginning of the study, and health habits (Berkman and Syme, 1979; see also Rutledge et al., 2004).

The positive impact of social contacts on health is as powerful or more powerful a predictor of health and longevity than well-established risk factors for chronic disease and mortality, with effect sizes on par with smoking, blood pressure, lipids, obesity, and physical activity (House et al., 1988). And as noted, in both animal and human studies, social isolation is tied to a significantly enhanced risk of mortality (House et al., 1988) and a heightened risk of both chronic and acute health disorders (Taylor, 2009).

Mechanisms underlying health benefits

Although not all the mechanisms explaining these strong relationships are known, one key pathway is via stress responses (Cacioppo and Hawkley, 2003). When humans are socially isolated, their sympathetic nervous system and HPA axis responses to stress may continue unabated. Consistent with the theory of allostatic load described earlier, to the

extent that contact with others in times of stress reduces sympathetic and HPA axis activity in response to threats, cumulative wear and tear on biological functioning is lessened. Social contact, then, may leave people less vulnerable to immunologic compromise in response to stress and to health disorders tied to the excessive or recurrent functioning of the sympathetic nervous system and HPA axis.

Whether the attenuation of stress responses by oxytocin and opioids contribute to these clinical effects of social support is, at present, unclear. However, animal research using a wound-healing paradigm suggests that this is a promising avenue for research (Detillion et al., 2004). In this study, female Siberian hamsters received a cutaneous wound and were then exposed to immobilization stress. The stressor increased cortisol concentrations and impaired wound healing, but only in socially isolated and not in socially housed animals. Thus, social housing acted as a stress buffer. Removing cortisol via adrenalectomy eliminated the impact of the stressor on wound healing, thereby implicating the HPA axis in the wound healing process. Of particular relevance to the current arguments, treating the isolated hamsters with oxytocin eliminated the stress-induced increases in cortisol and facilitated wound healing; treating socially housed hamsters with an oxytocin antagonist delayed wound healing. These data strongly imply that social contacts protect against the adverse effects of stress through a mechanism that implicates oxytocin-induced suppression of the HPA axis. Thus, there appear to be discernible clinical consequences (wound healing) of oxytocin in conjunction with social contact.

To summarize, evidence that social responses to threat in the form of tending and befriending are associated with beneficial mental and physical health outcomes is overwhelming. Both animal and human studies attest not only to the beneficial effects of social contact in times of stress, but also to the mechanisms that may underpin these relations.

SEX DIFFERENCES IN TEND AND BEFRIEND

Most, but not all, of the research demonstrating both the psychological effects of oxytocin and endogenous opioid peptides as well as their effects on downregulating stress responses has been conducted with female animals or with women. Although there is evidence that these processes may be implicated in men's reduced stress responses as well (e.g., Heinrichs et al., 2003), the research is less plentiful. Moreover, the biological underpinnings of the theory would appear to be more consistent with what is known about women's hormonal profiles than men's. For example, oxytocin's effects are enhanced in the presence of estrogen (see Taylor et al., 2000).

There is, however, a hormone, vasopressin, that is very similar in molecular structure to oxytocin and whose effects appear to be enhanced in the presence of androgens, and so it may play a parallel role in male social behavior (Panksepp, 1998). Vasopressin is important to stress responses because it is involved in the maintenance of plasma volume and blood pressure during shock, among other functions. In certain monogamous species, most notably the prairie vole, it has also been tied to males' prosocial responses to stress, for example, guarding and patrolling of territory, defense of mate, and defense of offspring against intruders. The vasopressin receptor gene has also been tied to pair bonding, monogamous behavior (Lim et al., 2004), empathy, and altruistic behavior (Anckarsäter and Cloninger, 2007; Knafo et al., 2008). In a recent test of the potential role of vasopressin in men's social behavior, Taylor et al. (2009) examined whether elevations in vasopressin and oxytocin were associated with dissatisfaction in the pair-bond relationship in men and women. Consistent with previous research described earlier, oxytocin was elevated in women experiencing distress, but vasopressin was not. Exactly the reverse pattern was found for

men, such that men who were in distressing pair bond relationships had elevated vasopressin but not oxytocin. Thus, elevated plasma vasopressin in men may act as a signal that the pair bond relationship is jeopardized, just as elevations in plasma oxytocin have been found to signal relationship distress in women. Whether vasopressin underlies additional aspects of men's affiliative responses to stress is currently unknown, but this issue merits additional exploration.

SOCIAL ISSUE IMPLICATIONS OF TEND AND BEFRIEND

What critical social issues and problems are raised or resolved by insights generated by the tend and befriend theory? One important social issue on which the theory helps to shed light is sex differences in life expectancy. Men are especially vulnerable to early mortality due to homicide, suicide, coronary heart disease, and disorders related to substance abuse for coping with stress. Women, on the other hand, enjoy a substantial advantage in mortality in most countries of the world. Only countries in which women are denied access to healthcare or those in which deaths during childbirth are still common show a reverse gap. Of interest is the fact that the causes of death that largely account for men's early mortality are those related to the fight or flight response, namely aggressive responses to stress, withdrawal in the form of substance abuse, and coronary artery disease, the risk for which is exacerbated by frequent or recurrent stress exposure. By contrast, women more reliably turn to their social contacts in times of stress, responses that are, as just noted, protective of health and longevity. The fact that men may be somewhat more likely to cope with stress via fight or flight and women to cope with stress via tend and befriend may help to explain the worldwide gender gap in mortality.

Building on this point, the theory may help to explain sudden increases in mortality rates

that are seen in countries experiencing substantial economic and political turmoil. For example, following the end of the Soviet Union in 1989, destabilization of the Eastern European social environment left many people socially unattached. Whereas unattached women often came together in informal groups with other women and children and shared the management of tasks of daily life, men more often coped with the same instability through alcohol abuse, smoking, and aggressive encounters with other males (Bobak and Marmot, 1996; Stone, 2000). Men in the former Eastern European nations subsequently experienced an abrupt decline of approximately seven years in life expectancy in a mere five years, worse than that sustained during World War II. This decline in life expectancy was explained substantially by deaths among unattached men (Bobak and Marmot, 1996; Stone, 2000). Having a theoretical basis for explaining events such as these, and the psychological and biological mechanisms that underpin them, constitutes an advance that may help governments to anticipate similar problems and intervene, so that similar upheaval does not result in similar carnage.

The theory and the evidence consistent with tend and befriend also point to the importance of making affiliative opportunities available to people when they are under stress. Affiliation is inherently comforting, even when no explicit efforts at social support are elicited or provided. Although questions have recently been raised about the necessity or value of making psychological counseling available to people in the immediate aftermath of a trauma or highly stressful event, providing people with opportunities for companionship which they may utilize in whatever ways are most comforting may be useful interventions.

A final issue on which the tend and befriend model sheds light concerns the relative benefits and costs of altruistic behavior. Conceptualizations of altruism and social support have been guided by the implicit assumption that support is beneficial for the

recipient but costly for the provider. This viewpoint has been shaped by the evolutionary perspective on altruism which addresses the paradox: how do we pass on our altruistic genes to future generations if those genes put us at potential risk? That is, when one person helps another in times of threat, the likelihood that the helper will be harmed can be high. Research on the physical and psychological costs of caregiving would seem to support the position that altruism is inherently costly. From an evolutionary perspective, altruism is largely rescued by the concept of reciprocal altruism (Hamilton, 1963; Trivers, 1971).

Evidence consistent with the tend and befriend model, however, demonstrates that the nurturant behavior of tending not only benefits the offspring but also benefits the tender. That is, tending behavior following a stressful encounter not only downregulates the stress systems of offspring, but also the stress systems of the mother; thus, to the extent that caregivers are providing nurturant behavior to others, their own stress systems may be benefited in cumulative fashion. Research by Brown et al. (2003a) found that death rates were significantly lower among people who provided instrumental support to friends, relatives and neighbors, and emotional support to their spouses. Receiving support did not affect mortality once giving support was statistically controlled. Thus, this study provides important evidence that the giving of support can promote health or retard illness progression. There are psychological benefits of giving support to others as well. Giving support may cement personal relationships, provide a sense of meaning or purpose, and signify that one matters, all of which have been found to promote wellbeing.

It is likely, then, that the benefits of providing support and the apparent absence of anticipated costs may work through some of the same physiological and neuroendocrine pathways whereby the receipt of support or perception of it from others achieves its benefits. In addition, the anxiolytic properties of oxytocin, coupled with its established

role in downregulating SNS and HPA axis responses to stress, may help to provide an understanding of the mental and physical health benefits of providing social support as well as receiving it.

THOUGHTS FOR THE FUTURE

Useful theories in social psychology are marked not only by their ability to explain a body of data and generate specific hypotheses, but also by their ability to expand and grow as new evidence emerges. The tend and befriend theory is no different. Among the extensions that have occurred since the original publication of the theory in 2000 (Taylor et al., 2000) is the accumulating evidence for a biologically based signaling system that may alert people to the need to enhance their social contacts and impel them to do so as well; the fact that plasma oxytocin in women and plasma vasopressin in men are elevated in conjunction with breaches in social contact is a relatively recent discovery.

A second extension involves the integration of evidence concerning the biological stress regulatory effects of maternal nurturance into the theory. Although insights regarding the importance of maternal nurturance for developing offspring's biological stress regulatory systems have been accumulating, the mechanisms underlying these effects remained unknown until relatively recently. Both animal and human evidence now points to some of the biological mechanisms, including epigenetic mechanisms, that help to explain why maternal nurturance in times of stress is so vital to the development of offspring's biological systems and socioemotional skills for managing stress. Integrating this evidence more fully with the literature on the biological and psychological bases of attachment is an important future step.

A third advance concerns potential genetic contributions to the processes detailed in the tend and befriend theory. At the time

the theory was developed, little was known about what specific genes might contribute to tending, befriending, and social behavior more generally. The only evidence for genetic contributions to social behavior was from twin studies indicating a large genetic contribution to the experience of social support (Kessler et al., 1992). Research has now enabled the identification of specific genes within the opioid, oxytocin, vasopressin, dopamine, and serotonin systems that may be implicated in the processes detailed here (see Way and Taylor, 2011, for a review).

Just as accumulating evidence has expanded the theory in certain ways, so has additional research identified certain problems to be resolved. For example, elevated plasma oxytocin is tied to social distress, whereas exogenous administration of oxytocin is tied to a sense of calm and relaxation. This paradox has yet to be fully resolved. As noted, biological underpinnings of men's social responses to stress have yet to be as rigorously explored, although some recent progress has been made (Taylor et al., 2009). No doubt the future will pose additional challenges for the theory.

Tend and befriend: biobehavioral redundancy?

Many of the benefits of social contact appear to result from sheer proximity and not necessarily from the socially supportive transactions that have typically been studied by social psychologists and health psychologists. For example, social ties are consistently found to be mentally and physically health protective in both stressful and non-stressful environments, whereas social support transactions appear to be most beneficial in situations of stress. Qualifying this last finding further, a large number of circumstances have been identified in which social support transactions are unsuccessful or have unintended negative consequences. Bolger and colleagues (e.g., Bolger and Amarel, 2007), for example, have suggested that

invisible support (i.e., support provided by a person without the recipient's awareness) is more beneficial to emotional functioning than social support efforts that are recognized by both the giver and the recipient as intended. Awareness that one is being supported may represent a threat to self-esteem. Research on the matching hypothesis, namely the idea that social support is most effective when it matches the need of the recipient (Cohen and McKay, 1984), also indicates a variety of circumstances under which mismatches between the type of support delivered or the person delivering it exacerbate stress (see Taylor, 2009, for a review). Social contacts during stressful times have the potential to be negative, and research has shown that negative interactions can have a worse effect on mental and physical health functioning than positive effects achieve beneficial effects (e.g., Rook, 1984). Other misfired efforts at social support have also been identified (see Taylor, 2009, for a review).

These findings suggest that simple proximity and the perception of support may be especially beneficial but not necessarily its use. That is, much of the benefit of social support may come from the perception that it is available and not necessarily its actual engagement (Thoits, 1995; Taylor, 2009). A biobehavioral theory of affiliation, as tend and befriend is, has little trouble with this paradox. Either the psychological benefits associated with social contact, the hormonal underpinnings of social contact, or both, may produce many of the beneficial biological and psychological consequences that consistently predict wellbeing both in stressful and nonstressful times.

A common observation in biological research concerns the redundancy that exists in the human being for sustaining vital biological functions. For example, there are five different mechanisms that ensure that the stomach can digest food. Similarly, people are endowed with two eyes, two hands, two legs, two kidneys, and the like. Not all vital functions are backed up through redundancy of course, the heart being an obvious

counterexample. However, many are, and it may be useful to think about psychological processes as contributing to the human beings' redundancy to ensure vital processes.

At the outset of this chapter, the point was made that social living in general and affiliation in response to threat in particular are essential to human survival. It is unlikely that these critical responses would be left to chance or even to a single underlying psychological or biological mechanism. Rather, the fact that there is both biological and psychological evidence for tending and befriending in response to stress suggests that these may be interrelated but semi-redundant pathways that ensure that social responses to threat take place, so as to protect human beings and ensure their survival.

Although a theory that focuses on affiliative responses to threat, as the tend and befriend theory does, stresses the important protective role that these processes have, it must be noted that tending and befriending also promote human growth as well. Through affiliation with others in times of stress and the ability to draw on mental representations of relationships, people acquire the resources to explore and grow both emotionally and intellectually in environments that ensure social connection.

ACKNOWLEDGMENT

Preparation of this manuscript was supported by a grant from NSF (SES-0525713).

REFERENCES

- Allen, K., Blascovich, J. and Mendes, W.B. (2002) Cardiovascular reactivity and the presence of pets, friends, and spouses: The truth about cats and dogs. *Psychosomatic Medicine*, 64, 727–739.
- Altemus, M.P., Deuster, A., Galliven, E., Carter, C.S. and Gold, P.W. (1995) Suppression of hypothalamic–pituitary–adrenal axis response to stress in

lactating women. *Journal of Clinical Endocrinology and Metabolism*, 80, 2954–2959.

Anckarsäter, H. and Cloninger, C.R. (2007) The genetics of empathy and its disorders. In T. Farrow and P. Woodruff (eds), *Empathy in Mental Illness*, pp. 261–288. New York: Cambridge University Press.

Berkman, L.F. and Syme, S.L. (1979) Social networks, host resistance, and mortality: A nine-year follow up study of Alameda County residents. *American Journal of Epidemiology*, 109, 186–204.

Bobak, M. and Marmot, M. (1996) East-West mortality divide and its potential explanations: Proposed research agenda. *British Medical Journal*, 312, 421–425.

Bolger, N. and Amarel, D. (2007) Effects of support visibility on adjustment to stress: Experimental evidence. *Journal of Personality and Social Psychology*, 92, 458–475.

Bowlby, J. (1982) *Attachment and Loss, Vol. 1, Attachment* (2nd edition). New York: Basic Books. (Original work published 1969).

Brown, S.L., Nesse, R.M., Vinokur, A.D. and Smith, D.M. (2003a) Providing social support may be more beneficial than receiving it: Results from a prospective study of mortality. *Psychological Science*, 14, 320–327.

Brown, J.L., Sheffield, D., Leary, M.R. and Robinson, M.E. (2003b) Social support and experimental pain. *Psychosomatic Medicine*, 65, 276–283.

Cacioppo, J.T. and Hawkley, L.C. (2003) Social isolation and health, with an emphasis on underlying mechanism. *Perspectives in Biology and Medicine*, 46, S39–S52.

Cannon, W.B. (1932) *The Wisdom of the Body*. New York: Norton.

Carlson, M. and Earls, F. (1997) Psychological and neuroendocrinological sequelae of early social deprivation in institutionalized children in Romania. *Annals of the New York Academy of Sciences*, 807, 419–428.

Carter, C.S. (1998) Neuroendocrine perspectives on social attachment and love. *Psychoneuroendocrinology*, 23, 779–818.

Carter, C.S., Lederhendler, I.I. and Kirkpatrick, B. (1999) *The Integrative Neurobiology of Affiliation*. Cambridge, MA: MIT Press.

Caspi, A., Sugden, K., Moffitt, T.E., Taylor, A., Craig, I.W., Harrington, H., McClay, J., Mill, J., Martin, J., Braithwaite, A. and Poulton, R. (2003) Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science*, 301, 386–389.

Champagne, F.A., Chretien, P., Stevenson, C.W., Zhang, T.Y., Gratton, A. and Meaney, M.J. (2004) Variations in nucleus accumbens dopamine associated with individual differences in maternal behavior in the rat. *Journal of Neuroscience*, 24, 4113–4123.

Chiodera, P. and Legros, J.J. (1981) Intravenous injection of synthetic oxytocin induces a decrease of cortisol plasma level in normal man. *Comptes rendus des séances de la Société de biologie et de ses filiales*, 175, 546–594.

Cohen, S. and McKay, G. (1984) Social support, stress, and the buffering hypothesis: A theoretical analysis. In A. Baum, S.E. Taylor, and J. Singer (eds), *Handbook of Psychology and Health*, 4, 253–268. Hillsdale, NJ: Erlbaum.

Cohen, S.D., Sharma, T., Acquaviva, K., Peterson, R.A., Patel, S.S. and Kimmel, P.L. (2007) Social support and chronic kidney disease: An update. *Advances in Chronic Kidney Disease*, 14, 335–344.

Collins, N.L., Dunkel-Schetter, C., Lobel, M. and Scrimshaw, S.C.M. (1993) Social support in pregnancy: Psychosocial correlates of birth outcomes and post-partum depression. *Journal of Personality and Social Psychology*, 65, 1243–1258.

Collins, N.L. and Feeney, B.C. (2000) A safe haven: An attachment theory perspective on support seeking and caregiving in intimate relationships. *Journal of Personality and Social Psychology*, 78, 1053–1073.

Detillion, C.E., Craft, T.K., Glasper, E.R., Prendergast, B.J. and DeVries, C. (2004) Social facilitation of wound healing. *Psychoneuroendocrinology*, 29, 1004–1011.

Eisenberger, N.I., Lieberman, M.D. and Williams, K.D. (2003) Does rejection hurt? An fMRI study of social exclusion. *Science*, 302, 290–292.

Fleming, R., Baum, A., Gisriel, M.M. and Gatchel, R.J. (1982) Mediating influences of social support on stress at Three Mile Island. *Journal of Human Stress*, 8, 14–22.

Fraley, R.C. (2002) Attachment stability from infancy to adulthood: Meta-analysis and dynamic modeling of developmental mechanisms. *Personality and Social Psychology Review*, 6, 123–151.

Francis, D., Diorio, J., Liu, D. and Meaney, M.J. (1999) Nongenomic transmission across generations of maternal behavior and stress responses in the rat. *Science*, 286, 1155–1158.

Grippo, A.J., Gerena, D., Huang, J., Kumar, N., Shah, M., Ughreja, R. and Carter, C.S. (2007) Social isolation induces behavioral and neuroendocrine disturbances relevant to depression in female and male prairie voles. *Psychoneuroendocrinology*, 32, 966–980.

Gunnar, M.R., Brodersen, L., Krueger, K. and Rigatuso, J. (1996) Dampening of adrenocortical responses during infancy: Normative changes and individual differences. *Child Development*, 67, 877–889.

Hamilton, W.D. (1963) The evolution of altruistic behavior. *The American Naturalist*, 97, 354–356.

Hart, J., Gunnar, M. and Cicchetti, D. (1996) Altered neuroendocrine activity in maltreated children related to symptoms of depression. *Development and Psychopathology*, 8, 201–214.

Heinrichs, M., Baumgartner, T., Kirshbaum, C. and Ehlert, U. (2003) Social support and oxytocin interact to suppress cortisol and subjective responses to psychological stress. *Biological Psychiatry*, 54, 1389–1398.

Holahan, C.J., Moos, R.H., Holahan, C.K. and Brennan, P.L. (1997) Social context, coping strategies, and depressive symptoms: An expanded model with cardiac patients. *Journal of Personality and Social Psychology*, 72, 918–928.

House, J.S., Landis, K.R. and Umberson, D. (1988) Social relationships and health. *Science*, 241, 540–545.

Insel, T.R. (1997) A neurobiological basis of social attachment. *American Journal of Psychiatry*, 154, 726–735.

Jalowiec, J.E., Calcagnetti, D.J. and Fanselow, M.S. (1989) Suppression of juvenile social behavior requires antagonism of central opioid systems. *Pharmacology, Biochemistry, and Behavior*, 33, 697–700.

Jamner, L.D., Alberts, J., Leigh, H. and Klein, L.C. (1998) Affiliative need and endogenous opioids. Paper presented to the Society of Behavioral Medicine annual meeting, March 1998 New Orleans, LA.

Kaufman, J., Yang, B.Z., Douglas-Palumberi, H., Houshyar, S., Lipschitz, D., Krystal, J.H. and Gelernter, J. (2004) Social supports and serotonin transporter gene moderate depression in maltreated children. *Proceedings of the National Academy of Sciences USA*, 101, 17316–17321.

Kendrick, K.M. and Keverne, E.B. (1989) Effects of intracerebroventricular infusions of naltrexone and phentolamine on central and peripheral oxytocin release and on maternal behaviour induced by vaginocervical stimulation in the ewe. *Brain Research*, 505, 329–332.

Kessler, R.C., Kendler, K.S., Heath, A.C., Neale, M.C. and Eaves, L.J. (1992) Social support, depressed mood, and adjustment to stress: A genetic epidemiological investigation. *Journal of Personality and Social Psychology*, 62, 257–272.

King, K.B., Reis, H.T., Porter, L.A. and Norsen, L.H. (1993) Social support and long-term recovery from coronary artery surgery: Effects on patients and spouses. *Health Psychology*, 12, 56–63.

Knafo, A., Israel, S., Darvasi, A., Bachner-Melman, R., Uzefovsky, F., Cohen, L. et al. (2008) Individual differences in allocation of funds in the dictator game associated with length of the arginine vasopressin 1a receptor RS3 promoter region and correlation between RS3 length and hippocampal mRNA. *Genes, Brain, and Behavior*, 7, 266–275.

Kosfeld, M., Heinrichs, M., Zak, P.J., Fischbacher, U. and Fehr, E. (2005) Oxytocin increases trust in humans. *Nature*, 435, 673–676.

Kulik, J.A. and Mahler, H.I.M. (1993) Emotional support as a moderator of adjustment and compliance after coronary artery bypass surgery: A longitudinal study. *Journal of Behavioral Medicine*, 16, 45–64.

Levine, S. and Wiener, S.G. (1988) Psychoendocrine aspects of mother–infant relationships in nonhuman primates. *Psychoneuroimmunology*, 13, 143–154.

Light, K.C., Smith, T.E., Johns, J.M., Brownley, K.A., Hofheimer, J.A. and Amico, J.A. (2000) Oxytocin responsiveness in mothers of infants: A preliminary study of relationships with blood pressure during laboratory stress and normal ambulatory activity. *Health Psychology*, 19, 560–567.

Lim, M.M., Wang, Z., Olazabal, D.E., Ren, X., Terwilliger, E.F. and Young, L.J. (2004) Enhanced partner preference in a promiscuous species by manipulating the expression of a single gene. *Nature*, 429, 754–757.

Lin, N., Ye, X. and Ensel, W. (1999) Social support and depressed mood: A structural analysis. *Journal of Health and Social Behavior*, 40, 344–359.

Liu, D., Diorio, J., Tannenbaum, B., Caldji, C., Francis, D., Freedman, A., Sharma, S., Pearson, D., Plotsky, P.M. and Meaney, M.J. (1997) Maternal care, hippocampal glucocorticoid receptors, and hypothalamic–pituitary–adrenal responses to stress. *Science*, 277, 1659–1662.

Mantella, R.C., Vollmer, R.R., Li, X. and Amico, J.A. (2003) Female oxytocin-deficient mice display enhanced anxiety-related behavior. *Endocrinology*, 144, 2291–2296.

Marteau, T.M., Bloch, S. and Baum, J.D. (1987) Family life and diabetic control. *Journal of Child Psychology and Psychiatry*, 28, 823–833.

Martel, F.L., Nevison, C.M., Rayment, F.D., Simpson, M.J.A. and Keverne, E.B. (1993) Opioid receptor blockade reduces maternal affect and social grooming in rhesus monkeys. *Psychoneuroimmunology*, 18, 307–321.

McCarthy, M.M. (1995) Estrogen modulation of oxytocin and its relation to behavior. *Advances in Experimental Medicine and Biology*, 395, 235–245.

McEwen, B.S. (1998) Protective and damaging effects of stress mediators. *New England Journal of Medicine*, 338, 171–179.

Meaney, M.J. (2001) Maternal care, gene expression, and the transmission of individual differences in stress reactivity across generations. *Annual Review of Neuroscience*, 24, 1161–1192.

Meaney, M.J. and Szyf, M. (2005) Environmental programming of stress responses through DNA methylation: life at the interface between a dynamic environment and a fixed genome. *Dialogues in Clinical Neurosciences*, 7, 103–123.

Modahl, C. and Newton, N. (1979) Mood state differences between breast and bottle-feeding mothers. *Emotion and Reproduction, Proceedings of the Serono Symposia*, 20B, 819–822.

Moles, A., Kieffer, B.L. and D'Amato, F.R. (2004) Deficit in attachment behavior in mice lacking the μ -opioid receptor gene. *Science*, 304, 1983–1985.

Nachmias, M., Gunnar, M.R., Mangelsdorf, S., Parritz, R.H. and Buss, K. (1996) Behavioral inhibition and stress reactivity: The moderating role of attachment security. *Child Development*, 67, 508–522.

Panksepp, J. (1998) *Affective Neuroscience*. New York: Oxford University Press.

Panksepp, J., Nelson, E. and Bekkedal, M. (1999) Brain systems for the mediation of social separation distress and social-reward. In C.S. Carter, I.I. Lederhendler, and B. Kirkpatrick (eds), *The Integrative Neurobiology of Affiliation*, pp. 221–244. Cambridge, MA: MIT Press.

Penninx, B.W.J.H., van Tilburg, T., Boeke, A.J.P., Deeg, D.J.H., Kriegsman, D.M.W. and van Eijk, J.T.M. (1998) Effects of social support and personal coping resources on depressive symptoms: Different for various chronic diseases? *Health Psychology*, 17, 551–558.

Petersson, M., Alster, P., Lundeberg, T. and Uvnäs-Moberg, K. (1996) Oxytocin causes a long-term decrease of blood pressure in female and male rats. *Physiology and Behavior*, 60, 1311–1315.

Repetti, R.L., Taylor, S.E. and Saxbe, D. (2007) The influence of early socialization experiences on the development of biological systems. In J. Grusec and P. Hastings (ed.), *Handbook of Socialization* (pp. 124–152). New York: Guilford.

Repetti, R.L., Taylor, S.E. and Seeman, T.E. (2002) Risky families: Family social environments and the mental and physical health of offspring. *Psychological Bulletin*, 128, 330–366.

Rook, K.S. (1984) The negative side of social interaction: Impact on psychological well-being. *Journal of Personality and Social Psychology*, 46, 1097–1108.

Rutledge, T., Reis, S.E., Olson, M., Owens, J., Kelsey, S.F., Pepine, C.J. et al. (2004) Social networks are associated with lower mortality rates among women with suspected coronary disease: The National Heart, Lung, and Blood Institute-sponsored women's ischemia syndrome evaluation study. *Psychosomatic Medicine*, 66, 882–888.

Sarason, B.R., Sarason, I.G. and Gurung, R.A.R. (1997) Close personal relationships and health outcomes: A key to the role of social support. In S. Duck (ed.), *Handbook of Personal Relationships*, pp. 547–573. New York: Wiley.

Seeman, T.E., Lusignolo, T.M., Albert, M. and Berkman, L. (2001) Social relationships, social support, and patterns of cognitive aging in healthy, high-functioning older adults: MacArthur Studies of Successful Aging. *Health Psychology*, 20, 243–255.

Simeon, D., Greenberg, J., Nelson, D., Schmeider, J. and Hollander, E. (2005) Dissociation and posttraumatic stress 1 year after the World Trade Center disaster: Follow-up of a longitudinal study. *Journal of Clinical Psychiatry*, 66, 231–237.

Sorkin, D., Rook, K.S. and Lu, J.L. (2002) Loneliness, lack of emotional support, lack of companionship, and the likelihood of having a heart condition in an elderly sample. *Annals of Behavioral Medicine*, 24, 290–298.

Spitz, R.A. and Wolff, K.M. (1946) Anaclitic depression: An inquiry into the genesis of psychiatric conditions in early childhood, II. In A. Freud et al. (eds), *The Psychoanalytic Study of the Child*, 2, 313–342. New York: International Universities Press.

Stone, A.A., Mezzacappa, E.S., Donatone, B.A. and Gonder, M. (1999) Psychosocial stress and social support are associated with prostate-specific antigen levels in men: Results from a community screening program. *Health Psychology*, 18, 482–486.

Stone, R. (2000) Stress: The invisible hand in eastern Europe's death rates. *Science*, 288, 1732–1733.

Suomi, S.J. (1987) Genetic and maternal contributions to individual differences in rhesus monkey biobehavioral development. In N.A. Krasnagor, E.M. Blass, M.A. Hofer, and W.P. Smotherman (eds), *Perinatal Development: A Psychobiological Perspective*, pp. 397–420. New York: Academic Press.

Suomi, S.J. (1999) Attachment in rhesus monkeys. In J. Cassidy and P. Shaver (eds), *Handbook of Attachment: Theory, Research, and Clinical Applications*, pp. 181–197. New York: Guilford Press.

Tamres, L., Janicki, D. and Helgeson, V.S. (2002) Sex differences in coping behavior: A meta-analytic review. *Personality and Social Psychology Review*, 6, 2–30.

Taylor, S.E. (2002) *The Tending Instinct: How Nurturing Is Essential to Who We Are and How We Live*. New York: Holt.

Taylor, S.E. (2009) Social support: A review. In H.S. Friedman (ed.), *Oxford Handbook of Health Psychology*. New York: Oxford University Press.

Taylor, S.E., Gonzaga, G., Klein, L.C., Hu, P., Greendale, G.A. and Seeman, T.E. (2006a) Relation of oxytocin to psychological stress responses and hypothalamic–pituitary–adrenocortical axis activity in older women. *Psychosomatic Medicine*, 68, 238–245.

Taylor, S.E., Klein, L.C., Lewis, B.P., Gruenewald, T.L., Gurung, R.A.R., and Updegraff, J.A. (2000) Biobehavioral responses to stress in females: Tend-and-befriend, not fight-or-flight. *Psychological Review*, 107, 411–429.

Taylor, S.E., Saphire-Bernstein, S. and Seeman, T.E. (2009) Plasma oxytocin in women and plasma vasoressin in men are markers of distress in primary relationships. *Psychological Science*, 21, 3–7.

Taylor, S.E., Way, B.M., Welch, W.T., Hilmert, C.J., Lehman, B.J. and Eisenberger, N.I. (2006b) Early family environment, current adversity, the serotonin transporter promoter polymorphism, and depressive symptomatology. *Biological Psychiatry*, 60, 671–676.

Thoits, P.A. (1995) Stress, coping and social support processes: Where are we? What next? *Journal of Health and Social Behavior*, (Extra Issue), 53–79.

Trivers, R.L. (1971) The evolution of reciprocal altruism. *The Quarterly Review of Biology*, 46, 35–57.

Turner, R.A., Altemus, M., Enos, T., Cooper, B. and McGuinness, T. (1999) Preliminary research on plasma oxytocin in normal cycling women: Investigating emotion and interpersonal distress. *Psychiatry*, 62, 97–113.

Turner-Cobb, J.M., Gore-Felton, C., Marouf, F., Koopman, C., Kim, P., Israelski, D. and Spiegel, D. (2002) Coping, social support, and attachment style as psychosocial correlates of adjustment in men and women with HIV/AIDS. *Journal of Behavioral Medicine*, 25, 337–353.

Uvnäs-Moberg, K. (1997) Oxytocin linked antistress effects – the relaxation and growth response. *Acta Psychologica Scandinavica*, 640, 38–42.

Uvnäs-Moberg, K., Ahlenius, S., Hillegaart, V. and Alster, P. (1994) High doses of oxytocin cause sedation and low doses cause an anxiolytic-like effect in male rats. *Pharmacology, Biochemistry, and Behavior*, 49, 101–106.

VanderPlate, C., Aral, S.O. and Magder, L. (1988) The relationship among genital herpes simplex virus, stress, and social support. *Health Psychology*, 7, 159–168.

Way, B.M. and Taylor, S.E. (2011) Genetic factors in social pain. In G. MacDonald and L.A. Jensen-Campbell (eds), *Social Pain: A Neuroscientific, Social, Clinical, and Developmental Analysis*, pp. 95–119. Washington, DC: American Psychological Association.

Wiklund, I., Oden, A., Sanne, H., Ulvenstam, G., Wilhemsson, C. and Wilhemsen, L. (1988) Prognostic importance of somatic and psychosocial variables after a first myocardial infarction. *American Journal of Epidemiology*, 128, 786–795.

Wills, T.A. (1991) Social support and interpersonal relationships. In M.S. Clark (ed.), *Prosocial Behavior*, pp. 265–289. Newbury Park, CA: Sage.

Zak, P.J., Kurzban, R. and Matzner, W.T. (2004) The neurobiology of trust. *Annals of the New York Academy of Sciences*, 1032, 224–227.

The Evaluative Space Model

John T. Cacioppo, Gary G. Berntson,
Catherine J. Norris, and
Jackie K. Gollan

ABSTRACT

The ability to differentiate hostile from hospitable stimuli is ubiquitous in animals. The evaluative space model (ESM) is a theory of the functional structure and operating characteristics of these evaluative processes across levels of the neuraxis, ranging from spinal cord reflexes to the executive functions of the frontal lobes (e.g., impulse control). According to the ESM, physical constraints limit behavioral expressions and incline behavioral predispositions toward a bipolar organization, but this bipolar organization is posited to be the consequence of multiple operations, including motivational activation function for positivity (appetition) and the activation function negativity (aversion). The partial segregation of positive and negative evaluative processes afforded evolution that opportunity to sculpt distinct activation functions for positivity and negativity, and permits greater flexibility of these evaluative processes such as reciprocal activation, uncoupled activation, or coactivation/coinhibition. The result is a much more flexible and adaptable affect system of evaluative processes than would be provided were evaluative processes characterized simply as a bipolar (positive–negative) activation function.

THE EVALUATIVE SPACE MODEL

The ABCs of the mind have been described as affect, behavior, and cognition. These were once thought to represent independent sets of psychological structures and processes, but this is no longer the case. Theory and research on emotional contagion, for instance, underscore the interplay between affect and behavior; work on embodied cognition stresses the overlap between behavior and cognition; studies of motivated attention and cognition focus on the interchange between affect and cognition; and research on attitudes emphasizes the interplay among all three. Nevertheless, there are sufficient distinctions among affect, behavior, and cognition that psychological theories of each are needed. Of the ABCs of the mind, the least well developed theoretically is the topic of affect.

The term “affect” in human studies has been used to refer to feelings beyond those of

the traditional senses, with an emphasis on the experience of emotions and variations in hedonic tone. Accordingly, the scientific study of human affect and emotion has tended to emphasize reportable feeling states. Studies of the conceptual organization of affect and emotion indicate that people represent feelings and emotions in terms of a circular order around the perimeter of the space defined by a bipolar valence dimension and an orthogonal dimension labeled activation (i.e., a circumplex; for example, see Russell, 1980) or, alternatively, by a space defined by two bipolar valence dimensions, one ranging from low positive affect to high negative affect and a second ranging from low negative affect to high positive affect (Green et al., 1999; Watson and Tellegen, 1985). As LeDoux noted, however:

It is widely recognized that most cognitive processes occur unconsciously, with only the end products reaching awareness, and then only sometimes. Emotion researchers, though, did not make this conceptual leap. They remained focused on subjective emotional experience ... The main lesson to be learned ... is that emotion researchers need to figure out how to escape from the shackles of subjectivity if emotion research is to thrive (LeDoux, 2000: 156).

There is an understandable appeal to settling for feelings as the appropriate data to model in the area of affect. It is these feelings that some theorists seek to describe, understand, and explain. The structure and processes underlying mental contents are not readily apparent, however, and most cognitive processes occur unconsciously with only selected outcomes reaching awareness. Over millions of years of evolution, efficient and manifold mechanisms have evolved for differentiating hostile from hospitable stimuli and for organizing adaptive responses to these stimuli. These are critically important functions for the evolution of mammals, and the integrated set of mechanisms that serve these functions can be thought of as an “affect system.” It is this affect system – its architecture and operating characteristics – that is the focus of the evaluative space

model (ESM) (Cacioppo and Berntson, 1994, 1999; Cacioppo et al., 1997, 1999).

We use the term “evaluative” rather than “affective” in the name of our theory to emphasize the focus on the affect system as the production mechanism for discrimination, valuation, feelings, and adaptive behavior. On the stimulus side, evaluative processes refer to the discrimination of a stimulus – or features of a stimulus – as appetitive or aversive, hostile or hospitable, pleasant or unpleasant, threatening or nurturing. Such discriminations, according to the ESM, extend beyond what can be verbalized, as spinal cord reflexes and affective priming amply demonstrate. On the response side, evaluative processes organize behavioral responses to promote appropriate approach or withdrawal, advancement or retreat, movement toward or away, attack or avoidance, nurturance or defense, acceptance or rejection. The evaluative operations intervening between these stimulus and response elements are also part of the affect system.

The ESM is a general theoretical formulation of the functional architecture and operating characteristics of the affect system, the postulates of which are summarized in Table 3.1. For instance, theorists have argued about whether the affect system should be construed in terms of discrete emotions or abstract dimensional structures. According to the ESM, these are not mutually exclusive but rather each is a theoretical representation at a specific level of organization (*level of analysis postulate*). There are important distinctions among the positive emotions and among the negative emotions, but the positive emotions are generally more similar to each other than they are to the negative emotions and vice versa. The ESM posits a superordinate dimensional structure representing appetitive predispositions, positive affects, and emotions (termed positivity), and a superordinate dimensional structure representing defensive predispositions, negative affects, and emotions (termed negativity). If appetitive predispositions and positive

Table 3.1 Table of ESM postulates

<i>Postulate</i>	<i>Definition</i>	<i>Additional detail</i>
Level of analysis	There are distinctions among both positive and negative emotions, but positive emotions are more similar to each other than they are to negative emotions, and vice versa.	A single valence continuum does not capture the structure and operating characteristics of affect system.
Functional separability	Positivity and negativity are not equivalent in their constitution, operations, or consequences.	There is a superordinate dimensional structure representing appetitive predispositions, positive affects, and emotions, as well as a superordinate dimensional structure representing defensive predispositions, negative affects, and emotions.
Heteroscedacity	The constellation of antecedents, emotions, expressions, and response is more diverse for negativity than positivity.	When an event elicits a positive emotion, staying the course is sufficient; when negative emotion is elicited, an adaptive response may vary greatly across eliciting events.
Energetic efficiency	Behavior in future encounters with target stimuli will tend to be more expected and stable when organized in terms of a bipolar evaluative dimension.	Behavioral and cognitive efficiency and a reduction in stress is served by mental representations of general action predispositions toward classes of stimuli.
Evaluative activation	Affect is a joint function of positively and negatively valent activation functions.	The resultant output disposition(s) is not necessarily the simple arithmetic mean of the separate functions. The net disposition, for example, may consist of ambivalence, vacillation, or active suppression of one or the other dispositions.
Monotonicity	Strength of the response varies as a function of the extremity of the stimulus.	The functions, however, appear to be negatively accelerating rather than linear across the full dynamic range.
Antagonistic effects	Directional response effects of positive affect (approach) are generally opposite to those of negative effect (withdrawal).	Exceptions exist, such as when aversive threat is met with aggressive defense, or when a pain stimulus (e.g., flu shot) is approached and solicited.
Functional separability	Activation of positivity and negativity is partially separable.	This separation confers additional adaptability and flexibility for learned dispositions.
Modes of evaluative activation	Positivity and negativity can be activated reciprocally, uncoupled, or nonreciprocally.	At high levels of coactivation (which minimizes the dynamic range, reduces response lability, and maximizes directional flexibility), energy expenditure is taxing over long periods of time; eliciting circumstances tend to be avoided.
Parallel evaluative processing	The ability to achieve coactivation of positivity and negativity by attending to positive and negative features of a stimulus simultaneously (e.g., bittersweet, disappointing wins).	The reciprocal activation postulate in prior models of affect and emotion is replaced by this postulate.
		Bivalent modes of evaluative action requires at least a three-dimensional approach: 1 for positivity; 1 for negativity; 1 representing the net behavioral predisposition or response orchestrated by the affect system.
		Although there may be some reciprocal inhibition between the evaluative dimensions, they are at least partly independently expressed.

Low-pass filtering	The ability to achieve coactivation of positivity and negativity by oscillating between positive and negative stimuli with sufficient speed that results in the sustained activation of positivity and negativity.	Even though there can be an oscillation between positive and negative activation, if the speed of the presentation of the contrasting stimuli is faster than the low-pass filter cutoff, the activation of each cannot follow the speed of the oscillations and coactivation results (ambivalence).
Distinct activation functions	The partial segregation of the positive and negative evaluative channels allows for distinctive activation functions for positivity and negativity.	From both a practical and theoretical perspective, a rapidly switching oscillation between positive and negative (or more generally, from the two ends of a bipolar system) is equivalent to a bivariate system in its formal properties.
Positivity offset	The offset (intercept) for the positive activation function is higher than that of the negative activation function.	Motivation to approach is stronger than the motivation to withdraw at very low levels of evaluative activation; this promotes exploratory behavior – without a positivity offset, a person in a neutral environment is unlikely to approach novel stimuli.
Negativity bias	The gain for the negative activation function is higher than that of the positive activation function.	Motivation to withdraw is stronger than the motivation to approach at very high levels of evaluative activation; it is more difficult to overcome a fatal (or near fatal) assault than to return to an opportunity unpursued.
Recalibration	The activation functions for positivity and negativity are capable of the same kind of recalibrations based on salient contextual and accessible stimuli as is seen in receptor mechanisms.	Both sensitivity to small variations among stimuli and a dynamic range suitable to detect a wide array of affective stimuli are preserved.
Affective dispositions Hierarchical organization	There are measurable individual differences in the positivity offset and negativity bias. Evaluative processes are implemented at multiple, re-representative levels of the neuraxis.	These individual differences may have both a biological and an experiential/psychological basis. There is a continuum of neuraxial organization that extends throughout the central nervous system in a heterarchical structure, ranging from the spinal cord to the frontal lobes. Rostral, in contrast to caudal, neurobehavioral organizations are slower, more serial like, susceptible to more contextual control; potentiate greater response flexibility; and manifest multiple modes of appetitive and aversive activation. The multiple levels of processing can result in coordinated synergistic outcomes, or can lead to conflicts.

affects and emotions were stars, the positivity dimension might be thought of as the galaxy constituted by these stars, and if defensive predispositions, negative affects, and emotions were stars, the negativity dimension could be conceived as a different galaxy comprising this very different cluster of stars. That is, these positive and negative dimensions are not treated as equivalent in their constitution, operations, or consequences (*functional separability postulate*; see Table 3.1).

The ESM is based on the premise that the affect system has been shaped by the hammer and chisel of natural selection to produce a range of generally adaptive responses, and it posits that there are more conjunctions of factors that produce negativity than positivity. As Tolstoy observed in *Anna Karenia*, all happy families are alike one another, but each unhappy family is unhappy in its own way. When an event elicits positive affect or emotion, staying the course generally is sufficient. When negative affect or emotion is elicited, what constitutes an adaptive response may vary greatly across eliciting events. The ESM, therefore, posits that there is a greater diversity of adaptive responses for negative than positive antecedents, including a greater diversity of negative than positive emotions. Thus, the *heteroscedacity postulate* of the ESM states that the constellation of antecedents, emotions, expressions, and response is more diverse for negativity than for positivity (Cacioppo et al., 1997).

The ESM does not deny that valence is a central dimension in the affect system. A fundamental premise of the ESM is that the affect system evolved to help organisms differentiate hostile from hospitable stimuli and to respond adaptively to these stimuli. The ESM, however, posits that the valence dimension does not provide a sufficient depiction of the architecture of the affect system. We will return to additional postulates of the ESM in a subsequent section, but we begin by discussing what led to the ESM.

PERSONAL NARRATIVE OF THE THEORY'S DEVELOPMENT

John Cacioppo and Richard Petty began working on attitudes in the early 1970s while they were graduate students together at Ohio State University. The field of attitudes at that time faced two challenges. First, independent variables appeared to have unreliable effects on attitude change; that is, variables like source credibility sometimes led to more attitude change, sometimes to less attitude change, and sometimes to no attitude change. Unspecified laboratory artifacts were thought to be a likely culprit. The implication was that carefully controlled laboratory research might have little or no bearing on the phenomenon theorists sought to understand – and that the voluminous body of attitude theory and research that had been spawned by careful experimental studies was of little scientific value. Second, attitudes were thought to not predict behavior. The implication was that attitudes and attitude change were not particularly important phenomena to understand because although attitudes might have a bearing on what people said, they putatively did not bear on what people actually did.

The Elaboration Likelihood Model (ELM) of attitudes and persuasion (Petty and Cacioppo, 1981, 1986) contributed to the resolution of these threats by specifying that attitude change was a multiply determined outcome; that is, there were different routes through which attitude development or change could be achieved, and that the same independent variable (e.g., source credibility) could produce various outcomes because its effects on attitude change and behavior depended on the specific route that was triggered by this variable in interaction with other variables (e.g., personal involvement, need for cognition). The ELM also specified the conditions under which these variables would operate to produce attitude change through the central or peripheral route. Moreover, the ELM predicted that attitudes

were more likely to be temporally stable and guide behavior when these attitudes were formed through the central than peripheral route. Thus, attitudes predicted behavior but only under certain specifiable conditions.

Two decades later, Petty had expanded his interests to address issues such as attitude strength, attitude confidence, attitude anchoring effects, and the metacognitive processes involved in attitudes and persuasion. Cacioppo, on the other hand, had become more interested in affect and stress, including the influence of affect and emotions on implicit social influence processes (e.g., emotional contagion; Hatfield et al., 1994) and the effect of somatovisceral processes on affect and attitudes (e.g., Cacioppo et al., 1993). During this period, Cacioppo returned to the psychology department at the Ohio State University as a faculty member in the social psychology program and in what is now called the psychobiology and behavioral neuroscience program. The director of the latter program was Gary Berntson, who similarly was interested in the topics of affect and stress. One of the first discussions between Cacioppo and Berntson was a paper the latter was writing on the inadequacy of the notion that the sympathetic and parasympathetic branches of the autonomic nervous systems were wired to be reciprocally activated. This insight led to discussions and collaborations that continue to this day, as well as to a theory of the structure and function of the autonomic nervous system (Berntson et al., 1991, 1993b) and to a companion theory of the structure and function of the affect system, the ESM (e.g., Berntson and Cacioppo, 2008; Cacioppo and Berntson, 1994, 1999; Cacioppo et al., 1997, 1999, 2004).

Theoretical developments in the ESM have been further fueled over the years by remarkable students and colleagues. Each has made many different contributions but it may be informative to mention a primary focus of each. Steve Crites performed a series of studies showing that evaluative and nonevaluative processing differed in the patterns of associated brain activation (Crites and

Cacioppo, 1996; Cacioppo et al., 1996). Wendi Gardner (Cacioppo et al., 1997; Gardner and Cacioppo, 1995) and Jon Krosnick (Holbrook et al., 2001) provided early evidence for the functional separability of positive and negative affect. Tiffany Ito showed that the activation functions for positive and negative affect differed even when evaluative processing was spontaneous and unintended (Ito and Cacioppo, 2000; Ito et al., 1998b). Kyle Smith provided evidence that differences in the activation functions for positive and negative affect emerged early in the information processing stream to manifest as an attentional bias (Smith et al., 2003, 2006). Jeff Larsen has shown in various studies that positive and negative affect are not invariably reciprocally activated and helped articulate the conditions under which reciprocal activation would or would not occur (Larsen et al., 2001, 2004, 2009). Beth Crawford demonstrated that differences in the activation functions of positive and negative affect have implicit effects on behavior, specifically that they influence the implicit learning of where appetitive and aversive stimuli are likely to be within a spatial context (Crawford and Cacioppo, 2002). Catherine Norris established that the nomothetic descriptions of the activation functions for positive and negative affect are robust, and that there are stable individual differences in the operating characteristics of the affect system – individual differences which predict subtle aspects of behavior (Norris and Cacioppo, 2009; Norris et al., in press). And Jackie Gollan has led the investigation of the possible clinical implications – specifically for understanding and treating affective disorders – of stable individual differences in the operating characteristics of the affect system (Gollan et al., 2009).

Work on the ESM has been funded primarily by the National Institute of Mental Health through a multi-institutional grant to the Center for the Study of Emotion and Attention led by Peter Lang at the University of Florida. Peter, Margaret Bradley at the University of Florida, Mike Davis at Emory

University, Chris Patrick at the University of Minnesota, Arne Ohman at the Karolinska Institute, and the various affiliated investigators associated with this center have contributed to theoretical and empirical developments by being such constructive critics and colleagues over the years (e.g., Ito et al., 1998a; Schupp et al., 2000).

INTELLECTUAL HISTORY OF THE THEORY

People have speculated about human affect and emotions for millennia, but the topic largely eluded scientific scrutiny until Louis Thurstone (Thurstone, 1928, 1931; Thurstone and Chave, 1929) adapted his quantitative methods in psychophysics to address the problem of measuring affect:

It has been stated by economists and by other social scientists that affect cannot be measured, and some of the fundamental theory of social science has been written with this explicit reservation. Our studies have shown that affect can be measured. In extending the methods of psychophysics to the measurement of affect we see the possibility of a wide field of application by which it will be possible to apply the methods of quantitative scientific thinking to the study of feeling and emotion, to aesthetics, and to social phenomena (Thurstone, 1931: 269).

Thurstone's methods included paired comparisons and rating scales, and his pioneering work led to the rating scales used in many surveys and experimental studies of affect today.

According to Thurstone's model, positive and negative affect fall on the opposite ends of a bipolar valence continuum, and this valence dimension is the central theoretical structure of the affect system:

The affect about an object may be of strong intensity or it may be weak. The positive and negative affect therefore constitutes a linear continuum with a neutral point or zone and two opposite directions, one positive and the other negative. Measurement along this affective

continuum is of a discriminatory character with the discriminant error as a unit of measurement (Thurstone, 1931: 261).

More formally, affect in Thurstone's model is specified to be a joint function of positively and negatively valent activation functions (*evaluative activation postulate*); the strength of the response varies as a function of the extremity of the stimulus (*monotonicity postulate*); the responses to an equally spaced series of valent stimuli produces a linear output (*linearity postulate*); the directional response effects of positive affect (e.g., approach/withdrawal) is generally opposite to that of negative affect (*antagonistic effects postulate*); positively and negatively valent activation functions are reciprocally controlled (*reciprocal activation postulate*); and an increase in positive affect and a decrease in negative affect elicited by a stimulus produce the same movement and positioning along the valence continuum and therefore are indistinguishable in terms of the affect system (*functional equivalence postulate*). The valence structure in Thurstone's model could be viewed as operating analogously to a balance knob on a stereo to produce movements along this continuum ranging from "very positive and not at all negative" at one end to "very negative and not at all positive" at the other. Because the simple valence theory casts positive and negative affect as functionally equivalent – that is, the activation of positive affect is equivalent to the reduction in the activation of negative affect and vice versa – the theory further predicts that the activation function for positive affect is the same as the activation function for negative affect (*equivalent activation functions postulate*).

Twenty years later, Neal Miller (1951, 1961) proposed that approach and withdrawal were behavioral manifestations with generally antagonistic effects on behavior that could come from *distinguishable* motivational substrates. His behavioral theory was enriched by conceptualizing approach and withdrawal separately, investigating their

unique antecedents and consequences, and examining the behavioral constraints that led to approach and withdrawal tendencies. The division of motivational substrates into appetitive/nurturant/approach and aversive/defensive/withdrawal subsets and the view that appetition and aversion generally have antagonistic effects are central tenets of several contemporary theories of emotion, including Gray's (1987), Lang et al. (1990), Kahneman and Tversky's (1979), and Carver's (2001), as well as the ESM. All except the ESM, however, assumed or explicitly posited that appetition/approach and aversion/withdrawal were invariably reciprocally activated.

A theoretical model of the architecture and operating characteristics of the affect system makes it possible to generate new and falsifiable predictions, and this applies to the simple valence model of the affect system as well. For instance, it follows from this valence model that a neutral stimulus, one that arouses neither positive nor negative affect, is positioned in the middle of the valence continuum, but so is a stimulus that arouses strong positive and negative affect in equal measure – that is, an ambivalent stimulus. If the valence continuum were a sufficient description of the structure of the affect system, people would have difficulty differentiating neutral and ambivalent stimuli. Empirical evidence is inconsistent with this prediction (e.g., Edwards and Ostrom, 1971; Grabenhorst et al., 2007), which raises the possibility that a single valence continuum does not adequately capture the structure and operating characteristics of the affect system (cf. Cacioppo and Berntson, 1994).

Physical constraints generally restrict *behavioral* manifestations to bipolar actions (approach/withdrawal), but the ESM posits that evolution favors the animal that can learn, represent, and access rapidly whether approach or withdrawal is adaptive when confronted by a stimulus. There is a behavioral efficiency, a conservation of limited cognitive resources, and a reduction in physiological stress that is served by mental representations of general and enduring net

action predispositions toward classes of stimuli. Accordingly, the ESM posits that guides for one's actions in future encounters with the target stimuli, such as those provided by affective or emotional responses to the stimuli, will tend to be more expected and stable when organized in terms of a bipolar evaluative dimension (*energetic efficiency postulate*). The bipolar (positive/negative) structure may represent a stable endpoint, but it is not sufficient to model the structures or operations that preceded this adaptive endpoint.

Evidence that a person's mixed affective reactions migrate toward a bipolar organization dates back more than half a century ago. Brehm (1956) found that individuals, following a selection between two alternatives, spread the appeal of these alternatives by some combination of amplifying the positive features of the chosen alternative, diminishing the negative features of the chosen alternative, magnifying the negative features of the unchosen alternative, and minimizing the positive features of an unchosen alternative. This motivational push toward affective bipolarity was especially strong when subjects initially regarded the alternatives to be similarly appealing.

The fact that positive and negative affects have antagonistic effects on approach or withdrawal behavior does not mean that the affective inputs are invariably reciprocally activated. If you press your palms together with your forearms parallel to the floor, the force you exert with your right and left arms has antagonistic effects on the movement of your palms, but as you increase the force you exert with your right and left arms, your palms will not move as long as the forces you exert are equal. In this example, you have coactivated the muscles in your right and left arms rather than reciprocally activated them. The effects of activating the muscles of your right and left arms in this position are still antagonistic – pushing with your right arm moves your palms to the left and pushing with your left arm moves your palms to the right. But these antagonistic effects do not

mandate that you necessarily exert reciprocal activation – increasing the force with which you are pushing with one arm while simultaneously decreasing the force with which you are pushing with the other.

As we noted in the opening of the chapter, the ESM posits two abstract dimensions, positivity and negativity, each of which can range from zero to high levels of activation as a result of the activational status of its constituent elements. In addition, the ESM posits that affect is a joint function of positively and negatively valent activation functions (*evaluative activation postulate*); the strength of the response varies as a function of the extremity of the stimulus (*monotonicity postulate*); the directional response effects of positive affect (e.g., approach/withdrawal) are generally opposite to that negative affect (*antagonistic effects postulate*); the activation of positivity and the activation of negativity are partially separable (*functional separability postulate*), and positivity; and negativity can be activated reciprocally (e.g., mutually exclusive, incompatible), uncoupled (e.g., singularly activated), or nonreciprocally (e.g., coactivational or coinhibitory) (*modes of evaluative activation postulate*). Thus, the ESM expands the principle of reciprocal evaluative activation to accommodate: (1) the separable activation of positivity and negativity, (2) the investigation of their unique antecedents and consequences, and (3) the concept and formal properties of modes of evaluative activation. Accordingly, the reciprocal activation postulate in prior models of affect and emotion is replaced by the modes of evaluative activation postulate.

The introduction of bivalent modes of evaluative activation requires at least a three-dimensional representation, one each to represent activation functions for positivity and negativity and a third “valence” dimension that represents the net behavioral predisposition or response orchestrated by the affect system. A two-dimensional representation of the activation of positivity and negativity, depicted as the bottom plane in Figure 3.1, is termed the evaluative space and may provide

a more comprehensive formulation for depicting these operations. For instance, this bivariate evaluative space accommodates all possible combinations of positive and negative evaluative activation: (1) the reciprocal mode of evaluative activation is represented as one diagonal vector that ranges from maximal positivity/minimal negativity to maximal negativity/minimal positivity, (2) the nonreciprocal mode of evaluative activation is represented as the alternate diagonal that ranges from minimal positivity and negativity to maximal positivity and negativity, and (3) the uncoupled modes of evaluative activation are represented as vectors lying along the axes. The family of vectors parallel to those above represent the general categories, or modes, of evaluative activation expressed from varying starting points within the two-dimensional plane depicted in Figure 3.1.

Given the affect system evolved to guide behavior, information processed by the affect system does not stop with its registration on the evaluative space. Instead, the antagonistic effects of the activation of positivity and negativity are integrated into a net affective predisposition or action, which can be represented as an overlying bipolar response surface (see Figure 3.1). The positivity \times negativity evaluative space depicted in Figure 3.1 represents the level of activation of the underlying positive and negative inputs to a bivalent affective response. The resulting behavioral predisposition to approach or withdraw can be represented in terms of an overlying surface whose projection on the z-axis constitutes the bipolar valence dimension. It is possible to derive this overlying surface for all combinations of positivity and negativity. Note that the mapping from the evaluative (positivity \times negativity) space to the valence dimension is many-to-one. The dashed lines in the evaluative space are illustrative, as the points constituting any of the dashed lines (i.e., *iso-affective contours*) map into the same point on the bipolar valence dimension. Thus, knowledge of where one is in the evaluative space permits mapping to

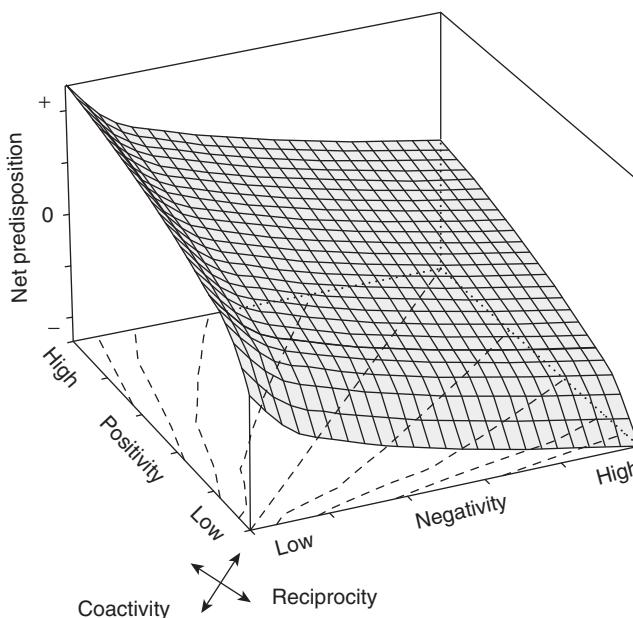


Figure 3.1 Illustrative bivariate evaluative space and its associated affective response surface. This surface represents the net predisposition of an individual toward (+) or away from (-) the target stimulus. This net predisposition is expressed in relative units and the axis dimensions are in relative units of activation. The point on the surface overlying the left axis intersection represents a maximally positive predisposition, and the point on the surface overlying the right axis intersection represents a maximally negative predisposition. Each of the points overlying the dashed diagonal extending from the back to the front axis intersections represent the same middling predisposition. Thus, the nonreciprocal diagonal on the evaluative plane – which represents different evaluative processes (e.g., neutral to ambivalence) – yields the same middling expression on the affective response surface. Dashed lines (including the coactivity diagonal) represent isocontours on the evaluative plane, which depict many-to-one mappings between the affective response surface and the underlying evaluative space. These isocontours are illustrative rather than exhaustive. (Adapted from Cacioppo and Berntson, 1994.)

the valence dimension, but the reverse is not the case – except at the endpoints of the valence continuum (not at all positive and very negative; not at all negative and very positive). Therefore, the range of indeterminism is smallest at the anchors of the reciprocal diagonal (“very negative/not at all positive” and “very positive/not at all negative”) and largest along the coactivity diagonal.

Evolution can genetically endow only limited fixed adaptive responses relative to the potential range of circumstances an

organism could encounter. Therefore, there is an evolutionary pressure to maximize flexibility and learning. According to the ESM, the partial segregation of the positive and negative evaluative channels in the affect system confers the additional flexibility of orchestrating appetitive and aversive motivational forces via modes of evaluative activation, which in turn affords greater flexibility for learned dispositions.

For instance, the concept of the “mode of evaluative activation” was derived from the architecture we posited for the affect system.

These modes, in turn, produce formal response properties. The behavioral effects of positivity and negativity are generally antagonistic, so reciprocal activation produces a large dynamic range, high response lability, and high directional stability – the very qualities that are adaptive once a stimulus is determined to be hospitable or hostile. Contrary to the reciprocal mode of evaluative activation, coactivation minimizes the dynamic range, minimizes the response lability, and maximizes the directional flexibility. Uncoupled modes of evaluative activation produce intermediate dynamic range, response lability, and directional stability/flexibility. The energy expenditure in highly coactivated states is further posited to be taxing over long periods of time so, as noted above, coactivated states tend to be resolved or the eliciting circumstances tend to be avoided or denied. Consistent with this postulate, ambivalence has been associated with evaluative instability (e.g., Hass et al., 1991).

To illustrate, a thirsty animal on the Savannah would be motivated to go to the water hole for water (appetitive stimulus), where predators also lurk (aversive stimulus). In such circumstances, coactivation may be more adaptive, at least momentarily, as the animal remains in a highly energetic (i.e., prepared) state while it stoops to drink and scans for predators ready to lurch from behind a bush or from beneath the murky water. Coactivation of positivity and negativity permits the thirsty animal to approach and consume some of the needed water while also maximizing the speed with which it can respond (reciprocally activate withdrawal and inhibit approach) when the predator approaches.

The stochastic and functional independence of positive and negative affect have been demonstrated (e.g., see Berntson and Cacioppo, 2008; Cacioppo and Gardner, 1999; Cacioppo et al., 1997, 2004). Some bipolar theorists have argued that positive and negative should not be expected to exhibit strong negative correlations due, for

instance, to random and nonrandom measurement error (e.g., Green et al., 1993). However, these explanations for the stochastic separability documented in some studies cannot explain why the same measures and procedures would produce a bipolar structure in some circumstances and a bivalent structure in others – as suggested by our story of the thirsty animal on the Savannah. Is there empirical evidence of this sort?

Larsen et al. (2001) measured people's feelings of happiness and sadness – two emotions at opposite ends of the bipolar circumplex – during a normal day and a day characterized by complex feelings and emotions. For instance, they found that individuals were more likely to report feeling both happy and sad immediately after watching the film *Life Is Beautiful* (study 1), moving out of their college dormitories (study 2), and graduating from college (study 3) than in more typical situations (e.g., a typical day on campus). In addition to including standard emotion items such as happy and sad, in study 3, Larsen et al. included a more intriguing emotion: bittersweet. The term "bittersweet" implies a commingling of positive and negative feelings and is therefore difficult to place in the circumplex and other bipolar frameworks. Yet its inclusion in the lexicon suggests that it can sometimes characterize individuals' feelings. Consistent with this possibility, graduates were not only more likely to report feeling both happy and sad than nongraduates, they also reported feeling more bittersweet than nongraduates. In sum, Larsen et al. replicated the typical finding that happiness and sadness are largely mutually exclusive in routine, steady-state conditions (i.e., a normal day), but demonstrated that these two seemingly opposite emotions can co-occur under specifiable circumstances.

Additional research has shown that mixed feelings can occur not only in situations as rich as those studied by Larsen et al. (2001), but in situations as simple as a game of chance. In a laboratory experiment with a gambling task, Larsen et al. (2001) presented

participants with 50/50 chances to win one of two amounts (e.g., \$7 or \$11) or lose one of two amounts. On winning trials, participants were more likely to feel both good and bad about winning the smaller outcome (i.e., a *disappointing win*) than the larger. Even though winning felt good, missing an opportunity to win an even greater amount felt bad. Similarly, participants were more likely to report mixed feelings after they lost the smaller of two amounts (i.e., *relieving losses*) rather than the larger. Thus, losing felt bad, even as avoiding a larger loss felt good.

The ESM is concerned with coactivation of underlying mechanism more so than the resulting experience of mixed feelings. One means of achieving this coactivation of positivity and negativity is by attending to positive and negative features (*parallel evaluative processing postulate*). Though attention is limited (e.g., Kahneman, 1973), the extant evidence supports the notion that attention can be directed to at least two channels of information (de Gelder and Vroomen, 2000; Spelke et al., 1976).

A second means of achieving coactivation posited by the ESM is oscillation between the positive and negative stimuli with sufficient speed that the low-pass filtering properties of the activation functions result in the sustained activation of positivity and negativity (*low-pass filtering postulate*). When one colloquially says that affect is sticky, the reference is to the low-pass filtering properties of the activation functions for positivity and negativity relative to cognition. A stimulus whose rate of oscillation exceeds a low-pass filter cutoff results in sustained activation. The theoretical implication is that the oscillation of positive and negative stimuli can produce a coactivation of positivity and negativity.

How the speed of the oscillation between two bipolar opposites can modulate activation and perception can be illustrated with a flip book, the pages of which alternate between solid black and solid white. If one stares for ten seconds at a black page and flips one page to the white page, the

brightness is enhanced by a contrast effect. If one flips the pages slowly, the black and white pages are clearly visible. When the pages are flipped more quickly (e.g., 60 Hz), we no longer see the alternating black and white pages but instead we perceive the streaming pages to be gray. That is, even though there is an oscillation between high and low luminance, the speed of their presentation is too fast to follow each presentation so we perceive a fusion of the two. There is a range of speeds for which this coactivation results. Too slow an oscillation, and the beginning and end of each stimulus is clearly perceived; too fast an oscillation, and nothing is perceived.

To determine whether the oscillation of univalent positive and univalent negative stimuli could produce evaluative coactivation, Cacioppo et al. (2009) presented brief, serial presentations of photographs of political figures that elicited univalent positive affect or univalent negative affect. Four conditions were constructed: (1) one of the oscillating pictures elicited univalent positive affect and the second elicited univalent negative affect (bivalent condition); (2) each of the two oscillating pictures elicited neutral affect (neutral condition); (3) each of the two oscillating pictures elicited negative affect (univalent negative condition); and (4) each of the two oscillating pictures elicited positive affect (univalent positive condition). The results, which are summarized in Figure 3.2, confirmed that brief, serial presentations of photographs of political figures that elicited univalent positive affect or univalent negative affect led to the coactivation of positivity and negativity. These results suggest that one mechanism by which ambivalence can be aroused is oscillating attention between univalent positive and univalent negative features of stimuli (Figure 3.2, upper left panel and bottom panels), and that the level of coactivation is not determinable from the valence dimension per se (Figure 3.2, upper right panel).

As shown in Figure 3.3, the ESM further posits that the partial segregation of the

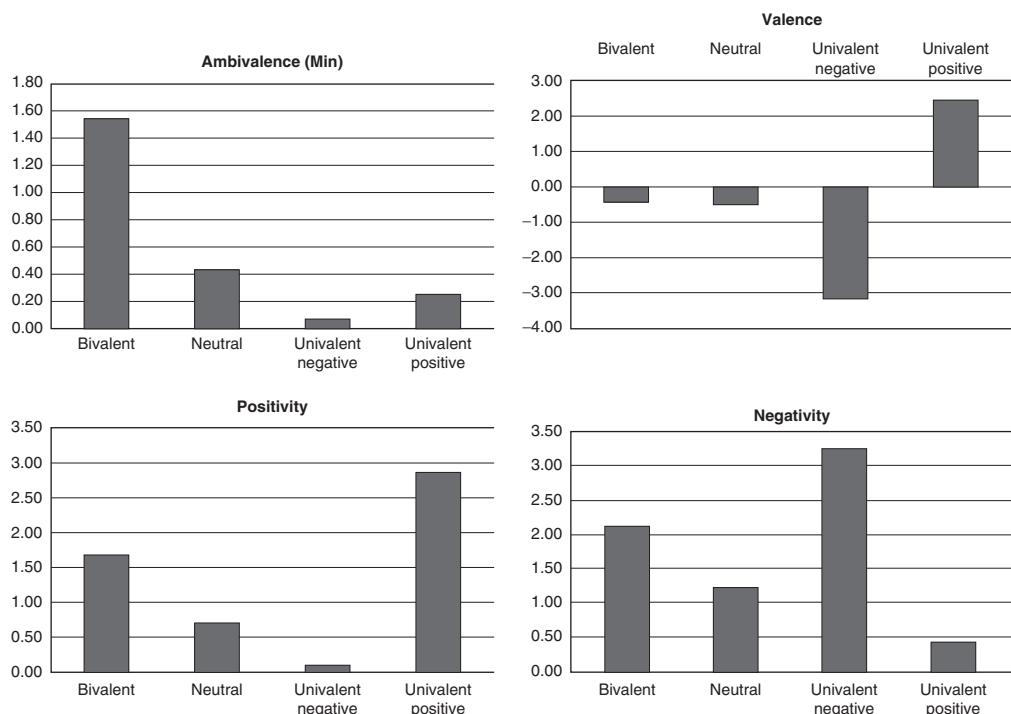


Figure 3.2 Participants evaluated stimulus blocks consisting of pairs of targets they rated as positive, negative, or neutral. The alternating presentation of positive and negative stimuli (bivalent blocks) evoked higher levels of ambivalence than the neutral and univalent blocks. The bivalent blocks, however, show valence scores indistinguishable from those of the neutral blocks. This shows that conventional bipolar valence measures do not convey information about underlying positive and negative affect and may therefore mask ambivalent states. (From Cacioppo et al., 2009.)

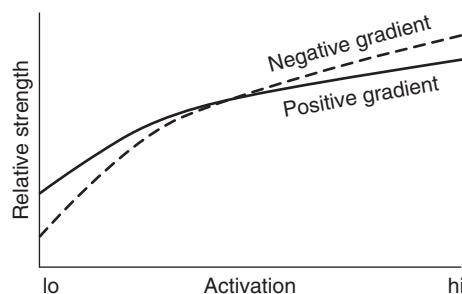


Figure 3.3 Activation functions for positive and negative dimensions of affective processing; the x-axis represents affective input, whereas the y-axis represents output of the system. The ESM proposes that there are two asymmetries in affective processing: the **positivity offset** is the result of greater positive than negative affect at low levels of emotional input; the **negativity bias** is the result of stronger responses to negative than to equally extreme positive input. (Adapted from Cacioppo and Berntson, 1994.)

positive and negative evaluative channels in the affect system affords evolution of the opportunity to sculpt distinctive activation functions for positivity and negativity (*distinct activation functions postulate*), that both activation functions are negatively accelerating (*nonlinearity postulate*), that the intercept for the positive activation function (i.e., the approach motivation at zero input) is higher than the intercept for the negative activation function (*positivity offset postulate*), and that the gain for the negative activation function is higher than that for the positivity activation function (*negativity bias postulate*). The consequence of the positivity offset is that the motivation to approach is stronger than the motivation to withdraw at very low levels of evaluative activation, whereas the consequence of the negativity bias is that the motivation to withdraw is stronger than the motivation to approach at high levels of evaluative activation.

The theoretical rationale for the positivity offset is that it produces exploratory behavior. Without a positivity offset, an organism in a neutral environment may be unmotivated to approach novel objects, stimuli, or contexts. The neophobic response to foreign stimuli that characterizes most species permits an initial period of observation. With no negative outcomes, this exposure allows the initial neophobic response to habituate, thereby allowing exploratory behavior to manifest. In the absence of such a motivation to explore, organisms would learn little about novel or neutral appearing environments and their potential reward value. With a positivity offset, however, an organism facing neutral or unfamiliar stimuli would be weakly motivated to approach, and with the quick habituation of the initial fear response, to engage in exploratory behavior. Such a pairing of initial neophobic and subsequent exploratory tendencies may have important survival value, at least at the level of a species. A positivity offset also fosters social cohesion even in the absence of other information about conspecifics.

The term “positivity bias” in the literature may be something of a misnomer. It has been

used to refer to the finding that people about whom only neutral information is known are nevertheless rated positively. The term “bias” in engineering refers to the gain of an amplifier or activation function. We, therefore, reserved the use of the term “bias” to the “negativity bias,” which refers to differences in the gain of the activation functions for positivity and negativity, and we instead use the term “positivity offset” to refer to differences in the thresholds (i.e., intercepts) for these activation functions. The research on the “positivity bias” in person perception, however, may reflect the operation of the positivity offset in the affect system. For instance, Cacioppo et al. (1997) investigated the robust “positivity bias” in impression formation to determine whether it was limited to diagnosticity (cf. Skowronski and Carlston, 1989) or if it reflects a more general positivity offset in the affect system. The results of a series of studies confirmed that the positivity offset in impression formation was not limited by the social desirability concerns of the participants, by the type or diagnosticity of the neutral behaviors used, nor by the similarity between target and participant. Indeed, the positivity offset was observed not only in impressions of human targets, but with impressions of novel fish and insects as well. The positivity offset demonstrated in this work, then, could not have been a result of the neutral behaviors, implying the absence of negative attributes (e.g., “Sam is susceptible to the laws of gravity”). Neither could it have merely reflected the process of similarity leading to attraction, as Sears’ “person positivity bias” (1983) would have predicted. Instead, the positivity offset appeared to be a more general operating characteristic of the affect system.

The theoretical rationale for the negativity bias is that it is more difficult to overcome a fatal (or even a near-fatal) assault than to return to an opportunity unpursued, so it is more adaptive to err on the side of caution as threats get nearer. The negativity bias, therefore, provides an adaptive response function

that complements the positivity offset. Human taste buds respond to sweet, salty, sour, and bitter stimuli. Most can detect sweetness in approximately one part in 200, saltiness in one part in 400, sourness in one in 130,000, and bitterness in one in 2,000,000. From the perspective of the affect system, a given amount of a negative or threat-related gustatory stimulus (e.g., most poisons taste bitter) activates a stronger affective response than the same amount of a positive (e.g., sweet) gustatory stimulus. This may be more than an epicurean curiosity; it may represent differences in the activation functions for positive and negative affective processing (see reviews by Baumeister et al., 2001; Cacioppo and Gardner, 1999; and Cacioppo et al., 1997). Moreover, the combination of spatial and affective information is essential for many approach and avoidance behaviors, and thus for survival. As predicted by the ESM, Crawford and Cacioppo (2002) found that the incidental learning of the likely spatial location of affective stimuli is greater for negative than positive stimuli.

According to the ESM, these distinct activation functions have evolved because they produce a neurobehavioral organization that is generally adaptive. There are two additional implications of this formulation that should be noted here. First, the activation functions for positivity and for negativity must balance the trade-off between dynamic range and sensitivity (precision). To illustrate, let us specify the dynamic range of the activation function for negativity to be an arbitrary and small amount, x . A weak negative stimulus (e.g., a predator at a safe distance) would produce a weak activation of negativity, but if the distance to the predator were to be reduced, the increase in negativity would soon reach the maximum activation for negativity. This might produce very precise changes in activation with slight changes in distance to the predator (sensitivity), but it would also mean a predator that was nearby or one some distance away would produce maximal and comparable activation. This would mean the animal would avoid

predators even when at a safe distance, dramatically reducing their access to food and water. Another alternative might be for the dynamic range of the activation function to be a very large amount, say $x \times 10^6$. While this would produce an activation function that would respond to changes in the distance from predators over a very large range, it would also decrease the sensitivity of these changes in that changes in negativity would require large rather than small changes in the distance to a predator. While such a range might permit distinctive activations to all or nearly all the negative stimuli an animal might encounter, this comes at a cost. Such a design could be fatal for an animal faced with multiple dangers in the same setting or when the animal's survival depends on its sensitivity to the proximity of a predator.

One of the means by which nature has solved this problem is adaptation/recalibration. If you have been sitting in a dimly lit room, you are able to see reasonably well once you have adapted. It is as if the dynamic range for luminance has been reduced so that you can see small variations in luminance that previously had been difficult to detect before your eyes adjusted (i.e., before the dynamic range for luminance was reduced). If someone suddenly turns on bright lights, the higher levels of luminance of the objects around you all produce comparably maximal levels of activation, making everything seem like an undifferentiated bright light, until your eyes again adapt to increase the dynamic range for luminance. The ESM posits that the activation functions for positivity and negativity are capable of the same kind of recalibrations based on the salient contextual and accessible stimuli (*recalibration postulate*). As a result of the recalibration of these activation functions, both sensitivity to small variations among stimuli and a dynamic range suitable to detect a wide array of affective stimuli are preserved, and so too is the positivity offset and the negativity bias.

Second, given that individual variation is the engine of natural selection, there should

be measurable individual differences in the positivity offset and negativity bias (*affective dispositions postulate*). The underlying structure and operation of the affect system is generally outside people's awareness, and these dispositional tendencies are similarly conceived as generally lying outside awareness, but like the affect system itself these dispositional tendencies should be measurable through people's responses to affective stimuli.

Consistent with this reasoning, stable and predictive individual differences in the positivity offset and the negativity bias have been identified (Ito and Cacioppo, 2005; Norris et al., in press). Participants in the Norris et al. study were exposed to three different sets of stimuli (pictures, sounds, and words), and during each set they were exposed to 66 stimuli, 6 of which were neutral and low in arousal, and 30 each of which vary in their extremity of pleasant or unpleasant and arousal but which were matched on these two dimensions. Ratings of each were made using the affect matrix – a 5 (positivity: zero to maximum) by 5 (negativity: zero to maximum) matrix on which participants rate each stimulus (Larsen et al., 2009). The positivity offset was indexed by the difference between the positivity and negativity ratings of the six neutral stimuli, and the negativity bias was gauged as the difference in rating of the six most extreme unpleasant stimuli minus the rating of the six most extreme (and initially matched on extremity and arousal) pleasant stimuli. Results revealed that individual differences in the positivity offset and negativity bias were uncorrelated, temporally stable, and generalizable across ratings of pictures, sounds, and words (see Figure 3.4). Furthermore, individual differences in the positivity offset predicted the spatial learning for positive stimuli, whereas individual differences in the negativity bias predicted the spatial learning for negative stimuli. In sum, although most individuals exhibit both a positivity offset and a negativity bias, stable individual differences in these constructs predict what we learn about the world.

Finally, whereas many theories of affect assume evaluative processing is performed by a singular, or perhaps dual, processing mechanism, the ESM posits that positivity and negativity each represent the cumulative activation of multiple processing mechanisms. The ESM posits that affective states and responses are mediated by a network of distributed, often recursively connected, interacting neural regions in the central nervous system (including spinal cord reflexes), with the different areas making specific, often task-modulated contributions. Specifically, the ESM posits a continuum of neuraxial organization relevant to evaluative processing – a continuum that extends throughout the central nervous system in a heterarchical structure ranging from the frontal lobes to the spinal cord (*heterarchical organization postulate*).

The nineteenth-century neurologist John Hughlings Jackson emphasized the hierarchical structure of the brain, and the representation of functions at multiple levels within this neural hierarchy (Jackson, 1958/1884). The notion was that information is processed at multiple levels of organization within the nervous system. Primitive protective responses to aversive stimuli are organized at the level of the spinal cord, as is apparent in flexor (pain) withdrawal reflexes that can be seen even after spinal transection. These primitive protective reactions are expanded and embellished at higher levels of the nervous system (see Berntson et al., 1993a). More elaborate defensive/protective behaviors are organized within the brain-stem, and decerebrate organisms (with no cerebral hemispheres) can display organized escape and aggressive responses to noxious stimuli. The evolutionary development of even higher neural systems, such as the limbic system and cerebral cortex, endowed organisms with a further-expanded behavioral and motivational repertoire, that can capitalize on experience-dependent associative knowledge, information-processing networks, and cognitive strategies that anticipate and prepare for or avoid aversive encounters.

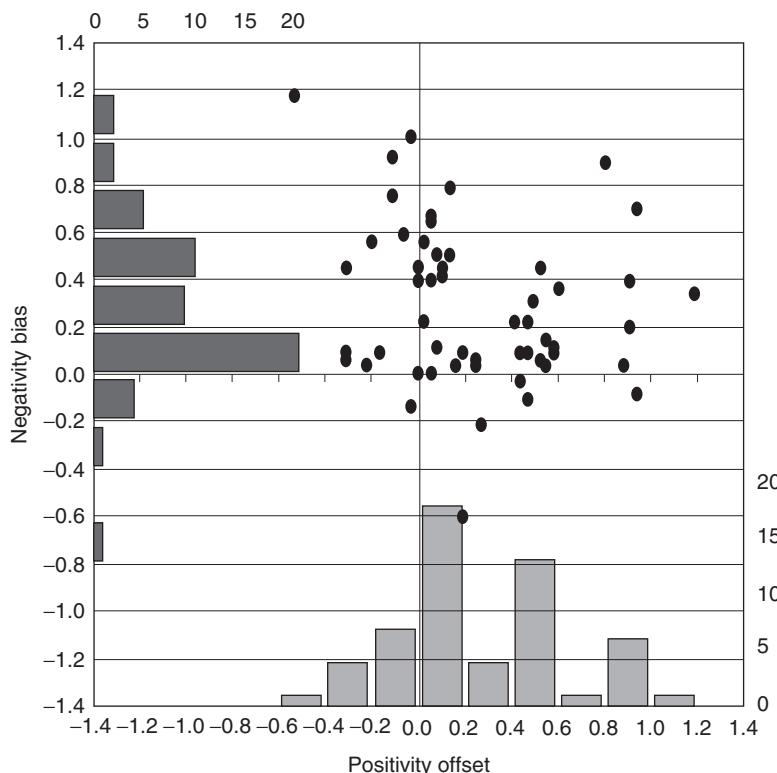


Figure 3.4 Histograms of the aggregate measures of the negativity bias and positivity offset and a scatterplot depicting their relationship. Most participants exhibited a negativity bias ($M = 0.26$, $SE = 0.04$) and a positivity offset ($M = 0.27$, $SE = 0.05$), $t(64) = 6.46$ and 5.89 , respectively, $ps < 0.001$, and the negativity bias and positivity offset were uncorrelated, $r(64) = -0.18$, ns. (From Norris et al., in press.)

Evolution not only endowed us with primitive, lower-level adaptive reactions, but it sculpted the awesome information processing capacities of the highest levels of the brain. Thus, neurobehavioral mechanisms are not localized to a single level of organization within the brain, but are represented at multiple levels of the nervous system. At progressively higher levels of organization, there is a general expansion in the range and relational complexity of contextual controls and in the breadth and flexibility of discriminative and adaptive responses (Bernston et al., 1993a). Although higher-level systems confer greater behavioral variability and adaptive flexibility, they do not replace lower neurobehavioral mechanisms.

Adaptive flexibility of higher-level systems has costs, given the finite information-processing capacity of neural circuits. Greater flexibility implies a less rigid relationship between inputs and outputs, a greater range of information that must be processed, and a slower serial-like mode of processing. Consequently, the evolutionary layering of higher processing levels onto lower substrates has adaptive advantage in that lower and more efficient processing levels may continue to be utilized, and may be sufficient in some circumstances. For example, pain withdrawal reflexes, mediated by inherent spinal circuits, can manifest in rapid protective responses to pain stimuli. However, ascending pain pathways also convey information to higher

levels of the neuraxis that subserve integrative aspects of affective, cognitive, and behavioral reactions such as fear, anxiety, avoidance, and/or aggression. Reflex responses provide a rapid low-level response, but they are not immutable, as higher neurobehavioral processes can come to suppress or bypass pain withdrawal reflexes (e.g., self-injecting insulin or recovering a billfold from a fire). These organizational features are not unique to defensive/protective behaviors, but rather reflect general neuroarchitectural principles that characterize generally hierarchical neural systems. Consequently, the defensive system and its re-representative organization across neuraxial levels offers a model system for conceptualizing neurobehavioral processes generally (e.g., see Berntson et al., 1998).

Although we have emphasized the features at the extremes, the heterarchical organization postulate of a continuum of neuraxial organization relevant to evaluative processing is in keeping with the principle of re-representation. Consider the architecture of spinal cord reflexes – the so-called final common pathway for behavior – where activation of flexor reflexes reciprocally inhibits extensor antagonists and vice versa. At the spinal cord per se, appetitive and aversive responses are reciprocally organized.

Recall that we defined the affect system as the efficient and manifold mechanisms that have evolved for differentiating hostile from hospitable stimuli and for organizing adaptive responses to these stimuli. The heterarchical nature of the central nervous system means that the functional architecture of the affect system extends downward from the bivariate evaluative plane at rostral levels to the reciprocal diagonal at the spinal cord wherein appetitive spinal cord reflexes and defensive spinal cord reflexes are reciprocally activated (see Figure 3.5). The ESM posits that affective states and responses are mediated by a network of distributed interacting neural regions in the central nervous system (including spinal cord reflexes). For instance, the spinal cord reflexes introduce a reciprocal bias in motor outputs, but this peripheral organization does not preclude the activation of both flexors and extensors (e.g., isometric contractions) via input from rostral brain areas (e.g., through voluntary efforts). Although the activities of the components of the affect system are generally integrated into a coherent cognitive/behavioral stream, the existence of multiple processing levels affords considerable flexibility in behavioral action as well as the potential for interference and conflict. In cases where relatively low-level processing is sufficient,

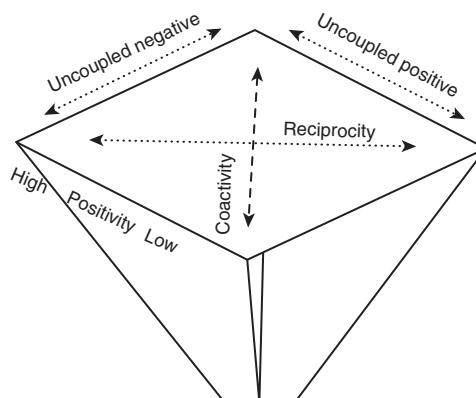


Figure 3.5 The bivariate evaluative space at the rostral levels extends downward to a bipolar line (the reciprocal diagonal) at the level of the spinal cord

or higher-level processing is precluded, lower substrates may predominate in behavioral expression. On the other hand, if higher cognitive computations, perhaps based on prior experience, raise the specter of a serious undesirable outcome, higher-level processes may predominate (e.g., recovering the billfold in the fire). In fact, although integrated to some extent, the multiple levels of processing may allow response conflicts, with different levels of processing each disposing toward differing behavioral responses. These conflicts may facilitate an outcome (e.g., retrieval of the billfold followed by a rapid, spinally mediated withdrawal from the fire), or may interfere (as in hesitancy, vascillation, or indecision). In part, the latter arises from the fact that physical constraints preclude both actions, as the arm and hand cannot extend (reach out) and flex (pull back) at the same time. Physical constraints may belie the complexity of the underlying dispositions, however. Although the limb may not be able to extend and flex at the same time, the distinct underlying flexor and extensor muscles can, in fact, both contract at the same time. This may lead to inaction, but that inaction is not sufficient evidence for a lack of response dispositions. The limb response may be constrained along a single bipolar continuum of flexion or extension, but the underlying neuromuscular machinery is not so constrained and may reveal a broader fundamental bivariate structure to neurobehavioral control. Thus, the framework provided by the ESM has the potential to promote theory and research on the affect system both at the level of the individual components and at the level of the integrated network.

THE THEORY'S APPLICABILITY TO SOCIAL ISSUES

The ESM has a variety of applications to social issues. Models of affect based on the valence continuum treat a reduction in negative affect as equivalent to an increase in positive affect. For instance, donor attitudes,

intentions, and behaviors have typically been conceptualized as organized along a bipolar continuum. Research on blood, organ, and bone-marrow donor behavior suggests that negative beliefs and emotions may constitute a particularly difficult obstacle to inducing donor behaviors, and that increasing people's positive beliefs and emotions toward donor behavior is not sufficient to disinhibit donor behavior. The ESM, which does not treat positive and negative affect as functionally equivalent, provides a framework and measurement methodology within which to study such phenomena.

The heterarchical structure proposed by the ESM also provides a framework for understanding implicit and explicit affective processes that may be playing a role in a variety of social problems, including racism. Briefly, latent inhibition in classical conditioning predicts that pre-exposure to a stimulus lessens the potency of classical conditioning, a prediction that has been confirmed in the human classical conditioning of affect (Cacioppo et al., 1992). If one assumes that children are exposed to majority more than to minority members, then salient distinguishing features of majority and minority members (e.g., skin coloration) of minority members should serve as a more powerful conditioning stimulus. Through the operation of latent inhibition, aversive depictions of whites and blacks in the news and media, even when equal, can lead to much stronger conditioned aversive responses to blacks than whites as a social category. These differential conditioned responses have no factual basis beyond the operation of latent inhibition, and the evocation of such feelings in the absence of a clear rationale promotes the search for an explanation and confabulation. When explicit rationalizations of this kind are challenged, the explicit beliefs can become less biased but this cognitive updating of facts and beliefs is not sufficient to fully eradicate conditioned emotional responses. That is, there are a multiplicity of evaluative representations within the affect system that reflect more than simple redundancy.

Finally, the ESM may offer unique contributions to our understanding of disorders of emotion such as depression, which, by its nature, is often profoundly interpersonal. Loss of reinforcement from the social environment, weakened social connections, reduced mastery of social situations, and uncontrollable social disruption are related to the onset of depression and its prediction of recurrence (Gortner et al., 1998; Lewinsohn et al., 1979). Moreover, social factors such as marital conflict, high levels of expressed negative emotion by family members, and low social support operate both as causes and consequences of depression (Hammen and Brennan, 2002). In this instance, initiating new affective responses to social cues, strengthening existing ones, or facilitating cognitive reappraisal to override attentional capture may modify depressive symptoms. However, it appears that these approaches and modification of social networks and behaviors alone are not sufficient to modify depression (Jacobson et al., 1996). The ESM may direct new integrative models that identify individual differences in affective processes, which, in their own right, may identify social triggers of depression.

The independence of the positivity offset and negativity bias indices, as well as the separability of positive and negative affective processes outlined by the ESM, may identify the social processes associated with depression. Research indicates that although both depression and anxiety are characterized by increased negative affect, they can be distinguished by the addition of decreased *positive* affect in depression (Watson et al., 1988). For example, it is theoretically possible that individuals who exhibit greater relative negativity bias and lower positivity offset activity may experience more negative affect, respond to a negative mood induction with more negative affect, and report a greater propensity for behavioral/social inhibition. That is, behavioral inhibition interferes with efforts to cultivate socially reinforcing environments and networks, which may inadvertently increase the risk for depression onset.

Moreover, depressed individuals may be distinguishable from nonpsychiatric healthy control subjects in their relatively higher endorsement of negativity bias and lower endorsement of positivity offset. Conversely, through the lower activation of the negativity bias and higher activation of positivity offset, an individual may seek to generate novel and appealing social contexts that are sufficient to encourage mastery and pleasure. These individuals may experience more positive affect, respond with more positive affect to social stimuli, and have a greater propensity for environmental and social exploration. These initiatives may reap dividends in promoting social support and warding off depressive symptoms. Research has yet to determine whether the relative difference in positivity offset and negativity bias changes with symptom remission, but the ESM offers a theoretical blueprint for distinguishing affective processes within affective disorders.

CONCLUSION

Most readers will be familiar with Occam's Razor – that theoretical entities should not be multiplied unnecessarily. Perhaps not all readers will be familiar with Einstein's Razor, however, which states that theorists should make everything as simple as possible, but not simpler. Our goal in the ESM is to begin to delineate details of the functional architecture and operating characteristics of the affect system. For instance, the ESM posits that the simple bipolar valence continuum, even when complemented by an arousal dimension as in the circumplex model, is not a comprehensive model of the affect system. The ESM instead posits that the outputs of the evaluative processors comprising the affect system determine bivalent action tendencies and actions. Such an organization fosters function: free and swift approach to appetitive stimuli and rapid and unfettered withdrawal from aversive stimuli.

ACKNOWLEDGMENTS

Preparation of this article was supported by grants from the National Institute of Mental Health (Grant No. P50 MH72850).

REFERENCES

Baumeister, R.F., Bratslavsky, E., Finkenauer, C. and Vohs, K.D. (2001) Bad is stronger than good. *Review of General Psychology*, 5(4), 323–370.

Berntson, G.G., Boysen, S.T. and Cacioppo, J.T. (1993a) Neurobehavioral organization and the cardinal principle of evaluative bivalence. *Annals of the New York Academy of Sciences*, 702, 75–102.

Berntson, G.G. and Cacioppo, J.T. (2008) The neuroevolution of motivation. In J. Shah and W. Gardner (eds), *Handbook of Motivation Science*, pp. 188–200. New York: Guilford.

Berntson, G.G., Cacioppo, J.T. and Quigley, K.S. (1991) Autonomic determinism: The modes of autonomic control, the doctrine of autonomic space, and the laws of autonomic constraint. *Psychological Review*, 98(4), 459–487.

Berntson, G.G., Cacioppo, J.T. and Quigley, K.S. (1993b) Cardiac psychophysiology and autonomic space in humans: Empirical perspectives and conceptual implications. *Psychological Bulletin*, 114(2), 296–322.

Berntson, G.G., Sarter, M. and Cacioppo, J.T. (1998) Anxiety and cardiovascular reactivity: The basal forebrain cholinergic link. *Behavioural Brain Research*, 94(2), 225–248.

Brehm, J.W. (1956) Postdecision changes in the desirability of alternatives. *The Journal of Abnormal and Social Psychology*, 52(3), 384–389.

Cacioppo, J.T. and Berntson, G.G. (1994) Relationship between attitudes and evaluative space: A critical review, with emphasis on the separability of positive and negative substrates. *Psychological Bulletin*, 115(3), 401–423.

Cacioppo, J.T. and Berntson, G.G. (1999) The affect system: Architecture and operating characteristics. *Current Directions in Psychological Science*, 8(5), 133–137.

Cacioppo, J.T., Crites, S.L., Jr. and Gardner, W.L. (1996) Attitudes to the right: Evaluative processing is associated with lateralized late positive event-related brain potentials. *Personality and Social Psychology Bulletin*, 22(12), 1205–1219.

Cacioppo, J.T. and Gardner, W.L. (1999) Emotion. *Annual Review of Psychology*, 50, 191–214.

Cacioppo, J.T., Gardner, W.L. and Berntson, G.G. (1997) Beyond bipolar conceptualizations and measures: The case of attitudes and evaluative space. *Personality and Social Psychology Review*, 1(1), 3–25.

Cacioppo, J.T., Gardner, W.L. and Berntson, G.G. (1999) The affect system has parallel and integrative processing components: Form follows function. *Journal of Personality and Social Psychology*, 76(5), 839–855.

Cacioppo, J.T., Larsen, J.T., Smith, N.K. and Berntson, G.G. (2004) The affect system: What lurks below the surface of feelings? In A.S.R. Manstead, N.H. Frijda and A.H. Fischer (eds), *Feelings and Emotions: The Amsterdam Conference*, pp. 223–242. New York: Cambridge University Press.

Cacioppo, J.T., Marshall-Goodell, B.S., Tassinary, L.G. and Petty, R.E. (1992) Rudimentary determinants of attitudes: Classical conditioning is more effective when prior knowledge about the attitude stimulus is low than high. *Journal of Experimental Social Psychology*, 28(3), 207–233.

Cacioppo, J.T., Priester, J.R. and Berntson, G.G. (1993) Rudimentary determinants of attitudes. II: Arm flexion and extension have differential effects on attitudes. *Journal of Personality and Social Psychology*, 65(1), 5–17.

Cacioppo, J.T., Rasinski, K.A., Sauerbrei, B.A. and Brehm, J.J. (2009) Ambivalence in the political arena. Unpublished manuscript.

Carver, C.S. (2001) Affect and the functional bases of behavior: On the dimensional structure of affective experience. *Personality and Social Psychology Review*, 5(4), 345–356.

Crawford, L.E. and Cacioppo, J.T. (2002) Learning where to look for danger: Integrating affective and spatial information. *Personality and Social Psychology Review*, 13(5), 449–453.

Crites, S.L.J. and Cacioppo, J.T. (1996) Electro cortical differentiation of evaluative and nonevaluative categorizations. *Psychological Science*, 7(5), 318–321.

de Gelder, B. and Vroomen, J. (2000) The perception of emotions by ear and by eye. *Cognition and Emotion*, 14(3), 289–311.

Edwards, J.D. and Ostrom, T.M. (1971) Cognitive structure of neutral attitudes. *Journal of Experimental Social Psychology*, 7(1), 36–47.

Gardner, W.L. and Cacioppo, J.T. (1995) Multi-gallon blood donors: Why do they give? *Transfusion*, 35(10), 795–798.

Gollan, J.K., Norris, C., Heyne, M. and Cacioppo, J.T. (2009) Translating basic science to the clinical context: Improving outcome in behavioral treatment of depression. Unpublished manuscript.

Gortner, E.T., Gollan, J.K., Dobson, K.S. and Jacobson, N.S. (1998) Cognitive-behavioral treatment for depression: Relapse prevention. *Journal of Consulting and Clinical Psychology*, 66(2), 377–384.

Grabenhorst, F., Rolls, E.T., Margot, C., da Silva, M.A.A.P. and Velazco, M.I. (2007) How pleasant and unpleasant stimuli combine in different brain regions: Odor mixtures. *Journal of Neuroscience*, 27(49), 13532–13540.

Gray, J.A. (1987) *The Psychology of Fear and Stress* (2nd edition). Cambridge: Cambridge University Press.

Green, D.P., Goldman, S.L. and Salovey, P. (1993) Measurement error masks bipolarity in affect ratings. *Journal of Personality and Social Psychology*, 64(6), 1029–1041.

Green, D.P., Salovey, P. and Truax, K.M. (1999) Static, dynamic, and causative bipolarity of affect. *Journal of Personality and Social Psychology*, 76(5), 856–867.

Hammen, C. and Brennan, P.A. (2002) Interpersonal dysfunction in depressed women: Impairments independent of depressive symptoms. *Journal of Affective Disorders*, 72(2), 145–156.

Hass, R.G., Katz, I., Rizzo, N., Bailey, J. and Eisenstadt, D. (1991) Cross-Racial Appraisal as Related to Attitude Ambivalence and Cognitive Complexity. *Personality and Social Psychology Bulletin*, 17, 83–92.

Hatfield, E., Cacioppo, J.T., and Rapson, R.L. (1994). *Emotional Contagion*. New York: Cambridge University Press.

Holbrook, A.L., Krosnick, J.A., Visser, P.S., Gardner, W.L. and Cacioppo, J.T. (2001) Attributes toward presidential candidates and political parties: Initial optimism, inertial first impressions, and a focus on flaws. *American Journal of Political Science*, 45(4), 930.

Ito, T. and Cacioppo, J. (2005) Variations on a human universal: Individual differences in positivity offset and negativity bias. *Cognition and Emotion*, 19(1), 1–26.

Ito, T.A. and Cacioppo, J.T. (2000) Electrophysiological evidence of implicit and explicit categorization processes. *Journal of Experimental Social Psychology*, 36(6), 660–676.

Ito, T.A., Cacioppo, J.T. and Lang, P.J. (1998a) Eliciting affect using the international affective picture system: Trajectories through evaluative space. *Personality and Social Psychology Bulletin*, 24(8), 855–879.

Ito, T.A., Larsen, J.T., Smith, N.K. and Cacioppo, J.T. (1998b) Negative information weighs more heavily on the brain: The negativity bias in evaluative categorizations. *Journal of Personality and Social Psychology*, 75(4), 887–900.

Jackson, J.H. (1958/1884) Evolution and dissolution of the nervous system (Croonian Lectures). In J. Taylor (ed.), *Selected Writings of John Hughlings Jackson* (Vol. 2). New York: Basic Books.

Jacobson, N.S., Dobson, K.S., Truax, P.A., Addis, M.E., Koerner, K., Gollan, J.K. et al. (1996) A component analysis of cognitive-behavioral treatment for depression. *Journal of Consulting and Clinical Psychology*, 64(2), 295–304.

Kahneman, D. (1973) *Attention and Effort*. Englewood Cliffs, NJ: Prentice-Hall.

Kahneman, D. and Tversky, A. (1979) Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 263–292.

Lang, P.J., Bradley, M.M. and Cuthbert, B.N. (1990) Emotion, attention, and the startle reflex. *Psychological Review*, 97(3), 377–395.

Larsen, J.T., McGraw, A.P. and Cacioppo, J.T. (2001) Can people feel happy and sad at the same time? *Journal of Personality and Social Psychology*, 81(4), 684–696.

Larsen, J.T., McGraw, A.P., Mellers, B.A. and Cacioppo, J.T. (2004) Research article the agony of victory and thrill of defeat mixed emotional reactions to disappointing wins and relieving losses. *Psychological Science*, 15(5), 325–330.

Larsen, J.T., Norris, C.J., McGraw, A.P., Hawkley, L.C. and Cacioppo, J.T. (2009) The evaluative space grid: A single-item measure of positivity and negativity. *Cognition and Emotion*, 23, 453–480.

LeDoux, J.E. (2000) Emotion circuits in the brain. *Annual Review of Neuroscience*, 23(1), 155.

Lewinsohn, P.M., Youngren, M.A. and Grosscup, S.J. (1979) Reinforcement and depression. In R.A. Dupue (ed.), *The Psychobiology of Depressive Disorders: Implications for the Effects of Stress*, pp. 291–316. New York: Academic Press.

Miller, N.E. (1951) Learnable drives and rewards. In S.S. Stevens (ed.), *Handbook of Experimental Psychology*, pp. 435–472. New York: Wiley.

Miller, N.E. (1961) Analytical studies of drive and reward. *American Psychologist*, 16(12), 739–754.

Norris, C.J. and Cacioppo, J.T. (2009) An fMRI investigation of neural mechanisms underlying the negativity bias. Unpublished manuscript.

Norris, C.J., Larsen, J.T., Crawford, L.E. and Cacioppo, J.T. (in preparation) Better (or worse) for some than others: Individual differences in the positivity offset

and negativity bias. *Journal of Research in Personality*. Unpublished manuscript.

Petty, R.E. and Cacioppo, J.T. (1981) Issue involvement as a moderator of the effects on attitude of advertising content and context. *Advances in Consumer Research*, 8, 20–24.

Petty, R.E. and Cacioppo, J.T. (1986) The elaboration likelihood model of persuasion. *Advances in Experimental Social Psychology*, 19, 123–205.

Russell, J.A. (1980) A circumplex model of affect. *Journal of Personality and Social Psychology*, 39(6), 1161–1178.

Schupp, H.T., Cuthbert, B.N., Bradley, M.M., Cacioppo, J.T., Ito, T. and Lang, P.J. (2000) Affective picture processing: The late positive potential is modulated by motivational relevance. *Psychophysiology*, 37(2), 257–261.

Sears, D.O. (1983) The person-positivity bias. *Journal of Personality and Social Psychology*, 44(2), 233–250.

Skowronski, J.J. and Carlston, D.E. (1989) Negativity and extremity biases in impression formation: A review of explanations. *Psychological Bulletin*, 105(1), 131–142.

Smith, N.K., Cacioppo, J.T., Larsen, J.T. and Chartrand, T.L. (2003) May I have your attention, please: Electrocortical responses to positive and negative stimuli. *Neuropsychologia*, 41(2), 171–183.

Smith, N.K., Larsen, J.T., Chartrand, T.L., Cacioppo, J.T., Katafiasz, H.A. and Moran, K.E. (2006) Being bad isn't always good: Affective context moderates the attention bias toward negative information. *Journal of Personality and Social Psychology*, 90(2), 210–220.

Spelke, E., Hirst, W. and Neisser, U. (1976) Skills of divided attention. *Cognition*, 4(3), 215–230.

Thurstone, L.L. (1928) Attitudes can be measured. *American Journal of Sociology*, 33(2), 529–554.

Thurstone, L.L. (1931) The measurement of attitudes. *Journal of Abnormal Psychology*, 26, 249–269.

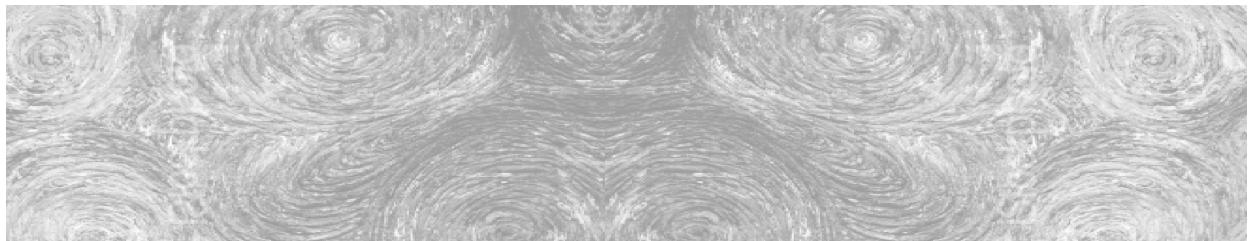
Thurstone, L.L. and Chave, E.L. (1929) *The Measurement of Attitude*. Chicago: University of Chicago Press.

Watson, D., Clark, L.A. and Tellegen, A. (1988) Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54(6), 1063–1070.

Watson, D. and Tellegen, A. (1985) Toward a consensual structure of mood. *Psychological Bulletin*, 98(2), 219–235.

PART II

Cognitive Level of Analysis



Accessibility Theory

E. Tory Higgins

ABSTRACT

Over 30 years have passed since I began studying the priming and accessibility of social constructs. What has made the journey since then so enjoyable are all the colleagues and friends who collaborated in deepening and broadening our knowledge of the nature and functions of priming and accessibility. Writing this chapter helped me to appreciate how much we have learned over the years about the sources and the consequences of accessibility as a basic principle of psychology. We know, for example, that when recent and frequent priming are combined with chronic accessibility (the additivity principle), even extremely vague information about someone will be used as a basis for forming an impression of that person (the compensation principle). We know that when people are aware of the priming event, they will often correct for its influence, thereby producing an opposite judgmental bias (contrast). We know that subliminally priming social category knowledge will produce actions in line with that knowledge (even in the absence of a member of the category), but the specific action will also depend on someone's attitude toward category members. This latter lesson is part of a recent and developing story that priming and accessibility effects depend on the *relevance* of a stimulated construct. Priming and accessibility is not just about cognition. It is about motivation.

ACCESSIBILITY THEORY

When I was a social psychology graduate student at Columbia in the early 1970s, most of the seminars I took were taught by Stanley Schachter, but I ended up doing most of my research in collaboration with Janellen Huttenlocher who was then a professor at Columbia Teachers College. Together we did research in psycholinguistics; more specifically, research on verbal reasoning. My dissertation advisor was Robert Krauss and my dissertation was on developmental and social class differences in communication. Given all this, when I applied for a job at Princeton, the faculty members were not quite sure what I was. Because of my background with Schachter and Krauss, I was partly a social psychologist, but my overarching research issue – the relations among language, thought, and society – was not what interested most social psychologists. Because of my research on psycholinguistics and communication, I was partly a cognitive psychologist, but my formal graduate training was in social psychology.

So what did Princeton do? Princeton's Solomon-like decision was to hire me in their "open" slot (rather than in either social psychology or cognitive psychology) where I would teach developmental psychology courses, without my ever having taken even one course in developmental psychology. The rationale was simple, however. After all, wasn't Janellen Huttenlocher herself a developmental psychologist (even though our research together was *not* developmental)? And wasn't my dissertation on developmental differences in communication? Well, I guess they had me there!

I relate this story because I believe that it is precisely my background as a part social psychologist and part cognitive psychologist that was critical to my early research on priming and accessibility. Indeed, before going to Columbia for my PhD I had worked on a social psychology MA at the London School of Economics and Political Science. "Social psychology MA" you say? Aha! That means that you *were* a social psychologist! Actually, it is not that simple. My advisor at LSE, Norman Hotopf, was a PhD student of Bertrand Russell and a postdoctoral student with Frederick Bartlett. He is best known for his work on eye movement and on the relation between language and thought. His hero was Roger Brown. Once again, he was not a typical social psychologist and was, if anything, more of a psycholinguist. But before I left LSE in 1968 he had said to me that what I should do in the future is to try to combine social psychology and cognitive psychology: "You know, something like *social cognition*."

Well, it took several years before I followed his advice, but once I reached Princeton in 1972, I decided to try to do exactly what he had told me – to combine social psychology and cognitive psychology. I had been interested in person perception for some time and was familiar with the "New Look" perspective on perception. The New Look literature had reported how individuals' needs and expectancies could influence their judgments of things, such as a greater need for money

being associated with judging coins to be bigger in size. At the general level, what this work seemed to demonstrate ("seemed" because the conclusions were still somewhat controversial) was that individual difference factors could influence judgments separately from the properties of the target of judgment. In Bruner's (1957a, b) terms, individuals can be set or *ready* to perceive certain things because they want them or expect them to be there, and this readiness results in their *going beyond the information given* in their judgments; that is, they make a judgment of the target that is not based solely on the target's properties.

According to Bruner (1957a: 133), categories varied in their accessibility, which is "the readiness with which a stimulus input with given properties will be coded or identified in terms of a category." To use Bruner's (1957a: 130) example, if the category "apple" has high accessibility, then "apples will be more easily and swiftly recognized, a wider range of things will be identified or misidentified as apples, and in consequence the correct or best fitting identity of these other inputs will be masked." What this means is that the likelihood that some input will be categorized in terms of a given category depends not only on the overlap between the sensory input and the category's specifications but also on the accessibility of that category.

Where does accessibility come from? This, of course, was a key question. According to Bruner (1957a), the relative accessibility of a category depends upon two factors. First, accessibility depends on the *expectancies* of a person regarding the likelihood of a *type of event that will be encountered* in the environment, with higher accessibility for events of higher likelihood. High accessibility from such frequency of past exposure functions to *minimize the surprise value of the environment*. Second, accessibility depends on the *search requirements* of the person imposed by that person's *needs and ongoing enterprises*. High accessibility in this case functions to *maximize the attainment of sought-after objects and events*.

It occurred to me that there was something about Bruner's conceptualization of accessibility that was incomplete. According to this theory, a category was accessible in the present when we expected or wanted an instance of that category to appear. When a category was accessible, we looked for and searched for an instance of that category in order to avoid surprises or attain desired end-states. I was not satisfied with this account of accessibility. I was not satisfied because it did not cover the case of accessibility which Janellen Huttenlocher and I had uncovered a few years earlier.

After arriving at Columbia in 1968, I began working with Huttenlocher on explaining why different kinds of three-term syllogisms varied in difficulty – the psychology of verbal reasoning. The so-called “three-term series problem” involves two premises and a question, such as: “Tom is lighter than Bob”; “Dick is heavier than Bob,” “Who is heaviest?” The difficulty of solving the problems depends on the linguistic form of each premise (e.g., contains a marked adjective [light] versus an unmarked adjective [heavy]; is a regular form versus a negative equative [“Bob is not as heavy as Dick”]). What Janellen and I most cared about was to explain how the information from the two premises was combined, and how the linguistic form of each premise influenced the process of combining the information. But we got involved in a debate with Herb Clark about verbal reasoning that forced us to provide some explanation for the effect of the adjective in the question – “Who is heaviest?” versus “Who is lightest?” Janellen and I found that participants solved the problem faster when the adjective in the question matched the active in the *second* premise, that is, the most recent premise. We proposed that activation of the category meaning of the adjective in the second premise made that meaning temporarily more accessible, which in turn made it easier to activate that meaning again when the same adjective appeared in the question (Huttenlocher and Higgins, 1971) – what we would now call a *recent priming* effect on *temporary accessibility*.

In our studies, which involved many trials for each participant, the likelihood that the adjective in the question would be a mismatch to the adjective in the second premise was the same as the likelihood that it would be a match. Given this equal likelihood, having higher accessibility of the adjective in the second premise would not function to “minimize the surprise value of the environment.” In addition, the goal of the task was to search for the *person* who was the answer to the question (e.g., “Tom”) and not to satisfy some need associated with the adjective in the second premise. That is, comprehending the question faster because its adjective matched the adjective in the second premise was not due to the question adjective being a desired end-state. What was clear to me was that something about accessibility was going on in the three-term series problem which was not captured in Bruner's account. But what was it?

Much of the accessibility studied by Bruner and others in the New Look concerned relatively long-term differences in accessibility, such as differences in the value of a coin as a function of personal wealth. Some of the phenomena examined, such as food deprivation effects on perceiving food-related items, were more temporary, but even here it concerned hours of deprivation and not seconds. In contrast, the effect of the second premise adjective on comprehending the question adjective would be an accessibility effect after seconds of activation. Indeed, we knew that the effect of adjective match was reduced if participants had to say three digits between the second premise and the question (see Huttenlocher and Higgins, 1971). This is why we called it *temporary accessibility*. How might such *temporary accessibility* be conceptualized?

Huttenlocher and I had suggested that activating categorical meaning from exposure to a word (i.e., priming) made that meaning temporarily more accessible, which in turn made that meaning easier to activate upon later exposure to a meaning-related stimulus. In our case, the later meaning-related stimulus

was the same adjective appearing again. But what if something more general were going on? What if recent activation of category meaning from priming a word made that category meaning temporarily more accessible, which in turn made it easier to activate upon later exposure to *behavioral information* that related to that meaning? If so, then verbal priming of a category in one situation could increase the likelihood of that category being used subsequently to categorize someone's behavior in a separate situation – without any awareness of the first situation having influenced the subsequent social judgment. And what if the target person's behavior was evaluatively ambiguous; that is, had features that fit two different categories with opposite valence (e.g., a persistent person versus a stubborn person)? If so, then perceivers could end up with either a positive or negative *impression* of the target person depending on whether – in an earlier totally separate situation – they had been verbally primed with one word or another (i.e., the word “persistent” versus the word “stubborn”). This would be a very different kind of accessibility effect!

MY FIRST PRIMING AND ACCESSIBILITY STUDY: THE “DONALD” STUDY

Psychologists had long recognized that categorizing an object or person in a certain way had subsequent effects on how that person would later be remembered and evaluated (e.g., Carmichael et al., 1932; Kelley, 1950). But where did such categorization come from? The standard answer was some combination of the subjective perception of the target's properties (i.e., the perceived overlap between those perceived target properties and the properties of different stored categories), plus some motivational biases or preferences for applying one or another category to a particular target. Temporary accessibility of a category meaning from recent verbal

priming was a *very different* kind of source. This source had nothing to do with the target's properties or the perceivers' needs. It concerned simply prior activation of the category meaning, which could occur in an *incidental* situation that had nothing to do with the target *and* could impact how the target is categorized without the perceiver being aware of the influence. It was the kind of scary source that reminded me of Freud's unconscious motives as a biasing factor in judgment – but it was *situational priming* rather than the id.

Now the question was whether temporary accessibility from verbal priming could really do all of this. In 1973 I began to develop a way to study this issue with an undergraduate at Princeton, Carl Jones, and then continued to work on it with Steve Rholes when he arrived as a graduate student in 1974. The psycholinguistic roots of the project were apparent in the title of Carl Jones' senior thesis, “An experiment on the effect of language on nonlinguistic behavior.” The combining of cognitive psychology *and* social psychology was more evident in the final title of the published version of the research, “Category accessibility and impression formation” (see Higgins et al., 1977).

The first question was how to accomplish the verbal priming in a way that would increase a category's accessibility for more than just a few seconds, *and* in a manner that would psychologically separate the verbal priming event itself from the target person information used to form an impression. We decided to use an *unrelated studies* paradigm. The participants were told that there were two separate studies conducted by two different experimenters that took place in two separate rooms (separated by a long hallway). Given our procedure, the increased accessibility from priming in “the first study” had to last several minutes to have an effect on judgment in “the second study.” I should note that some of these details do *not* appear in the published paper, but it influenced our decisions when designing subsequent priming studies. The participants were told that

the purpose of the first study was “to examine the effects of information processing on perception,” and the purpose of the second study was to study “verbal comprehension.” It was within the first study that the verbal priming took place, and it was within the second study that participants were exposed to ambiguous target person information that they could use to form an impression of the target person.

In the supposed first study, the participants were told to name, as quickly as possible, the background color of ten different slides. Ostensibly to make the task more difficult, after naming the color of a slide the participants also had to repeat a word they had received aurally before the slide appeared. The verbal priming was accomplished by embedding the critical words among the other “memory” words, which meant that these primed words had to be remembered for several seconds. In this way, the verbal priming was strengthened, but the words themselves were secondary to the focal task. Indeed, by the time of the second study they did not consciously remember these “memory” words very well as episodic events.

There were four priming conditions involving four priming words and six other words:

- 1 *applicable, positive*: “adventurous,” “self-confident,” “independent,” “persistent”;
- 2 *applicable, negative*: “reckless,” “conceited,” “aloof,” “stubborn”;
- 3 *nonapplicable, positive*: “obedient,” “neat,” “satirical,” “grateful”;
- 4 *nonapplicable, negative*: “disrespect,” “listless,” “clumsy,” “sly.”

Note that in half of the conditions the priming words were “applicable” and in the other half they were “nonapplicable.” This was done to control for the possibility that somehow the positive or negative priming words would induce a positive or negative mood that, in turn, would influence impressions. The “applicable” and “nonapplicable” words were selected to be equally positive or equally negative to control for this possibility.

The verbal priming effect should occur only for the “applicable” words whose categorical meaning could be used to characterize the input information. They should *not* occur for the “nonapplicable” words. Only later did *applicability* become an important principle of knowledge activation theory in its own right. But this was the beginning.

Following the “perception” study, the participants went down the hallway and did the “reading comprehension” study. They were given a paragraph about a target person called “Donald” – who subsequently became infamous for overappearance in experimental studies – and they were told simply to familiarize themselves with the paragraph because later they would have to answer questions about it. It took subjects about two minutes to read the paragraph, which was composed of evaluatively ambiguous descriptions of Donald’s behaviors, such as the following *persistent-stubborn* description:

Once Donald made up his mind to do something it was as good as done no matter how long it might take or how difficult the going might be. Only rarely did he change his mind even when it might well have been better if he had.

After the participants had read the paragraph, they were asked to characterize Donald’s behaviors. As predicted, the study found that in the “applicable” conditions, but not in the “nonapplicable” conditions, participants generally characterized Donald’s behaviors in terms of the positive categories in the positive verbal priming condition and in terms of the negative categories in the negative verbal priming condition. In order to check for awareness of the priming words, one set of participants were told that we wished to know “whether *anything* about the ‘perception’ study interfered with or affected their behavior in the ‘reading comprehension’ study, as it was still easy for us to change our procedure to avoid such problems.” In other words, they would *help* us by telling us about the priming words. Only one participant even noticed a connection between the priming words and how he characterized Donald.

The participants returned 10 to 14 days later for “another comprehension study.” (They were only debriefed after this second session.) Although the positive and negative “applicable” participants during the first session had not differed in their ratings of Donald’s desirability as a person, they did differ *now*, rating him as more desirable if they had been positively primed than if they had been negatively primed. Importantly, this effect on attitudes toward Donald was found both for those participants who had been asked explicitly in the first session to characterize Donald (overt characterization) *and* for those who had not received such explicit instructions (no overt characterization). For the latter participants, their attitudes toward Donald were spontaneous and evident two weeks after priming. Thus, for them, the priming effect on categorization was *covert* in the first session and had a lasting, if not increasing, effect on attitudes.

To my knowledge, this was the first study to demonstrate that verbal priming in one situation could influence how target information was categorized several minutes later in a totally separate situation, without any awareness of this happening. Moreover, this impact on categorization influenced attitudes toward the target two weeks later. Although I had predicted that this would happen, I was still surprised that it did. In fact, it was surprising to many other people as well and was considered a fluke by some (i.e., a type II error). It was not fully accepted until it was replicated several times by us and by others (e.g., Srull and Wyer, 1979, 1980). Of course, by now it has been replicated hundreds of times in many different forms in many different labs. It is no longer considered surprising. Thanks to Thomas Mussweiler, it was even given a thirtieth birthday party, the “Donald Symposium,” at the Society of Experimental Social Psychology in 2007.

I should say, however, that I was not just surprised that the priming really worked. I was disturbed by it. Indeed, I am still disturbed by it. Bruner’s accessibility is

not disturbing. It is reasonable and even beneficial for current input information to be captured by an accessible category because it is expected or because it satisfies some need. But it is neither reasonable nor beneficial for a category to be accessible simply because it happened to be, incidentally, verbally primed in a prior situation; and then determine, several minutes later, how a person’s behavior is evaluatively categorized in a separate situation; and then, two weeks later, make that person seem more desirable or undesirable as a person. I will return to this problem at the end of this chapter.

GENERALIZING PRIMING AND ACCESSIBILITY EFFECTS: THE “DUNCKER CANDLE” STUDY

My first question after the “Donald” study was about the potential breadth of priming and accessibility effects. If priming could do this to people’s impressions of a target person, what else could it do? Working with Huttenlocher on “three-term series” problems had sparked my interest in problem-solving more generally. I became especially fascinated with “creativity” problems (e.g., Duncker, 1945). I wondered whether priming and accessibility could be used to get people to see something differently and thereby facilitate creative insight.

In 1976, I began working on this problem with William Chaires, another undergraduate at Princeton. To study this issue, we chose Duncker’s (1945) famous candle problem:

Subjects are seated at a table on which there is a cardboard wall, a candle, a full book of matches, and a box filled with thumbtacks. Subjects are told that their task is to affix the candle to the cardboard wall so that the candle burns properly but does not drip wax on the table. The difficult part of the problem is to think of using the box as a platform for the candle, rather than just as a container for the tacks. The critical factor in solving the problem appears to be how subjects encode the box filled with thumbtacks.

What is difficult about this problem is that the most natural way for people to categorize the stimulus is *box of tacks*, which makes people think of the box as just a *container* for the tacks (see Glucksberg and Weisberg, 1966). The “of” categorization associates the box with the tacks in an *undifferentiated* manner, which highlights its function as a container. To solve the problem requires changing set and thinking of the box instead as a potential platform – a platform for the candle! If the stimulus were categorized as a “box *and* tacks,” this could *differentiate* the box from the tacks, establishing the box as more of an independent object.

The trick, then, would be to make the “and” construction temporarily more accessible than the “of” construction by priming the “and” construction. To accomplish this, we again used the “unrelated studies” paradigm. The participants were told that we wished to study the effects of interference on long-term memory; they would be shown a series of objects to remember and, in the “interference” condition, they would have to work on a problem before recalling the objects. The “interference” was working on the Duncker candle problem.

The “objects to be remembered” were various different objects, such as a banana, a football, a pair of scissors, plus some objects that were a container with content, such as a bowl containing cereal and a carton containing eggs. The experimenter labeled the objects when they appeared on a slide. In the “Of” priming condition, the slides with container-content objects were labeled as “bowl *of* cereal,” “carton *of* eggs,” and so on. In the “And” priming condition, these same slides were labeled as “bowl *and* cereal,” “carton *and* eggs,” and so on. After priming, all participants were given the instructions for the “interference” candle problem. There was a time limit of ten minutes to solve the problem.

In the “Of” priming condition, the average time working on the problem was nine minutes and only 20 percent solved the problem, which is quite typical for this problem. In the “And” condition, on the other hand, the

average time was 4.5 minutes and 80 percent solved the problem – a dramatic improvement! These studies (Higgins and Chaires, 1980) demonstrated that priming and accessibility had a potential power over thought and behavior that was far greater than had been appreciated before. They were the first studies to demonstrate that priming could influence someone’s behavior on a task, including creative behavior on a difficult problem-solving task. They showed that priming and accessibility effects are not just restricted to influencing people’s recognition of *what things are* (e.g., someone’s behavior is “persistent” versus “stubborn”); they can also influence people’s recognition of *how things work*.

BEYOND ACCESSIBILITY FROM PRIMING

In 1978 I left Princeton for the University of Western Ontario in London, Ontario. For the first time I was in a department with colleagues who identified themselves as personality psychologists, such as Dick Sorrentino. Up to this point I had thought of accessibility only in terms of priming and temporary accessibility. Now I began to consider the possibility of their being individual differences in the *chronic accessibility* of stored constructs that was separate from the temporary accessibility produced by priming. But there was no available measure of individual differences in chronic accessibility. How might such differences be measured?

INDIVIDUAL DIFFERENCES IN CHRONIC ACCESSIBILITY

I wanted to distinguish accessibility from availability. Previous theories had emphasized individual differences in the kinds of constructs that are actually present in memory to be used to process the world, which involves a difference in construct *availability*

(see Kelly, 1955; Markus, 1977; Mischel, 1973). In contrast, I wanted a measure of individual differences in construct *accessibility*. At Western, Gillian King and I ultimately developed a measure where we asked participants to list the traits (maximum of ten) for: (1) a type of person whom you like; (2) a type of person whom you dislike; (3) a type of person whom you seek out; (4) a type of person whom you avoid; and (5) a type of person whom you frequently encounter. There was a four- to five-minute delay between completing one question and receiving the next question – time that was filled with a nonverbal task to wipe out working memory. To capture the *accessibility* of a construct, the measure selected traits that were listed *first* in response to a question – *output primacy*.

Now that we had a measure of chronic accessibility, the question was whether individual differences in chronic accessibility influence how target person information is processed. To examine this question, we first gave participants our chronic accessibility measure. Then, in an “unrelated study” that took place about a week later, they read information about a target person. The essay with the target person information was individually tailored to each participant. Half of the trait-related behavioral descriptions exemplified one of that participant’s chronically accessible constructs, and the other half of the descriptions exemplified a trait that was not accessible to that participant but was an accessible construct for another participant in the study – a *quasi-yoking* design. For example, the trait construct “friendly” might be a chronically accessible construct for one participant but not for another participant – although it was clearly available for the latter participant – and the essay description of the target person which *both* participants received would be as follows: “Person A is the kind of person who spontaneously strikes up conversations with others and who goes out of their way to say hello to someone.”

The first thing that we discovered was that individuals’ chronically accessible constructs are surprisingly idiosyncratic (see Higgins

et al., 1982). Looking at all possible pairs of participants, we found that the average percentage of overlap of accessible constructs was *less than 10 percent*, with more than half of the pairs having *no overlap at all* in their chronically accessible constructs! After reading the essay, there was a ten-minute delay filled with a nonverbal interference task, and then the participants were asked to reproduce the essay exactly, word for word (recall). Next they were asked to give their impression of the target person. For both participants’ recall and impression of the target person, a trait description of the target person from the original essay was *less* likely to be deleted if it exemplified a participant’s chronically accessible constructs than if it exemplified an inaccessible construct. There was also a high positive correlation between the evaluative tone of participants’ impressions of the target and their *liking* for the target ($r = 0.70$), which means that individual differences in chronic accessibility can influence attitudes toward others as well.

Around the time that the new research on chronic accessibility was being conducted (1979), I went to the University of Michigan as a visiting professor. It was during my stay there when I began to develop a more general model of knowledge activation (see Higgins and King, 1981). In addition to distinguishing between available versus accessible constructs, chronic versus temporary accessibility, and applicable versus nonapplicable constructs, the model introduced two additional distinctions: active versus passive processing (i.e., controlled versus automatic); accessibility as “top of the storage bin” (what became Wyer and Srull’s “storage bin” model) versus accessibility as “energy potential” (what became my “synapse” model).

ACTIVE (CONTROLLED) VERSUS PASSIVE (AUTOMATIC) PROCESSING

In the early 1970s when I was a graduate student at Columbia, one of Michael Posner’s

collaborators, Robert Warren, came to Columbia as a new assistant professor. Posner and Warren (1972) had distinguished between *active*, conscious processes that are deliberate and controlled, versus *passive*, unconscious processes that occur automatically and are uncontrolled – a distinction that became more popularly known as *controlled* versus *automatic* processing (Shiffrin and Schneider, 1977). Whereas *set* is an active process in which conscious attention is deliberately directed toward the expected event, priming effects, such as those found in the Higgins et al. (1977) study, involve passive, automatic activation of constructs: they occur without intention and without conscious awareness (see Bargh and Pietromonaco, 1982; Higgins and King, 1981; Smith et al., 1992). This distinction was highlighted in Higgins and King (1981) for good reason – it is the fact that accessibility effects can occur without intention or awareness which makes them so intriguing and disturbing.

ACCESSIBILITY AS ENERGY POTENTIAL (THE SYNAPSE MODEL) VERSUS TOP OF THE STORAGE BIN

In Higgins and King (1981) I also addressed for the first time what it meant for something to be accessible. I probably could have continued doing research on accessibility without explicitly addressing this question. After all, I could do priming studies and measure chronic accessibility without having to make an explicit claim about what it meant for something to be accessible. But two events happened that motivated me to think about this basic issue. The first event was a visit to the University of Alberta in the spring of 1978 where I gave a talk about the “Donald” study. An animal learning psychologist listened carefully to the talk and then, during the discussion period, asked me – in a thoughtful and polite manner – what I meant by something being “accessible.” It was an innocent, “pardon my ignorance” question

that completely floored me. I had no idea how to answer his question, and simply said that I needed time to answer it and could we talk about it later – which we did. And it was true – I did need time to answer his question.

The second event was in the summer of 1978 at the Ontario Symposium on Social Cognition that took place in London, Ontario when I had a conversation with Bob Wyer. What we talked about – among many other things late into the night – was his new “storage bin” model of accessibility that he presented at the conference (Wyer and Srull, 1981). According to Wyer and Srull’s storage bin model, priming a construct made it more accessible because when a construct is recently activated it is replaced on top of a layered storage bin, and constructs are subsequently used as a function of their position in the bin, beginning from the top of the bin. With time, other constructs are likely to be used and then deposited on top of the bin, thereby accounting for the decline of priming effects over time (i.e., accessibility decay). A construct that is frequently activated, that is, frequent priming, has a greater likelihood of having been activated recently and then deposited on top of the bin, thereby accounting for why decay effects over time are reduced when there is frequent priming (see Srull and Wyer, 1979, 1980).

Thanks to the Alberta question from months before, I had been thinking about the concept of “accessibility” for a while and I had a different metaphor in mind than a storage bin. My preferred metaphor was an energy cell whose energy or action potential is increased whenever the cell is activated, and this energy slowly dissipates with time. This was a more dynamic and less structural metaphor. Later on when I was at New York University (NYU), I realized that there were different possible “energy transmission” models, and my preference was a more biological system model than a battery-like model. I went to speak to a colleague of mine who was a neuroscientist, Tony Movshon. After explaining what I wanted my model to

accomplish, he suggested that I read the work of Eric Kandel at Columbia (e.g., Kandel, 1976). I became convinced that the functioning of the “synapse” was my best metaphor, and the model became the *synapse model* (see Higgins et al., 1985). But, to return to my story, when I was talking to Bob in 1978, I had only a vague energy cell model. However, even then, it was different from Wyer and Srull’s storage bin model.

An unusual thing happened in my conversation with Bob that night. We actually discussed what the essential differences between the two models were, and whether an experimental study could be designed to provide a critical test whose results would support one model but disconfirm the other. This was the first time when I had ever had such a conversation. It was the kind of conversation that philosophers of science claim that scientists should do – find ways to disconfirm a theory. Instead, most scientists look for ways to support their theory. But Bob and I tried to find an experiment that would disconfirm one of the two models. One critical difference between the models was that, according to the storage bin model, recent activation could maintain high accessibility, even more than an alternative frequently activated construct, as long as the recently activated construct (rather than the frequently activated construct) remained on top of the bin. The key was to make sure that no competing construct was activated during the delay period.

The full design and execution of such a critical test would take a few years – until after I left Western and went to NYU. At NYU I was blessed with a new colleague, John Bargh, who had very similar “burning issues” to me. Luckily for me, he was willing to collaborate, together with a graduate student Wendy Lombardi, to fully flesh out the design and deal with the issue of ensuring that no competing construct was activated following recent priming. The solution was to have participants perform a counting backward task immediately after the last prime, that is, the most recently activated construct. The positive and negative primes

were the same trait terms used in Higgins et al. (1977), plus synonyms of those trait terms. For each positive–negative set of trait-related terms (e.g., persistent, determined, steadfast versus stubborn, obstinate, headstrong), either the positive or negative trait was primed most frequently and its opposite was primed only once but most recently (i.e., the last prime), *followed immediately by the counting backward task*. Thus, frequency of activation was pitted against recent activation and no competing construct was activated after recent priming.

Because there was no competing construct after the last prime, the storage bin model predicts that the construct primed most recently will dominate categorization of Donald’s behaviors after both a short and a long delay filled with counting backward. But the synapse model makes a different prediction. After a very brief delay, the excitation level of the recently primed construct should be higher than that of opposite frequently primed construct because recent activation brings excitation to its maximal level and the frequently primed construct has already decayed to some extent. This would result in Donald’s behaviors being categorized in terms of the recently primed construct. But after a long two-minute delay, the frequently primed construct will have decayed less than the recently primed construct because, consistent with the synapse metaphor, the decay function for a construct that is multiply primed should be slower than the decay function for a construct primed only once. Given this, Donald’s behaviors should now be categorized in terms of the frequently primed construct.

The synapse model, then, predicts a reversal over time: recent priming dominating categorization after a brief delay, and frequent priming dominating after a long delay. Again, unusual in science, I telephoned Bob Wyer, described the final design, and asked him whether he agreed with the competing predictions. He did, which made us happy and grateful, and we could then proceed with the study. The results of the study

supported the predictions of the synapse model rather than the storage bin model (see Higgins et al., 1985), and this recency-frequency reversal effect over time was replicated by Lombardi et al. (1987). The same team later conducted a conceptual replication substituting chronic accessibility for frequent priming, still pitted against recent priming, and we replicated the reversal effect as a function of delay time (Bargh et al., 1988). This research program provided strong support for the synapse model, and it led to a revision of the storage bin model to account for this reversal effect (Wyer and Srull, 1989). It is also a fond memory that I will always cherish of collaborating with another scientist, Bob Wyer, on testing the competing predictions of two different theories.

KNOWLEDGE ACTIVATION THEORY: THE NEXT GENERATION

Explicitly distinguishing between accessibility from recent priming versus from frequent priming, and between accessibility from recent priming versus from chronic individual differences, turned out to be important not only for testing the unique predictions of the synapse model and the storage bin model, but also for appreciating a significant characteristic of accessibility – *its additive nature*. What the Higgins et al. (1985) and Bargh et al. (1988) studies demonstrated was that the accessibility of a construct lasted longer when its source was frequent priming or chronic accessibility than when its source was just recent priming. Moreover, the comparable findings in these two sets of studies demonstrated that accessibility from chronic accessibility functioned like accessibility from frequent priming. Together, this suggests that combining different sources of accessibility would heighten accessibility and make it last longer. It was as if you could combine Bruner's expectancy source and his need source and make accessibility even stronger – something which had never

been suggested. Was accessibility additive in this way? If it were, then people could not possibly know where a construct's current level of accessibility comes from – how much from chronic accessibility and how much from priming (an “aboutness” problem; see Higgins, 1998b).

ADDITIVITY IN ACCESSIBILITY

The storage bin model does not predict additivity of accessible sources. Instead, the accessibility of a construct from frequent priming or from chronic accessibility simply increases the likelihood that the construct will have been activated recently and then placed on top of the bin. That is, multiple sources simply increase the likelihood of obtaining an effect of recent priming. But there is evidence that accessibility *is* additive (Bargh et al., 1986; Higgins and Brendl, 1995). And this evidence includes a demonstration by Higgins and Brendl (1995) of an additivity effect on intensity of judgment from different levels of chronic accessibility combined with a short delay after recent priming plus prior frequent priming (two additional times) of the same recently primed construct. The condition of frequent priming plus recent priming with short delay should itself have placed the construct on top of the storage bin. From the perspective of the storage bin model, an additional factor of varying levels of chronic accessibility should not matter much, but in fact it mattered a lot.

Higgins and Brendl (1995) used a revised version of the standard measure of chronic accessibility described earlier which now took into account both primacy of output *and* frequency of output in order to distinguish different *levels of chronic accessibility* rather than just chronic versus nonchronic. Participants with varying levels of chronic accessibility for the construct “conceited” were (or were not) both frequently and recently primed with “conceited” and then immediately read about a target person “Sue”

who displayed behaviors which pilot testing had shown were extremely *vague* (i.e., they did *not* elicit spontaneous impressions that Sue was conceited). The participants were simply asked to give their impression of the target person. Like the pilot participants, even those experimental participants who were both recently and frequently primed and had only a short delay before judging Sue did *not* characterize her as being conceited. Nothing happened with this extremely vague input, with one exception. For those participants who had relatively high levels of chronic accessibility for the construct “conceited,” suddenly Sue appeared to be conceited. And the higher the participant’s level of chronic accessibility for “conceited,” the *stronger* their impressions that Sue was conceited were. Compared to the other experimental participants and the pilot participants, it was as if these participants were *hallucinating* – they were seeing something that wasn’t there. The strength of this additivity effect was, once again, disturbing in its implications, which I discuss later in the section on “Applicability to Social Issues.”

ASSIMILATION VERSUS CONTRAST EFFECTS AND THE PRINCIPLE OF JUDGED USABILITY

The additivity effect means that it is not possible for anyone to know where the current level of a construct’s accessibility is coming from. It could be from recent priming, frequent priming, chronic accessibility, or any combination of these. This makes it difficult – indeed, impossible – to know how much accessibility we should correct for if we believe that recent priming might bias our current judgments. What happens when people do try to correct for potential bias from priming? This became another major issue in the next generation of knowledge activation theory.

The phenomenon of correcting for bias from priming was first introduced by Martin (1986). What was highlighted early on was

that the priming event itself was an episodic event that could be remembered later at the time that the judgment of the target information was to be made (see Lombardi et al., 1987). Because the priming event itself was irrelevant to what the target was like, it would be inappropriate for it to influence the judgment of the target. As I noted earlier, it is neither reasonable nor beneficial for incidental priming to affect judgment. Naturally, then, people would be motivated to correct for this possible source of bias. Unfortunately, it is impossible to calibrate exactly how much accessibility to correct for and people often *overcorrect*. This produces a *contrast* effect where the target is not only not judged in terms of the primed construct but it is judged in terms of the *opposite*, competing construct, as when someone primed with “persistent” then judges Donald to be “stubborn.” Notice the correction is an overcorrection because the original ambiguous information is neither just persistent *nor* just stubborn.

The conditions that produce the standard assimilation effect of priming versus a contrast effect of priming have been studied extensively (for reviews, see Higgins, 1996; Mussweiler, 2003). Early on, knowledge activation theory emphasized the principle of *judged usability* to account for these effects (Higgins, 1996). Just because a construct is activated does not mean that it will necessarily be used subsequently to process input information. The construct could be judged to be irrelevant to use or judged to be inappropriate to use, as in the example above. Thus, the factors that determine whether a construct is activated must be distinguished from the factors that determine whether an activated construct is used or not.

Judged usability – which need not be a conscious process – can be a rather subtle process. The Higgins and Brendl (1995) study, for example, also had another condition with the usual ambiguous stimulus, in addition to the condition with the vague stimulus. For the ambiguous stimulus, the usual assimilation effect of priming was

found when participants were not aware of the priming event. The question was, what would happen when participants *were* aware of the priming event? What the study found was that among these participants there was still an assimilation effect as a function of chronic accessibility – the stronger the chronic accessibility, the greater the assimilation effect. What this suggests is that when overall accessibility is higher than what would be expected from the priming event alone (i.e., higher because the level of accessibility derives from chronic accessibility and not just priming), then people are more likely to judge the accessible knowledge as being usable and less in need of correction even when they are aware of the priming event.

KNOWLEDGE ACTIVATION: ACCESSIBILITY AND APPLICABILITY WORKING TOGETHER

What determines whether stored knowledge is activated? According to knowledge activation theory, both the accessibility of stored knowledge and its applicability to an input determine whether stored knowledge is activated, where *applicability* refers to the amount of overlap between the features of stored knowledge and the attended features of the input. Accessibility and applicability work together to activate knowledge according to a *compensation rule*: the greater the accessibility of stored knowledge, the less applicability is needed for knowledge activation to occur, and the greater the applicability, the less accessibility is needed for knowledge activation to occur. The first half of this compensation rule is demonstrated in the Higgins and Brendl (1995) study described above in which strong chronic accessibility of the construct “conceited” increased the likelihood that the vague input, which was input for which the construct “conceited” had very low applicability, would be categorized in terms of the primed

construct “conceited.” To appreciate the second half of the compensation rule, the difference between vague, ambiguous, and unambiguous input information needs to be appreciated (Higgins, 1996).

Vague versus ambiguous versus unambiguous input information concerns applicability. Input, such as a behavioral description, is vague when no stored construct has more than weak applicability to it. Input is ambiguous when there are least two alternative constructs with moderate to strong applicability to it – usually two alternatives with strong applicability in most past studies. Input is unambiguous when there is only one construct that is moderately or strongly applicable to it. To return to the second half of the compensation rule, with unambiguous input little accessibility is needed for knowledge activation because applicability is high and there are no competing alternatives. With ambiguous input, more accessibility of one of the alternatives is needed in order for that alternative to win over the competition. With vague input, much more accessibility is needed because the applicability is so low. Indeed, in the Higgins and Brendl (1995) study, even recent and frequent priming together were not enough to produce a categorization effect without the addition of chronic accessibility. For an ambiguous input, in contrast, recent priming is sufficient to produce a categorization effect.

KNOWLEDGE ACTIVATION THEORY: THE ROLE OF MOTIVATION

When I began doing research on priming and accessibility, I thought of it as part of cognitive psychology. As I discussed earlier, it began as part of work that I was doing in psycholinguistics, and then later I extended it to impression formation. To me, this became a cognitive psychology/social psychology interface, the “social cognition” that I had been advised to pursue several years before. But this social cognition was grounded in

cognitive psychology. It was part of the *information processlytizing* that I have described in an earlier paper with John Bargh (Higgins and Bargh, 1987). Perhaps the biggest change that has occurred for me over the last ten years – and, I believe, for others as well – is to rethink the *role of motivation* in knowledge activation principles.

IMPORTANCE = ACCESSIBILITY

Some relation between accessibility and importance has been recognized for a long time. For Bruner (1957), current needs such as hunger or poverty were postulated as increasing the accessibility of need-related objects such as food or coins (see also Higgins, 1981, for a discussion of motivation as a source of accessibility). In my self-discrepancy theory of how socialization creates strong ideal and ought self-guides in children (Higgins, 1989, 1991; see also my Regulatory Focus chapter in this volume), the motivational importance of a caretaker's response to a child's behavior was emphasized as a factor contributing to a child acquiring highly accessible (i.e., strong) ideal and ought self-guides. This developmental theory expanded the “accessibility = importance” proposal beyond current need satisfaction to trait-related constructs made important from past interactions with significant others (for later work building on this proposal, see Shah, 2003).

Later, this general “accessibility = importance” proposal provided the conceptual rationale for using individuals’ response times in reporting their personal ideal and ought goals as a measure of the relative importance of the promotion system (ideals) or the prevention system (oughts) in their self-regulation (Higgins et al., 1997). It was assumed, as before, that stored trait-related constructs which were important for self-regulation, such as being “friendly,” “athletic,” or “hardworking,” would have high chronic accessibility. Given this, it was

now assumed that trait-related constructs that had high chronic accessibility (as measured by response times for reporting them) must be important for self-regulation. Subsequent studies testing the effects of promotion and prevention strength (i.e., the effects of highly accessible ideals and oughts) on emotions, decision making, and performance supported this assumption (see Higgins, 1998a; Higgins et al., 1997; see also my Regulatory Focus chapter in this volume).

CONSTRUCT ACTIVATION FROM PRIMING = ACTION

During the same period that this “accessibility = importance” proposal was being developed and tested, another perspective on the role of motivation in knowledge activation was also being developed and tested. John Bargh, inspired by James’ (2007/1890) “ideo-motor action” notion, proposed that priming a stored construct would activate that construct and its associated behaviors, such as priming “elderly” would activate “walks slowly,” which would result in direct expression of the activated behavior. Indeed, in a now classic study, Bargh et al. (1996) did find that people walked slowly down a hallway after being primed with “elderly” – *in the absence of any elderly person*. (Priming effects on other kinds of behavior had been reported earlier [Carver et al., 1983; Higgins and Chaires, 1980], but under conditions where the current situation also contributed to what happened behaviorally.)

The “ideo-motor action” notion as proposed by James (2007/1890) is that activation of a bare idea can be sufficient to prompt action: “We think the act, and it is done” (see James, 2007/1890: 522). For James, the idea prompted action relatively immediately. For Bargh, there could be a delay between construct activation and carrying out a construct-related action. This greatly increases the significance of the ideo-motor action proposal. It is as if the construct activated by

priming becomes an *intent* to directly express the construct in action. By becoming an intent, it is like Lewin's "quasi-need" where there is a tension created that cannot be reduced until an action fulfills the quasi-need (see Lewin, 1951). Thus, the tension can remain for a considerable period until the construct-related action is expressed.

This James–Bargh "ideo-motor action" notion of direct expression in action of an activated construct, where activation can occur from subliminal priming, inspired countless studies after the seminal Bargh et al. (1997) research. Construct activation from priming was shown to affect all kinds of behaviors, from aggressive behavior to achievement behavior to cooperative behavior (for reviews, see Bargh, 2005; Dijksterhuis and Bargh, 2001). What fascinated researchers was the implication that behavior could be determined by situational priming occurring outside of a person's awareness that activated a construct which then "needed" to be behaviorally expressed – "bypassing the will" (Bargh, 2005). The behavior was not directed by a person's preferences or desires or chronic goals but by construct activation itself, *by the activated idea itself*.

CONSTRUCT ACTIVATION FROM PRIMING + VALENCE = GOAL-DIRECTED ACTION

Over the last five years or so, it has become increasingly clear that, generally speaking, construct activation from priming is *not* sufficient, by itself, to determine action. It is not the case that we think the act and it is done. A construct-related behavior activated by priming need not result in the direct expression of that behavior. Indeed, a behavior *opposite* to the primed behavior can be expressed depending on the valence of the activated construct.

To illustrate, consider a study by Cesario et al. (2006) which replicated the classic "elderly" study of Bargh et al. (1997).

Cesario et al. (2006) argued that walking slowly would only make strategic sense for individuals who *like* the elderly because they would walk slowly in order to interact with them better. Some of the participants in the Cesario et al. (2006) study liked the elderly, and when the concept "elderly" was subliminally primed with pictures of elderly men, these participants afterward walked slowly down the hallway (in the absence of any elderly person) – just as Bargh et al. (1997) had found. But for those participants who *disliked* the elderly, they walked *quickly* down the hallway after "elderly" was subliminally primed – as if they were trying to avoid interaction with an elderly person (even though there was, again, no elderly person present).

In earlier studies, Plaks and Higgins (2000) activated stereotypes for a teamwork partner that were associated with task performance, such as females being good on verbal tasks, and again found pragmatic behavioral choices rather than mimicry or direct expression of the stereotyped behavior; for example, *loafing* on a verbal task when preparing to working on that task with a female team partner. What these and the Cesario et al. (2006) studies show is that activation of categorical behavior, even from *subliminal* priming, does not necessarily result in direct expression of that behavior. Instead, people use the activated information to prepare for future interaction with a category member, and the behavior they prepare depends on a variety of factors, including their attitude toward category members (see Cesario et al., 2006 for a fuller discussion of this "motivated preparation to interact" proposal).

It seems, then, that activating a stored idea, even subliminally, does not necessarily produce direct expression of the activated behavior. It can produce the *opposite* behavior. Indeed, there is evidence that whether construct activation from priming even affects subsequent construct-related behavior can depend on whether value is attached to the activated construct. If we think of an activated construct as a state, such as the state of

achievement or the state of cooperation, then attaching positive value to that state, either momentarily or chronically, transforms it into a *desired end-state*, that is, a goal to be fulfilled. If the current state of a person is discrepant from this desired end-state, then the person is motivated to take action to attain the desired end-state (see Custers and Aarts, 2005). Thus, the enacted behavior related to a primed construct is *behavior directed toward attaining a personal goal*. It becomes comparable to the behavior directed toward fulfilling personal ideals and oughts that have high chronic accessibility – chronically accessible goals that can unconsciously direct behavior (Higgins, 1998a).

There is now substantial evidence that behavioral effects of construct activation from priming vary as a function of the value (momentary or chronic) that is attached to the construct as an end-state (e.g., Ferguson and Bargh, 2004; for reviews, see Custers and Aarts, 2005; Eitam and Higgins, 2010). There is evidence, for example, that participants who are deprived of water will drink more when the construct “drink” is primed than when it is not primed, but priming has no effect when participants are not water deprived (e.g., Strahan et al., 2002). There is also evidence that when subliminal priming increases the positivity of goal-supporting objects, it does so most strongly for individuals for whom the goal currently is more important (Ferguson, 2008).

IMPORTANCE = ACCESSIBILITY REVISTED: ACCESSIBILITY AS RELEVANCE

According to the traditional version of the “importance = accessibility” notion, when there is a goal (or need), goal-related constructs will be activated, which will make these constructs accessible. Construct accessibility depends on the frequency and recency of the activation and the time since the last activation. Motivation affects the frequency

and recency of activation. It does not play a role, postactivation, in accessibility changes over time. Motivation, such as the need for accuracy or the need to avoid closure, can still affect judgment postactivation (e.g., Ford and Kruglanski, 1995), but this is through affecting the use of the accessible construct rather than the accessibility of the construct per se.

This traditional perspective on the relation between importance and accessibility did not change until very recently. Throughout the twentieth century, accessibility remained basically a cognitive variable that can be influenced by motivational factors. But more and more accessibility itself is becoming a motivational variable; or, more accurately, a *motivated cognition* variable. What is becoming increasingly clear is that accessibility depends on motivation. *Accessibility increases and decreases as a function of its relevance to current self-regulation.*

What determines the relevance of accessibility in current self-regulation? If accessibility levels serve current self-regulatory concerns, then we would expect the accessibility of goal-related constructs to be higher when a goal is higher in importance. Classic determinants of a goal’s importance are the value of goal attainment and the likelihood of goal attainment. Does the accessibility of a goal-related construct increase as the *value* and *likelihood* of goal attainment increase? Studies by Förster et al. (2005) indicate that indeed it does. Another classic determinant of a goal’s importance is goal fulfillment or completion. The importance of a goal decreases after it is completed. Does the accessibility of a goal-related construct decrease after the goal is completed? There is evidence that this also happens (e.g., Förster et al., 2005). Notably, there is also evidence that goal completion *inhibits* the accessibility of goal-related constructs rather than just suppressing construct expression (see Liberman et al., 2007). The importance of a goal also decreases if it is fulfilled through a substitutable task (Lewin, 1951). Does the accessibility of a goal-related construct

decrease after the goal is fulfilled through a substitutable task? There is evidence for this as well (Cesario et al., 2006).

The decrease in accessibility following goal completion which I described above can be understood in functional terms. Generally speaking, a completed goal that remains accessible could interfere with subsequent goal pursuit. The more important the goal, the more the risk of interference from goal-related constructs remaining accessible. This leads to the intriguing prediction that the more important a goal, the more goal-related constructs should decrease following goal completion. The Förster et al. (2005) studies supported this prediction as well.

But is it always the case that right after goal completion, the accessibility of goal-related constructs quickly decreases? Might such accessibility decay also depend on motivational relevance? Indeed, there is evidence for this from a study by Hedberg and Higgins (in press; see also Hedberg, 2007). Using the paradigm of Förster et al. (2005), the participants' task was to view a series of pictures and find when a picture of a pair of scissors was immediately followed by a picture of eyeglasses. Once this goal was completed, the accessibility of constructs related to "eyeglasses" was measured using a lexical decision task that appeared at different post-completion delay times. Hedberg and Higgins (in press) predicted that the accessibility decay function would be different for individuals with a strong promotion focus on accomplishment than for individuals with a strong prevention focus on security. In order to accomplish something new, promotion-focused individuals need to reduce the accessibility of "eyeglasses" in order that this old construct will not interfere with new accomplishments – precisely the kind of motivation discussed earlier. The more individuals are promotion-focused, the *faster* should be the rate of accessibility decay. Prevention-focused individuals, on the other hand, want to maintain the status quo until it is necessary to change. Thus, the more individuals are prevention-focused, the *slower* should be the

rate of accessibility decay – the opposite of promotion-focused individuals. These predicted *opposite* patterns of *accessibility decay* were found for promotion-focused and prevention-focused individuals.

The notion of accessibility as motivational relevance has been recently developed more fully and formally in a paper by Eitam and Higgins (2010). In a new framework called Relevance of a Representation (or ROAR for short), it is proposed that not all stimulated representations become active in the sense of being functionally available to mental processes such as categorization, planning, and effort allocation that impact judgment and behavior. ROAR proposes that a representation's impact on judgment and behavior over time does *not* derive from the maintenance of the representation's accessibility *per se*. Instead, the effects depend on the continuing relevance of the representation. Whereas the traditional notion is that priming produces accessibility that then decays at a certain rate as a function of factors like frequency of priming, ROAR suggests that what changes or is maintained is the motivational relevance of a representation. This means that the likelihood that stimulation of a representation would produce activation with judgmental and behavioral effects is independent of frequent or recent priming except for the effects that frequent or recent priming can have on motivational relevance.

APPLICABILITY TO SOCIAL ISSUES

People's evaluations and decisions are influenced by the accessibility of their attitudes, beliefs, and past experiences (for reviews, see Förster and Liberman, 2007; Higgins, 1996). People's actions, from their voting decisions to their choice of activities, are influenced by the accessibility of action-related constructs (e.g., Fazio, 1989; Lau, 1989). This has important implications for social issues because the decisions and actions associated with accessible constructs

include social discrimination and hostile actions toward others.

What is troubling about the relation between accessibility and action? If the current accessibility of a construct or goal for a person were related solely to its present utility for that person, or to its accuracy in representing the world that person is facing, then there would not be a problem. What is troubling, however, is that this is not the case. There can be sources of accessibility that are independent of current utility or accuracy. This fact raises two separate troubling issues: the issue of *control* and the issue of *truth*. People want to be effective in making something happen – to be effective at control. People also want to be effective at establishing what's real, to be accurate, correct, and right – to be effective at truth (Higgins, 2011). The results of priming studies raise questions regarding people's effectiveness at control and at truth.

Recently, the control problem has received the most attention, inspired especially by the work of John Bargh and his collaborators. The issue is whether it is people themselves who make things happen, that is, who are in charge, or whether instead they are directed (without their intent or awareness) by situational priming of goals and constructs. This is clearly an important issue with significant implications for social issues, such as controlling behavioral discrimination and stereotyping (e.g., Devine and Monteith, 1999). It should be noted, however, that the above discussion of motivational factors underlying accessibility effects from priming offers some hope regarding the control problem. It suggests that individuals who are motivated, for example, to behave nonprejudicially could do so – even unconsciously. Indeed, there is research on individuals with chronically accessible egalitarian beliefs that supports this conclusion (Moskowitz et al., 1999).

If there is some hope for the control issue, what about the truth issue? The truth issue was always my “burning issue.” *Accessibility does not know where it comes from.*

A current level of accessibility can be the result of any combination of recent situational priming (after a short delay or a long delay), frequent situational priming, and personal chronic accessibility. Knowledge activation itself is even more complex because all of these sources of accessibility combine with applicability to a target that can be ambiguous, vague, or relatively unambiguous. A person does not know where the activation comes from. It is impossible to calibrate accurately the contributions of these different sources. And the dominant tendency of people is to infer that the source of the activation is the contribution of the target, that is, that the activation is *about* the target's stimulus features (the “aboutness principle”; see Higgins, 1998b). The Higgins and Brendl (1995) study demonstrates this tendency to treat the target as the source of activation – to perceive Sue as “conceited” despite the vague input when there is strong accessibility. Poor Sue was hardly conceited, but she was judged as such under conditions that maximized accessibility (i.e., frequent and recent priming with a short delay plus high chronic accessibility).

What does this say about the truth of our judgments and evaluations (and memories)? What does it mean for our ability to be accurate, correct, veridical (see also Kruglanski, 1989; Popper, 1959)? Does this false “aboutness” from uncalibrated sources of knowledge activation contribute to delusions and hallucinations, to false eyewitness testimony, to stereotypic and ethnocentric evaluations, to the conviction that others who disagree with our beliefs must be either lying or crazy? I believe that it does. It is a serious problem. And this problem cannot be solved by somehow ridding ourselves of the principles of knowledge activation. After all, the same principles underlie human education and learning. The knowledge activation principles are among the great trade-offs of being human. They necessarily produce errors of judgment, memory, decision making, and behavior, while at the same time they are a necessity if we are to learn from experience.

They free us from the confines of the here-and-now, but it comes with a costly downside regarding truth (Higgins, 2000).

So what can we do about it? We can't surgically remove the functioning of knowledge activation principles without losing the benefits along with the costs. Perhaps this is where other people come in – the importance of not being an “island unto yourself.” We could, for example, compare our evaluation of a target to those of other people. This should improve validity, reliability, and objectivity because other people's sources of accessibility are likely to be different from ours while the target's actual properties remain constant (generally speaking). Indeed, all things being equal, this is probably a good way to reduce error. The fly in the ointment is that people tend to engage in such comparisons with in-group members – often a small number of significant others – with whom they want to create a *shared reality* (Echterhoff et al., 2009; Hardin and Higgins, 1996; Higgins, 1992; see also the Shared Reality chapter in this volume). By working for a shared reality, they bring their accessibility sources in line with one another, thereby undermining the benefits of having *independent* evaluations of a target. This is a strong argument in favor of having true *diversity* among individuals when a target is being evaluated or remembered. What's the bottom line? Introduce group diversity whenever possible to reduce the downside of knowledge activation principles on undermining the truth – increased benefits with reduced costs.

Diversity, then, is one solution to the truth problem created by accessibility and the other knowledge activation principles. But the truth problem has very broad significance that needs other kinds of solutions as well. The significance of the issue is evident in the important role that accessibility has been given in models of attitudes and behavior (e.g., Ajzen and Sexton, 1999; Fazio, 1990), interpersonal relations (e.g., Andersen and Chen, 2002), and motivated cognition more generally (e.g., Kruglanski et al., 2006). In the theory of planned behavior (see Ajzen

and Sexton, 1999; see also the Planned Behavior chapter in this volume), for example, the functioning of each of the major sources of behavioral intentions (i.e., attitudes, norms, perceived control) depends on what is accessible. Thus, whether or not beneficial health behaviors, environmental behaviors, intergroup behaviors, and so on actually occur depends on what knowledge is accessible and activated when the behavioral choices are made. This has profound implications for interventions targeted at increasing beneficial behaviors. Methods need to be devised that increase the likelihood that stored knowledge which supports beneficial behavioral intentions is accessible when the choices are actually being made. Interventions targeted at accessibility could be as important for dealing with social issues as interventions targeted at any other contributing factor to behavioral intentions.

REFERENCES

Ajzen, I. and Sexton, J. (1999) Depth of processing, belief congruence, and attitude-behavior correspondence. In S. Chaiken and Y. Trope (eds), *Dual-Process Theories in Social Psychology*, pp. 117–138. New York: Guilford Press.

Bargh, J.A. (2005) Bypassing the will: Toward demystifying the nonconscious control of social behavior. In R.R. Hassin, J.S. Uleman, and J.A. Bargh (eds), *The New Unconscious*, pp. 37–58.

Bargh, J.A., Bond, R.N., Lombardi, W.J. and Tota, M.E. (1986) The additive nature of chronic and temporary sources of construct accessibility. *Journal of Personality and Social Psychology*, 50, 869–878.

Bargh, J.A., Chen, M. and Burrows, L. (1996) Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology*, 71, 230–244.

Bargh, J.A., Gollwitzer, P.M., Lee-Chai, A., Barndollar, K. and Trotschel, R. (2001) The automated will: Nonconscious activation and pursuit of behavioral goals. *Journal of Personality and Social Psychology*, 81, 1014–1027.

Bargh, J.A., Lombardi, W.J. and Higgins, E.T. (1988) Automaticity of chronically accessible constructs in

person x situation effects on person perception: It's just a matter of time. *Journal of Personality and Social Psychology*, 55, 599–605.

Bargh, J.A. and Pietromonaco, P. (1982) Automatic information processing and social perception: The influence of trait information presented outside of conscious awareness on impression formation. *Journal of Personality and Social Psychology*, 43, 437–449.

Bruner, J.S. (1957a) On perceptual readiness. *Psychological Review*, 64, 123–152.

Bruner, J.S. (1957b) Going beyond the information given. In H. Gruber, G. Terrell, and M. Wertheimer (eds), *Contemporary Approaches to Cognition*. Cambridge, MA: Harvard University Press.

Carmichael, L., Hogan, H.P. and Walter, A.A. (1932) An experimental study of the effect of language on the reproduction of visually perceived form. *Journal of Experimental Psychology*, 15, 72–86.

Carver, C.S., Gannellen, R.J., Froming, W.J. and Chambers, W. (1983) Modeling: An analysis in terms of category accessibility. *Journal of Experimental Social Psychology*, 19, 403–421.

Cesario, J., Plaks, J.E. and Higgins, E.T. (2006) Automatic social behavior as motivated preparation to interact. *Journal of Personality and Social Psychology*, 90, 893–910.

Chen, S. and Chaiken, S. (1999) The heuristic-systematic model in its broader context. In S. Chaiken and Y. Trope (eds), *Dual-process Theories in Social Psychology*, pp. 73–96. New York: Guilford Press.

Custers, R. and Aarts, H. (2005) Beyond priming effects: The role of positive affect and discrepancies in implicit processes of motivation and goal pursuit. *European Review of Social Psychology*, 16, 257–300.

Devine, P.G. and Monteith, M.J. (1999) Automaticity and control in stereotyping. In S. Chaiken and Y. Trope (eds), *Dual Processes*, pp. 339–360. New York: Guilford.

Dijksterhuis, A. and Bargh, J.A. (2001) The perception-behavior expressway: Automatic effects of social perception on social behavior. In M.P. Zanna (ed.), *Advances in Experimental Social Psychology*, 33, 1–40. San Diego: Academic Press.

Duncker, K. (1945) On problem solving. *Psychological Monographs*, 58, 5 (Whole No. 270).

Echterhoff, G., Higgins, E.T. and Levine, J.M. (2009) Shared reality: Experiencing commonality with others' inner states about the world. *Perspectives on Psychological Science*, 4, 496–521.

Eitam, B. and Higgins, E.T. (2010) Motivation in *mental accessibility: Relevance Of A Representation (ROAR)* as a new framework. *Social and Personality Psychology Compass*, 4, 951–967.

Fazio, R.H. (1989) On the power and functionality of attitudes: The role of attitude accessibility. In A.R. Pratkanis, S.J. Breckler, and A.G. Greenwald (eds), *Attitude Structure and Function*, pp. 153–179. Hillsdale, NJ: Erlbaum.

Fazio, R.H. (1990) Multiple processes by which attitudes guide behavior: The mode model as an integrative framework. In M.P. Zanna (ed.), *Advances in Experimental Social Psychology*, 23, 75–109. New York: Academic Press.

Ferguson, M.J. (2008) On becoming ready to pursue a goal you don't know you have: Effects of nonconscious goals on evaluative readiness. *Journal of Personality and Social Psychology*, 95, 1268–1294.

Ferguson, M.J. and Bargh, J.A. (2004) Liking is for doing: Effects of goal pursuit on automatic evaluation. *Journal of Personality and Social Psychology*, 88, 557–572.

Ford, T.E. and Kruglanski, A.W. (1995) Effects of epistemic motivations on the use of accessible constructs in social judgment. *Personality and Social Psychology Bulletin*, 21, 950–962.

Förster, J. and Liberman, N. (2007) Knowledge activation. In A.W. Kruglanski and E.T. Higgins (eds), *Social Psychology: Handbook of Basic Principles* (2nd edition), pp. 201–231. New York: Guilford.

Förster, J., Liberman, N. and Higgins, E.T. (2005) Accessibility from active and fulfilled goals. *Journal of Experimental Social Psychology*, 41, 220–239.

Glucksberg, S. and Weisberg, R.W. (1966) Verbal behavior and problem solving: Some effects of labeling in a functional fixedness problem. *Journal of Experimental Psychology*, 71, 659–664.

Hedberg, P.H. (2007) What remains on your mind? Motivated up and down regulation of accessible knowledge. PhD dissertation, Columbia University.

Hedberg, P.H. and Higgins, E.T. (in press) What remains on your mind after you are done? Flexible regulation of knowledge accessibility. *Journal of Experimental Social Psychology*.

Higgins, E.T. (1989) Continuities and discontinuities in self-regulatory and self-evaluative processes: A developmental theory relating self and affect. *Journal of Personality*, 57, 407–444.

Higgins, E.T. (1991) Development of self-regulatory and self-evaluative processes: Costs, benefits, and trade offs. In M.R. Gunnar and A. Sroufe (eds), *Minnesota Symposium on Child Psychology*, pp. 125–165. Hillsdale, NJ: Erlbaum.

Higgins, E.T. (1996) Knowledge activation: Accessibility, applicability, and salience. In E.T. Higgins and A.W. Kruglanski (eds), *Social Psychology: Handbook of Basic Principles*, pp. 133–168. New York: Guilford.

Higgins, E.T. (1998a) Promotion and prevention: Regulatory focus as a motivational principle. In M.P. Zanna (ed.), *Advances in Experimental Social Psychology*, 30, 1–46. New York: Academic Press.

Higgins, E.T. (1998b) The aboutness principle: A pervasive influence on human inference. *Social Cognition*, 16, 173–198.

Higgins, E.T. (2000) Social cognition: Learning about what matters in the social world. *European Journal of Social Psychology*, 30, 3–39.

Higgins, E.T. (in press) *Beyond Pleasure and Pain: How Motivation Works*. New York: Oxford University Press.

Higgins, E.T. and Bargh, J.A. (1987) Social cognition and social perception. *Annual Review of Psychology*, 38, 369–425.

Higgins, E.T., Bargh, J.A. and Lombardi, W. (1985) The nature of priming effects on categorization. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 11, 59–69.

Higgins, E.T. and Brendl, M. (1995) Accessibility and applicability: Some 'activation rules' influencing judgment. *Journal of Experimental Social Psychology*, 31, 218–243.

Higgins, E.T. and Chaires, W.M. (1980) Accessibility of interrelational constructs: Implications for stimulus encoding and creativity. *Journal of Experimental Social Psychology*, 16, 348–361.

Higgins, E.T. and King, G. (1981) Accessibility of social constructs: Information processing consequences of individual and contextual variability. In N. Cantor and J. Kihlstrom (eds), *Personality, Cognition, and Social Interaction*, pp. 69–121. Hillsdale, NJ: Lawrence Erlbaum Associates.

Higgins, E.T., King, G.A. and Mavin, G.H. (1982) Individual construct accessibility and subjective impressions and recall. *Journal of Personality and Social Psychology*, 43, 35–47.

Higgins, E.T., Rholes, W.S. and Jones, C.R. (1977) Category accessibility and impression formation. *Journal of Experimental Social Psychology*, 13, 141–154.

Higgins, E.T., Shah, J. and Friedman, R. (1997) Emotional responses to goal attainment: Strength of regulatory focus as moderator. *Journal of Personality and Social Psychology*, 72, 515–525.

Huttenlocher, J. and Higgins, E.T. (1971) Adjectives, comparatives, and syllogisms. *Psychological Review*, 78, 487–504.

James, W. (2007) *The Principles of Psychology*, Vol. 2. New York: Cosimo (Original work published 1890).

Kandel, E.R. (1976) *Cellular Basis of Behavior: An Introduction to Neurobiology*. San Francisco: W.H. Freeman.

Kelley, H.H. (1950) The warm-cold variable in first impressions of persons. *Journal of Personality*, 18, 431–439.

Kelly, G.A. (1955) *The Psychology of Personal Constructs*. New York: W.W. Norton.

Kruglanski, A.W. (1989) *Lay Epistemics and Human Knowledge: Cognitive and Motivational Bases*. New York: Plenum.

Kruglanski, A.W., Erbs, H-P., Pierro, A., Mannetti, L. and Chun, W.Y. (2006) On parametric continuities in the world of binary either ors. *Psychological Inquiry*, 17, 153–165.

Lau, R.R. (1989) Construct accessibility and electoral choice. *Political Behavior*, 11, 5–32.

Lewin, K. (1951) *Field Theory in Social Science*. New York: Harper.

Liberman, N., Förster, J. and Higgins, E.T. (2007) Completed versus interrupted priming: Reduced accessibility from post-fulfillment inhibition. *Journal of Experimental Social Psychology*, 43, 258–264.

Lombardi, W.J., Higgins, E.T. and Bargh, J.A. (1987) The role of consciousness in priming effects on categorization. *Personality and Social Psychology Bulletin*, 13, 411–429.

Markus, H. (1977) Self-schemata and processing information about the self. *Journal of Personality and Social Psychology*, 35, 63–78.

Martin, L.L. (1986) Set/reset: Use and disuse of concepts in impression formation. *Journal of Personality and Social Psychology*, 51, 493–504.

Mischel, W. (1973) Toward a cognitive social learning reconceptualization of personality. *Psychological Review*, 80, 252–283.

Moskowitz, G.B., Gollwitzer, P.M., Wasel, W. and Schaal, B. (1999) Preconscious control of stereotype activation through chronic egalitarian goals. *Journal of Personality and Social Psychology*, 77, 167–184.

Mussweiler, T. (2003) Comparison processes in social judgment: Mechanisms and consequences. *Psychological Review*, 110, 472–489.

Plaks, J.E. and Higgins, E.T. (2000) The pragmatic use of stereotypes in teamwork: Motivational tuning to inferred partner/situation fit. *Journal of Personality and Social Psychology*, 79, 962–974.

Popper, K.R. (1959) *The Logic of Scientific Discovery*. New York: Basic Books. (Original work published 1934).

Posner, M.I. and Warren, R.E. (1972) Traces, concepts, and conscious constructions. In A.W. Melton and E. Martin (eds), *Coding Processes in Human Memory*. Washington, DC: Winston.

Sedikides, C. and Skowronski, J.J. (1991) The law of cognitive structure activation. *Psychological Inquiry*, 2, 169–184.

Shah, J. (2003) The motivational looking glass: How significant others implicitly affect goal appraisals. *Journal of Personality and Social Psychology, 85*, 424–439.

Shiffrin, R.M. and Schneider, W. (1977) Controlled and automatic human information processing: II. Preceptual learning, automatic attending, and a general theory. *Psychological Review, 84*, 127–190.

Smith, E.R., Stewart, T.L. and Buttram, R.T. (1992) Inferring a trait from a behavior has long-term, highly specific effects. *Journal of Personality and Social Psychology, 62*, 753–759.

Srull, T.K. and Wyer, R.S. (1979) The role of category accessibility in the interpretation of information about persons: Some determinants and implications. *Journal of Personality and Social Psychology, 37*, 1660–1672.

Srull, T.K. and Wyer, R.S. (1980) Category accessibility and social perception: Some implications for the study of person memory and interpersonal judgments. *Journal of Personality and Social Psychology, 38*, 841–856.

Strahan, E.J., Spencer, S.J. and Zanna, M.P. (2002) Subliminal priming and persuasion: Striking when the iron is hot. *Journal of Experimental Social Psychology, 38*, 556–568.

Wyer, R.S. and Srull, T.K. (1981) Category accessibility: Some theoretical and empirical issues concerning the processing of social stimulus information. In E.T. Higgins, C.P. Herman, and M.P. Zanna (eds), *Social Cognition: The Ontario Symposium, 1*, 161–197. Hillsdale, NJ: Erlbaum.

Wyer, R.S. and Srull, T.K. (1989) *Memory and Cognition in Its Social Context*. Hillsdale, NJ: Erlbaum.

A Theory of Impulse and Reflection

Fritz Strack and Roland Deutsch

ABSTRACT

The aim of the present chapter is to outline our theory of the dual determination of human behavior (Strack and Deutsch, 2004) in a larger context. Specifically, we attempt to demonstrate that impulse and reflection have long been identified as antagonistic forces. Although this distinction has emerged in various cultural and religious contexts as well as in different scientific disciplines, there exists no coherent account that describes the interaction of these influences along a sequence of information processing. Thus, it seemed justified to build a Reflective–Impulsive Model (RIM) that links thinking and feeling with behavior and allows connecting different research traditions and applied phenomena with the described duality. We begin by laying out the larger context in which we see the model to be embedded.

THE DUAL DETERMINATION OF BEHAVIOR

Cultural manifestations

In Greek philosophy, Plato created an allegory in his Phaedrus (Rowe, 1998) in which

he used two horses and a chariot to illustrate the different underlying mechanisms. Specifically, the chariot is meant to symbolize the human soul (in love) that is jointly pulled forward by both a bright and a dark horse. The two horses differ in their personality such that the bright horse is temperate, obedient and in no need of the whip while the dark horse acts tempestuously and does not yield to whip or spur. However, despite the dark horse's predominantly negative characteristics, Plato assigns it a crucial role in spontaneously approaching the other person and experiencing the joys of love. Thus, the lesson of this allegory seems to be that both modes of operation are necessary under different circumstances and that impulsive mechanisms are important for emotional action.

Quite a different lesson comes from the Christian religion, which has made the two modes of behavior a central theme. In particular, the Christian doctrine (e.g., Sorabji, 2002) holds that virtuous behavior may be in conflict with temptations. In this case, the “spirit” may be willing to guide one’s action

toward a positively evaluated end, the “flesh,” however, may be weak and cause the person to engage in sinful behavior. Despite its negative evaluation, the Christian religion acknowledges that “sinning” (i.e., deviating from the path of virtue) will repeatedly manifest itself in human behavior. Therefore, it is suggested that people purge themselves from the sins by verbalizing and repenting their sinful behaviors. Thus, the Christian opposition of spirit and flesh explicitly acknowledges the duality of human behavior and recognizes the emergence of conflicts. Moreover, it capitalizes on their inevitability and suggests solutions that allow actors to remain good despite their occasionally bad behaviors. It is important to note that the Christian religion subscribes to a philosophically dualistic doctrine in which virtuous and more reflective processes are seen as mental or spiritual phenomena whereas sinful and more impulsive mechanisms are assumed to reside in humans’ physical endowment.¹

Finally, most legal systems make a distinction between criminal actions that have been conducted with the intention to achieve a given outcome and those that are the consequence of mere negligence or of affective influences. Because punishment is determined by a person’s guilt and guilt is related to the intentionality of sinning, premeditated crimes are believed to require a harsher punishment than the other two types of criminal behavior. This differentiation also seems to be based on assumptions about different determinations of behavior, such that more responsibility is assigned to one than to the other. Responsibility, in turn, is related to reflective processes where the anticipated outcome determines the action. A lack of responsibility is more likely to be found for impulsive behaviors.

These examples demonstrate that the idea of a dual determination of human behavior is part of Western cultural knowledge that manifests itself in different ways. Of course, this is more a social acknowledgment of the existing duality than a precise analysis of its structure and dynamics.

Personality differences

As for any human characteristic, the question arises if, how and why humans differ on a relevant dimension. Not surprisingly, the distinction between reflective and impulsive processes has long been the focus of psychological theorizing.

One of the earliest and best-known approaches that focus on the described duality is psychodynamic theory (Freud, 1927). Specifically, impulses were assumed to originate from the id and to seek immediate behavioral expression. The ego, however, has the function to restrain the impulses and to reconcile their behavioral consequences with the requirement of the situation. While the id is governed by the pleasure principle, the ego obeys the reality principle. These two instances of the “psychic apparatus” along with the superego that contains social norms are assumed to evolve as a function of early socialization. Obviously, the id is in charge of expressing behavioral impulses whereas the ego reflects about their situational and normative appropriateness.

The idea of a psychic instance that controls impulses is also part of more recent theories of personality (e.g., Carver, 2005). Under the label of “ego psychology,” Block and Block (1980) have emphasized ego control as an important variable. High ego control can be understood as the dominance of reflective tendencies, low ego control as that of impulsive reactions. Not surprisingly, individuals who are high on this dimension are more likely to delay gratification (Funder and Block, 1989).

Because a moderate level of ego control is often considered to be most desirable, it would imply an adaptive balance between constraint and hedonistic enjoyment. Empirically, this claim has been supported by the outcomes of research on ego resilience, which describes the flexibility with which a person reacts to requirements of a situation in terms of more reflective or more impulsive responses. Interestingly, individuals found to be high on ego resilience had

moderate levels of ego control (Asendorpf and van Aken, 1999).

Within the more conventional trait approaches, the five-factor model (Costa and McCrae, 2009) plays a central role. In this framework, the balance between reflection and impulse is best represented by the factor called “conscientiousness,” which juxtaposes careful and planned deliberation with rash and unorganized reactions. Also, the factor “agreeableness” may be somewhat related to the reflective–impulsive dimension (see Carver, 2005) in that it stands for overcoming egotistic tendencies.

Interestingly, the dimensions of “ego control” on the one hand and “conscientiousness” and “agreeableness” on the other seem to be empirically related. That is, people who were low in ego control also had low scores in both conscientiousness and agreeableness (Asendorpf and van Aken, 1999).

Another scale to be mentioned comes from Zuckerman’s (1994a) work on sensation seeking. He identified a factor labeled “impulsive unsocialized sensation seeking” (IUSS) that plays a role in antisocial personality disorders (Zuckerman, 1994b) and is negatively correlated with conscientiousness and agreeableness (Zuckerman, 1996).

Within the five-factor model, Whiteside and Lynam (2001) have identified four distinct personality facets that are associated with impulsive behavior. They included the presence of both urgency and sensation seeking and the absence of both premeditation and perseverance.

THEORETICAL APPROACHES

Approaches from economics

A central feature of impulsive responses is that they are not future-oriented. Specifically, impulsive characteristics or processes are understood to cause behaviors to be immediately expressed whereas their reflective counterparts allow for temporal postponement

(McClure et al., 2004). This dimension of temporal delay is deeply rooted in economic thinking where the immediate consumption stands in conflict with saving, investment and consumption at a later time (Loewenstein and Thaler, 1997, 2005). In economics, this topic has come up under the name of “inter-temporal choice” and has long been explained by the concept of “temporal discounting” (e.g., Ainslie, 2005). This idea implies that individuals differ in the degree to which a given temporal distance decreases the utility of a given outcome. To compensate for this devaluation, the delayed outcome must be increased by the same degree. This rule, however, was oversimplified and the discount rate was shown to vary within one person. For example, the same compensation that would be necessary to get people to accept a given delay that would start from the present would be lower if the same delay would begin ten years from now. This hyperbolic discounting was seen as a way of explaining impulsive behavior (Ainslie, 1975) and a means of describing it. More recently, however, behavioral economists have abandoned the hope of fully understanding the delay of reward by identifying the right discount rate or function. Instead, insights from psychology are believed to provide some important progress toward an understanding of the underlying processes (Frederick et al., 2003).

The role of impulse in social psychology

Early approaches

It is interesting to note that in the first systematic approach to social psychology, social behavior was explained to be impulsively determined. In 1895, the French medical doctor Gustave Le Bon (1895) published a book titled *The Crowd* in which individuals were described as slaves of impulses if their actions occurred in a social context. When by themselves, their behavior was guided by

reason and responsibility. However, while in groups, they were assumed to lose their “conscious personality” and be suggestively influenced by the behavior of others. Resisting these suggestive impulses requires a “strong personality”. Otherwise, social behavior was seen to be driven by impulsivity, irritability, a lack of critical judgment and logical thinking and an exaggeration of emotions. According to Le Bon, these characteristics were localized in a “group mind” that was seen to be in charge of emotional behaviors. Thus, social action was understood not to be a function of its anticipated consequences but the result of uncontrolled forces from outside the individual. Moreover, the internal mechanisms that are elicited in social settings are assumed to differ from those that are operating under individual conditions. Thus, Le Bon’s analysis has invoked two systems with different operating principles that determine behavior under different circumstances. The two minds that he had been identified may be understood as two sets of processes of which either one or the other was operating. As a consequence, the author’s theory does not imply the possibility of a conflict between the impulsive “group mind” and its reflective “individual” counterpart.

Such a conflict, however, was explicitly admitted by Sigmund Freud (1993/1921) who has criticized Le Bon’s notion of a group mind. While denying the society as the source of impulsiveness, Freud acknowledged the existence of different operating principles which he identified *within* the individual. Specifically, the more reflective ego is often seen to be in conflict with the more impulsive id that is mainly located in the unconscious. Unlike in Le Bon’s theory, human thinking, feeling and acting was believed to be the product of the joint operation of the two instances. However, both theories agree that behavior is not only determined by rational considerations about its outcome but also by emotional forces. Also, both approaches agree that affective impulses may undermine a more reflective control of behavior.

The re-emergence of impulse in social psychology

During the last 30 years, the notion of impulse in social psychology has re-emerged in various forms (Strack and Deutsch, 2007), perhaps the most important ones being the distinction between automatic and controlled processes (e.g., Bargh, 1994; Schneider and Shiffrin, 1977), the distinction between rules and associations (e.g., Sloman, 1996), and the distinction between heuristic and systematic processing (e.g., Chaiken, 1980; Tversky and Kahneman, 1974). The belief that there is more than one way in which people can think, feel and act has paved the way to a view of human behavior that is not exclusively guided by the anticipation of its outcomes.

From a social psychological perspective, various duality theories have been summarized in the contributions to a book by Chaiken and Trope (1999). Instead of reviewing each of them, we shall merely point at some landmark theories and findings that we think have had a particularly strong influence (see also Deutsch and Strack, 2006a). However, it is important to note that notions of duality have been prevalent not only in social psychology, but also in other subfields of psychology. In cognitive psychology, for example, there has been a discussion about the possibility that the world may not only be represented in linguistically defined concepts but also in an analogue fashion. In particular, it was argued (e.g., Kosslyn, 1980) and demonstrated (e.g., Farah, 1988) that perceptual representations may play an important role alongside their conceptual counterparts. This dual representation implies that people may simultaneously possess a perceptual experience on the one hand and knowledge or beliefs about what is the case on the other (Sloman, 1996; Strack, 1992). Typically, the two representations are in line with one another. At times, however, they may be inconsistent. Extending findings from judgmental heuristics (Greifeneder and Bless, 2007), one would suspect that motivation and cognitive capacity determine if the

experiential or the noetic representation will dominate further processing. That is, the knowledge may override the experience.

Visual illusions may serve as an example (cf. Sloman, 1996). Being asked to judge the length of the two lines in the Müller-Lyer illusion, people may perceive the line between the open arrows to be shorter than that between the closed ones. At the same time, they may know that this impression is not caused by the actual length of the line but by contextual circumstances. As a consequence, they may not base their judgment solely on their perceptual experience but may use their knowledge to correct it (Strack, 1992; Koriat and Levy-Sadot, 1999; Sloman, 1996), which suggests different processes are connected with different contents.

The same is true for internally elicited experiences such as feelings. As Schwarz and Clore (e.g., Schwarz and Clore, 1983, 1996) have repeatedly demonstrated, feelings may enter judgments "as information." This means the effects that subjective experiences exert on judgments may be qualified by knowledge about their origin. Thus, positive affect that comes, for example, from the current weather conditions will not be used as a basis for judgments of happiness with one's life as soon as the person making the judgment becomes aware of the source. The same reasoning applies to "cognitive feelings" that are not accompanied by strong affect. Most prominently, the impact of feelings of fluency may be qualified if they are not representative for the judgments but originate from some irrelevant facet of the task or its context (e.g., Schwarz et al., 1991). Again, it should be noted that the described correction requires cognitive capacity and is less likely to occur under suboptimal circumstances (Strack et al., 1982).

Duality in stereotypes and attitudes

The impact of duality models on modern psychology was mediated by its influence in the domain of attitudes. At first, duality models had focused on the processes suggested as routes to persuasion. In the spirit of

Tversky and Kahneman's (1974; Kahneman and Tversky, 1982) notion of heuristics as ways of simplifying judgments under conditions of uncertainty, the process of persuasion was increasingly understood to be based on two different types of psychological processes (e.g., Chaiken, 1980; Petty and Cacioppo, 1986). On the one hand, the systematic route to persuasion will draw on the strength of arguments to come to a response to a persuasive message. Taking this more effortful route would require the appropriate motivation and cognitive capacity. On the other hand, the heuristic route to attitude change does not require scrutinizing the quality of the arguments but merely a less effortful check of peripheral cues such as the attractiveness of the communicator or the mere length of the persuasive message. Variants of these basic strategies and their consequences have been elaborated in different theories of attitude change. However, it has also been argued that the different routes to persuasion are not distinct in their psychological processes but in the type of information that enters into the judgments (Kruglanski and Thompson, 1999).

In a seminal paper, Devine (1989) offered a new perspective on prejudice by distinguishing between automatic and controlled processes in stereotyping. Specifically, she observed that although high and low prejudiced people may differ in their stereotyped beliefs about characteristics of group members, they did not differ in the accessibility of stereotypic contents in memory. Instead, low prejudice individuals seemed to be more motivated to prevent these activated thoughts from influencing judgment and action – at least if motivation and capacity to control prejudice are high (cf. Dunton and Fazio, 1997). These findings provide a strong hint that prejudice may be represented in two ways, as an evaluative belief and as a habitual tendency to activate certain thoughts.

The distinction between automatic and controlled components has also been applied to attitudes in general. Specifically, it has been argued that evaluations may exist not

only as momentary beliefs about something being good or bad, which may be constructed on the spot by reflection, but also as over-learned object-evaluation associations in memory (e.g., Fazio et al., 1986; Greenwald and Banaji, 1995; Wilson et al., 2000). Such evaluative associations may be activated automatically when relevant objects are perceived and may then immediately trigger behavioral tendencies (cf. Chen and Bargh, 1999; Solarz, 1960). Among these behavioral manifestations are the facilitations and inhibitions of reactions that are either compatible or incompatible with a given valence.

This reasoning has led to a duality in the domain of attitudes that has exerted an enormous influence on thinking in social psychology. It has caused the distinction between explicit and implicit attitudes (Greenwald and Banaji, 1995; Wilson et al., 2000; Wittenbrink and Schwarz, 2007), and has elicited a methodological discussion on the various procedures of measuring the implicit variant (De Houwer and Moors, 2007). Some procedures of “indirect” assessments of attitudes are based on pre-existing reactions that signal valence (like a smile or a frown), which are more easily executed in response to stimuli of compatible valence (Dimberg et al., 2002; Neumann et al., 2005). The response speed is thus an index of evaluative processes, which can be mapped even when participants are instructed to respond to a criterion other than valence such as the mere occurrence of the stimulus (Neumann et al., 2005). Other indirect attitude measures do not rely on pre-existing valence–behavior links, but create these links experimentally. For example, participants in Fazio and colleagues’ evaluative priming task (Fazio et al., 1986) are trained to categorize clearly positive and negative words according to their valence by making previously neutral responses (i.e., key presses). In a later priming phase, these responses are facilitated when the target stimuli are preceded by evaluatively congruent prime stimuli. In a similar vein, the Implicit Association Test (IAT) (Greenwald et al., 1998) requires

participants to categorize targets toward which the participants are suspected to have an implicit attitude with two keys. These keys, however, are also used to categorize clearly positive or negative stimuli, and after some training, the evaluation trials occur intermixed with the categorization trials. Depending on conditions, the category of interest shares the key with positive or negative stimuli. Depending on whether categorizations are quicker when the concept is mapped on the positive or negative key, the target concept is assumed to trigger a positive or negative automatic evaluation. For example, if participants are faster at pressing a particular key to categorize people as “old” if the same key is also used to indicate objects with a negative (compared to a positive) valence then this person is attributed a negative implicit attitude toward the elderly.

Despite some controversies about the concept of implicit attitudes and the techniques for their assessment (e.g., Blanton and Jaccard, 2006; De Houwer et al., 2009; Fiedler et al., 2006; Rothermund and Wentura, 2004), implicit attitudes as measured by the IAT and related procedures have proved to be valid predictors of a wide range of behaviors that occur under suboptimal circumstances (e.g., Friese et al., 2009). They reach from the tendency to shoot at minorities under stress (Unkelbach et al., 2008) to overeating under the influence of alcohol (Hofmann and Friese, 2008). At times, implicit and explicit attitudes are positively related to one another, but under specified conditions, they dissociate (Gawronski and Strack, 2004). It is exactly these conditions that allow deeper insights into the different operating mechanisms. Although attitudes are understood as predictors of behavior, the proposed duality has strongly directed the field’s attention toward the dual representation of evaluations.

Duality in behavior

Although many duality theories in social psychology focus on cognitive precursors of behavior such as attitude activation, person perception, stereotyping, and so on, at least

some address the issue of the dual determination of behavior more directly. In fact, the idea that behavior can be determined by multiple factors which may sometimes be in conflict is at the heart of the implicit social cognition research program (e.g., Greenwald and Banaji, 1995; Wilson et al., 2000). Moreover, a great number of studies have addressed the dynamics of automatic social behavior and how it is controlled by environmental cues and inner regulatory mechanisms (e.g., Bargh, 2005; Bargh and Ferguson, 2000; Bargh et al., 1996). Conflicts between reflection and impulse have received most attention in the realm of research on self-control and impulsivity (e.g., Baumeister et al., 1998; Hofmann et al., 2009; Loewenstein, 1996; Metcalfe and Mischel, 1999).

In general, the concept of self-control implies that the execution of a behavior is preceded by a conflict between two forces. Economists were particularly concerned about conflicts between short-term and long-term consequences of decisions (Loewenstein and Elster, 1992) and in psychology, Walter Mischel and his colleagues (Metcalfe and Mischel, 1999; Mischel, 1974) have explored children's psychological dynamics in situations where they have the chance to increase their reward if they are able to wait. One of their determinants was the attention on the reward and the way it is mentally represented (Mischel et al., 1972). More recently, self-control has been identified as an activity that requires some unspecified type of energy (Baumeister et al., 1994). Using the analogy of a muscle, Baumeister and his colleagues suggest that self-control depletes this energy and replenishes it after some time. Although this notion has not been developed into a full-blown dual-process or system theory, it shares with them the assumption that there exist two types of activities that differ in the energy (or cognitive capacity) they require. Specifically, there is energy consuming self-control and, by implication, the less taxing behavior that needs to be controlled.

An elaborate behavioral dual-system theory was proposed by Metcalfe and Mischel

(1999) to account for the dynamics that undermine people's "willpower." Specifically, they postulate a "cool" system that operates on the basis of knowledge and is only indirectly related to behavior. The second, "hot," system is driven by emotions that serve as immediate determinants of impulsive behaviors. The model describes how the two systems interact and, more importantly, how they conflict when self-regulation breaks down. Thus, impulsive behavior is understood to be controlled not by the power of the will but by emotional forces that may sometimes be incompatible with willful intentions.

The Reflective–Impulsive Model

Most duality theories focus either on cognition (such as dual process models of persuasion or person perception) or on behavior (such as the strength model of self-regulation), but only very few models attempted to integrate cognitive and behavioral aspects (e.g., Epstein, 1990; Metcalfe and Mischel, 1999). Our own model (Strack and Deutsch, 2004) was developed with an integrative goal and particularly aims at specifying the interaction between the mechanisms of the systems that jointly determine behavior.

Building blocks of the model

The Reflective–Impulsive Model (RIM) describes the operating principles of two mental systems (reflective and impulsive) that jointly and interactively generate behavior. It is assumed that the two systems run in parallel and interact during the processing. The impulsive system is believed to be permanently engaged without consuming cognitive capacity; the reflective system is only active if enough cognitive capacity is available. Both systems, however, jointly determine how a perceptual input may eventually result in behavior.

The impulsive system The most fundamental link between perception and behavior was

originally described as the ideo-motor principle (James, 1890; Lotze, 1852). It posits that without an intention or a goal, behavior may be elicited by merely perceiving or imagining certain content. As more recent findings suggest, activating concepts that include a motor component may exert similar effects (Rizzolatti et al., 2002). In the RIM, complex sensory-motor associations are an integral part of the impulsive system and are termed behavioral schemata. These structures resemble the concept of habits (Wood and Neal, 2007), and serve as the basis of overt behavior.

Associative structures. More frequently, the link between perception and behavior is less immediate but will be influenced by information that is already stored in memory like concepts and images. Stereotypes may exert such a mediating influence. For example, being exposed to a stereotype of an elderly person may cause people to reduce people's speed of walking (e.g., Bargh et al., 1996; Cesario et al., 2006). However, this influence seems to crucially depend on the actor's goal in a given situation (Cesario, et al., 2006). The absence of a pertinent decision or goal suggests that the behavior is elicited by a more complex structure in the impulsive system, which is assumed to emerge through past operations that connect perceptual input and behavioral output. Specifically, the impulsive system can be understood as an associative memory that closes the temporal gap between past and present. Based on the associative principles of frequency and recency, links between elements are assumed to emerge and are strengthened by their simultaneous activation. The general assumption of a passive creation of associative links, however, has recently been questioned (Mitchell et al., 2009). Instead, such links may mainly occur after more thorough encoding. There is, however, little doubt that, once associations have been formed, the activation of a particular element may be caused by another element with which it is strongly associated (like "hot" and "cold") or by

several other elements that are jointly activated. In the previous example, having been exposed to the elderly stereotype may not have been sufficient to initiate the congruent walking behavior without inducing the person to get walking otherwise (Cesario et al., 2006).

Affect and valence. Beyond associative links to behavioral schemata, the RIM postulates a modifying role of valence and affect that is experienced in the course of the behavioral execution. First, the impulsive system is influenced by a mechanism that facilitates approach and avoidance. Specifically, it is subject to a "motivational orientation" (Cacioppo et al., 1993; Neumann et al., 2003) that is triggered in a bidirectional fashion. On the one hand, the experience of positive or negative affect or the processing of positive or negative information may facilitate behavioral approach versus avoidance. In turn, these behaviors may facilitate the experience of the type of affect or the processing of the type of information that is compatible. In other words, decreasing the distance to a target (e.g., a person) will ease the experiencing of positive affect and the processing of positive information whereas increasing its distance will do the same for negative affect and information. It is important to note that on a spatial dimension, this may be achieved in two ways. Specifically, the distance may be changed by either moving away from the target (flight) or by causing the target to be moved away (fight).

It is important to note that changes in distance may be brought about by actual locomotion, by virtual operations on a computer screen (e.g., De Houwer et al., 2001; Markman and Brendl, 2005), or by symbolic actions like facial expressions and head movements (Neumann et al., 2005). Also, it is important to note that it has long been recognized that behaviors may influence evaluations. Specifically, Bem (1972) has posited that self-perception may be a causal determinant of attitudes. However, the proposed mechanism is quite different from the

feedback of behavior on evaluations as outlined in the RIM. Specifically, self-perception theory proposes that people draw inferences from their behavior to the underlying attitudes. In contrast, the mechanisms of the RIM's impulsive system do not require inferences to be drawn. For example, it has been shown that the effects occurred even if experimental conditions prevented participants from identifying the meaning of their own behavior. The motivational orientation can therefore be seen as a general predisposition of the impulsive system that directs information processing and behavior by means of facilitation.

While approaching positive and avoiding negative targets has adaptive value, survival requires basic needs to be fulfilled. This suggests that the impulsive systems should also be responsive to specific needs such as food if the organism is in a state of deprivation. In fact, supporting evidence was found by Seibt et al. (2007) who found that the degree of approach motivation (as measured by lever movements) triggered by food stimuli was intensified in hungry compared to satiated participants. In summary, affording automatic and fast adjustments to an environment is the speciality of the impulsive system.

The reflective system As much as speed is an important quality in reacting adaptively, fast behavior often prevents the actor from doing what is best. That is, the impulsive system has serious shortcomings that result from its fast and effortless processing. Specifically, it occurs with great rigidity and often fails to perform certain tasks. For example, the associative principle of frequency prevents single pieces of information from exerting an impact. Also, consequences that are temporarily remote, in both the future and the past, have less of an impact than they often should. In addition, the impulsive system has no place to allow for the transmission of knowledge without immediate exposure. That is, adaptation would be improved if people could learn from the experiences of others without the need to repeat them. This is

particularly beneficial if such experiences have negative consequences.

The RIM assumes that primates are endowed with a second system that operates according to different principles to compensate for the shortcomings of the impulsive system. It is termed the “reflective system” and can be understood as being in charge of the generation and transformation of knowledge. The term “knowledge” is used in a relatively narrow sense in the RIM. More specifically, knowledge is defined as propositional statements about the world and the self, which contain truth-values. The cognitive processes of the reflective system assign and transform truth-values and generate new knowledge by inferences.

Returning to the previous example, being exposed to an old person, the impulsive system would automatically activate representations corresponding to the elderly stereotype. The reflective system, on the other hand, would form a proposition about the relation between this characteristic and the perceptual input, resulting in the knowledge that “this person is old.” Such knowledge can then be used for further transformations and inferences. Most simply, the truth-value can be reversed by applying the operation of negation. Thus, one might generate the subjective knowledge, or belief, that “this person is not old.” More important, the person may apply categorical knowledge about old people to draw inferences about characteristics of the target that are not accessible to immediate perception; for example, that “this person is hard of hearing.”

In other words, the reflective system generates an epistemic process that consists of several operations. It starts with linking a perceptual input to a name or a category (a process that we have referred to as “pointing and referring”; Strack and Deutsch, 2004). This results in a “propositional categorization” or “identification” which may then serve as a basis for a “noetic decision.” The nature of this decision may be factual and/or evaluative and lead to a “behavioral decision,” which is meant to reduce a

discrepancy between a current state of the self and a possibility that has previously been positively evaluated. Given that decisions typically do not immediately result in behavior, it is plausible to assume specific processes to be responsible for mediating between decision and action. This process is labeled “intending” in the RIM (e.g., Gollwitzer, 1999). It links a behavioral decision to the actual execution of the behavior. The process of “intending” is particularly important to bridge the temporal gap between the “behavioral decision” and the execution of the behavior and to prevent the person from being preoccupied by reiterating or rehearsing the decision. Instead, it delegates the realization of the decision to the environment that reactivates it when the opportunity for the behavior is available. The process is concluded by the execution of the behavior or if the decision has become obsolete.

Interactions between the two systems Thus far, the two systems have been described separately, and we have identified both their central elements and their operational principles. In the following section, the interaction between the reflective and the impulsive system will be outlined by explaining how they relate at various stages of processing.

Most importantly, the operations in the reflective system need an informational basis. That is, to generate factual, evaluative, or behavioral decisions, it is necessary that relevant information is available. If the information comes from the environment, it needs to be categorized to serve as a basis for inferences. These concepts must have been stored at the time of categorization. Alternatively, the information does not come from concurrent environmental stimuli but from past experiences or from previous judgments. In all cases, information must have been stored to afford further processing. Different types of memory are assumed to be part of the two systems.

A working memory with a limited capacity and the capability to directly address its contents (Baddeley, 1986) is part of the

reflective system, while the impulsive system possesses a long-term store of unlimited capacity (e.g., Johnson and Hirst, 1991) that consists of associative structures.

The reflective system crucially depends on the associative store in that it provides the contents that are needed for its syllogistic operations. At a fundamental level, the execution of a propositional categorization depends on a category that must be available in memory. Even more so, inferences that are based on general knowledge can only be drawn if an appropriate schema is retrieved. This can be driven by the incoming stimulus if a category is strongly associated with the input. This category will be automatically activated and used for further processing in both the reflective and the impulsive system. In the absence of a strong link between the input and the category, the activation may be diluted and disappear. However, because the syllogistic operations in the reflective system rely on the input to be categorized, appropriate concepts need to be retrieved from the associative store in the impulsive system. In this case, the outcome depends not only on search cue but also on the activation potential (Johnson and Hirst, 1991) of the stored information, the accessibility of which relies on recency and frequency of its prior activation. Thus, each piece of information in the associative store has its own probability with which it enters into the operations of the reflective system. Thus, processes in the impulsive system exert a strong impact on reflective operations. Drawing from the seminal work by Higgins and his colleagues (Higgins, 1996), a person may be categorized as “reckless” only because people had previously been exposed to this concept, even if this had occurred in a different context.

From this logic, it follows that the causal impact may also go in the opposite direction. That is, the reflective use of a piece of information will alter its activation potential in the impulsive system. That is, thinking about a specific content will increase the probability that the same or related information will be

retrieved at a later time. As trivial as this may sound, it has severe consequences for the generation of judgments. This becomes apparent in findings on anchoring effects in judgments under uncertainty (Higgins et al., 1977). In one of the experiments developed within this paradigm, people are required to generate two judgments in sequence. The first judgment is comparative in nature and uses the “anchor” as a standard of comparison. For example, judges may be asked to indicate if the Mississippi River is longer or shorter than 3,500 miles (high anchor) or 1,500 miles (low anchor). Subsequently, they have to estimate the river’s actual length. Typically, this succession of two judgment tasks causes the absolute judgment to be assimilated toward the standard of the preceding comparative judgment. In its original account, the effect was explained as the result of an “insufficient adjustment.” This explanation, however, begs many questions and fails to specify the underlying mechanisms (Tversky and Kahneman, 1974). From the vantage point of the RIM, the observed assimilation is the result of a selective accessibility of information at the time of the absolute judgment that is caused by the preceding comparative judgment. Specifically, we (Mussweiler et al., 1997) were able to demonstrate that to generate comparative judgments, people act as if the standard of comparison is a single hypothesis (Klayman and Ha, 1987) and seek confirmatory information. Even if this biased search results in a rejection of the hypothesis, and in a response that is qualified by the standard, it also leads to a selective accessibility of information at the time of the absolute judgment. As a consequence, the information that is then retrieved is biased and the resulting judgment is assimilated toward the standard of the comparative task. Assuming such an interaction between the two systems allows for an understanding of judgmental biases that is linked to more general psychological mechanisms that are responsible for a great variety of phenomena. This parsimony is an important achievement of dual-system

models that are sufficiently specified in its processes.

Two types of interaction between the reflective and the impulsive system are described by these mechanisms. On the one hand, the existing activation potential determines the likelihood with which a content is used for reflective operations; at the same time, reflective processes alter the accessibility of information for operations of the impulsive system.

Synergistic and antagonistic operations It is an important feature of the RIM that the final pathway to behavior is jointly used by both systems. In the reflective system, behavior is the result of a noetic decision about the desirability of an outcome and the feasibility with which it can be achieved (Ajzen, 1991; Mussweiler and Strack, 1999). As has been outlined previously, the information on which the decision is based may be influenced by the selective accessibility of information in the impulsive system. However, the translation of a behavioral decision into overt behavior may not occur in an immediate fashion. Instead, the impulsive system may activate behavioral schemata that are incompatible with the behavioral decision and thus interfere with its execution.

This suggests that the compatibility of the two systems determines the strength of a particular behavior. Specifically, if the activities in both systems lead to the activation of the same schema, it will be easier to execute the behavior. It may be further facilitated if the impulsive system creates fluency and a feeling of positive affect (Bandura, 1977; Topolinski and Strack, 2009).

In contrast, if the two systems activated incompatible schemata or if the reflective system inhibits impulsive behaviors, the two systems compete with one another. At the same time, feelings of temptation and conflict may arise. Which system will prevail depends on situational and dispositional characteristics. In the following, we shall describe such determinants along with pertinent empirical evidence.

Determinants of impulsive versus reflective behaviors

Situational determinants For the reflective system to operate, cognitive capacity is necessary. Thus, being preoccupied by another mental activity will decrease the probability that behaviors are controlled by reflective operations. Specifying this distracting influence, Baumeister and colleagues (e.g., Winkielman and Cacioppo, 2001) have invoked the metaphor of energy that exists as a limited resource and will be depleted by a mentally or physically taxing task. Also, they claim that it takes time until the energy is replenished. This assumption is largely consistent with the RIM, which provides a conceptual basis to account for mechanisms of self-control (Baumeister and Heatherton, 1996).

Combining Baumeister' energy model with the RIM, Hofmann et al. (2007) have depleted participants' cognitive resources by requiring them to suppress a negative emotion that had been induced by a movie. Subsequently, they took part in a taste test in which M&M candies had to be evaluated in several dimensions. However, the researchers' interest was not in participants' evaluations but in their eating behavior. Specifically, they wanted to find out if they were more likely to act against their dieting intentions under depletion. This was the case. Participants' eating behavior corresponded to their dietary restraint standards much more if they had not previously been depleted. If they had, however, eating behavior was related to their automatic affective reactions to M&Ms that was previously assessed by a variant of the IAT (Hofmann et al., 2007). This finding was replicated for other types of food and for alcohol consumption (Greenwald et al., 1998). In summary, these results suggest that behavior is controlled by reflection only if the necessary cognitive resources are available. Otherwise, it is more likely to be elicited by affectively loaded impulses which can be assessed by implicit measures capturing the affect that is associatively linked to the target of the behavior.

In a more direct way, Friese et al. (2008) depleted participants' cognitive capacity by giving and having them execute a secondary memory task. Specifically, while they were invited to choose between chocolate bars and fruits, they had to memorize either an eight-digit number (high cognitive load) or a one-digit number (low cognitive load). As expected, people's actual behavior corresponded more strongly with their explicit evaluation if the secondary task was easy than if it was difficult. Again, under the difficult, high-load condition, automatic affective reactions (measured by the IAT) were more likely to predict the choice.

Finally, the behavioral conflict between reflective and impulsive determination can be resolved by alcohol consumption. Specifically, impulsive influences are assumed to be stronger under alcoholic intoxication. This prediction was tested by Hofmann and Friese (2008) in a study that again employed a taste test in which participants had to evaluate different kinds of candy products. Unlike in the previous study, there was a preceding "product test" that involved orange juice that was either blended with vodka or not. Again, the amount of candy that was consumed was the dependent variable. The results conceptually replicate those of the experiment by Hofmann et al. (2007). That is, alcohol consumption was equivalent with the depletion (i.e., affect suppression) manipulation in that participants' dietary restraint standards were less and associative affect was more effective in determining eating behavior.

Dispositional determinants Of course, a considerable part of the variation in people's tendency toward one of the two conflicting behaviors is determined by personal characteristics that may have emerged as a function of inheritance or learning. For example, such interpersonal differences can be found for the domain of cognitive capacity. Specifically, working memory capacity has been assessed and studied in its relation to behaviors that often occur in an impulsive fashion. As a

consequence, people's actual behavior can be better predicted using explicit evaluative assessments if their cognitive capacity is high. If it is low, impulsive evaluative reactions should be better predictors.

This pattern was tested using erotic stimuli. The authors (Hofmann and Fries, 2008) recorded how long participants would look at erotic versus art pictures. In addition, their working memory capacity was assessed. The findings demonstrate that for those people who had a high memory capacity attention toward the erotic pictures was best predicted by explicit evaluations. In contrast, the IAT was the best predictor for those whose working memory capacity had been low.

APPLICABILITY TO SOCIAL ISSUES

Beyond providing a conceptual structure to integrate diverse psychological phenomena, the RIM is relevant to various issues in the real world. This is particularly apparent in the reception that the model received in several disciplines of applied psychology.

One outstanding example is the field of health psychology where interventions play an important role in preventing people from engaging in harmful behaviors. This has been particularly evident in the case of addiction. For example, the RIM implies that addictive behaviors are predominantly controlled by the impulsive system which does not have the ability to negate propositional contents (e.g., Deutsch et al., 2006; Gilbert, 1991). Instead, the affirmative content will be activated. As a consequence, persuasive communications with demands that include a negation (e.g., "Don't smoke!") should be less effective than communications that do not mention the harmful behavior at all. A study comparing media campaigns against smoking suggests that this is, in fact, the case (Farrelly et al., 2002). As a recommendation, the RIM suggests that in the case of impulsive behaviors, just saying no is not enough (Gawronski et al., 2008).

Implications of the model have been applied to the harmful consumption of alcoholic beverages and addictive behaviors in general (Deutsch and Strack, 2006b). Reinout Wiers and his colleagues have found that alcoholic beverages affected drinkers' motivational orientation. For these persons, behaviors approaching alcoholic beverages were found to be facilitated (Wiers et al., 2009). More importantly, counterimpulsive training proved effective even over an extended period of time. Specifically, heavy drinkers' consumption of alcohol was reduced if they had to engage in avoidance behaviors (pushing a lever away) while being exposed to pictures of alcoholic beverages (Wiers et al., 2010).

In the health domain, the conflict between dietary restraint and the impulse to eat caloric food is particularly obvious. The choice between an apple and a chocolate ice cream is not the same as between an apple and a pear, because impulses are more likely to be involved in the first case. Wilhelm Hofmann and his colleagues (for a review, see Hofmann et al., 2009) have convincingly demonstrated that reflective control of tempting behaviors breaks down if the psychological operating conditions of reflective control are undermined or not fulfilled; for example, when people are temporarily depleted of self-regulatory resources, under the influence of alcohol, or low in chronic working memory capacity. Thus, the colloquial "strength of willpower" in successfully resisting temptation appears to be grounded in specific antecedents of reflective operations related to concepts of central executive functioning. In cases where these functions are temporarily or chronically ineffective, impulsive processes appear to gain the upper hand in determining behavior (e.g., Allan et al., 2010; Hofmann et al., 2008; Hoefling and Strack, 2008).

The term "impulsivity" has also been used to describe behaviors in the domain of marketing. There, "impulse buying" (Vohs and Faber, 2007) is understood as a type of economic exchange that is not determined by

rational anticipation of the outcome of the action but by an immediate urge to appropriate a product. Of course, the RIM is particularly capable of explaining the interaction between rational choice and impulsive behavior which is at the core of economic phenomena (Strack et al., 2006).

More generally, the RIM seeks to integrate previous research and theorizing on stereotypes and attitudes, and to connect these constructs to the general assumptions about judgment, motivation, and behavior. We hope that this integration may be helpful in improving and stimulating research aimed at reducing intergroup conflict and discrimination.

CONCLUSION

The reported findings show that impulsive and reflective types of behavior are systematically related to determinants that can be situational in nature or may reside as a stable characteristic in the person. In both cases, the results strongly suggest that human behavior is not a unified phenomenon but is determined by different psychological mechanisms. They are assumed to belong to different systems whose operation depends on specific conditions.

The structure of the RIM allows studying the synergistic or antagonistic influences of the two systems and to explore the conditions of their operation. Some of these have been described in this chapter. Others are open to future explorations. For example, it might be possible to facilitate reflective versus impulsive processing using appropriate induction tasks. Specifically, this might be achieved by inducing concrete versus abstract representations that have been found to induce feelings versus evaluations (Strack et al., 1985). Recent research (Fujita and Han, 2009) suggests that this may be generalized across contents. As a consequence, one might predict that impulsive behaviors are caused by concrete representations whereas their

reflective regulations may be more successful if abstract thinking is prevalent.

Other programs of research may focus on the relationships between evaluations and goals (Markel, 2009), evaluative conditioning versus inferences (Corneille et al., 2009; Hofmann et al., 2008), frustration and behavioral orientation (Kriegelmeyer et al., submitted) or the role of counterimpulse training in the domain of addiction (Mitchell, et al., 2009).

In summary, it seems as if with new methodologies and theoretical models, the old theme of the dual determination of human behavior has become a central topic of psychological research, which has important basic and applied implications. This allows the achievement of a deeper understanding of those components that go beyond the anticipation and evaluation of outcomes and includes unplanned behaviors. In particular, it reconsiders impulsive elements that have long been neglected in social psychology. More importantly, it integrates them into a coherent model of human behavior in which the contribution of different components is explained as a function of situational and dispositional antecedents. As a consequence, human behavior is understood as a blend of constituents that is rarely “process pure.” Instead, planned and spontaneous components are part and parcel of any behavioral manifestation. To further disentangle their contributions and explain their operation under specified circumstances is a challenge for future endeavors.

APPENDIX: PERSONAL HISTORIES OF THE RIM

Reporting one's own intellectual history is not without dangers because it lends itself to various biases. Admitting one's own thinking as autobiographical evidence presupposes that our memory for our own cognitive activity is solely determined by what has actually happened. Unfortunately, this is not the case

and reconstructive endeavors (e.g., Strack and Bless, 1994) may greatly contribute to the recollection and distort the validity of the reminiscence for the benefit of a coherent narrative. However, these distortions tend to remain largely unnoticed by the protagonists and may require corrections from critical observers. Given these caveats, here is how we came to develop the Reflective–Impulsive Model.

F.S.: Quite some time ago, I was excited about the controversial discussion on the possibility of mental imagery (e.g., Paivio, 1969). Following this exchange, I was convinced that there exists something like a visual representation simply because people who try to visually imagine something often cover their eyes but never their ears. This suggested to me that relation between visual representation and visual perception is not metaphorical but psychologically substantive. That is, to “see” somebody with one’s mind’s eye is qualitatively similar to the result of the perceptual act of seeing. Luckily, this view was subsequently supported in a review article by Martha Farah (1988) in which evidence was collected demonstrating that similar cortical areas are involved in visual imagery and visual perception.

Thus, it seemed reasonable to assume that perceptual representations do not completely vanish once they have left the sensory register. Their quality may continue to exert its effect if people actively imagine a target or event. At the same time, categorical or symbolic representation may coexist and the world may be represented in two ways: as categorical knowledge or as a perceptual experience. Of course, in reality both types may coexist and supplement one another. Still, this distinction convinced me about the possibility that the world may be internally represented in different ways.

In some early studies in collaboration with Norbert Schwarz (Strack et al., 1985), I tried to manipulate the type of representation by inducing participants to think about past life events in either a concrete (“*How* did it happen?”) or in an abstract (“*Why* did

it occur?”) fashion. We found that in the concrete condition, participants experienced more affect that was congruent with the valence of the event than in the abstract condition. These findings suggest that perceptual experiences may be similar to immediate experiences that are elicited by imagery.

The two types of representations may sometimes be in conflict. This is apparent in visual illusions, like the Müller–Lyer Illusion. Here, we perceive the segment of the line between the open arrows to be shorter than the segment between the pointed arrows. At the same time our perceptual experience may be at conflict with our knowledge that the two lines are of equal length. Because they have divergent implications, we are aware of the dual representation of the target.

In a subsequent book chapter (Strack, 1992), I have made a case for a dual representation of attitudes. In particular, I have argued that formation of attitudes involves both an experiential and an informational route and that the inclusion of experiences will be subject to correction if they are not “representative” for the judgment.

Yet another type of duality fed into the theory that is described in the chapter. It comes from research on metacognition I conducted in collaboration with Herbert Bless (Strack and Bless, 1994) and Jens Förster (Strack and Förster, 1998). Here, it was the interaction between recollective experience and metacognitive knowledge that was shown to jointly contribute to judgments of previous occurrence.

To account for these observations in a systematic way, it has long been my plan to integrate these various insights into a model that focuses on the interaction of different representations and concomitant operating mechanisms in information processing. Moreover I wanted to link affective and cognitive processes with behavior as the ultimate outcome.

However, this ambitious goal could only be reached with a highly competent and motivated collaborator. Roland Deutsch, whom I knew since he was an undergraduate

student, turned out to be the ideal person with who to attempt the project. Luckily, I was able to persuade him to come aboard. This resulted in numerous discussions and many revisions of a manuscript that was finally published in 2004. Independent of how the end product is eventually evaluated, the joint way to get there was one of the most stimulating and challenging phases of my entire career.

R.D.: While working towards my degree in psychology, I chiefly focused on theories on animal and human behavior that were developed in the field of experimental psychology. Back then (and still), I was fascinated by connectionist models of behavioral performance, and by associative learning more broadly defined. At the same time, I delved into the rich literature on the regularities of social judgment and decision making.

Although both fields attracted me equally, I had the impression that theory and research in these fields were conducted largely in an independent manner – a state of affairs that I deemed highly undesirable. As a consequence, theories that apparently seemed to make a step toward reconciling the realms of judgment and decision making on one hand and basic learning mechanisms on the other hand strongly attracted my attention. One of these theories was the theory of a supervisory attentional system by Norman and Shallice (1986), which served as a working model for my own psychological thinking for quite a while.

After I decided to pursue my doctoral studies in the field of social psychology, the idea of integrating judgmental and behavioral theories was still a strong motive. And although social psychology already had a lot of “duality theories” to offer, most of these theories focused on different types of social judgments without thoroughly trying to integrate basic processes of motivation and behavior control.

Primed by this history of studies and interests, I entered a lab meeting in the fall of 1999, which I still remember very vividly.

I was still relatively new to the department, and I had experienced Fritz Strack mainly as a person who contributed to the lab meeting as supervisor and discussant. But for this special occasion, he had invited all members to attend an introductory talk to his plans on developing an integrative theory of social judgment and social behavior.

Much of these ideas seemed to represent what I had been looking for in the past. At the same time, Fritz pointed out that substantial work was still necessary to turn the ideas conveyed in this talk into an elaborate theory, and he invited the audience to join his efforts. I immediately knew that I wanted to be part of this project. The resulting hours of stimulating and challenging discussions, massive literature searches, and many fine-grained revisions of what was to become the final Reflective–Impulsive Model where a formative experience, which was and still is the basis for my academic work

NOTE

1 Another example for a cultural manifestation of the duality of human behavior is a passage from Goethe's *Faust* in which the protagonist complains that “two souls, alas!” reside within his “breast and each withdraws from and repels its brother” (*Faust I*, 2, 307). Here, the two principles are given equal status and the dissociation between them is depicted as a regrettable experience.

REFERENCES

- Ainslie, G. (1975) Specious reward: A behavioral theory of impulsiveness and impulse control. *Psychological Bulletin*, 82, 463.
- Ainslie, G. (2005) Precis of breakdown of will. *Behavioral and Brain Sciences*, 28, 635–673.
- Ajzen, I. (1991) The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50, 179–211.
- Allan, J.L., Johnston, M. and Campbell, N. (2010) Unintentional eating. What determines goal-congruent chocolate consumption. *Appetite*, 54, 422–425.

Asendorpf, J.B. and van Aken, M.A. (1999) Resilient, overcontrolled, and undercontrolled personality prototypes in childhood: Replicability, predictive power, and the trait-type issue. *Journal of Personality and Social Psychology*, 77, 815–832.

Baddeley, A.D. (1986) *Working Memory*. Oxford: Clarendon Press.

Bandura, A. (1977) *Social Learning Theory*. Oxford: Prentice-Hall.

Bargh, J.A. (1994). The four horsemen of automaticity: Awareness, intention, efficiency, and control in social cognition. In R.S. Wyer and T.K. Srull (eds), *Handbook of Social Cognition*, 1, 1–40. Hillsdale, NJ: Erlbaum.

Bargh, J.A. (2005) Bypassing the will: Toward demystifying the nonconscious control of social behavior. In R.R. Hassin, J.S. Uleman and J.A. Bargh (eds), *The New Unconscious*, pp. 37–58. Oxford: Oxford University Press.

Bargh, J.A., Chen, M. and Burrows, L. (1996) Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology*, 71, 230–244.

Bargh, J.A. and Ferguson, M.J. (2000) Beyond behaviorism: On the automaticity of everyday life. *Psychological Bulletin*, 126, 925–945.

Baumeister, R.F., Bratslavsky, M., Muraven, M. and Tice, D.M. (1998) Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology*, 74, 1252–1265.

Baumeister, R.F. and Heatherton, T.F. (1996) Self-regulation failure: An overview. *Psychological Inquiry*, 7, 1–15.

Baumeister, R.F., Heatherton, T.F. and Tice, D.M. (1994) *Losing Control: How and Why People Fail at Self-regulation*. San Diego: Academic Press.

Bem, D.J. (1972) Self-perception theory. In L. Berkowitz (ed.), *Advances in Experimental Social Psychology*, 6, 1–62. New York: Academic Press.

Blanton, H. and Jaccard, J. (2006) Arbitrary metrics in psychology. *American Psychologist*, 61, 27–41.

Block, J.H. and Block, J. (1980) The role of ego-control and ego-resiliency in the organization of behavior. In W.A. Collins (ed.), *Development of Cognition, Affect, and Social Relations*, 13, 39–101. Hillsdale, NJ: Lawrence Erlbaum.

Cacioppo, J.T., Priester, J.R. and Berntson, G.G. (1993) Rudimentary determinants of attitudes. II: Arm flexion and extension have differential effects on attitudes. *Journal of Personality and Social Psychology*, 65, 5–17.

Carver, C.S. (2005) Impulse and constraint: Perspectives from personality psychology, convergence with theory in other areas, and potential for integration. *Personality and Social Psychology Review*, 9, 312–333.

Cesario, J., Plaks, J.E. and Higgins, E.T. (2006) Automatic social behavior as motivated preparation to interact. *Journal of Personality and Social Psychology*, 90, 893–910.

Chaiken, S. (1980) Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology*, 39, 752–766.

Chaiken, S. and Trope, Y. (eds) (1999). *Dual-process Theories in Social Psychology*. New York: Guilford.

Chen, M. and Bargh, J. (1999) Consequences of automatic evaluation: Immediate behavioral predispositions to approach or avoid the stimulus. *Personality and Social Psychology Bulletin*, 25, 215–224.

Corneille, O., Yzerbyt, V.Y., Pleyers, G. and Mussweiler, T. (2009) Beyond awareness and resources: Evaluative conditioning may be sensitive to processing goals. *Journal of Experimental Social Psychology*, 45, 279–282.

Costa, P.T., Jr. and McCrae, R.R. (2009) The five-factor model and the NEO inventories. In J.N. Butcher (ed.), *Oxford Handbook of Personality Assessment*, pp. 299–322. New York: Oxford University Press.

De Houwer, J., Crombez, G., Baeyens, F. and Hermans, D. (2001) On the generality of the affective Simon effect. *Cognition and Emotion*, 15, 189–206.

De Houwer, J. and Moors, A. (2007). How to define and examine the implicitness of implicit measures. In B. Wittenbrink and N. Schwarz (eds), *Implicit Measures of Attitudes*, pp. 179–194. New York: Guilford Press.

De Houwer, J., Teige-Mocigemba, S., Spruyt, A. and Moors, A. (2009) Implicit measures: A normative analysis and review. *Psychological Bulletin*, 135, 347–368.

Deutsch, R., Gawronski, B. and Strack, F. (2006) At the boundaries of automaticity: Negation as reflective operation. *Journal of Personality and Social Psychology*, 91, 385–405.

Deutsch, R. and Strack, F. (2006a) Duality models in social psychology: From dual processes to interacting systems. *Psychological Inquiry*, 17, 166–172.

Deutsch, R. and Strack, F. (2006b) Reflective and impulsive determinants of addictive behavior. In R.W. Wiers and A.W. Stacy (eds), *Handbook of Implicit Cognition and Addiction*. Thousand Oaks, CA: Sage.

Devine, P.G. (1989) Automatic and controlled processes in prejudice: The role of stereotypes and

personal beliefs. In A.R. Pratkanis, S.J. Breckler and A.G. Greenwald (eds), *Attitude Structure and Function*, pp. 181–212. Hillsdale, NJ: Lawrence Erlbaum Associates.

Dimberg, U., Thunberg, M. and Grunedal, S. (2002) Facial reactions to emotional stimuli: Automatically controlled emotional responses. *Cognition and Emotion*, 16, 449–472.

Dunton, B.C. and Fazio, R.H. (1997) An individual difference measure of motivation to control prejudiced reactions. *Personality and Social Psychology Bulletin*, 23, 316–326.

Epstein, S. (1990) Cognitive-experiential self-theory. In L.A. Pervin (ed.), *Handbook of Personality. Theory and Research*, 165–192. New York: Guilford.

Farah, M.J. (1988) Is visual imagery really visual? Overlooked evidence from neuropsychology. *Psychological Review*, 95, 307–317.

Farrelly, M.C., Healton, C.G., Davis, K.C., Messeri, P., Hersey, J.C. and Haviland, M.L. (2002) Getting to the truth: Evaluating national tobacco countermarketing campaigns. *American Journal of Public Health*, 92, 901–907.

Fazio, R.H., Sanbonmatsu, D.M., Powell, M.C. and Kardes, F.R. (1986) On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, 50, 229–238.

Fiedler, K., Messner, C. and Bluemke, M. (2006) Unresolved problems with the 'I', the 'A', and the 'T': A logical and psychometric critique of the Implicit Association Test (IAT). In W.H.M. Stroebe (ed.), *European Review of Social Psychology*, 17, 74–147. New York: Psychology Press.

Frederick, S., Loewenstein, G.F. and O'Donoghue, T. (2003) Time discounting and time preference: A critical review. In G.-F. Loewenstein, D. Read, and R. Baumeister (eds), *Time and Decision: Economic and Psychological Perspectives on Intertemporal Choice*, pp. 13–86. New York: Russell Sage Foundation.

Freud, S. (1927) *The Ego and the Id*. London: Hogarth.

Freud, S. (1993) *Massenpsychologie und Ich-Analyse*. Frankfurt: Fischer Taschenbuch Verlag. (Originally published in 1921.)

Friese, M., Hofmann, W. and Schmitt, M. (2009) When and why do implicit reaction time measures predict behavior? Empirical evidence for the moderating role of motivation, opportunity, and process reliance. *European Review of Social Psychology*, 19, 285–338.

Friese, M., Hofmann, W. and Wänke, M. (2008) When impulses take over: Moderated predictive validity of explicit and implicit attitude measures in predicting food choice and consumption behaviour. *British Journal of Social Psychology*, 47, 397–419.

Fujita, K. and Han, H.A. (2009) Moving beyond deliberative control of impulses: The effect of construal levels on evaluative associations in self-control conflicts. *Psychological Science*, 20, 799–804.

Funder, D.C. and Block, J. (1989) The role of ego-control, ego-resiliency, and IQ in delay of gratification in adolescence. *Journal of Personality and Social Psychology*, 57, 1041–1050.

Gawronski, B., Deutsch, R., Mbirkou, S., Seibt, B. and Strack, F. (2008) When 'just say no' is not enough: Affirmation versus negation training and the reduction of automatic stereotype activation. *Journal of Experimental Social Psychology*, 44, 370–377.

Gawronski, B. and Strack, F. (2004) On the propositional nature of cognitive consistency: Dissonance changes explicit, but not implicit attitudes. *Journal of Experimental Social Psychology*, 40, 535–542.

Gilbert, D.T. (1991) How mental systems believe. *American Psychologist*, 46, 107–119.

Gollwitzer, P.M. (1999) Implementation intentions: Strong effects of simple plans. *American Psychologist*, 54, 493–503.

Greenwald, A.G. and Banaji, M.R. (1995) Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, 102, 4.

Greenwald, A.G., McGhee, D.E. and Schwartz, J.L.K. (1998) Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology*, 74, 1464–1480.

Greifeneder, R. and Bless, H. (2007) Relying on accessible content versus accessibility experiences: The case of processing capacity. *Social Cognition*, 25, 853–881.

Higgins, E.T. (1996) Knowledge activation: Accessibility, applicability, and salience. In E.T. Higgins and A.W. Kruglanski (eds), *Social Psychology: Handbook of Basic Principles*, pp. 133–168. New York: Guilford Press.

Higgins, E.T., Rholes, W.S. and Jones, C.R. (1977) Category accessibility and impression formation. *Journal of Experimental Social Psychology*, 13, 141–154.

Hoefling, A. and Strack, F. (2008) The tempting effect of forbidden foods: High calorie content evokes conflicting implicit and explicit evaluations in restrained eaters. *Appetite*, 51, 681–689.

Hofmann, W. and Friese, M. (2008) Impulses got the better of me: Alcohol moderates the influence of implicit attitudes toward food cues on eating behavior. *Journal of Abnormal Psychology*, 117, 420–427.

Hofmann, W., Friese, M. and Strack, F. (2009) Impulse and self-control from a dual-systems perspective. *Perspectives on Psychological Science*, 4, 162–176.

Hofmann, W., Gschwendner, T., Friese, M., Wiers, R.W. and Schmitt, M. (2008) Working memory capacity and self-regulation: Toward an individual differences perspective on behavior determination by automatic versus controlled processes. *Journal of Personality and Social Psychology*, 95, 962–977.

Hofmann, W., Rauch, W. and Gawronski, B. (2007) And deplete us not into temptation: Automatic attitudes, dietary restraint, and self-regulatory resources as determinants of eating behavior. *Journal of Experimental Social Psychology*, 43, 497–504.

James, W. (1890) *The Principles of Psychology* (Vol. 1). New York: Dover.

Johnson, M.K. and Hirst, W. (1991) Processing subsystems of memory. In R.G.W.H.J. Lister (ed.), *Perspectives on Cognitive Neuroscience*, pp. 197–217. New York: Oxford University Press.

Kahneman, D. and Tversky, A. (1982) The simulation heuristic. In D. Kahneman, P. Slovic and A. Tversky (eds), *Judgment Under Uncertainty: Heuristics and Biases*, pp. 201–208. Cambridge: Cambridge University Press.

Klayman, J. and Ha, Y. (1987) Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review*, 94, 211–228.

Koriat, A. and Levy-Sadot, R. (1999) Processes underlying metacognitive judgments: Information-based and experience-based monitoring of one's own knowledge. In S. Chaiken and Y. Trope (eds), *Dual-process Theories in Social Psychology*, pp. 483–502. New York: Guilford Press.

Kosslyn, S.M. (1980) *Image and Mind*. Cambridge, MA: Harvard University Press.

Krieglmeyer, R., Deutsch, R. and Strack, F. (submitted) The effect of frustration on behavioral tendencies of approach and avoidance.

Kruglanski, A.W. and Thompson, E.P. (1999) Persuasion by a single route: A view from the unimodel. *Psychological Inquiry*, 10, 83–109.

Le Bon, G. (1895) *The Crowd*. New York: Viking Press.

Loewenstein, G.F. (1996) Out of control: Visceral influences on behavior. *Organizational Behavior and Human Decision Processes*, 65, 272.

Loewenstein, G.F. and Elster, J. (1992) *Choice Over Time*. New York, NY: Russell Sage Foundation.

Loewenstein, G. and Thaler, R.H. (1997) Intertemporal choice. In W.M. Goldstein and R.M. Hogarth (eds), *Research on Judgment and Decision Making: Currents, Connections, and Controversies*, pp. 365–378. New York: Cambridge University Press.

Loewenstein, G. and Thaler, R.H. (2005) Anomalies: Intertemporal choice. In M. Bazerman (ed.), *Negotiation, Decision Making and Conflict Management*, pp. 428–440. Northampton, MA: Edward Elgar Publishing.

Lotze, H. (1852) *Medizinische Psychologie oder die Physiologie der Seele*. Leipzig: Weidmann'sche Buchhandlung.

Markel, P. (2009) If it's to be, it starts with me! The bidirectional relation between goals and the self. PhD dissertation, University of Würzburg.

Markman, A.B. and Brendl, C.M. (2005) Constraining theories of embodied cognition. *Psychological Science*, 16, 6–10.

McClure, S.M., Laibson, D.I., Loewenstein, G. and Cohen, J.D. (2004) Separate neural systems value immediate and delayed monetary rewards. *Science*, 306, 503–507.

Metcalfe, J. and Mischel, W. (1999) A hot/cool system analysis of delay of gratification: Dynamics of willpower. *Psychological Review*, 106, 3–19.

Mischel, W. (1974) Processes in delay of gratification. In L. Berkowitz (ed.), *Advances in Experimental Social Psychology*, 7, 249–292. San Diego: Academic.

Mischel, W., Ebbesen, E.B. and Zeiss, A.R. (1972) Cognitive and attentional mechanisms in delay of gratification. *Journal of Personality and Social Psychology*, 21, 204–218.

Mitchell, C.J., De Houwer, J. and Lovibond, P.F. (2009) The propositional nature of human associative learning. *Behavioral and Brain Sciences*, 32, 183–198.

Mussweiler, T., Förster, J. and Strack, F. (1997) Anchoring effects and the applicability of anchor-consistent information: A selective accessibility model. *Zeitschrift für Experimentelle Psychologie*, 44, 589–615.

Mussweiler, T. and Strack, F. (1999) Hypothesis-consistent testing and semantic priming in the anchoring paradigm: A selective accessibility model. *Journal of Experimental Social Psychology*, 35, 136–164.

Neumann, R., Förster, J. and Strack, F. (2003) Motor compatibility: The bidirectional link between behavior and evaluation. In J. Musch and K.C. Klauer (eds), *The Psychology of Evaluation*, pp. 371–391. London: Lawrence Erlbaum Associates.

Neumann, R., Hess, M., Schulz, S.M. and Alpers, G.W. (2005) Automatic behavioural responses to valence: Evidence that facial action is facilitated by evaluative processing. *Cognition and Emotion*, 19, 499–513.

Norman, D.A. and Shallice, T. (1986) Attention to action. Willed and automatic control of behavior. In R.J. Davidson, G.E. Schwartz and D. Shapiro (eds), *Consciousness and Self Regulation: Advances in Research*, pp. 1–18. New York: Plenum.

Paivio, A. (1969) Mental imagery in associative learning and memory. *Psychological Review*, 76, 241–263.

Petty, R.E. and Cacioppo, J.T. (1986) The elaboration likelihood model of persuasion. In L. Berkowitz (ed.), *Advances in Experimental Social Psychology*, 19, 123–205. San Diego: Academic Press.

Rizzolatti, G., Craighero, L. and Fadiga, L. (2002) The mirror system in humans. In M.I.G.V. Stamenov (ed.), *Mirror Neurons and the Evolution of Brain and Language*, pp. 37–59. Amsterdam: John Benjamins.

Rothermund, K. and Wentura, D. (2004) Underlying processes in the Implicit Association Test (IAT): Dissociating salience from associations. *Journal of Experimental Psychology: General*, 133, 139–165.

Rowe, C.J. (1998) *Plato: Phaedrus, with Translation and Commentary*. Warminster: Aris and Phillips.

Schneider, W. and Shiffrin, R.M. (1977) Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, 84, 1–66.

Schwarz, N., Bless, H., Strack, F., Klumpp, G., Rittenauer-Schatka, H. and Simons, A. (1991) Ease of retrieval as information: Another look at the availability heuristic. *Journal of Personality and Social Psychology*, 61, 195–202.

Schwarz, N. and Clore, G.L. (1983) Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology*, 45, 513–523.

Schwarz, N. and Clore, G.L. (1996) Feelings and phenomenal experiences. In E.T. Higgins and A.W. Kruglanski (eds), *Social Psychology. Handbook of Basic Principles*, pp. 433–465. New York: Guilford Press.

Seibt, B., Häfner, M. and Deutsch, R. (2007) Prepared to eat: How immediate affective and motivational responses to food cues are influenced by food deprivation. *European Journal of Social Psychology*, 37, 359–379.

Sloman, S.A. (1996) The empirical case for two systems of reasoning. *Psychological Bulletin*, 119, 3–22.

Solarz, A.K. (1960) Latency of instrumental responses as a function of compatibility with the meaning of eliciting verbal signs. *Journal of Experimental Psychology*, 59, 239–245.

Sorabji, R. (2002) *Emotion and Peace of Mind: From Stoic Agitation to Christian Temptation*. Oxford: Oxford University Press.

Strack, F. (1992) The different routes to social judgments: Experiential versus informational strategies. In L.L. Martin and A. Tesser (eds), *The Construction of Social Judgment*, pp. 249–275. Hillsdale, NJ: Erlbaum.

Strack, F. and Bless, H. (1994) Memory for nonoccurrences: Metacognitive and presuppositional strategies. *Journal of Memory and Language*, 33, 203–217.

Strack, F. and Deutsch, R. (2004) Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, 8, 220–247.

Strack, F. and Deutsch, R. (2007) The role of impulse in social behavior. In A.W. Kruglanski and E.T. Higgins (eds), *Social Psychology: Handbook of Basic Principles*, pp. 408–431. New York: Guilford.

Strack, F., Erber, R. and Wicklund, R.A. (1982) Effects of salience and time pressure on ratings of social causality. *Journal of Experimental Social Psychology*, 18, 581–594.

Strack, F. and Förster, J. (1998) Self-reflection and recognition: The role of metacognitive knowledge in the attribution of recollective experience. *Personality and Social Psychology Review*, 2, 111–123.

Strack, F., Schwarz, N. and Gschneidinger, E. (1985) Happiness and reminiscing: The role of time perspective, affect, and mode of thinking. *Journal of Personality and Social Psychology*, 49, 1460–1469.

Strack, F., Werth, L. and Deutsch, R. (2006) Reflective and impulsive determinants of consumer behavior. *Journal of Consumer Psychology*, 16, 205–216.

Topolinski, S. and Strack, F. (2009) The architecture of intuition: Fluency and affect determine intuitive judgments of semantic and visual coherence and judgments of grammaticality in artificial grammar learning. *Journal of Experimental Psychology: General*, 138, 39–63.

Tversky, A. and Kahneman, D. (1974) Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124–1131.

Unkelbach, C., Forgas, J.P. and Denson, T.F. (2008) The turban effect: The influence of Muslim headgear and induced affect on aggressive responses in the shooter bias paradigm. *Journal of Experimental Social Psychology*, 44, 1409–1413.

Vohs, K.D. and Faber, R.J. (2007) Spent resources: Self-regulatory resource availability affects impulse buying. *Journal of Consumer Research*, 33, 537–547.

Whiteside, S.P. and Lynam, D.R. (2001) The Five Factor Model and impulsivity: Using a structural model of personality to understand impulsivity. *Personality and Individual Differences*, 30, 669–689.

Wiers R.W., Rinck M., Dictus M. and van den Wildenberg E. (2009) Relatively strong automatic appetitive action-tendencies in male carriers of the

OPRM1 G-allele. *Genes, Brain and Behavior*, 8, 101–106.

Wiers, R.W., Rinck, M., Kordts, R., Houben, K. and Strack, F. (2010) Retraining automatic action-tendencies to approach alcohol in hazardous drinkers. *Addiction*, 105, 279–287.

Wilson, T.D., Lindsey, S. and Schooler, T.Y. (2000) A model of dual attitudes. *Psychological Review*, 107, 101–126.

Winkielman, P. and Cacioppo, J.T. (2001) Mind at ease puts a smile on the face: Psychophysiological evidence that processing facilitation elicits positive affect. *Journal of Personality and Social Psychology*, 81, 989–1000.

Wittenbrink, B. and Schwarz, N. (eds) (2007). *Implicit Measures of Attitudes: Procedures and Controversies*. New York: Guilford.

Wood, W. and Neal, D.T. (2007) A new look at habits and the habit-goal interface. *Psychological Review*, 843–863.

Zuckerman, M. (1994a) *Behavioral Expressions and Biosocial Bases of Sensation Seeking*. New York, NY: Cambridge University Press.

Zuckerman, M. (1994b) Impulsive unsocialized sensation seeking: The biological foundations of a basic dimension of personality. In J.E.W.T.D. Bates (ed.), *Temperament: Individual Differences at the Interface of Biology and Behavior*, pp. 219–255. Washington, DC: American Psychological Association.

Zuckerman, M. (1996) The psychobiological model for impulsive unsocialized sensation seeking: A comparative approach. *Neuropsychobiology*, 34, 125–129.

Construal Level Theory

Yaacov Trope and Nira Liberman

ABSTRACT

We started with how the value of outcomes changes over temporal distance and ended up with what we hope is a step toward a general theory of psychological distance. There are different psychologies for the different dimensions of psychological distance – temporal, spatial, social, and hypotheticality. Without denying the uniqueness of each of these dimensions and more specific aspects comprising those dimensions, we believe they all entail transcending the present through constructing mental models of what is not directly experienced. The farther removed an object is from me on any distance dimension, the higher (more abstract) the level of mental construal of that object. Psychological distance thus expands or contracts depending on level of construal. Consistent with this proposal, research conducted in the framework of construal level theory (CLT) suggests that (1) different distance dimensions are mentally associated, (2) that distance on any of these dimensions influences and is influenced by higher levels of mental construal, and (3) that the various distances are, to some extent, interchangeable in their effects on prediction, evaluation, and choice.

INTRODUCTION

From a bird's eye view, social psychology stands out as a discipline that, more than any

other discipline in the behavioral and social sciences, has sought to understand how people experience the immediate situation. The pursuit of this goal was inspired by Kurt Lewin's field-theoretic emphasis on the critical role played by the forces operating in the "here-and-now." Following Lewin's lead, classic experiments in social psychology provide dramatic demonstrations of the power of the immediate situation and how individuals get caught up in it. For example, Asch's (1956) conformity experiment, Milgram's (1965) obedience studies, and Zimbardo's (1971) prison study vividly illustrate how individuals become engrossed in accommodating group pressure, formal authority, and social role requirements. People have always been known to conform, obey, and adhere to their social roles. The contribution of these and other classic social psychology experiments is in showing how readily individuals become mentally immersed in the intricacies of the immediate situation and how readily their pre-existing perceptions, attitudes, and values can be crowded out and no longer predict behavior. Consistent with this metatheoretical orientation, experimental social psychology has

developed over the years an expertise, unique among the behavioral and social sciences, in how to mentally immerse people in the immediate situation.

Studying mental immersion in the present is important in its own right, but it is also important for another reason: It brings to the fore the question of what is the alternative psychological state. Bearing on this question, research and anecdotal evidence suggest that, perhaps unlike any other species, humans are sometime able to transcend the “here-and-now.” They are able to recollect themselves in the past, plan the future, take others’ perspective, cognize spatially remote places, and contemplate counterfactual alternatives to reality. In each case, a psychological distance from the self in the here-and-now is traversed. It is physically impossible to be in the past, future, or at different locations at the same time, as it is impossible to experience counterfactual alternatives to reality and other people’s states. Yet humans have evolved an ability to broaden their **spatial**, **temporal**, and **social** horizons beyond the present to include the nonpresent. For example, participants in the Milgram obedience experiment might imagine how they would feel if they were the learner, how others would view what they are doing, what they would do if they were in a similar situation in the future, whether they have acted in the same way in the past or in other places, what would they do if the experimenter was not present, and more.

It is hard to tell how many participants actually had these thoughts in the Milgram experiment – probably a precious few; after all, the experiment was set up to make mental travel difficult. Were those who did engage in mental travel more likely than others to disobey the experimenter’s commands? Perhaps. We still don’t know the answer, but this motivated our interest in the more general question of how mental travel occurs; that is, what are the psychological processes that enable people to mentally shift back and forth between being immersed in the here-and-now and removing oneself from it?

This is the question that the construal level theory (CLT) of psychological distance is addressing.

A TEAM NARRATIVE OF THE THEORY DEVELOPMENT

Temporal construal

We started with a more specific, and seemingly simpler, question of how temporal distance of an outcome affects judgment and choice with respect to that outcome. This question was not new. For many years, researchers across different behavioral science disciplines, including psychology, economics, and political science, have studied how people make choices for their immediate future versus distant future (e.g., Ainslie, 1975; Elster, 1977; Kirby and Herrenstein, 1995; LaPiere, 1934; Loewenstein and Prelec, 1992; Lovallo and Kahneman, 2000; Mischel et al., 1969; Rachlin, 1995; Schelling, 1984) and how they evaluate their past (e.g., Gilovich and Medvec, 1995; Nisan, 1972; Ross, 1989). This research has yielded a wealth of findings on a wide range of phenomena, including time discounting, delay of gratification, shifts in level of aspiration, future planning, future optimism, overconfident prediction, regret, hindsight bias, and biased autobiographical memory.

We were particularly intrigued by two sets of phenomena: the “planning fallacy” and temporal discounting. The planning fallacy refers to the tendency, not unfamiliar to academics, to overcommit oneself when making plans for the future (Kahneman and Tversky, 1979). This phenomenon was brought to our attention by Mike Ross, who has been studying it at the time and who has been a great source of inspiration in our thinking about temporal construal. Mike Ross, Tom Gilovich, and their colleagues (e.g., Buehler et al., 1994; Gilovich et al., 1993) have reported ingenious experimental demonstrations of people undertaking more activities than they

can possibly accomplish and which they come to regret only when they get closer to the time of enactment. These findings strongly resonated with our own personal experience and that of colleagues marveling at themselves for repeatedly making the mistake of committing far in advance to writing book chapters, journal reviews, letters of recommendation, and the like. An exception, which might actually prove the rule, is Amos Tversky who, hearing about our interest, responded self-contentedly that he has firmly established with the administration at Stanford that he could not be asked to do anything with a deadline beyond six weeks.

The other set of phenomena, time discounting, is people's tendency to attach greater value to immediate outcomes than to delayed outcomes. This phenomenon has attracted our attention because it has been studied across many disciplines in the neural, behavioral, and social sciences. In fact, we could not (and still cannot) think of any single behavioral phenomenon that has been studied across such a broad range of disciplines as time discounting. We were particularly fascinated and inspired by Walter Mischel's many years of programmatic investigation of delay of gratification in children and the many insights into the nature of self-regulation that this research has inspired. Research on time discounting has focused on identifying factors that determine the discount rate. For example, affect-dependent time discounting suggests that affective outcomes undergo steeper time discounting than cognitive outcomes (Loewenstein, 1996; Loewenstein et al., 2001; Metcalfe and Mischel, 1999). Conflict theories (Lewin, 1951; Miller, 1944) assume that negative outcomes undergo steeper time discounting than do positive outcomes. Behavioral economists have shown that the discount rate depends on the magnitude of the value of outcomes, such that small rewards are discounted at a faster rate than large rewards (Green et al., 1997; Thaler, 1981).

Each of these lines of research employed well-developed methodologies and offered

important descriptive and prescriptive implications for individuals and society. What seemed to be missing was an overarching framework that could integrate the diverse and sometime conflicting findings in the area as well as the specific hypotheses that have been offered to account for the findings. Our approach to this problem was based on the idea that people **rely on schematic mental models for making judgments about their past and future** (see Griffin and Ross, 1994; Kahneman and Lovallo, 1991; Ross, 1989). We thought that an integrative framework would have to take into account **how temporal distance from an object changes the way people mentally represent that object**. The effects of temporal distance would then emerge from an active online construction of mental representations of future objects; hence the term "temporal construal." Those effects, moreover, would involve a basic aspect of the way people represent objects but could nevertheless **take place without conscious awareness, so that it would repeatedly lead people to change their mind as they get closer to the objects without insight into the source of their temporal inconsistency**.

But what is that "basic aspect" of mental construal that changes with temporal proximity? Adopting a functional perspective, we asked what mental construal would be like in order for it to enable people to transcend the present and predict, evaluate, and take action with respect to the future. Our answer, which led to the development of CLT, built on past research in many areas of psychology, including memory, concept formation, categorization, causal reasoning, person perception, and goal hierarchies. We distinguished between high-level and low-level construals. **High-level construals, we proposed, are schematic, decontextualized representations that extract the gist from the available information, emphasizing a few superordinate core features of events. Low-level construals, in contrast, are relatively unstructured, contextualized representations that include subordinate and incidental features of events.** Consider, for example, two children playing

catch with a ball in a backyard. A low-level construal of this activity might include such details as the age of the children, the color of the ball, and the temperature outside. In contrast, a high-level construal of this activity might simply be “having fun.” The high-level construal, “having fun,” disregards the unique features of the event, and involves an implicit decision about which features are central to the event and which are peripheral. Moving to high-level construal involves omitting features that are perceived as less central to the abstract construct in question.

In line with our functional approach, we viewed **high-level construals as having evolved to represent distal events because, with distance, one needs to preserve the essential, invariant properties of the referent event**. High-level construals are capable of fulfilling this function because they preserve the essential properties of a stimulus across momentary changes in appearance and through changes in time. High-level construals function somewhat like perceptual constancies, abstracting stimuli to their invariant properties. Low-level construals, by contrast, preserve a stimulus in minute detail, emphasizing its uniqueness rather than its similarity to other stimuli.

Based on this reasoning, and as described in more detail below, we launched a research program on temporal construal. Our first set of studies addressed the two phenomena mentioned earlier: the planning fallacy and time discounting. We started with a simple study asking students to indicate the number of weekly hours they planned to engage in each of several routine academic activities (e.g., “attend classes”) and nonacademic activities (“sports and exercise”) either next week or a week a few months later. Two results of this study caught our attention. One was that the total number of weekly hours planned for next week was considerably smaller than that planned for a week a few months later. The other was that the number of hours planned for different activities was negatively correlated for the near future week, but uncorrelated for the distant future

week. Students seemed to make plans for the distant future as if there were no limits on their time resources and as if engaging in one activity didn’t come at the expense of engaging in another.

We understood this and related planning fallacy phenomena as resulting from the tendency to construe distant actions in terms of their high-level, superordinate aspects rather than low-level, subordinate features. **Desirability concerns, which involve the action’s end-state (i.e., the “why” aspect of the action) are superordinate aspects of actions, whereas feasibility concerns, which involve the means used to reach the end-state (i.e., the “how” aspect of the activity, the specific aspects of its enactment and its contextual constraints), are subordinate aspects of activities.** CLT thus predicts that desirability concerns should receive greater weight over feasibility concerns as psychological distance increases. From a temporally distant perspective, activities are represented in terms of their desirability aspects, but as one gets closer in time to actual enactment, feasibility aspects become more prominent. The explanation of the overcommitment findings from our little weekly planning study seemed straightforward: **temporal constraints and conflicts among different activities were feasibility aspects and therefore omitted from distant future plans. The time planned for each activity was based on its inherent attractiveness, irrespective of the time already allocated to other activities (Liberman and Trope, 1998).**

The same logic applies to time discounting, the second issue we addressed early in our temporal construal research. We reasoned that different construals may entail different evaluations. For example, the construal “running subjects in the lab” may foster a less positive evaluation of an activity than the higher-level construal “conducting a psychology study.” Because CLT assumes that people use higher-level construals for distant future events than for near future events, it predicts that the **value associated with low-level construals would be more prominent in evaluating near future events, whereas the**

value associated with high-level construals would be more prominent in evaluating distant future events. Time discounting, which has been commonly assumed to be a general principle in psychology, decision science, and economics, should hold only for the value associated with low-level construals. For value associated with high construals, the opposite may obtain, namely, time augmentation.

CLT predicts, then, that when the value associated with low-level construals is more positive than that associated with high-level construals, time delay would discount the attractiveness of an option. In contrast, when the value associated with low-level construals is less positive than that associated with high-level construals, time delay would augment the attractiveness of an option. Consistent with this prediction, we found that the value of the core, superordinate aspects of options were more influential in making a choice for the distant future, whereas the value of the secondary, subordinate aspects of options were more influential in making a choice for the near future (Trope and Liberman, 2000). For example, the extent to which an option promised to fulfill the individual's superordinate goal was more influential in choosing a distant future option, whereas incidental, goal-irrelevant considerations were more influential in choosing a near future option.

Psychological distance

While conducting this research, we suspected that our conceptualization might be overly narrow. If, as we have assumed, high-level construals serve to transcend the present self and afford future time travel, couldn't they also serve to transcend the self in other respects? These construals may enable one to retrospect, imagine oneself in remote locations, take the perspective of other people, and contemplate counterfactual alternatives to reality. In each case, one traverses a psychological distance from oneself be it

temporal distance (prospective and retrospective), spatial distance, social distance, or hypotheticality. High-level construal may expand not only one's temporal horizons, but also one's spatial horizons, social horizons, and the realm of possibility. Correspondingly, low-level construals may enable one to become immersed in one's own direct experience of the here-and-now.

Switching between levels of construal may thus afford mental travel back and forth between the proximal and the distal across the various psychological distances. Time travel is extremely interesting and for many years has fascinated scholars across the social, cognitive, and neural sciences (e.g., Ainslie, 1975; Gilbert and Wilson, 2007; Schacter and Addis, 2007; Schelling, 1984). However, it might be only one instance of mental travel. As social psychologists, we were particularly intrigued by the idea that social distance might be thought of as one of the dimensions of psychological distance. Distinguishing between self and others, mine and yours, and taking others' perspective seemed fundamental to traversing psychological distance. The possession of this capability and its extension to dissimilar others, strangers, and outgroups, we thought may have played a key role in social life and in the development of civilization.

As social psychologists with interest in social cognition and self-regulation, we were also intrigued by the idea that the many distinctions that comprise what we call hypotheticality can be all thought of as aspects of psychological distance. These distinctions include real versus imaginary, factual versus counterfactual, true versus false, probable versus improbable, and expected versus unexpected. For example, perhaps people intuit improbable events as psychologically distant, as events one has to wait long or go far in order to encounter, and perhaps people employ high-level construals to incorporate such events into their judgments and choices.

In general, the functional approach we have taken suggests that the relationship between construal level and psychological

distance is intrinsic. The ability to traverse psychological distance confers an obvious evolutionary advantage, and mental construal levels may have evolved, and may continue to evolve, in order to enable traversing increasingly greater psychological distances. This may be why level of construal and the expansion of mental horizons may have been linked throughout human evolution and why they appear to be linked in personal development. We felt that in relating CLT to these issues, we were responding to Lewin's (1951) call for developing a unified and general theory of psychological distance.

QUESTIONS CONSTRUAL LEVEL THEORY ADDRESSES

Thus far we have recounted how trying to understand specific time perspective effects has led us to develop a general CLT of psychological distance. But what are the new research questions the theory has led us to explore? What are its empirically testable hypotheses? There are three sets of questions that CLT has led us to ask: First, do all distances affect and are affected by level of construal? Second, are the various distances interrelated? That is, does distancing an object on one dimension lead people to expect the object to be distant on other dimensions? Third, do the various distances from an object similarly affect prediction, evaluation, and action regarding that object? The overarching, metatheoretical question was whether the level of generality of CLT is matched by the breadth of the phenomena this theory might predict. In this section we very briefly describe how we and our colleagues have addressed these questions.

Psychological distance and mental construal

Our initial research examined the effect of temporal distance on the level of categorization

of actions. However, to establish the generality of CLT, it was important to find out whether other aspects of construal, not only action categorization, were related to other psychological distances and whether the association was bidirectional. This led us to explore construal levels in perception, categorization, and inference. The following investigations illustrate these lines of research.

Perception

In a series of studies, participants completed what they believed to be sample items of a task that required abstraction of coherent images from fragmented or noisy visual input (the Gestalt Completion Test). Participants' performance improved when they anticipated working on the actual task in the more distant future (Förster et al., 2004), when they thought the actual task was less likely to take place (Wakslak et al., 2006), and when social distance was enhanced by priming of high social status (Smith and Trope, 2006). A psychologically distant perspective thus seems to enable people to better see the “big picture.”

While abstraction improves the ability to perceive the gestalt in a visual array, it should have the opposite effect on performance when the task is to recognize missing elements within a gestalt. Distance should therefore have a detrimental effect on the ability to identify a missing element within a coherent whole. Indeed, participants did worse on sample items of this task when they believed they were less likely to later complete it (Wakslak et al., 2006).

Categorization

Our initial temporal construal research was based on the assumption that actions may be construed in high-level terms, which link them to a superordinate purpose (why one performs them), or in low-level terms, which link them to subordinate means (how one performs them). Consistent with this assumption, we found that participants tended to describe more distant future activities

(e.g., studying) in high-level terms (e.g., “doing well in school”) rather than in low-level terms (e.g., “reading a textbook”) (Liberman and Trope, 1998). Similar effects emerged when actions were to take place in a spatially distant location (Henderson et al., 2006), when the actions were framed as unlikely to actually take place (Wakslak et al., 2006), and when the actor was described as dissimilar to the perceiver (Liviatan et al., 2008).

Going beyond action categorization, we then examined breadth of categorization of objects by asking participants to imagine an event (e.g., a camping trip) occurring in either the near or the distant future. For each event, participants grouped a set of related objects (e.g., tent, ball, snorkel) into as many groups as they deemed appropriate. Consistent with the idea that distance promotes the use of more abstract terms, participants who thought of a more distant event created fewer, broader groups of objects (Liberman et al., 2002). Reduced likelihood and social distance had the same effect (Smith and Trope, 2006; Wakslak et al., 2006). For example, objects that pertained to less likely events (e.g., a trip that had a high probability of being canceled) were grouped into broader categories.

Bi-directional relationships between construal level and distance

Do high-level construals, compared to low-level construals, lead individuals to think about psychologically more remote possibilities? Our research on this issue used a variety of manipulations of construal level (e.g., global versus local perceptual processing, generating categories versus exemplars, using broad versus narrow categories, identifying means versus ends, etc.). The results consistently showed that high-level construals lead individuals to think about more distant times, spatial locations, people, as well as relatively unlikely and rare events (Liberman and Förster, 2009; Wakslak and Trope, 2009).

The association between distance and level of construal was also demonstrated with

implicit measures. In a series of studies using the Implicit Association Test (Bar-Anan et al., 2006) participants were presented with words from four categories: high-level construal (e.g., category names such as “drinks”), low-level construal (e.g., exemplar names such as “coke”), small psychological distance (e.g., socially proximal words such as “ours”, “friend”), and large psychological distance (e.g., socially distant words such as “theirs,” “stranger”). Participants mapped words from these four categories on two responses, pressing either a left key or a right key on the computer keyboard. On CLT-congruent trials, high-level stimuli were paired with distant stimuli and low-level stimuli were paired with proximal stimuli, whereas on CLT-incongruent trials high-level stimuli were paired with proximal stimuli and low-level stimuli were paired with distal stimuli. With all four dimensions of psychological distance, participants were faster with congruent than with incongruent pairings, suggesting that they implicitly associated psychological distance with high-level construal and psychological proximity with low-level construal.

It seems, then, that as psychological distance increases, construals become more abstract, and as level of abstraction increases, so too do perceptions of psychological distance. This supports the basic tenet of CLT that abstract thinking is used to transcend the present and expand one’s mental horizon by thinking farther into time and space and considering remote social targets and unlikely possibilities. At a metatheoretical level, we felt that the range of perceptual and cognitive phenomena CLT has led us to explore justified the level of generality of the theory.

The association among distance dimensions

We have found that different distances are similarly related to construal and similarly. But the question still remained: do the distances in themselves have something in

common? If the different distance dimensions have a shared meaning, as CLT contends, then these dimensions should be mentally associated. We felt that it is important for the theory to address this issue empirically. The idea was that different objects whose distance on one dimension is congruent with their distance on any other dimension would be associated even when they have little or nothing else in common. For example, remote locations would bring to mind the distant rather than the near future, other people rather than oneself, and unlikely rather than likely events.

Initial support for this hypothesis comes from a set of studies that assessed implicit association between spatial distance and other distance dimensions (Bar-Anan et al., 2007). Using a picture-word version of the Stroop paradigm (Stroop, 1935), participants discriminated between cues of one psychological distance dimension while ignoring cues of another psychological distance dimension. We predicted that it would be easier to perform the task when the relevant and irrelevant cues are congruent (as opposed to incongruent) in terms of psychological distance. Participants viewed perspective pictures containing an arrow that was pointing to either a proximal or distal point on the landscape shown in the picture; inside the arrow was printed a word denoting either a psychologically proximal entity (e.g., tomorrow, we, sure) or a psychologically distal entity (e.g., year, others, maybe). In a spatial discrimination version of the task, participants had to indicate whether the arrow pointed to a proximal or distal location. In a semantic discrimination version, participants had to indicate what the word printed on the arrow was. In both tasks, and across the distance dimensions, participants were faster when responding to distant congruent stimuli (e.g., “we” printed on proximal arrow) than distant incongruent stimuli (e.g., “we” printed on distal arrow). The results suggest that people access the psychological distance of stimuli automatically, even when it is not directly related to people’s current goals.

We concluded from these and related findings that the different psychological distance dimensions are associated and that psychological distance is an important aspect of meaning, common to spatial distance, temporal distance, social distance, and hypotheticality. It is also possible that cues of distance from events on one dimension affect the perceived distance from those events on other dimensions. For example, the spatial distance from an event may depend not only on its location relative to that of the perceiver but also on whether it is expected in the near or distant future, whether it occurred recently or a long time ago, whether it is probable or improbable, and whether it is expected to be experienced by oneself or another person. In this respect, the different psychological distances may be to some extent interchangeable.

THE SOCIO-PSYCHOLOGICAL APPLICABILITY OF CONSTRUAL LEVEL THEORY

Having established that construals depend not only on the actual attributes of the objects but also on their psychological distance, we turned to the psychological consequences of this association. The question was whether all distances similarly affect (via their effect on construal) predictions, evaluation, and self-regulation. We have also started to explore the consequences of the distance-construal association for human relations.

Psychological distance and prediction

The function of high-level construals is to enable individuals to mentally transcend the here-and-now by forming a structured representation of the invariant features of the available information and projecting it onto distal objects. Consequently, predictions of future experiences would be more schematic

than the actual experiences, giving rise to a variety of prediction biases that stem from underweighting contextual and incidental features.

Two studies exemplify our research on this issue. One study, examining the effect of temporal distance on prediction of scientific findings, found that students were more confident that the exact same experiment would yield theory-confirming results when they expected the experiment to take place in a more distant point in time (Nussbaum et al., 2006). Apparently, when making predictions for the more distant experiments, participants gave more weight to the theory (high-level construal) and less weight to incidental noise factors (low-level construal). Another study investigated the effect of spatial distance on the tendency to base predictions on global rather than local information (Henderson and Fujita, 2006). New York University (NYU) participants viewed a series of graphs depicting information from the years 1999 to 2004 (e.g., average number of photocopies per student). The information was said to pertain to the NYU campus in Manhattan (spatially near condition) or to the NYU campus in Florence, Italy (spatially distant condition). Each graph showed either an upward or a downward trend, with the final year (2004) always deviating from that global trend. Participants estimated the likelihood that the year 2005 would be consistent with the general trend or with the more recent local deviation. In terms of CLT, global trends convey a high-level construal, whereas deviations, being local exceptions, should receive more weight in low-level construals. Consistent with this reasoning, spatial distance enhanced the tendency to predict on the basis of the global trend rather than the local deviation.

Psychological distance and evaluation

As noted earlier, a common assumption in the behavioral sciences is that the value of an

outcome is discounted as temporal distance from the outcome increases. We proposed, in contrast, that temporal distance, as any psychological distance, should shift the overall attractiveness of an outcome closer to its high-level construal value and away from its low-level construal value. Two sets of studies addressed this issue.

The first set of studies examined desirability and feasibility concerns (Liberman and Trope, 1998). Desirability concerns involve the value of the action's end-state (a high-level construal), whereas feasibility concerns involve the means used to reach the end-state (a low-level construal). Therefore, desirability concerns should receive greater weight over feasibility concerns as psychological distance increases. Consistent with this prediction, we found that as temporal distance from an activity (e.g., attending a guest lecture) increased, the attractiveness of the activity depended more on its desirability (e.g., how interesting the lecture was) and less on its feasibility (e.g., how convenient the timing of the lecture was). Similar results emerged with other psychological distance dimensions: probability (Todorov et al., 2007) and social distance (Liviatan et al., 2008). These findings suggest that distance increases the attractiveness of alternatives that are desirable but hard to obtain, but decreases the attractiveness of alternatives that are less desirable but easy to obtain. Extending this effect to the realm of risky choice, we found that people take higher risks (i.e., favoring bets with a low probability of winning a high amount over those that offer a high probability to win a small amount) in decisions about more distal bets (Sagristano et al., 2002).

The second set of studies concerned central and peripheral sources of value. According to CLT, central, goal-related features of outcomes constitute a high-level construal of these outcomes, whereas peripheral, goal-irrelevant features of outcomes constitute a low-level construal. Distancing an outcome should therefore increase the weight of central features relative to peripheral features. Support for this prediction was found in a

study in which participants imagined buying a radio set in order to listen to morning programs either the next day or in one year (Trope and Liberman, 2000). In one version, participants read that the sound quality of the radio set was good, but that the clock that was incidentally included was relatively useless. In a different version, participants read that the sound quality of the radio set was poor, but that the clock turned out to be quite useful. As predicted, thinking about the radio set in the more distant future increased satisfaction when the sound quality was good and the clock was useless but decreased satisfaction when the sound quality was poor and the clock was good, indicating that time delay increased the weight of central features and decreased the weight of peripheral features. It seems, then, that people's overriding goals are more likely to guide their choices for psychologically distant than psychologically near situations.

Psychological distance and self-regulation

Like predictions and evaluations, self-regulation should be increasingly based on high-level construal aspects as psychological distance increases. **As outcomes seem more temporally, spatially, or socially remote or unlikely, actions should be guided more by one's central, global concerns and less by secondary, local concerns.** Our research on self-control exemplifies our approach to this issue. Exercising self-control requires acting in line with ones' central, superordinate, and global considerations in the presence of more locally tempting alternatives. Because such considerations naturally relate to high-level construals, psychological distance should facilitate self-control. Indeed, **people seem to be better able to choose delayed but valuable outcomes for the distant than for the near future** (Loewenstein and Prelec, 1992). As another example, choosing a negative but diagnostic assessment of one's abilities rather than a flattering but nondiagnostic

assessment requires prioritizing the long-term benefits of self-improvement over subordinate concerns about feeling good. Consistent with this prediction, participants were more likely to prefer the negative but diagnostic assessment when it was expected in the more distant future (Freitas et al., 2001). A recent series of studies has directly linked construal level to self-control, showing that forming a high-level construal of situations enables better self-control (e.g., choosing a delayed reward, enduring painful but valuable diagnostic procedures). In the same vein, research on delay of gratification in children showed that an abstract representation of the temptation increases delay relative to a more concrete representation (Fujita et al., 2006).

Human relations across psychological distance

Much of our research has examined how people traverse psychological distance. Recently, we have started to explore what is arguably the most important aspect of this phenomenon, namely how people traverse psychological distances in their social relations. Many of the key issues in social psychology – how we connect with others, how our sense of self is influenced by them, how we give to, take from, and reciprocate with other people – are inherently about crossing the gap between oneself in the here-and-now, and another person existing outside of that current direct experience. Our research on interpersonal similarity, social power, politeness, and social conflict illustrate the implications of CLT for social relations.

Interpersonal similarity

In CLT, interpersonal similarity is a form of social distance (Liviatan et al., 2008). The less similar someone is to oneself, the more socially distant they typically seem; therefore, we hypothesized and found that behavior performed by a dissimilar other was represented at a higher level of construal than

behavior performed by a similar other (Liviatan et al., 2008). Given this association between social similarity and mental construal, attraction to similar compared to dissimilar others should reflect the valence of low-level features more than the valence of high-level features. Indeed, Liviatan et al. (2008) showed that participants' liking of a similar (versus dissimilar) target was based to an increasing degree on the positivity of the target's behaviors and secondary help (i.e., low-level construals) over the positivity of the target's traits and primary help (i.e., high-level construals). It seems, then, that **high-level construals expand our social horizons enabling us to relate to socially diverse people, whereas low-level construals guide our response to people who are like us.**

Social power

Social power may engender a sense of distance from others. Indeed, individuals who have power see themselves as less similar to and thus more distant from other people than individuals who have less power. This perception might be due to the fact that groups, organizations, and societies ordinarily have a pyramidal structure with fewer individuals occupying high-power positions than low-power positions. There is therefore greater similarity in the positions held by individuals with low power than individuals with high power. If **social power makes people feel distant from others, then it should also predispose them to construe information abstractly, focus on the central aspects of situations, disregard secondary aspects, and establish clear priorities.** Consistent with this view, Smith and Trope (2006) research suggests that the distal perspective activated by the possession of social power promotes going beyond the information given, detecting the underlying structure, and abstracting from its superordinate, central features. Power-related construal may expand people's mental horizons, enabling them to transcend the immediate circumstances and take into account the past, future, a broad range of people, and unlikely possibilities. Do we

think of individuals as suitable for power positions, or as actually holding such positions, when they articulate and enact high-level construals of the situation? Do we prefer power to be held by individuals who articulate plans that transcend the present and extend to the distant future, remote places, diverse groups, and unusual circumstances? These questions await future research.

Politeness

The theory of politeness (Brown and Levinson, 1987) views politeness as a signifier and producer of social distance. CLT views social distance as a specific case of psychological distance, and as such expects it to be related to level of construal and to other distances. Indeed, research by Stephan et al. (in press) shows that **greater politeness is associated with higher-level construals and with greater temporal and spatial distance.** For example, participants were more polite when they addressed a person whom they construed in terms of abstract goals and dispositions, when they expected the target to receive the message in the relatively distant future, and when they addressed individuals in relatively distant locations. Examining the opposite direction of influence, the research also showed that a request to generate polite statements prompted participants to use abstract verbs, that polite utterances were judged as pertaining to the relatively distant future and as directed to addressees in relatively remote locations.

These findings help to understand why different languages and cultures seem to signify politeness in similar ways. For example, standing very close to an interlocutor is generally considered to be less polite than standing farther away (i.e., greater spatial distance is associated with more politeness). Also, tense shifts in verbal communication tend to affect politeness, with the use of present tense being less polite as compared to past or future tenses (i.e., greater temporal distance is associated with more politeness). Likewise, requests and remarks are considered more polite when presented in a more indirect,

abstract form rather than a direct, concrete form (i.e., higher construal is associated with greater politeness). These examples demonstrate that greater politeness may be achieved in social relations and communication when they entail greater spatial or temporal distancing and abstraction. This is neither a coincidence nor an arbitrary convention, but rather an expression of an underlying regularity wherein politeness, as a regulator of social distance, is expressed by a high level of construal and greater distance on various dimensions. Of course, different cultures might use different distance and construal cues to convey politeness (e.g., the use of plural as a polite way of addressing a person does not exist in many languages). However, CLT predicts that such cues would always use higher-level construals and greater distances to communicate greater politeness, rather than less politeness. This prediction awaits empirical cross-cultural examination.

Social conflict

Issues within an interpersonal negotiation can differ in their centrality and worth. If a pair of negotiators can trade off their lowest and highest priority issues (e.g., give in on secondary issues in exchange for getting what they want on high-priority issues; a process called logrolling), they are more likely to succeed in “expanding the pie,” maximizing both individual and joint outcomes. Because negotiators should be expected to focus more on central concerns and less on peripheral concerns as distance increases, we would expect to see more logrolling agreements in a distant future than near future context. Examining this idea within the context of a live negotiation, Henderson et al. (2006) found that while 91 percent of dyads with a temporally distant perspective reached a fully logrolling agreement, only 50 percent of dyads with a temporally near perspective did so. The enhanced reciprocal concessions made by dyads with the temporally distant perspective culminated in better negotiated individual and joint outcomes. Moreover, research on the role of

construal levels in the negotiation process has shown that negotiators who construed issues abstractly rather than concretely were more likely to discover integrative agreements (Henderson and Trope, 2009).

These findings have implications for how negotiators handle conflicts over minor versus major issues and specific interests versus broad values and ideological differences in situations that do not allow for trade-offs (Harinck and De Dreu, 2004). Specifically, CLT suggests that the resolution of minor issues and specific interests should be hindered when individuals adopt a psychological proximal perspective or a lower-level construal and facilitated by having a more distal perspective and abstract construal. Conversely, the resolution of conflict over major issues, values, and ideological differences should be hindered when individuals adopt psychological distant perspective or a higher-level construal and facilitated by a less abstract construal and a psychologically proximal perspective.

In summary, a range of studies suggests that people rely on high-level construals to a greater extent when predicting, evaluating, and taking action with respect to more distant situations. Once again, the broad range of distance-dependent responses uncovered by this line of CLT research seemed to justify the level of generality at which the theory was formulated. Moreover, because these responses involved potentially consequential judgments and choices in the social context, not only construals, the scope of the findings we have obtained is testimony to the practicality of the theory and its applicability to the functioning of individuals in groups and society.

WHAT CONSTRUAL LEVEL THEORY IS NOT

In the preceding two sections, we have discussed the questions addressed by CLT. In this final section we discuss questions that

one might address to CLT. Throughout the history of social psychology, three questions have often been asked of theories of social cognitive processes: Do the processes lead to accurate judgments about reality? Do they facilitate goal attainment? Do they promote social accommodation? Applied to CLT's distinction between high-level and low-level mental construals, these questions become: Does one construal level lead to more accurate judgments than the other? Is one construal level more conducive to goal attainment? Is one construal level more conducive to social accommodation? We consider these three questions in turn.

Is one construal level more conducive to accurate predictions than the other?

It is debatable whether there is a single objective criterion for accuracy. But supposing there is some verifiable event – in one's personal life or in the social or physical reality – the question is which construal process leads to more accurate prediction of that event. Our answer is: it depends. In reality, an event might be determined by low-level or high-level factors. Quite obviously, then, when high-level factors determine an event, high-level construals are likely to be more accurate, and when low-level factors determine an event, low-level construals are likely to be more accurate. In other words, any given level of construal would lead to more accurate predictions of future events to the extent that it matches the level of factors that actually determine the occurrence of those events.

Of particular interest are predictions of one's own evaluative, emotional, or behavioral reaction to events (e.g., evaluations of a course, job, or vacation). Research on affective forecasting, for example, has shown that people's predictions of their affective reactions to a future event are inaccurate (Wilson and Gilbert, 2003). We think that this is, in part, because in affective forecasting studies

the levels of construal underlying people's predictions and their actual reactions do not match. Specifically, reactions are assessed soon after the event occurs and are therefore based on a low-level construal of the event. In contrast, predictions of one's reactions to an event are made from a more distant perspective and therefore likely to be based on high-level construals of the event. The mismatch between the construal levels underlying reactions and predictions should make the predictions inaccurate, as affective forecasting research shows. For example, a scenic (high-level aspect) but uncomfortable (low-level aspect) bike trip is likely to seem more enjoyable in prospect than during the trip.

However, although the construal levels of predictions and *immediate* reactions to an event are mismatched, the construal levels of predictions and subsequent *retrospective* reactions to the event might be matched, rendering predictions relatively more accurate. For example, both predictions and distant retrospective reactions are likely to be based on ends rather than means to those ends. To use the bike trip example, both prospective and retrospective evaluations of a scenic but uncomfortable bike ride are likely to be positive, and predictions of retrospective evaluations would thus be relatively accurate. In short, failure to predict one's immediate reactions to a future event is one of many possibilities: predictions from a distal perspective may better predict long-term than short-term reactions to events, whereas predictions from a proximal perspective may better predict short-term than long-term reactions to events.

Is one construal level more conducive to goal attainment than the other?

Again, our answer to this question is that it depends. First, in reality, the attainment of a goal may depend on high-level or low-level factors. Therefore, high-level or low-level construals are more conducive to the attainment

of a goal if they match the level of factors that actually determine the attainment of that goal. For example, abstract thinking would promote the attainment of goals that require detection of global patterns, whereas concrete thinking would promote the attainment of goals that require detection of specific details (see Wakslak et al., 2006). Second, and perhaps more interesting, the goal itself may be a high-level or low-level goal. High-level goals are superordinate, global, and central, whereas low-level goals are subordinate, local, and secondary. Therefore, high-level construals and a distal perspective might promote high-level goal pursuit (and thus also self-control), whereas low-level construals and a proximal perspective might promote low-level goal pursuit. For example, abstract thinking would promote pursuit of one's high-priority goals, whereas concrete thinking would promote pursuit of one's low-priority goals (see Henderson et al., 2006; Eyal et al., 2009).

Is one construal level more conducive to social accommodation than the other?

Yet again, the answer is that it depends. **Taking the perspective of others is a form of traversing psychological distance and therefore likely to be facilitated by high-level construals.** Taking others' perspective, in turn, promotes forming a socially shared reality and might therefore also promote social alignment with the group opinion. However, the group opinion might be entirely incidental to the decision the individual is facing. For example, the political opinions voiced by strangers one happens to meet in a local bar are likely incidental to one's decision how to vote in an upcoming election. Individuals at a high level of construal are therefore less likely to align their opinion to that of this group than are individuals at a low level of construal. For example, when the election is temporally or spatially distant, thus activating a high-level construal of the

vote, individuals are likely to resist the group opinion and, instead, rely on their own pre-existing attitudes. However, when the election is temporally or spatially proximal, thus activating low-level construals, individuals' opinions are more likely to shift toward that of the group. In general, a person's judgments are more likely to conform to those espoused by an incidental group when psychological distance is small rather than large. Recent research on attitude alignment with incidental strangers provides support for this hypothesis (Ledgerwood et al., 2009). For example, participants' voting intentions aligned with those of an incidental stranger when the policy was going to be implemented in the near future. In contrast, participants were unaffected by the incidental stranger's views when the policy was going to be implemented in the distant future. Moreover, this pattern was replicated when construal (abstract versus concrete) was directly manipulated via a mindset induction.

CONCLUSION

We started with how the value of outcomes changes over temporal distance and ended up with what we hope is a step toward a general theory of psychological distance. There are different psychologies for the different dimensions of psychological distance – temporal, spatial, social, and hypotheticality. Examples are the psychologies of time perspective, propinquity, social perspective taking, ingroup and outgroup perception, counterfactual thinking, and source monitoring. Without denying the uniqueness of each of these dimensions and the even more specific aspects comprising those dimensions, we believe they all entail transcending the present through constructing mental models of what is not directly experienced. The farther removed an object is from me on any distance dimension, the higher (more abstract) the level of mental construal of that object. Psychological distance thus expands or

contracts depending on level of construal. Consistent with this proposal, research conducted in the framework of CLT suggests that (1) **different distance dimensions are mentally associated**, (2) **distance on any of these dimensions influences and is influenced by higher levels of mental construal**, and (3) **the various distances are, to some extent, interchangeable in their effects on prediction, evaluation, and choice**.

Let us note in conclusion that CLT is a theory of cognition, motivation, and self-regulation. Levels of construal refer not only to how people represent and process information about objects, it also refers to the goals people want to attain and the plans they make to actually attain them. **High-level construals entail pursuing superordinate goals** (e.g., being healthy), whereas **low-level construals involve pursuing subordinate goals** (e.g., avoiding fatty food). **High-level construals entail prioritizing one's primary and global goals over secondary and local ones**, whereas **low-level construals do not sharply distinguish between these types of goals**. **High-level construals entail forming general and structured plans for achieving one's goals**, whereas **low-level construals entails forming contextualized and ad-hoc plans**. Most important, **high-level construals enable us to care about and act upon what is not present**, whereas **low-level construals enable us to care about and act upon what is present**.

REFERENCES

Ainslie, G. (1975) Specious reward: A behavioral theory of impulsiveness and impulse control. *Psychological Bulletin*, 82, 463–496.

Asch, S.E. (1956) Studies of independence and conformity: A minority of one against a unanimous majority. *Psychological Monographs*, 70, Whole No. 416.

Bar-Anan, Y., Liberman, N., Trope, Y. and Algom, D. (2007) Automatic processing of psychological distance: Evidence from a Stroop task. *Journal of Experimental Psychology: General*, 136, 610–622.

Bar-Anan, Y., Liberman, N. and Trope, Y. (2006) The association between psychological distance and construal level: Evidence from an Implicit Association Test. *Journal of Experimental Psychology: General*, 135, 609–622.

Brown, P. and Levinson, S. (1987) *Politeness: Some universals in language usage*. Cambridge, England: Cambridge University Press.

Buehler, R., Griffin, D. and Ross, M. (1994) Exploring the 'planning fallacy': Why people underestimate their task completion times. *Journal of Personality and Social Psychology*, 67, 366–381.

Elster, J. (1977) Ulysses and the sirens: A theory of imperfect rationality. *Social Science Information/sur les sciences sociales*, 16, 469–526.

Eyal, T., Liberman, N., Sagristano, M. and Trope, Y. (2009) Resolving value conflicts in making choices for the future. Unpublished manuscript, Ben Gurion University, Israel.

Förster, J., Friedman, R.S. and Liberman, N. (2004) Temporal construal effects on abstract and concrete thinking: Consequences for insight and creative cognition. *Journal of Personality and Social Psychology*, 87, 177–189.

Freitas, A.L., Salovey, P. and Liberman, N. (2001) Abstract and concrete self-evaluative goals. *Journal of Personality and Social Psychology*, 80, 410–412.

Fujita, K., Trope, Y., Liberman, N. and Levin-Sagi, M. (2006) Construal levels and self-control. *Journal of Personality and Social Psychology*, 90, 351–367.

Gilbert, D.T. and Wilson, T.D. (2007) Prospection: Experiencing the future. *Science*, 317, 1351.

Gilovich, T. and Medvec, V.H. (1995) The experience of regret: What, when, and why. *Psychological Review*, 102, 379–395.

Gilovich, T., Kerr, M. and Medvec, V.H. (1993) Effect of temporal perspective on subjective confidence. *Journal of Personality and Social Psychology*, 64, 552–560.

Green, L., Meyerson, J. and McFadden, E. (1997) Rate of temporal discounting decreases with amount of reward. *Memory and Cognition*, 25, 715–723.

Griffin, D.W. and Ross, L. (1994) Subjective construal, social inference, and human misunderstanding. In M. Zanna (ed.), *Advances in Experimental Social Psychology*, 26, 319–360. New York: Academic Press.

Harinck, F. and De Dreu, C.K.W. (2004) Negotiating interests or values and reaching integrative agreements: The importance of time pressure and temporary impasses. *European Journal of Social Psychology*, 34, 595–612.

Henderson, M.D., Fujita, K., Trope, Y. and Liberman, N. (2006) Transcending the 'here': The effect of spatial distance on social judgment. *Journal of Personality and Social Psychology*, 91, 845–856.

Henderson, M.D., Trope, Y. and Carnevale, P. (2006) Negotiation from a near and distant time perspective. *Journal of Personality and Social Psychology*, 91, 712–729.

Henderson, M.D. and Trope, Y. (2009) The effects of abstraction on integrative agreements: When seeing the forest helps avoid getting tangled in the trees. *Social Cognition*, 27, 399–415.

Kahneman, D. and Lovallo, D. (1991) Timid choices and bold forecasts: A cognitive perspective on risk taking. *Management Science*, 39, 17–31.

Kahneman, D. and Tversky, A. (1979) Prospect theory: An analysis of decisions under risk. *Econometrica*, 47, 263–291.

Kirby, K.N. and Herrnstein, R.J. (1995) Preference reversals due to myopic discounting of delayed reward. *Psychological Science*, 6, 83–89.

LaPiere, R.T. (1934) Attitudes versus actions. *Social Forces*, 13, 230–237.

Ledgerwood, A., Trope, Y. and Chaiken, C. (2009) Attitudes in their social context: Malleability, stability, and the role of construal. Unpublished manuscript, University of California, Davis.

Lewin, K. (1951) *Field Theory in Social Science*. New York: Harpers.

Liberman, N. and Förster, J. (2009) Distancing from experienced self: How global versus local perception affects estimation of psychological distance. *Journal of Personality and Social Psychology*, 97, 203–216.

Liberman, N. and Trope, Y. (1998) The role of feasibility and desirability considerations in near and distant future decisions: A test of temporal construal theory. *Journal of Personality and Social Psychology*, 75, 5–18.

Liberman, N., Sagristano, M. and Trope, Y. (2002) The effect of temporal distance on level of construal. *Journal of Experimental Social Psychology*, 38, 523–535.

Liviatan, I., Trope, Y. and Liberman, N. (2008) Interpersonal similarity as a social distance dimension: Implications for perception of others' actions. *Journal of Experimental Social Psychology*, 44, 1256–1269.

Liviatan, I., Trope, Y. and Liberman, N. (2010) A construal level analysis of the relationship between similarity and attraction. Unpublished manuscript, New York University.

Loewenstein, G.F. (1996) Out of control: Visceral influences on behavior. *Organizational Behavior and Human Decision Processes*, 65, 275–292.

Loewenstein, G. and Prelec, D. (1992) Anomalies in intertemporal choice: Evidence and an interpretation. *Quarterly Journal of Economics*, 107, 53–587.

Loewenstein, G.F., Weber, E.U., Hsee, C.K. and Welch, N. (2001) Risk as feelings. *Psychological Bulletin*, 127, 267–286.

Lovallo, D. and Kahneman, D. (2000) Living with uncertainty: Attractiveness and resolution timing. *Journal of Behavioral Decision Making*, 13, 179–190.

Metcalfe, J. and Mischel, W. (1999) A hot/cool-system analysis of delay of gratification: Dynamics of willpower. *Psychological Review*, 106, 3–19.

Milgram, S. (1965) Some conditions of obedience and disobedience to authority. *Human Relations*, 18, 57–76.

Miller, N.R. (1944) Experimental studies of conflict. In McV. Hunt (ed.), *Personality and the Behavior Disorders*, pp. 431–465. New York: Ronald Press.

Mischel, W., Grusec, J. and Maters, J.C. (1969) Effects of expected delay time on the subjective value of rewards and punishments. *Journal of Personality and Social Psychology*, 11, 363–373.

Nisan, M. (1972) Dimension of time in relation to choice behavior and achievement orientation. *Journal of Personality and Social Psychology*, 21, 175–182.

Nussbaum, S., Liberman, N. and Trope, Y. (2006) Predicting the near and distant future. *Journal of Experimental Psychology: General*, 135, 152–161.

Rachlin, H. (1995) Self control: Beyond commitment. *Behavioral and Brain Sciences*, 18, 109–159.

Ross, M. (1989) Relation of implicit theories to the construction of personal histories. *Psychological Review*, 96, 341–357.

Sagristano, M.D., Trope, Y. and Liberman, N. (2002) Time-dependent gambling: Odds now, money later. *Journal of Experimental Psychology: General*, 131, 364–536.

Schacter, D.L. and Addis, D.R. (2007) The cognitive neuroscience of constructive memory: Remembering the past and imagining the future. *Philosophical Transactions of the Royal Society*, 362, 773–786.

Schelling, T. (1984) Self command in practice, in theory and in a theory of rational choice. *American Economic Review*, 74, 1–11.

Smith, P.K. and Trope, Y. (2006) You focus on the forest when you're in charge of the trees: Power priming

and abstract information processing. *Journal of Personality and Social Psychology*, 90, 578–596

Stephan, E., Liberman, N. and Trope, Y. (2010) Politeness and social distance: A construal level perspective. *Journal of Personality and Social Psychology*, 98, 268–80.

Stroop, J.R. (1935) Studies of interference in serial verbal reactions. *Journal of Experimental Psychology: General*, 18, 643–662.

Thaler, R.H. (1981) Some empirical evidence on dynamic inconsistency. *Economic Letters*, 8, 201–207.

Todorov, A., Goren, A. and Trope, Y. (2007) Probability as a psychological distance: Construal and preferences, *Journal of Experimental Social Psychology*, 43, 473–482.

Trope, Y. and Liberman, N. (2000) Temporal construal and time-dependent changes in preference. *Journal of Personality and Social Psychology*, 79, 876–889.

Wakslak, C.J., Trope, Y., Liberman, N. and Alony, R. (2006) Seeing the forest when entry is unlikely: Probability and the mental representation of events. *Journal of Experimental Psychology: General*, 135, 641–653.

Wakslak, C.J. and Trope, Y. (2009) The effect of construal-level on subjective probability estimates. *Psychological Science*, 20, 52–58.

Wilson, T.D. and Gilbert, D.T. (2003) Affective forecasting. In M.P. Zanna (ed.), *Advances in Experimental Social Psychology*, 35, 345–411. San Diego: Academic Press.

Zimbardo, P. G. (1971) The power and pathology of imprisonment. *Congressional Record*. (Serial No. 15, 1971-10-25). Hearings before Subcommittee No. 3, of the Committee on the Judiciary, House of Representatives, Ninety-Second Congress, *First Session on Corrections, Part II, Prisons, Prison Reform and Prisoner's Rights: California*. Washington, DC: U.S. Government Printing Office.

An Attribution Theory of Motivation

Bernard Weiner

ABSTRACT

Expectancy value theory as modified by Atkinson, locus of control research initiated by Rotter, and the naïve psychology of achievement performance advanced by Heider provided the stepping stones for the formulation of the attribution theories of motivation presented in this chapter. Two theories are reviewed – an intrapersonal conception that primarily addresses achievement striving, and an interpersonal theory that focuses on social behaviors including help-giving, aggression, and reactions to the stigmatized. Causal beliefs, their underlying properties, and distinctive links to emotions form the foundation for these theories, which suggest a thinking–feeling–action sequence as the “deep structure” for motivated behavior. Application of the theories to achievement striving, physical and mental health, and social support are discussed.

INTRODUCTION

To understand the particular version of attribution theory presented in this chapter,

I must first introduce my mentor, John W. Atkinson. Atkinson took part in the pursuit of a “grand theory” of motivation. This goal characterized the study of motivation for about half a century, from roughly 1930 to 1975. For the experimental motivation psychologist, the aim was to identify the determinants of action and specify their mathematical and/or sequential (temporal) relations. The most influential of these approaches was associated with Clark Hull and Kenneth Spence (Hull, 1943; Spence, 1956), the prime creators of drive theory. Their conception specified that behavior is determined by $\text{Drive} \times \text{Habit}$ and other factors such as incentives. A competing theoretical approach was linked to Edward Tolman (1932), Julian Rotter (1954), and Atkinson (1957), who instead argued that behavior is directed by $\text{Expectancy} \times \text{Value}$ and, for Atkinson, motives as well. Kurt Lewin (1938), who also may be considered an $\text{Expectancy} \times \text{Value}$ theorist, additionally proposed a temporal link connecting the behavioral determinants. He postulated

that an object acquires motivational properties (a positive or negative incentive valence) only after there is a need within the organism. This results in a motivation sequence of: need → incentive → force (behavior tendency).

These grand theories were presumed to account for all behavior, regardless of content domain. If not content-free, then at least the conceptions could be called content independent (although their experimental instantiation was often in quite narrow settings). As an example, for Expectancy/Value theorists motivated behavior across contexts is a product of what one expects to receive multiplied by the perceived probability of attaining that goal, with motivated behavior conceived as a rational, hedonic choice. Hence, the theories are called “grand,” perhaps a shortened version of grandiose, for their aspirations were formidable and their hopes embraced a unifying theory for all behavior. It is reasonable to propose that Einstein was the guiding light, with recognition of some limitations imposed by animate as opposed to inanimate objects of study. Indeed, while I was a graduate student, Atkinson was immersed in Isaac Newton and attempted to incorporate Newton’s principles of motion into motivation theory (see Atkinson and Birch, 1970).

My use of attribution-related concepts, which I impose on issues in motivation rather than (or, in addition to) social perception, imitates the grand theory pursuits. I attempt to specify the determinants of behavior (as the Drive and Expectancy/Value theorists) and identify their sequential arrangement (as did Lewin), albeit with causal beliefs as the theoretical bedrock. This theoretical approach distinguishes my version of attribution theory from other attribution theories and from many current motivation and social psychological theories as well. That is primarily because the theoretical structure includes multiple components in specified interrelations and transcends specific content domains, which I regard as the *sine qua non* for the label of “theory.”

There is another historical antecedent, most clearly associated with E.L. Thorndike and his law of effect, which guides my work. Thorndike (1911) stated that behaviors previously rewarded will be repeated, whereas those that were punished will be avoided or extinguished. In this manner, Thorndike incorporated the past into the present. As an attribution theorist, I also ask how the past makes its way into the present. But rather than having reinforcement as the mechanism, it is contended that the interpretation of the past, that is, the perceived causes of prior events, determine what will be done in the future. Why one was rewarded or punished, not the outcome *per se*, directs the future, so that both reward and punishment may have positive or negative motivational consequences.

THEORETICAL BEGINNINGS

Atkinson’s version of Expectancy/Value theory embraced three idiosyncratic components:

- 1 As already revealed, motivation is also determined by motives, so that the equation for strength of motivation is Motive \times Expectancy \times Value. The motive Atkinson was wed to was the need for achievement.
- 2 In achievement settings only, incentive (value) was conceptualized as an affect, pride in accomplishment (here I consider only approach motivation and positive affect).
- 3 In achievement settings only, incentive was postulated to be inversely related to expectancy so that pride was presumed to be greater given success at a difficult task (low expectancy of success) rather than at an easy task (high expectancy of success).

A number of perils awaited as I continued in this research tradition following graduate school. First, achievement needs were assessed with a projective instrument, the Thematic Apperception Test (TAT). Scoring and motive classification decisions were time-consuming. Then, those designated as

high or low in achievement needs were called back for an experimental manipulation, typically involving success or failure at tasks varying in subjective expectancy of success. But by the time I completed the individual difference assessment, some participants had fulfilled their experimental requirements. Others would not take part in the second phase of the research. One consequence of this procedure and the loss of subjects was that I could only complete a few experiments yearly, hardly a fountain of empirical research for an assistant professor.

Furthermore, the theory was very restricted in its predictions, in spite of the “grand” foundation. The main theoretically generated hypothesis was that individuals high in achievement needs are especially attracted to intermediate difficulty tasks relative to persons low in achievement needs. Unfortunately, this and other hypotheses related to motive-group differences often were not confirmed in experimental studies (see Weiner, 1992). In short, I had reason to fear the outcome of any later tenure decision.

I therefore began to search for other predictors and predictions and more economical experimental procedures that would facilitate my research output. About that time, Julian Rotter (1966) published a monograph examining “locus of control.” I was previously unaware of this body of work, which grew from studies manipulating skill versus luck tasks and their influence on expectancies. It seemed logical to pursue the idea that persons with high achievement needs view the world as controllable by them, more so than persons low in achievement needs. That is, persons high in achievement needs attribute outcomes to internal factors such as skill, whereas those low in achievement needs perceive luck (external control) to be the major determinant of their success and failure. I devised a correlation study to test these ideas, administering achievement and locus of control measures to a large subject population (Weiner and Potepan, 1970). This study did yield some positive findings. I find it amusing, or frightening,

that two measures (the TAT and locus of control scale) and two constructs (need for achievement and locus of control) that I now reject launched my research direction in attribution theory. I am unsure if scientific progress is better made by standing on the shoulders of giants to see further or by jumping off these giants to attain a different perspective.

After examining the locus of control construct and measure, a few enigmas became evident. For example, if one fails because of a perceived lack of aptitude, the locus is internal yet the cause is not subject to volitional change. That is, there can be internality without control. In addition, if one succeeds at a task because of perceived high ability, then subsequent failure at that task would not be ascribed to low ability. But this is not the case given effort as the perceived cause – success or failure at the same task could be regarded as due to the degree of effort expenditure. Furthermore, the perceived determinants of success are not the same as failure – evidence of personal enhancement is pervasive, so that success more than failure elicits beliefs that the self played a role. In sum, distinctions between locus and control, ability and effort, and success and failure have to be made when considering causal ascriptions. These conclusions promoted my skepticism that locus of control is a trait and/or a unitary construct (just as I questioned whether the achievement motive is a trait inasmuch as one could be motivated, e.g., at tennis but not ping-pong). It also pushed me to think harder about perceived causality, the gateway to attribution theory. And it raised a lifelong question that is implicit in all my work, namely: “In what ways do ability and effort differ?”

I was mulling over these issues when reading a book by Richard de Charms (1968) entitled *Personal Causation*. De Charms offered a distinction quite similar to Rotter, labeling some individuals “origins” and others “pawns.” In his book he also reviewed an experiment by Schmitt (1964) in which

moral judgments are made about individuals who do not repay their debts either because of lack of ability to work (e.g., because of illness) or insufficient effort (e.g., not working because of laziness). I recognized that ability and effort also could be varied as causes of success and failure in an achievement context. In a simulation experiment (Weiner and Kukla, 1970), we described schoolchildren as succeeding or failing and factorially varied whether they had ability or not, and exerted effort or not. The participants were instructed to evaluate (reward or punish) these students. These studies were easy to conduct, the variables easy to manipulate, the results easy to evaluate, and the findings systematic and replicable with no individual differences involved. Among the results, the data revealed that the pattern of low ability, high effort, and success produces the greatest reward, whereas the combination of high ability, low effort, and failure gives rise to the greatest punishment. Hence, lack of ability is a positive facilitator of achievement evaluation (holding outcome and effort constant), and achievement evaluation shares characteristics with other moral judgments. I felt more at ease regarding my future empirical prospects.

I submitted a manuscript of about 20 pages for publication, containing three experiments, to the *Journal of Personality and Social Psychology*. The editor at that time was extremely critical, insightful, smart, and wordy. He wrote a ten-page single-spaced editorial response, with an invitation to resubmit. By the time I read and understood his comments, I had conducted a fourth experiment, which was included in the resubmission. This was answered with about a five-page editorial response, again asking for a resubmission; I resent a manuscript with five experiments. The editor responded in his usual manner, this time with an abbreviated two-page comment. Finally, a six-experiment article was accepted and published (Weiner and Kukla, 1970). It was a key turning point in my research directions.

THEORETICAL DEVELOPMENT

I then rediscovered a field called attribution theory and that the seminal figure was Fritz Heider. I had skimmed Heider's (1958) book while a graduate student at the University of Michigan but did not take his "common sense" approach to motivation seriously. After all, Atkinson was studying Newton and seeking a mathematical representation for the universal laws of motivated behavior, whereas Heider was citing Shakespeare and Ibsen and relying on the everyday vocabulary of laypersons.

But reading Heider with the background knowledge of Rotter's research and my thoughts separating locus from control and ability from effort now placed his ideas within a more compatible framework. In Heider's (1958) naïve analysis of action, outcomes are ascribed to $\text{Can} \times \text{Try}$. Can , in turn, captures the relation of ability to the difficulty of the task. Thus, Heider specified three determinants of behavior: ability, task difficulty, and effort. Two of these (ability and effort) he considered internal to the person, whereas task difficulty is an external cause of an outcome.

Heider and Rotter did not cite one another, although both were concerned with the perceived causes of success and failure. Rotter acknowledged one internal and one external cause, respectively skill (ability) and luck (chance), whereas Heider intuited three causes (ability, effort, and task difficulty). I combined these two lists and hypothesized four main perceived causes of achievement outcomes – **ability, effort, task difficulty, and luck** (see Weiner et al., 1971) Two of these are internal to the person (ability and effort) and two are external (task difficulty and luck).

Within Rotter's Expectancy/Value framework, locus of control is related to expectancy, with greater expectancy shifts (increments after success, decrements following failure) anticipated given internal rather than external control beliefs. For example, following a

win at a tennis match he anticipated greater increments in expectancy of success than following a win at a coin toss, and similarly greater decrements after failure were hypothesized given lack of skill rather than bad luck ascriptions. Yet, closer examination of these predicted associations raised doubts. For example, if an exam is failed because the teacher is extremely harsh, it seems reasonable to hypothesize that expectancies will fall although the cause is external to the student, just as they had been hypothesized to drop when failure is attributed to the internal cause of low aptitude. Furthermore, if failure is ascribed to lack of studying and the student wants to do better in the future, then expectancy may not drop even though this cause is internal, just as expectancy had been anticipated to remain relatively unchanged if failure is attributed to the external cause of bad luck. That is, locus of control is not systematically related to expectancy and expectancy shifts because some internal causes of failure lead to large expectancy shifts (e.g., low ability) whereas others do not (e.g., lack of effort), just as some external ascriptions for failure produce major expectancy decrements (e.g., a harsh teacher) while others have little effect on subjective likelihood (e.g., bad luck).

Could there be another property or characteristic of causes rather than locus that accounts for expectancy shifts? I argued that this property is **causal stability** (see Weiner et al., 1971). If a cause is subject to future change, such as lack of effort (internal to the person) or bad luck (external causality), then failure would not produce downward shifts in expectancy – hope could be maintained. On the other hand, if the cause of failure is regarded as unchanging or stable, such as aptitude (internal) or a harsh teacher (external causality), then there would be an expectation of future failures and a state of hopelessness. Causal stability, I concluded, rather than causal locus, is the basis of expectancy shifts. Simply put, if the cause will prevail in the future, then the prior effect will be anticipated to recur, regardless

of causal locus. I later confirmed this in experimental investigations (see Weiner et al., 1976). In this manner I bridged the past to the present and future, as had Thorndike, but with a cognitive rather than a behavioral variable.

But then what is the function or consequence of causal locus? Here I returned to Expectancy/Value theory. If expectancy is accounted for by causal stability, then only the value (incentive, affect) component of Expectancy/Value theory remained, which could then be related to causal locus. Guided by Atkinson's analysis of pride, I reasoned that **attributions of success to internal factors give rise to more pride than do external ascriptions** – one feels greater pride after success at a tennis match attributed to high ability or high effort than following success at a match determined by chance or by the prior quality of the opponent (task ease). Hence, perception of locus of causality was the mediator accounting for Atkinson's assumption and the supporting data that pride is related to task difficulty. The harder the task, the more likely success is ascribed to the self (rather than to the ease of the task) and thus the greater the pride in accomplishment (Figure 7.1).

This line of reasoning resulted in Figure 7.2 (see Weiner et al., 1971), which is a 2×2 representation including four determinants of behavioral outcomes (ability, effort, task difficulty, and luck), their two properties or dimensions (causal stability and causal locus), and the linkages of expectancy to

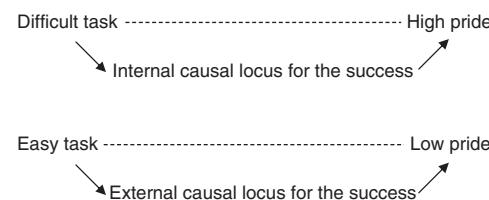


Figure 7.1 The harder the task, the more likely success is ascribed to the self (rather than to the ease of the task) and thus the greater the pride in accomplishment

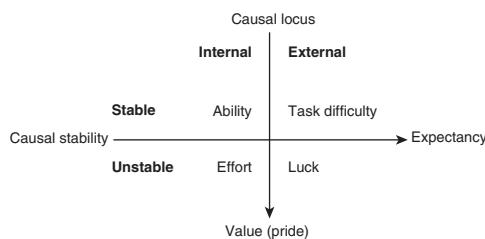


Figure 7.2 A 2 × 2 representation of perceived causality and linked consequences

causal stability and value to causal locus. Figure 7.2 reveals that when Rotter compared expectancy changes following success or failure given ability (internal and stable) versus luck (external and unstable) ascriptions, two dimensions or properties of phenomenal causality (locus and stability) were confounded. Rotter had incorrectly related expectancy to causal locus rather than to causal stability. The simple structure shown in Figure 7.2 provided the guide for my causal thinking and was the foundation for my subsequent theory building.

A more complete theory, I believed, needed to specify the antecedents that influence causal beliefs; all the perceived causes; additional properties or dimensions of causes; the connections between these dimensions and expectancy, affect, and other determinants of motivated behavior; and then action itself, operationalized with the usual motivational indicators of choice, intensity, and persistence of behavior. I was beginning to develop a structure that transcended content domains, included interrelated component parts, and provided a temporal sequence for a motivational episode, while retaining Expectancy/Value theory (see Weiner, 1985, 1986). These were the goals of the so-called “grand” theorists. My immediate aim was to progress in this direction.

THEORETICAL ELABORATION

I proceeded to amend the structure in Figure 7.2 by “filling in the blanks” mentioned

above. Perhaps the most important addition was to include a third causal dimension, **causal control**, guided by my prior reasoning that locus and control must be distinguished. The control construct created some difficulties because it apparently is not orthogonal to locus. An external cause is by definition not controllable by the actor, whereas some internal causes are controllable (e.g., effort) whereas others are not (e.g., aptitude). To address the orthogonality problem, I regarded some external causes as also controllable, but by others. For example, failure due to teacher bias is external to the pupil and uncontrollable by him or her, but bias is perceived by the student as controllable by the teacher. Hence, by shifting focus, a cause can be external yet controllable. Chance, on the other hand, is an external cause not controllable by anyone.

All causes, then, are locatable within a three-dimensional taxonomic space. Considering the four main determinants of achievement outcomes: aptitude is internal, stable, and uncontrollable; effort is internal, unstable, and controllable; task difficulty generally is regarded external, stable, and controllable (by the teacher); and luck is external, unstable, and uncontrollable. Dimensions provide the meaning or connotation of the cause. For example, aptitude “is” a property or characteristic that is internal to the person and stable that cannot be volitionally altered. But this is a phenomenological system, so that dimensional placement depends on “how it seems to me.” Thus, if success is attributed to being a “lucky person,” then luck is classified internal, stable, and uncontrollable, thereby having the same genotypic properties as aptitude. Similarly, if effort is expended by an industrious person, then effort might be considered not only internal and controllable but also stable rather than unstable.

A second important addition was to expand the list of emotions linked with causal beliefs. I became acquainted with appraisal approaches to emotion and adopted the position that **feelings are directed by thoughts** (see Weiner et al., 1978, 1979). Thus, for

example: attributions of success to the self generate pride; internal controllable causes for failure (e.g., lack of effort) give rise to guilt and regret; internal uncontrollable ascriptions for failure (e.g., lack of aptitude) produce shame and humiliation; stable causes of failure give rise to hopelessness; success due to an external cause controllable by others evokes gratitude; failure ascribed to an external cause controllable by others elicits anger; luck as the cause of either success or failure generates surprise; and so on (see Weiner, 2007). That is, I realized that a great deal of emotional life is guided by beliefs about causality.

In addition, it appears that some emotions, specifically happiness and unhappiness, are not linked to attributions but are tied to task outcomes. These are outcome-dependent, attribution-independent feelings. Hence, one is happy as well as surprised when succeeding because of good luck, while pride is not experienced nor is self-esteem enhanced because the cause is external to the actor. In a similar manner, failure to reach a desired goal attributed to lack of aptitude is hypothesized to produce unhappiness (outcome-dependent), a lowering of self-regard (locus-related), and shame (a consequence of internal, uncontrollable causality), along with a low expectation of future success and hopelessness and/or helplessness (stability-linked). Thus, two classes of emotion, **attribution-independent and attribution-dependent, are differentiated on the basis of their cognitive antecedents (task outcome versus causal attribution for the outcome)**, and multiple emotions are hypothesized to coexist.

Other theory-pertinent facts meanwhile were accumulating from many researchers interested in attribution processes. Foremost among the contributors was Harold Kelley (1967), who formulated the well-known "Kelley cube." His analysis identified social norms and past personal history among the key antecedents of causal beliefs. Kelley (1972) also examined causal schemata, or rules relating causes to effects, and these became recognized among the important

determinants of causal conclusions. It also became apparent that causal understanding is not sought in all instances and **is most fostered by the unexpected nonattainment of an important goal** (see Gendolla and Koller, 2001).

These additions were incorporated into the formulation of an attribution-based theory of personal motivation, shown in Figure 7.3. The temporal sequence in a motivated episode can be illustrated as follows. Assume a student fails an important exam. The initial experience following the failure is unhappiness. Assuming the outcome is negative and/or unexpected, there is a search for causality. Presume this person failed in the past even though she studied many hours, while others succeeded on this exam. This pattern of information gives rise to the belief that the current failure is due to lack of aptitude. Aptitude is an internal, stable, uncontrollable cause, so there is a lowering of self-esteem, low expectancy of future success, hopelessness and helplessness, and shame and humiliation. Low expectancy accompanied by these negative affects is predicted to promote the decision to drop out of school.

Conversely, imagine that another fails the same task. This person also initially experiences unhappiness. But she has been successful in the past and laments that the night before the exam she was partying rather than studying. Hence, her current failure is ascribed to insufficient effort. This internal ascription to an unstable, controllable cause lowers personal regard but also gives rise to the maintenance of expectancy, hope, guilt, and regret, all of which are positive motivators. Hence, motivation is anticipated to increase and she is predicted to try harder in the future.

This same line of reasoning can be applied to affiliative contexts as well. For example, assume Bill calls Jane for a date and is rejected. This causes unhappiness and elicits attributional search; Bill asks: "Why won't you go out with me?" Jane may be unlikely to tell some truths because she does not want to lower Bill's self-regard by providing

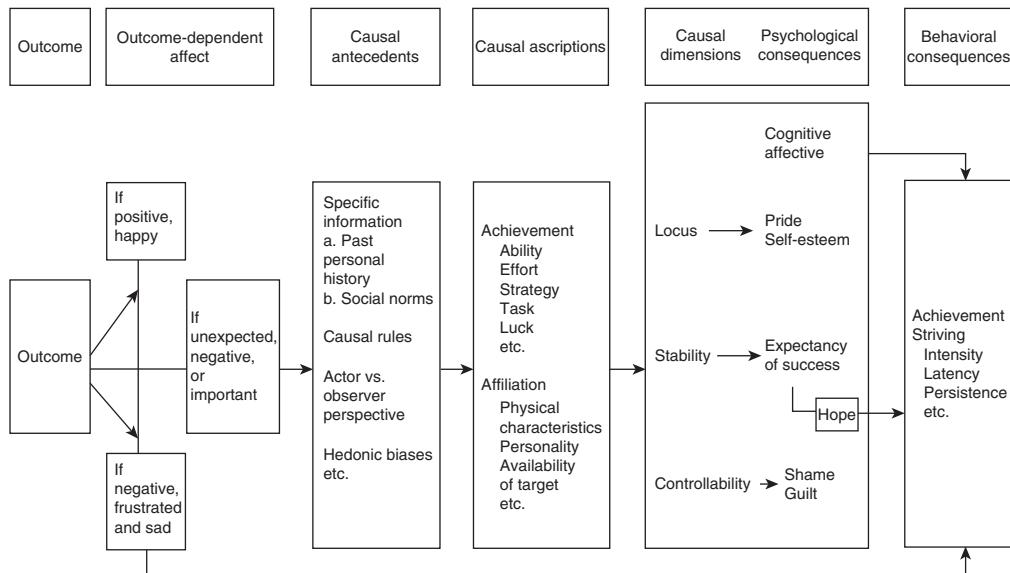


Figure 7.3 An attributional theory of intrapersonal behavior

a cause internal to him such as: “You are boring” (see Folkes, 1982). But in spite of her replying: “I have to study” (an excuse or causal substitution), Bill ruminates that he has been refused many times, whereas others have girlfriends. He concludes: “I am a boring person.” This ascription has the same genotypic or conceptual properties as does failing math because of perceived low ability. That is, the cause is internal (lowering esteem), uncontrollable (raising humiliation), and stable (lowering expectancies and hope). Hence, he does not seek further dates. The theory is therefore applicable across motivational domains, primarily because there is genotypic representation of causal characteristics rather than a phenotypic description of causes. That is, low math aptitude and being a boring person are conceptually identical.

positive and negative outcomes respectively give rise to general positive and negative emotions; that undesired and/or unexpected events promote causal search; that the cause selected is dependent on a variety of antecedents, including social norms, past history, schematic beliefs, and hedonic biases; and that locus, stability, and control are key causal properties (although there may be other causal characteristics as well, such as globality or causal generality across situations). Causal locus relates to pride and self-esteem; causal stability in part determines expectancy shifts and the extent of hope, hopelessness, and helplessness; and causal control is linked with affects including shame (to uncontrollable causality), as well as guilt, and regret (to controllable causality). There are voluminous literatures supporting these assertions (see Weiner, 1985, 1986, 1995).

However, in spite of this array of evidence and theoretical generality, I would be remiss, perhaps even intellectually dishonest, if I did not share a misgiving. I am very concerned about the last links in the theory going from expectancy and affect to performance. Perhaps I am overly worried about these

EMPIRICAL VALIDATION AND EVALUATIVE CONCLUSION

Does this theoretical structure receive empirical support? There is little doubt that

associations given the secure place of Expectancy/Value theory in motivation history.

What is unclear are the associations of expectancy and attribution-linked emotions to achievement behaviors such as academic performance and dropping out of school. With regard to expectancy, motivational theorists have offered contradictory predictions. It may be that low expectancy of success or a difficult task generates most effort (see Locke and Latham, 1990); or, as Atkinson (1957) intuited, perhaps motivation is greatest when tasks are of intermediate difficulty; or perhaps being close to a goal (high expectancy) enhances performance (Lewin, 1938). In sum, the relation between expectancy and intensity of performance and choice behavior is uncertain. In regards to affect, similar difficulties arise. It is possible that high guilt impedes rather than facilitates performance, or that high shame is a positive motivator in certain settings, or even that shame and guilt cannot be readily distinguished inasmuch as they correlate quite highly in experimental research. Few studies have tested the entire theory with appropriate path analytic techniques (see Weiner, 1986), which in part accounts for my fear that the whole theory may be less than the sum of its parts because of the weak final links to motivated behavior.

There is another theoretical message in the massive support for the component associations in the theory accompanied by the uncertain validity of the entire conception. That is, it is not possible to construct a complete theory of achievement motivation. Achievement striving is impacted by causal beliefs but also is determined by other factors including anticipated costs and benefits (e.g., a promise of increased allowance from a parent could raise grades); affiliative concerns (e.g., studying with a potential mate may increase library time); the need to earn money (which impedes performance by decreasing the time available for study); and on and on. Overdetermination renders it exceedingly difficult to significantly relate

causal beliefs to molar achievement indices such as grade point average and dropping out of school.

What are the best predictors of school dropout and other signs of achievement striving in school settings? My guess is that these variables are school identification and feelings of belonging, peer-group norms, parental guidance and goals, and other determinants tied to societal values and socioeconomic class. This is not to imply that attributions have no effects. After all, attribution change programs have even proven effective in altering the performance of some students, as discussed later in this paper. In addition, attributions do explain aspects of emotional life in the classroom and expectations of success – no small feats.

This same general analysis is applicable to other motivation domains as well. Hence, the implicit position of this theoretical approach is that one set of principles (attribution-consequence relations) has been isolated that contribute to successful prediction in a wide variety of settings, yet these associations often are not the sole determinants of the behavior in question. Many principles stand side by side that are pertinent to the understanding of achievement strivings, which is why attribution thinking alone cannot be predictive of (or will only weakly predict) school performance. This conclusion must be taken into account by all motivation-based theories of school performance.

A THEORETICAL TURNING POINT

I now shift focus and offer a different, albeit related, theoretical development. As just indicated, I was disturbed by the difficulty in predicting school-related achievement performance and about the possibility that most variance in predictions of school motivation and dropout might be accounted for by broad demographic, socioeconomic indices rather than by hypothesized psychological mediators. I thus explored opportunities to shift

research from achievement to other human arenas. In doing so, I drifted further and further away from my historical roots and identification as an achievement theorist, thereby making comparisons with Atkinson's theory more problematic.

When thinking about future theoretical and empirical directions, I recognized that I was approaching achievement motivation from an intrapersonal perspective. The pertinent thoughts and emotions did not presume the presence or influence of others (with small exceptions such as the use of social norms as an antecedent of causal conclusions or social comparisons affecting emotional experiences such as shame). A research participant could be tested in an isolated room, with success and/or failure manipulated, and it was believed that attributions for these outcomes, as well as linked expectancies and emotions, would be generated that influence subsequent achievement striving. This is an asocial theory.

Yet in my first extended attribution research discussed earlier (Weiner and Kukla, 1970), participants assuming they were teachers evaluated students following hypothetical success or failure ascribed to various combinations of ability and effort. This judgment research concerned reactions toward others. Although the findings were essential to the generation of an intrapersonal theory, they in fact had little to do with personal achievement strivings but related to motivated behavior toward others, or social motivation. I had been thinking of asocial motivation but at times engaged in social motivation research; that is, I failed to recognize an intrapersonal/interpersonal distinction in the study of motivated behavior. My subsequent theory building explicitly turned from intrapersonal to interpersonal behavior and from achievement strivings to social motivation, particularly to help-giving (which I discuss here) and, to a lesser extent, aggression, impression management strategies such as giving excuses, and punishment (which I do not discuss here because of space limitations; see Weiner, 1995, 2006). I hoped that findings in these

areas would not only verify the associations between the already discovered components within the theory but also would provide support for the entire conception. And this proved to be the case!

THE EMPIRICAL DEPARTING POINT TO SOCIAL MOTIVATION

The initial studies explicitly in the domain of social motivation that I conducted returned to the ability/effort distinction explored by Weiner and Kukla (1970) and the earlier research by Schmitt (1964). It was evident from these publications that **lack of effort as the cause of a personal or social failure** (doing poorly on an exam; nonrepayment of a debt) **results in greater social disapproval than achievement or social failure because of broadly defined lack of ability** (e.g., low school capability; no money because of illness). Viewing these findings from the intrapersonal theory perspective shown in Figure 7.3, it appeared that ability and effort give rise to different evaluations because of their contrasting placements on the causal dimension of control. Lack of effort is a controllable cause; it "could have been otherwise." On the other hand, the absence of aptitude (or illness) is uncontrollable and not subject to volitional change. Thus, the same theoretical analysis is applicable in these two diverse contexts (see Weiner, 1995, 2006), again because phenotypically distinct causes share genotypic representations regarding causal controllability. But remember that in this case the behavior of the other, and not the self, is being considered. Hence, self-directed affects such as guilt and shame are inappropriate to include. How, then, should emotions be represented in an interpersonal context?

Incorporating emotions

The emotion appraisal literature provides extensive support for the hypothesis that

control beliefs regarding others have emotional consequences. Specifically, if another commits a transgression and the cause is under personal control, whether school failure due to lack of effort or coming late for an appointment because it was forgotten, then involved observers react with anger. One is angry when one's child fails at school because of not studying! On the other hand, if the cause of a transgression was not subject to volitional change, whether it is school failure attributed to low aptitude or being tardy for an appointment because the bus broke down, then observers react with sympathy (see, for example, Weiner et al., 1982). Consider your own reactions to the academic failure of a child with a mental handicap. A more complete social motivation analysis incorporating emotional reactions to causal thoughts regarding control by others is represented as follows, as in the intrapersonal theory:

Event → Cause → Causal controllability →
Emotion → Action

More specifically, considering observer reactions to the achievement striving of others, this sequence can be depicted as shown in Figure 7.4.

Refining controllability

The concept of causal control appears similar to personal responsibility. In earlier paragraphs the phrase “could have been otherwise” was used to elucidate or define

this construct. One is responsible (able-to-respond) for not expending effort, but not for lacking aptitude, which cannot be changed merely by willpower. On the other hand, the philosophical literature and writings on criminal justice clearly point out that control is distinct from responsibility. This is in part because there are controllable actions for which the individual is either not held responsible or responsibility is diminished. For example, responsibility is negated or weakened if a social transgression is in service of a higher moral goal (labeled a justification). One is justified in coming late for an appointment if a roommate became ill and had to be driven to the hospital. In this case, coming late is not regarded a “moral failure.” In a similar manner, if a crime is committed because the individual is mentally unstable, or does not understand the difference between right and wrong, then again the act may be controllable although the person is perceived as not responsible (see Weiner, 1995, 2006).

This reasoning promoted the inclusion of an additional path within the theory, one connecting control to responsibility. The **control–responsibility separation provided the opportunity** to incorporate moderators of this relation, even though on most occasions control and responsibility have the same value and meaning. The elaborated theory, including only the causal dimension of control, is depicted as:

Event → Causal antecedents → Cause → Causal dimension (control) → Personal responsibility → Emotion → Action

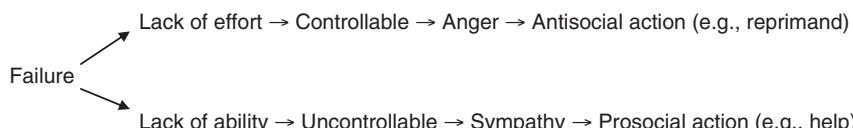


Figure 7.4 Social motivation analysis incorporating emotional reactions to causal thoughts regarding control by others, considering observer reactions to the achievement striving of others

TESTING THE INTERPERSONAL THEORY

The testing of the interpersonal theory is next documented in some detail to present a contrast to the discussion of the intrapersonal theory. Research attention was first given to help-giving because altruism and prosocial behavior form the core of social motivation. Two specific research contexts were examined: (1) help-giving in achievement situations (e.g., lending class notes); and (2) charity aid to the needy (e.g., financial help to various stigmatized individuals).

The logic of the experimental manipulations in these two settings was identical. Information was varied that created disparate beliefs about causal control/personal responsibility. In the achievement-related studies, pupils were portrayed as asking to borrow class notes either because they “went to the beach” (which elicits beliefs in personal control and responsibility) or because of eye problems (which gives rise to attributions of uncontrollability and inferences of nonresponsibility; see Schmidt and Weiner, 1988; Weiner, 1980). In the charity research, reactions toward individuals with stigmas that are apparently produced by controllable or volitionally undertaken behaviors (e.g., alcoholism, obesity) were compared with reactions to stigmas associated with uncontrollable states or conditions (e.g. blindness, Alzheimer’s disease). An alternative to offering contrasting stigmas was to present research participants with identical stigmas accompanied by different causes of the conditions (e.g., obesity because of overeating versus obesity caused by a thyroid problem; AIDS due to promiscuous sexual behavior versus AIDS caused by a transfusion with contaminated blood; see Weiner et al., 1988).

Research participants then rated the controllability of and/or personal responsibility for the cause of the need for class notes or the cause of the stigma, their affective reactions of anger and sympathy, and their prosocial

tendencies (likelihood of lending the notes; amount of charity allocation). Because hypothetical situations were created, behavioral determinants such as real time available to help and the actual amount of money that one has were not anticipated to affect the judgments. That is, other determinants of “behavior” were rendered irrelevant, thus overcoming some problems that plagued research regarding achievement strivings. Indeed, the studies in which I was involved often used simulation procedures – research participants “pretended” to be enrolled in the same class as the student in need or “pretended” to be members of a charity board dispensing funds. This simulation procedure raises hackles among some psychologists for both valid and invalid reasons that cannot be addressed here. Thankfully, others engaging in similar pursuits performed “real” manipulations (assuming that an experimental intervention can be labeled “real”), while some studies involved observations without personal intervention.

In contrast to the paucity of research investigations examining the full intrapersonal theory, many scientists conducted help-related research that assessed attributions as well as emotions and/or behaviors. Although the reported research studies did not manipulate or measure all the components in the hypothesized theory, the most essential mediational concepts were taken into account. The motivation sequences typically tested are shown in Figure 7.5.

Along with three colleagues (Rudolph et al., 2004), we conducted a meta-analysis of these research studies. To be included in the analysis, the investigation had to have at least one attribution variable, such as controllability; at least one emotion synonymous with anger or sympathy; and a relevant behavior. Also, first-order correlations were required. Our search found 39 usable studies, including nearly 8,000 research participants. A number of path models were examined, two of which are presented here – Model 1, which has direct links between thinking as well as affect and doing; and Model 2,

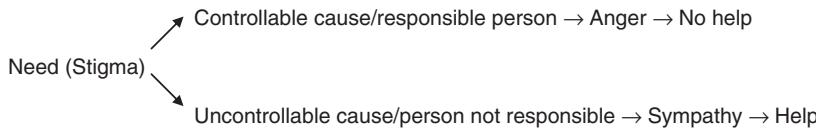


Figure 7.5 Typically tested motivational sequences

which has no thinking–action path so that affect is the only proximal determinant of behavior. Thus, the two models differ in whether thinking directly or only indirectly affect action.

Table 7.1 shows the path coefficients between the attribution–affect–behavior variables, all of which are significant. The table reveals that controllability is negatively related to sympathy ($\beta = -0.45$) and positively to anger ($\beta = 0.52$) for both models. That is, the more controllable the need or stigma (e.g., poverty because of laziness as opposed to a lack of available jobs, or AIDS because of promiscuous sexual behavior rather than a blood transfusion), the less the sympathy and the more the anger. In addition, sympathy positively relates to help-giving (average $\beta = 0.38$), while anger relates negatively to help (average $\beta = -0.08$). Table 7.1 also shows that the path between thoughts about control and behavior is relatively weak ($\beta = -0.05$); including this linkage does not improve the fit of the models.

In short, the full pattern of data is very consistent with the theory and, guided by the rule of parsimony, supports a thinking → feeling → acting motivation sequence – thoughts direct feelings and

feelings guide action. Thoughts do play a role in helping behavior, but only as distal determinants through their influence on emotion. I regard the thought (attribution) → affect → action sequence as the “deep structure,” or the basic genotypic representation, of a motivation episode. However, I certainly recognize that not all emotions require causal antecedents and that affect may influence thinking (i.e., there is an affect–cognition sequence). These beliefs nonetheless are relatively peripheral to this theoretical approach.

In related social motivation research examining aggression, the main findings reported in the studies of help-giving were replicated. However, a meta-analysis of aggression studies also found a direct link between attributions and action (see Rudolph et al., 2004; Weiner, 2006). Thus, although the basic structure of a motivation sequence is captured by the theory, much remains to be determined regarding the dynamics of behavior in specific motivation settings.

THE COMPLETE INTERPERSONAL THEORY

What, then, is the structure of the interpersonal theory and the contexts in which it has been validated? In addition to achievement judgments, help-giving, and aggression, the interpersonal theory has been extended to an analysis of the basis of social power (Rodrigues, 1995; Rodrigues and Lloyd, 1998) and to broader examinations of stigmas, although these areas of research are not reviewed here (see Weiner, 1995, 2006). They are, however, included in the complete

Table 7.1 Path coefficients for two help-giving models (from Rudolph et al., 2004)

	Model 1	Model 2
Control–Sympathy	-0.45	-0.45
Control–Anger	0.52	0.52
Control–Help-giving	-0.05	—
Sympathy–Help-giving	0.37	0.39
Anger–Help-giving	-0.07	-0.09

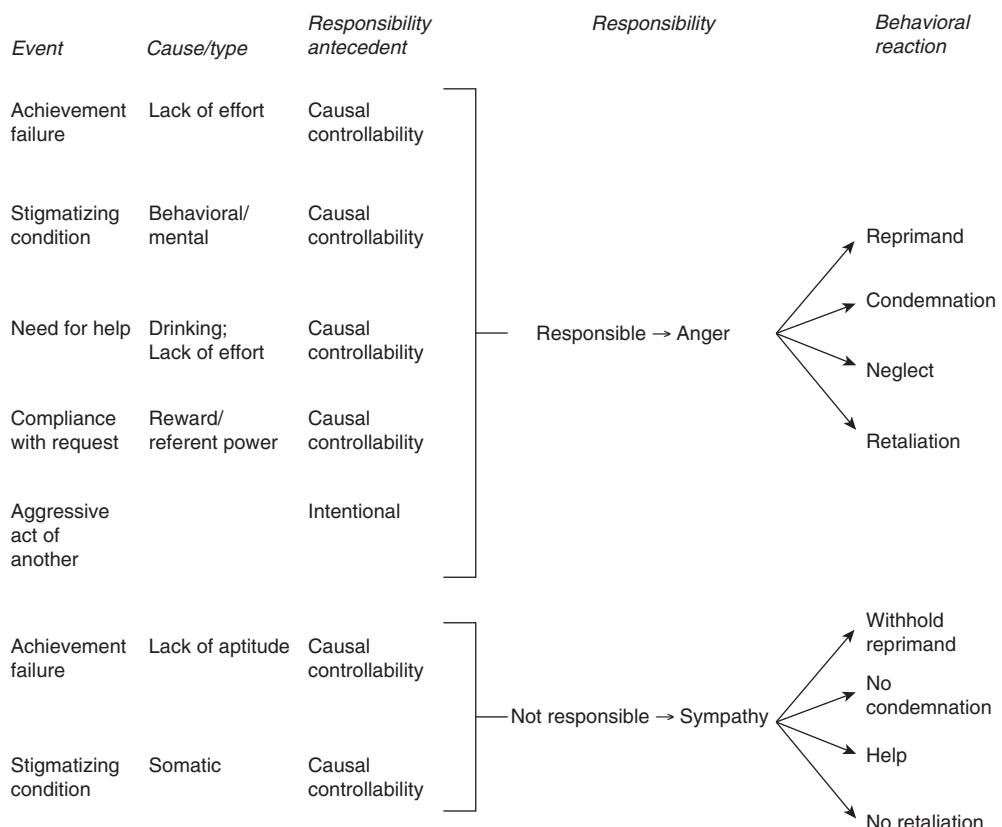


Figure 7.6 An attributional theory of interpersonal behavior

theory, which is shown in Figure 7.6. The upper portion of Figure 7.6 portrays situations in which the actor is personally responsible for a **negative outcome**. In the first row, the domain is achievement and the failed outcome is ascribed by an observer to lack of effort. Effort is a controllable cause and, in the absence of mitigating information, the person is held responsible for the failure. If there also is some personal involvement (e.g., the failing individual is your child), then the reaction elicited by low effort is anger. Anger, in turn, promotes an antisocial reaction such as punishment or rejection.

In contrast, the bottom portion of Figure 7.6 portrays instances in which the actor is not held responsible for the negative outcome. In achievement settings, the uncontrollable cause of failure is lack of aptitude.

Since aptitude is not controllable by the failing person, that individual is not held personally responsible for the negative outcome. Lack of responsibility for failure gives rise to sympathy and pity. These emotions, in turn, promote prosocial responding such as help-giving and comfort. This same general analysis applies to the other content domains included in Figure 7.6, where the logic of the theory is extended to explain help-giving, aggression, and social power.

In sum, this theory has interrelated component parts, generalizes to wide-ranging motivational contexts, and receives extensive empirical support. At the most molar level, the theory states that **motivated behavior is a function of cognition and affect**. This perhaps provides one among the many alternatives to the equally broad Lewinian statement

that behavior is a function of the person and the environment. A somewhat more specific formulation of this general law is that behavior is determined by causal reasoning and responsibility inferences, along with their linked emotions. In addition, emotions are proximal in accounting for action, while the proximal versus distal role of thinking remains to be determined and appears to differ in altruistic versus aggressive contexts.

COMPARING AND CONTRASTING THE INTRAPERSONAL AND INTERPERSONAL THEORIES

The intrapersonal and interpersonal attribution theories of motivation portrayed respectively in Figures 7.3 and 7.6 have much in common. For both, a motivational episode is initiated with a prior outcome; that is, the analysis is postperformance yet prior to the next action. Given an unexpected, negative, and/or important event, there is a search for causality. The cause selected is then analyzed according to its causal properties. This categorization gives rise to emotions, which are proximally linked to action. Hence, the structures of the intrapersonal and interpersonal theories are virtually identical and, for both, causal beliefs provide the foundation for theoretical analysis.

On the other hand, the two theories have important theoretical and empirical differences. At the conceptual level, the intrapersonal theory includes outcome-dependent affects not necessarily part of the interpersonal conception; the information used to make causal inferences about others differs from information used regarding the self inasmuch as personal causal antecedents are more readily available and there is the possibility of hedonic biasing; in the interpersonal theory the dimension of importance is controllability, while locus and stability are also included in the intrapersonal theory; because stability and expectancy are not part

of the interpersonal theory, the Expectancy/Value framework is maintained only in the intrapersonal conception (although these other dimensions could be incorporated within the interpersonal theory); in the interpersonal but not the intrapersonal conception, there is an added step linking control to responsibility, permitting the inclusion of mitigating factors; and given the intrapersonal theory, the emotions elicited are directed toward the self, whereas in the interpersonal theory the emotions involve others.

In addition to these theoretical contrasts, the theories differ in their empirical support. The intrapersonal approach has been primarily tested in achievement contexts, whereas the interpersonal framework has been empirically examined in a wide variety of motivational domains including help-giving, aggression, and social power. Because the interpersonal theory often is concerned with the judgments of others, the theory is more readily testable and has found greater support than the intrapersonal theory, which has attempted to predict the more elusive achievement performance. Thus, for the interpersonal theory, the whole indeed is more than the sum of the parts, whereas this may not be the case given the intrapersonal perspective. But both provide rich and complementary conceptual frameworks for the study of human motivation. They qualify for inclusion among the grandiose attempts to provide a unifying perspective to understand motivated behavior.

THEORETICAL APPLICATIONS

My motivational hero, Kurt Lewin, stated somewhere that “there is nothing as practical as a good theory.” It is one of the few things he wrote that I do not believe. In fact, there is nothing as practical as a simple engineering principle, without theoretical burden. For example, applications of psychology have perhaps been most fostered by Skinnerian rules regarding the consequences of reward

and punishment. Getting a pigeon (or a human) to respond by reinforcing a response with a desirable incentive is a very effective method for behavior change. It is a proven technique in many clinical studies and has been used to investigate a variety of problems. Yet it is known that Skinner hated theories. On the other hand, there have been few (if any) real-life applications or consequences resulting from Drive or Expectancy/ Value theory, and the same lack of applied value can be concluded about Lewin's formal theory of motivation.

I must admit that as a motivational theorist, I have had little interest in practical application. I played a game of chess and struggled with the creation of a viable theory – parsimonious, generalizable, empirically supported. Nevertheless, it is very satisfying when attribution theory is successfully applied to a problem area, particularly when the theoretical extension had not been foreseen.

A number of applications are ongoing and promising. It must be said, however, that the utilization potential far exceeds the applications in progress. In the current theoretically originated programs, the focus is on altering dysfunctional causal attributions to those that are more adaptive. The underlying principle guiding utilization of the theory is that if there is a change in perceived causality, then there also will be a change in subsequent behavior. This rule has been applied to achievement strivings; health maintenance among the elderly; treatment of the mentally ill and other stigmatized groups; help-giving; aggression; and even to behaviors in the airport, the hospital, and the classroom, thereby capturing the vast range of attribution theory.

Corresponding to the intra- and interpersonal theories, the applications will be grouped according to their goal of changing self-perception versus altering ascriptions about others. There is a wonderful paradox in the goals of these programs; when altering self-perceptions, the typical aim is to shift the cause of nonattainment of a goal so it no

longer is perceived as uncontrollable. For example, in achievement settings, the aim is to shift attributions for failure from lack of ability to lack of effort. On the other hand, when changing other perceptions, the desire often is to have the cause no longer be perceived as controllable. For example, when considering stigmas the aim is to reduce the blaming of the victim.

Self-perception: pupils regarding classroom performance and the elderly regarding health

There are demotivating (dysfunctional) and motivating (adaptive) causal beliefs, particularly in regard to achievement failure. As already extensively discussed, perhaps the least adaptive ascription to failure is lack of aptitude. On the other hand, lack of effort as the inferred cause of failure is perhaps most adaptive. The contrasting consequences of ability versus effort ascriptions for failure have given rise to a variety of attribution change programs that attempt to shift attributions for failure from low ability to insufficient effort; that is, from stable and uncontrollable to unstable and controllable.

These programs have varied formats, but the technique with the widest usage shows video clips of students discussing their prior failures and communicating the realization that their failure was not due to low ability but rather was caused by insufficient effort or poor strategy. The filmed students also state that this causal shift greatly improved their school performance. Perry et al. (1993) document that this intervention, along with related information, has lasting positive consequences on grade point average in a college setting. This effect has been found in a large number of research studies (see Van Overwalle et al., 1989).

In contrast to shifting causal beliefs to the internal, controllable causes of effort and strategy, some treatments have followed a different path, attempting to alter attributions

for poor performance among freshmen to the initially harsh grading policies of the college, which are described as becoming more lenient over school years. That is, manipulating the perceived cause of failure from stable to unstable is the goal of the experimental treatment, which theoretically increases expectancy of success and hope without asking the students to alter their behavior. Positive treatment effects also have been reported from this intervention (see Wilson et al., 2002).

Health maintenance

Achievement is typically associated with school or job performance. Yet success and failure have a variety of indicators and meanings. Another application of attribution change programs targets health maintenance and walking among the elderly. It is known that exercise has great benefits for all, and particularly for the elderly. Yet the aged often do not engage in physical activities, attributing their "failure" (inactivity) to old age, which is seen as internal, stable, and uncontrollable. This attribution has the same conceptual properties and dysfunctional consequences as an attribution for failure in the classroom to lack of aptitude. Programs thus have been initiated to persuade the elderly to ascribe their inactivity to insufficient effort or "not trying" rather than to their age. A variety of techniques are used to accomplish this change, particularly via information directly transmitted by health professionals (see Sarkisian et al., 2007). This intervention has proven effective in increasing exercise and walking, extremely important changes for health maintenance. However, the research supporting this claim is yet minimal.

In sum, attribution change programs enhance student performance in the classroom and perhaps elderly "performance" regarding exercise. The attribution intervention techniques are simple, require little cost, and promise great benefit. This is an important direction of application that needs more attention.

Teacher-training programs

It has been documented that some prevalent teacher practices result in students making low ability attributions for their poor performance, the most harmful of the attributions for failure (see Graham, 1990). Such practices include not punishing and conveying sympathy given failure at an easy task, overly praising success at an easy task, and providing help when it is not sought. When reprimand for failure at an easy task is withheld, the attribution for the poor performance is perceived as something other than a lack of effort. Furthermore, the conveyance of sympathy provides a cue that the failure was uncontrollable. Given these communications, the student is likely to conclude that the teacher believes his or her failure was caused by an absence of ability. This causal communication then increases the pupil's own belief that he or she is unable. In a similar manner, pre-emptive help-giving is a cue that the pupil "cannot," while praise for success at an easy task conveys the cause of this outcome was high effort (which implicates a low level of ability).

In sum, there are a variety of apparent teacher practices in need of alteration from an attribution perspective. Communication of low ability is not the intent of these teacher behaviors and the actions appear "kind" (e.g., withholding blame, giving praise, and providing unsought help). These principles and other extrapolations from attribution theory presently are being incorporated into some teacher training programs to hopefully alter teacher behavior and, in turn, the potentially harmful low ability self-perceptions of their pupils. But the effects of this inclusion remain to be systematically examined and the research is scanty.

Other perception

As already noted, rather than altering causal ascriptions from uncontrollable and stable (ability) to controllable and unstable (effort), thereby affecting expectancy and affect, the

goals of programs regarding other perception often are to shift attributions (e.g., for illness) from controllable (e.g., lifestyle) to uncontrollable (e.g., genetics). Theoretically, this change reduces perceptions of responsibility and anger, increases sympathy, and eliminates observer antisocial behavior (or promotes prosocial behavior). Many populations have been the targets of this type of attribution change in research studies. Again, however, there has not been a sufficient number of programs that apply this knowledge to reach any scientific closure. The rather sparse change attempts are the focus of the following discussion.

The mentally ill

In the current climate of the growth of neuropsychology, biological models of major mental disorders such as schizophrenia and depression predominate, along with drug treatments. Yet the majority of the public may still perceive the mentally ill as morally weak and responsible for their problems (see Neff and Husaini, 1985). This is reflected in critical and hostile comments (labeled expressed emotion, or EE) directed by family members toward the victims of these mental disorders. Furthermore, it is well established that "living in a high EE-home environment more than doubles the baseline relapse rate for schizophrenic patients" (Barrowclough and Hooley, 2003: 849).

Inasmuch as EE is elicited by controllable beliefs regarding the illness of others and harms patients by promoting relapse, it then follows that programs altering ascriptions of mental illness from controllable to uncontrollable should reduce recidivism. From the viewpoint of attribution theory, this intervention reduces inferred responsibility and the expression of anger, while increasing sympathy, warmth, and prosocial actions. Some interventions have been undertaken to convince caretakers of the uncontrollability of the problems of their mentally ill family members (see Medvene and Krauss, 1989), and it has been reported that decreasing

EE does lessen relapse (see Pitschel-Walz et al., 2001). What is needed is a rash of such intervention programs.

Other stigmas

Those with major mental disorders are not the only stigmatized individuals who additionally suffer because they are regarded as moral failures. Across a variety of stigmas, the most defining characteristic that promotes rejection and social distancing is the perception of personal responsibility (Feldman and Crandall, 2007). Many illnesses are believed to be caused by lifestyle choices that are amenable to personal control. For example, those with lung cancer are blamed because of the association of this disease to smoking (see Cooper, 1984; Kim and Shanahan, 2003), which produces family conflict. And many stigmatized are not given opportunities to gain jobs and rent apartments because they are perceived as "immoral" (see Corrigan et al., 2003). The antisocial reactions toward those perceived as responsible for their stigmas has even been observed in the actions of health professionals, including less favorable medicine-dispensing decisions among doctors (see Brewin, 1984; Brewin and Antaki, 1982) and reduced caring behavior of nurses (Harborne, 1996). Unfortunately, intervention attempts targeting these punitive attribution-driven reactions have not been reported.

Aggression

As documented earlier in this chapter, aggression is one response following an inference of other-person responsibility for a negative outcome and the anger that this elicits. Hence, one method to reduce aggression is to lessen perceptions of other-person responsibility for negatively valenced events.

In a business-related utilization of the principle, Folkes (1984) demonstrated that consumer reactions to product failure are linked with the perceived cause of their plight. Often, for example, flight delays are perceived by passengers as caused by

company shortcomings, which generates anger and consumer complaints. Hence, company adoption of simple communication principles, such as announcing the flight delay is due to weather problems (an excuse shifting attributions from controllable to uncontrollable) will reduce anger and passenger complaints (one indicator of aggressive retaliation).

Evidence of these attribution → affect → behavior relations also is displayed in more serious aggressive acts. For example, there are some data that abusive mothers tend to overattribute controllability to their children regarding acts such as the accidental spilling of milk, which then causes anger and abusive actions (see Bradley and Peters, 1991). Treatment programs for abusive parents often focus on anger management, but an attributional perspective suggests that the cognitive antecedents of anger also should be targeted.

An overattribution of intention and responsibility for negative actions also has been documented among aggressive adolescents (see Dodge and Crick, 1990). Guided by this finding, cognitive intervention programs have been designed to reduce intentionality inferences ("he stepped on my toe on purpose") among aggressive-prone children (see Graham and Hudley, 1992). Part of the program includes training to discriminate between intentionally caused versus unintentional negative actions in situations of causal ambiguity. Such training programs do decrease subsequent aggression. In addition, Graham et al. (1995) report that aggressive children benefit from learning the adaptive value of providing excuses so that their own negative behavior is not perceived by others as controllable or intentional. As has been true throughout this discussion, it is unfortunate that the body of work targeted to reduce aggression is not sufficient to draw conclusions about the "success" of attribution interventions. On the positive side, the available research is certainly reasonable and encouraging.

SOME CONCLUDING THOUGHTS

Carved on the National Archives Building in Washington, DC is a saying from William Shakespeare's *The Tempest*: "What's past is prologue." For an attribution theorist, this means that the interpretation of the past, and particularly the perceived causes of prior events, determines what will be undertaken in the future. I have attempted to create conceptual structures to capture this truism. The defining characteristics of these structures include testable hypotheses, generalization to a wide range of behavioral domains and both intrapersonal and interpersonal actions, parsimony, incorporation of thoughts and emotions, and potential for application. These qualities make me happy (outcome-dependent), grateful (attribution of success to help from others), and proud (attribution of success to a great deal of effort). Thus, the theory accurately predicts the feelings and actions of at least one of its creators!

REFERENCES

Atkinson, J.W. (1957) Motivational determinants of risk-taking behavior. *Psychological Review*, 64, 359-372.

Atkinson, J.W. and Birch, D. (1970) *The Dynamics of Action*. New York: Wiley.

Barrowclough, C. and Hooley, J.M. (2003) Attributions and expressed emotion: A review. *Clinical Psychology Review*, 23, 849-880.

Bradley, E. and Peters, R. (1991) Physically abusive and nonabusive mothers' perceptions of parenting and child behavior. *American Journal of Orthopsychiatry*, 61, 455-460.

Brewin, C.R. (1984) Perceived controllability of life-events and willingness to prescribe psychotropic drugs. *British Journal of Social Psychology*, 23, 285-287.

Brewin, C.R. and Antaki, C. (eds) (1982) *Attributions and Psychological Change*. London: Academic Press.

Cooper, E.T. (1984) A pilot study of the effects of the diagnosis of lung cancer on family relationships. *Cancer Nursing*, 7, 301-308.

Corrigan, P.W., Markowitz, F.E., Watson, A.C., Rowan, D. and Kubiak, M.A. (2003) An attribution model of public discriminations towards persons with mental illness. *Journal of Health and Social Behavior*, 44, 162–179.

de Charms, R. (1968) *Personal Causation*. New York: Academic Press.

Dodge, K.A. and Crick, N. (1990) Social information-processing biases of aggressive behavior in children. *Personality and Social Psychology Bulletin*, 16, 8–22.

Feldman, D.B. and Crandall, C.S. (2007) Dimensions of mental illness stigma: What about mental illness causes social rejection? *Journal of Social and Clinical Psychology*, 26, 137–154.

Folkes, V.S. (1982) Communicating the causes of social rejection. *Journal of Experimental Social Psychology*, 18, 235–252.

Folkes, V.S. (1984) Consumer reactions to product failure: An attributional approach. *Journal of Consumer Research*, 10, 398–409.

Gendolla, G.H.E. and Koller, M. (2001) Surprise and causal search: How are they affected by outcome valence and importance? *Motivation and Emotion*, 24, 237–250.

Graham, S. (1990) On communicating low ability in the classroom: Bad things good teachers sometimes do. In S. Graham and V.S. Folkes (eds), *Attribution Theory: Applications to Achievement, Mental Health, and Interpersonal Conflict*, pp. 17–36. Hillsdale, NJ: Lawrence Erlbaum.

Graham, S. and Hudley, C. (1992) An attributional approach to aggression in African-American children. In D. Schunk, and J. Meece (eds), *Social Perceptions in the Classroom: Causes and Consequences*, pp. 75–94. Hillsdale, NJ: Erlbaum Press.

Graham, S., Weiner, B. and Benesh-Weiner, M. (1995) An attributional analysis of the development of excuse giving in aggressive and nonaggressive African American boys. *Developmental Psychology*, 31, 274–284.

Harborne, A. (1996) Challenging behavior in older people: nurses' attitudes. *Nursing Standard*, 11, 39–43.

Heider, F. (1958) *The Psychology of Interpersonal Relationships*. New York: Wiley.

Hull, C.L. (1943) *Principles of Behavior*. New York: Appleton-Century-Croft.

Kelley, H.H. (1967) Attribution theory in social psychology. In D. Levine (ed.), *Nebraska Symposium on Motivation*, pp. 192–238. Lincoln: University of Nebraska Press.

Kelley, H.H. (1972) Causal schemata and the attribution process. In E.E. Jones, D.E. Kanouse, H.H. Kelley, R.E. Nisbett, S. Valins, and B. Weiner (eds), *Attribution: Perceiving the Causes of Behavior*, pp. 151–174. Morristown, NJ: General Learning Press.

Kim, S. and Shanahan, J. (2003) Stigmatizing smokers: Public sentiment toward cigarette smoking and its relationship to smoking behaviors. *Journal of Health Communications*, 8, 343–367.

Lewin, K. (1938) *The Conceptual Representation and the Measurement of Psychological Forces*. Durham, NC: Duke University Press.

Locke, E. and Latham, G. (1990) *A Theory of Goal Setting and Task Performance*. Englewood Cliffs, NJ: Prentice-Hall.

Medvene, L.J. and Krauss, D.H. (1989) Causal attributions and parent-child relationships in a self-help group for families of the mentally ill. *Journal of Applied Social Psychology*, 19, 1413–1430.

Neff, J.A. and Husaini, B.A. (1985) Lay images of mental health: Social knowledge and tolerance of the mentally ill. *Journal of Community Psychology*, 13, 3–12.

Perry, R.P., Hechter, F.J., Menec, V.H. and Weinberg, L.E. (1993) Enhancing achievement motivation and performance in college students: An attributional retraining perspective. *Research in Higher Education*, 34, 687–723.

Pitschel-Walz, G., Leucht, S., Bauml, J., Kissling, W. and Engel, R.R. (2001) The effect of family interventions on relapse and rehospitalization in schizophrenia – meta-analysis. *Schizophrenia Bulletin*, 27, 73–92.

Rodrigues, A. (1995) Attribution and social influence. *Journal of Applied Social Psychology*, 25, 1567–1577.

Rodrigues, A. and Lloyd, K.L. (1998) Reexamining bases of power from an attributional perspective. *Journal of Applied Social Psychology*, 28, 973–997.

Rotter, J.B. (1954) *Social Learning and Clinical Psychology*. Englewood Cliffs, NJ: Prentice-Hall.

Rotter, J.B. (1966) Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs*, 80, 1–28.

Rudolph, U., Roesch, S.C., Greitemeyer, T. and Weiner, B. (2004) A meta-analytic review of help giving and aggression from an attributional perspective. *Cognition and Emotion*, 18, 815–848.

Sarkisian, C.A., Prohaska, T.R., Davis, C. and Weiner, B. (2007) Pilot test of an attributional retraining intervention to raise walking levels in sedentary older adults. *Journal of the American Geriatrics Society*, 55, 1842–1846.

Schmidt, G. and Weiner, B. (1988) An attribution-affect-action theory of behavior: Replications of judgments of help giving. *Personality and Social Psychology Bulletin*, 14, 610–621.

Schmitt, D.R. (1964) The invocation of moral obligation. *Sociometry*, 27, 299–310.

Spence, K.W. (1956) *Behavior Theory and Conditioning*. New Haven, CT: Yale University Press.

Thorndike, E.L. (1911) *Animal Intelligence*. New York: Macmillan.

Tolman, E.C. (1932) *Purposive Behavior in Animals and Men*. New York: Appleton-Century-Croft.

Van Overwalle, F., Segebarth, K. and Goldchstein, M. (1989) Improving performance of freshmen through attributional testimonies from follow students. *British Journal of Educational Psychology*, 59, 75–85.

Weiner, B. (1980) A cognitive (attribution)–emotion–action model of motivated behavior: An analysis of judgments of help giving. *Journal of Personality and Social Psychology*, 39, 186–200.

Weiner, B. (1985) An attribution theory of achievement motivation and emotion. *Psychological Review*, 92, 548–573.

Weiner, B. (1986) *An Attributional Theory of Motivation and Emotion*. New York: Springer-Verlag.

Weiner, B. (1992) *Human Motivation: Metaphors, Theory, and Research*. Newbury Park, CA: Sage.

Weiner, B. (1995) *Judgments of Responsibility: A Foundation for a Theory of Social Conduct*. New York: Guilford.

Weiner, B. (2006) *Social Motivation, Justice, and the Moral Emotions*. Mahwah, NJ: Erlbaum Press.

Weiner, B. (2007) Examining emotional diversity in the classroom: An attribution theorist considers the moral emotions. In P.A. Schutz and R. Pekrun (eds), *Emotion in Education*, pp. 75–88. San Diego: Academic Press.

Weiner, B., Frieze, I.H., Kukla, A., Reed, L., Rest, S. and Rosenbaum, R.M. (1971) *Perceiving the Causes of Success and Failure*. Morristown, NJ: General Learning Press.

Weiner, B., Graham, S. and Chandler, C.C. (1982) Pity, anger, and guilt: An attributional analysis. *Personality and Social Psychology Bulletin*, 8, 226–232.

Weiner, B. and Kukla, A. (1970) An attributional analysis of achievement motivation. *Journal of Personality and Social Psychology*, 15, 1–20.

Weiner, B., Nierenberg, R. and Goldstein, M. (1976) Social learning (locus of control) versus attributional (causal stability) interpretations of expectancy of success. *Journal of Personality*, 44, 52–68.

Weiner, B., Perry, R.P. and Magnusson, J. (1988) An attributional analysis of reactions to stigmas. *Journal of Personality and Social Psychology*, 55, 738–748.

Weiner, B. and Potepan, P.A. (1970) Personality correlates and affective reactions toward exams of succeeding and failing college students. *Journal of Educational Psychology*, 61, 144–151.

Weiner, B., Russell, D. and Lerman, D. (1978) Affective consequences of causal ascriptions. In J.H. Harvey, W.J. Ickes and R.F. Kidd (eds), *New Directions in Attribution Research*, 2, 59–88. Hillsdale, NJ: Lawrence Erlbaum Press.

Weiner, B., Russell, D. and Lerman, D. (1979) The cognition-emotion process in achievement-related contexts. *Journal of Personality and Social Psychology*, 37, 1211–1220.

Wilson, T.D., Damiani, M. and Shelton, N. (2002) Improving the academic performance of college students with brief attributional interventions. In J. Aronson (ed.), *Improving Academic Achievement: Impact of Psychological Factors on Education*, pp. 89–108. San Diego: Academic Press.

A Theory of Social Information Processing

Robert S. Wyer, Jr.

ABSTRACT

A general theory of social information processing was developed 25 years ago in an attempt to integrate diverse phenomena that I and others had identified in our research on social cognition. After several iterations, the model continues to provide a useful theoretical framework for conceptualizing and integrating the deliberate and automatic processes that occur at different stages of cognitive functioning, including comprehension, memory storage and retrieval, inference, judgment, decision making, and output generation. More specific formulations of belief organization and change, person impression formation, dual processing, and the interplay of goal-directed and unconscious processing can be viewed in terms of the cognitive operations that are governed by the general model I proposed.

INTRODUCTION

My research during much of the past 25 years has been guided by a general theoretical formulation of social information

processing. The model specifies several stages of processing en route to a judgment or behavioral decision (comprehension, organization in memory, inference, integration, and the generation of an overt response). The processing at these stages can be either deliberative or automatic. More specific theoretical formulations of comprehension, inference, judgment, and decision making can often be conceptualized in terms of the cognitive operations that occur at different stages of processing and the factors that influence their activation and use.

The theory has gone through several iterations (Wyer and Srull, 1980, 1986, 1989) en route to its present version (Wyer, 2004). In this chapter, I summarize the essential features of the model and discuss its potential for integrating the implications of other, more circumscribed formulations of information processing. The question is what motivated me to develop such a grandiose and all-encompassing model. The answer has its roots in my graduate work nearly 50 years ago.

A PERSONAL HISTORY

Initial influences

Having been discouraged by my father from careers as a sportswriter and jazz musician and not knowing what else to do, I wound up going to an engineering school. However, I felt intellectually stifled by my first job at Bell Laboratories and wanted to get a liberal arts education. I saw psychology as an indirect means of attaining this objective. When I applied to graduate school in psychology at the University of Colorado, I knew nothing at all about the field.

Research at Colorado during the early 1960s, largely stimulated by O.J. Harvey and Bill Scott, was focused on global structural characteristics of the cognitive system that might influence responses to social experience. During a summer job at Hughes Aircraft Company, I developed a model of artificial intelligence that provided measures of these constructs, and in my dissertation I attempted to relate the measures to various indices of personality and behavior. The research was actually published. However, my first job was at the University of Iowa. The Hull-Spence learning tradition that pervaded the atmosphere at Iowa during that period, coupled with some blistering reviews of some articles I had submitted for publication, convinced me that the approach I had been taking to understanding cognitive structure and function was a dead end.

The predominant view at Iowa during those early years was that thinking was an epiphenomenon of little relevance to a scientific investigation of human behavior. This view temporarily dampened my enthusiasm for studying cognitive processes. After leaving Iowa for the University of Illinois at Chicago Circle, however, four quite unrelated experiences rekindled my interest. First, I read a paper by Fishbein and Hunter (1964) on additivity and averaging processes in impression formation. Although I was unimpressed by this particular paper, I believed that the authors' theory was correct

and conducted an even less impressive, never-to-be-published study to establish this fact. However, I sent the results of the study to Norman Anderson for comments and, to my surprise, he was quite encouraging. Stimulated by this reinforcement, I submitted a grant proposal to the National Science Foundation that was probably reviewed by Anderson and, therefore, was approved for funding. This resulted in a number of studies in the Anderson tradition which attempted to understand how people integrated the implications of personality traits into an overall person impression.

Second, a graduate course I was teaching stimulated me to read about W.J. McGuire's (1960) syllogistic model of belief organization. I realized that a slight modification of this model was consistent with the more general assumption that beliefs might be organized in memory in a manner that conformed to the laws of mathematical probability. Furthermore, if this were so, quantitative predictions could be made of the effects of change in one belief on other, unmentioned beliefs that were theoretically related to it. A flurry of experiments (for summaries, see Wyer and Hartwick, 1980; Wyer and Srull, 1989) evaluated implications of this possibility and of McGuire's formulation more generally.

Third, I happened to have an office next to Harry Upshaw. I became enamored of his conceptualization of how the perspective that people brought to bear on the stimuli they were judging affected their subjective positioning of the response scale they used to report their judgments (Upshaw, 1965). This conceptualization, which had implications for the effects of one's own attitude on responses to attitude-related messages, led to research in this area as well (Wyer, 1969).

Finally, I was asked to review the *Handbook of Personality Theory and Research*, which included another awe-inspiring chapter by McGuire (1968). In this chapter, McGuire conceptualized the different stages of processing that underlie responses to a persuasive message and how the processing at each

stage combined to influence the message's impact. This chapter led me to reflect upon my own work and to realize that, in fact, I had done work bearing on four different phases of processing: the organization of cognitions in memory, the use of previously formed cognitions to make an inference, information integration processes, and output (response) processes. I had simply not thought of them together. The outcome was my first book (Wyer, 1974). This book argued that human information processors were analogous to electronic processors (computers). Thus, they could simply be understood in terms of the "programs" they used to attain objectives specified by the user (others to whom they communicated in the social environment). My approach was rather naïve. However, it set the stage for my general interest in several different phases of information processing and how they might fit together.¹

The advent of social cognition

Although the aforementioned experiences contained the seeds of my general interest in social information processing, it was not until the mid-1970s that this interest began to bear fruit. The conditions that stimulated the information processing model that I ultimately developed were quite unexpected and somewhat ironic.

In 1974, Norman Anderson and Seymour Rosenberg held a three-week workshop at the University of California, San Diego. Its purpose was to bring together research and theory on cognitive structure (exemplified by multidimensional scaling models) and process (represented by Anderson's formulations of functional measurement and information integration; see Anderson, 1971). The workshop was attended by several researchers, notably Tom Ostrom, Dave Hamilton and myself, who had been conducting impression formation research within the framework that Anderson proposed, but were frustrated

by our failure to come to grips with the processes we intuitively believed to be involved in the formation of person impressions outside the laboratory. Reid Hastie and Ebbe Ebbesen, who had quite different perspectives on impression formation processes, were also participants.

Although we had all been reviewing one another's work for journals, few of us had ever met. We nevertheless found ourselves involved in daily discussions, often extending into the early hours of the morning, which culminated in our agreement that the research we had been doing was going nowhere. To make progress, we would need to understand how complex bodies of information were organized and stored in memory and how this information was later retrieved and operated upon in order to make a judgment.

To this end, the five of us, eventually joined by Don Carlston, began meeting informally twice a year over a long weekend, exchanging research findings and their implications.² Several findings, although now widely accepted, were nonintuitive at the time. For example, people were better able to recall someone's behaviors if they had received them for the purpose of forming an impression of the person than if they had been told explicitly to learn and remember the behaviors (Hamilton et al., 1980). People who had formed a trait-based impression of a person had better recall of the person's behaviors that were inconsistent with this impression than behaviors that confirmed its validity (Hastie and Kumar, 1979). This was not the case, however, when people formed impressions on the basis of observed behavior (Cohen and Ebbesen, 1979). Finally, once people had made one judgment of a person on the basis of trait descriptions, subsequent judgments of him were based on their first judgments regardless of the implications of the original information they had received (Lingle and Ostrom, 1979). Furthermore, the effects of initial judgments on later ones *increased* with the time interval between the two judgments (Carlston, 1980; Srull and

Wyer, 1980). These and other findings were sufficiently provocative that we decided to publish them in a book on person memory (Hastie et al., 1980).

The birth of a theory

I was assigned to write an integrative chapter for the Hastie et al. volume that would bring others' diverging findings together within a single conceptual framework. I embarked on this project with Thom Srull, a graduate student at the time. The challenge was rather daunting, as it required an integration of not only phenomena such as those mentioned earlier but also findings that had begun to emerge on the impact of cognitive heuristics (e.g., Nisbett and Ross, 1980), knowledge accessibility (Higgins et al., 1977), and the earlier work we had done within the frameworks proposed by Anderson and McGuire. To account for Hamilton et al.'s (1980) findings, for example, we needed to take into account the effects of different processing goals on the way that information was organized in memory and was subsequently retrieved. To account for the effects of making an initial judgment on later ones (Carlston, 1980), we needed to develop a conceptualization of memory and judgment that would specify when and why one type of information took priority over another. We had to account for more specific encoding and organizational processes of the sort identified by Hastie and Kumar (1979). Finally, we needed to take into account the process of encoding ongoing behavior (Cohen and Ebbesen, 1979) as well as verbal information and the use of these nonverbal encodings on memory and judgment.

Our main challenge was to specify a common set of memory storage and retrieval processes that would apply at different stages of processing (comprehension, integration, inference, etc.) and would account for phenomena that were localized at these stages. Although Don Carlston and I had proposed

an associative network model of the representation of social information in memory (Wyer and Carlston, 1979), it failed to capture the sort of findings that Hamilton, Hastie, Ostrom, and others were identifying. At the end of one of our early discussions, Srull incidentally made an observation that proved to be the central feature of our conceptualization: "If anyone is going to develop a viable conceptualization of social memory," he remarked, "they will have to drop the associative network metaphor." Having struggled with the Wyer and Carlston model, I quickly agreed. The result was the development of a "storage bin" metaphor of memory. That is, we conceptualized long-term memory as consisting of a number of content-addressable storage bins, into which different representations of knowledge could be deposited in the order they were formed and used, and from which representations could be retrieved on the basis of a probabilistic top-down search. The "bin" metaphor proved to be a powerful conceptual tool.

With this construct, we developed a metaphorical flow diagram consisting of a number of processing units (each pertaining to a different stage of processing) and memory storage units, all of which were interconnected by pathways that indicated the transmission of information from one unit to another. This was more challenging than we had expected. Any given pathway between two units had implications for the type of information that could be transmitted and the conditions in which it could occur. At the same time, the failure to specify a path placed constraints on system functioning. Several iterations were required to come up with a model that could account for the phenomena we needed to explain without creating problems elsewhere in the processing system. As a consequence, what was originally intended solely as a descriptive device for integrating the phenomena reported elsewhere turned out to be a full-blown theory of social information processing with testable empirical implications.

Further iterations

The 1980 model was updated several times to account for new phenomena and to eliminate ambiguities that existed in its initial formulation (Wyer and Srull, 1986, 1989). After proposing the 1989 version of the model, however, I had a choice between continuing to test specific implications of the formulation and exploring new empirical issues that I found to be intrinsically interesting without considering their relevance to the model we had developed. As Bob Zajonc once observed, theories have many attackers, and unless someone is willing to defend them, they are likely to die a premature death. Often, however, the only person who is motivated to defend a theory is the individual who proposed it.

Nevertheless, I decided to let the model defend itself and to explore new horizons. This led me into research and theorizing on the antecedents of humor elicitation (Wyer and Collins, 1992), the influence of affect in information processing (Wyer et al., 1999), the representation of event sequences (Wyer et al., 2002), and comprehension processes more generally (Wyer and Gruenfeld, 1995; Wyer and Radvansky, 1999). In pursuing this latter line of research, however, it became clear that the original Wyer and Srull (1989) model could not account for the comprehension of information about familiar persons and events without substantially modifying its assumptions. This created a need (in my own mind, if not others') to modify the conceptualization in ways that would provide this account. The "final" (or, at least, most recent) version of the model, reported in Wyer (2004; Wyer and Radvansky, 1999) was the result.³

THE THEORY

The conceptualization of information processing that we developed is multifaceted, and a detailed exposition of it is beyond the scope of this chapter. Although it is capable

of generating many specific predictions (e.g., Wyer and Srull, 1989), its primary value lies in its ability to provide a framework for integrating the phenomena identified in other, more specific areas of social information processing research. We hoped to account for the activities that occur at different stages of cognitive processing (comprehension, organization of information in memory, inference and integration, and the generation of overt responses) and to specify the processes of storing and retrieving information required at each of these stages. We further needed to address (1) the interplay of deliberative and automatic processing, (2) the unconscious influences of information on judgments, (3) the effects of externally imposed and internally generated goals on information processing at different stages, (4) the cognitive activities that occur in the absence of any conscious goal-directed activity (e.g., the free flow of thought), and (5) the theoretical relation between memory and judgment.

System architecture

As shown schematically in Figure 8.1, the model is composed of three major memory units, four special purchase processing units, and an Executor that directs the flow of information between these units. The memory units include the *Work Space* (analogous to working memory), the *Permanent Storage Unit* (long term memory), and a *Goal Specification Box*. The latter unit is a temporary repository of goal schemas, or sequences of cognitive steps that are involved in the pursuit of a particular objective.

The processing units include a *Comprehender*, which is an initial comprehension device that interprets input information in terms of general verb, noun, and adjective concepts, an *Encoder/Organizer*, which performs high-order interpretations of information and forms mental representations consisting of several features, an *Inference Maker*, which combines the implications of information to form a subjective

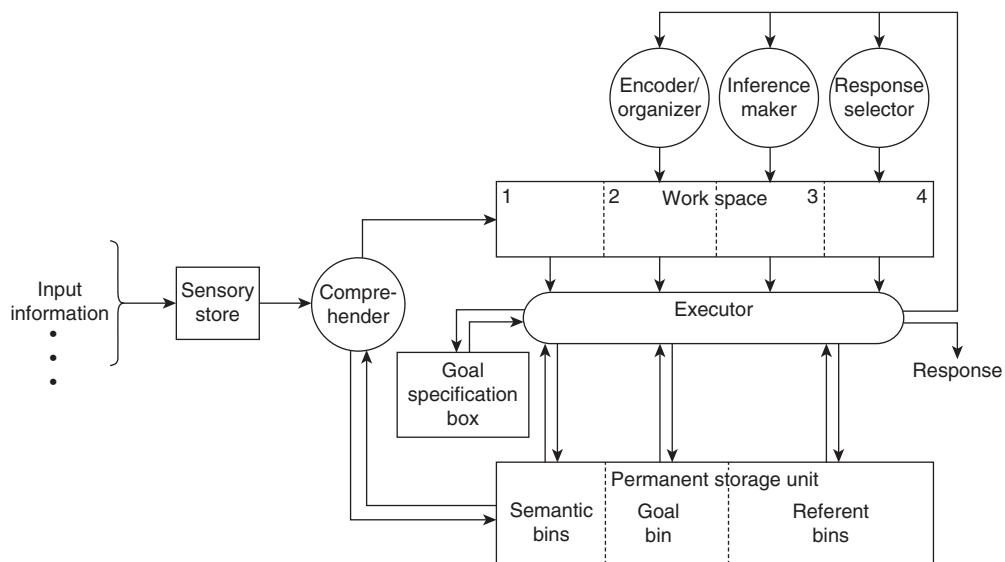


Figure 8.1 Metaphorical representation of the Wyer and Srull (1989) model. Rectangles denote storage units. Ovals and circles denote processing units. Arrows denote the direction of transmission of cognitive material between these units

judgment, and a *Response Selector*, which transforms subjective inferences into an overt response. The *Comprehender* comes into play in the processing of all of the information one receives and performs its functions automatically. The other units come into play only in the course of more specific, goal-directed processing. The activation of these latter units is governed by an *Executor*, which directs the flow of information between processing units and storage units. In doing so, it takes instructions from a goal schema that specifies the sequence of operations to be performed and the type of information required.

Systems operation

The general operation of the processing system can be described briefly.

- 1 Input information is initially encoded by the *Comprehender* in terms of semantic concepts that are retrieved from Permanent Storage.

The results of this processing are then transmitted to the Work Space.

- 2 If the new information contains a goal specification, the *Executor* retrieves a relevant goal schema from Permanent Storage and deposits it in the *Goal Specification Box*. Based on the steps specified in the schema, it retrieves features of the input information along with previously acquired concepts and knowledge retrieved from Permanent Storage, and transmits them to a relevant special-purpose processing unit with an instruction to negotiate the step in question. The results of this processing are then returned to the Work Space.
- 3 The *Executor* again consults the relevant goal schema and, if no further processing is required, transmits the output of the initial processing to an additional processing unit along with the necessary goal-relevant knowledge required to perform its function. This continues until either the sequence of steps required to attain the objective is completed or other processing demands interfere.
- 4 *Simultaneous goal-directed processing.* More than one goal schema can be contained in the *Goal Specification Box* at any one time. This means that more than one goal can be

pursued simultaneously. However, its capacity is limited. Consequently, when the schema that is required to attain a particular objective is complex and detailed, there is less room for other schemas and so fewer other goals can be simultaneously pursued.

5 *Free flow of thought.* If no goal schema is activated, the system enters into a default routine. Specifically, the Executor samples a subset of features that are present in the Work Space and uses them as retrieval cues to identify a previously acquired knowledge representation. It then repeats this process, sampling a second set of features (which may include those of the representation just retrieved), and continuing until a goal specification is identified (based either on external input information or material retrieved from memory), in which case steps 1 and 2 are performed. Thus, this activity accounts for the free flow of thought that occurs in the absence of any specific goal-directed activity.

One other aspect of the model is important. That is, *consciousness resides in the Executor*. The Executor has access to (1) the information transmitted by the Comprehender, (2) previously acquired knowledge that is retrieved from Permanent Storage, (3) the steps involved in attaining a particular goal that are contained in a goal schema, and (4) the output of special-purpose processing units. Consequently, this material is subject to conscious awareness. However, the activities of the Comprehender and other processing units, which are governed by routines that are stored in the libraries of these units, are *not* subject to awareness. Furthermore, note that the Comprehender, unlike other processing units, is not under the control of the Executor (see Figure 8.1). Thus, the system is not only unaware of the cognitive processes that underlie the comprehension of new information, but the processes are themselves not under conscious control. Thus, for example, it is impossible *not* to comprehend the statement “the boy kicked the ball” in terms of verb and noun concepts that are retrieved from permanent memory.⁴

Similarly, the actions performed by special-purpose processing units are not subject

to awareness. Only the more general sequences of cognitive steps contained in a goal schema are conscious. Thus, suppose a goal is to form an impression of someone on the basis of information about the person’s behaviors. The goal schema retrieved for use in attaining this objective might contain instructions to interpret the behaviors in terms of more general trait concepts and to form an evaluative concept of the individual on the basis of the behaviors’ evaluative implications. These trait encodings and the overall evaluation, which are outputs of these processes, are subject to awareness. However, the specific cognitive mechanisms that are necessary to accomplish these steps are specified in the libraries of the processing units involved and are not subject to awareness.

Storage and retrieval processes

Three memory units are postulated. One, the Goal Specification Box, has already been discussed. The Work Space and Permanent Storage Unit require some elaboration.

Work Space

This unit is a temporary store of (1) information that has recently been transmitted to it by the Comprehender, (2) previously acquired concepts and knowledge that have been used by special-purpose processing units to attain specific processing objectives, and (3) the results of processing by these units. Information is retained in the Work Space as long as it is likely to be involved in attaining goals that exist at the time. However, if an objective has been attained, or if information has not been employed for some time, the Work Space may be cleared to facilitate the identification of material that is of more immediate relevance. In this case, features of the original input information that have not been involved in higher order goal-directed processing (and, therefore, are not contained in a knowledge representation

that is transmitted to Permanent Storage) are irretrievably lost.

Permanent Storage Unit

In the 1989 version of the model, permanent storage was assumed to consist of a number of content-addressable "storage bins." One, *semantic* bin, contains general semantic concepts that are not specific to any particular referent and are drawn upon by the Comprehender in interpreting input information at an early stage of processing. A second, *goal* bin contains goal-schemas that are retrieved by the Executor for use in guiding goal-directed activity. The remaining, *referent* bins contain material pertaining to specific referents of goal-directed processing (persons, objects, events, or situations). Each referent bin is assigned a header, or set of verbal and nonverbal features that identify its referent. These features can include the name of the referent, a visual image, and attributes that have become strongly associated with the referent through learning and, can serve to identify it. A bin's referent can be at different levels of generality (e.g., George W. Bush, incompetent US Presidents, US Presidents, etc.). The contents of a bin can also vary. Thus, it can include single trait concepts that have been assigned to the referent, a verbally or nonverbally coded sequence of events that could describe either a prototypic experience (e.g., eating at a restaurant) or a specific one, a cluster of traits or interrelated behaviors, and a prior judgment or evaluation. Each such representation constitutes a separate unit of knowledge that is stored independently of other representations.

Five principles govern the storage and retrieval of these representations.

- 1 Units of knowledge are stored in a bin that refers to the stimulus to which processing objectives pertain. They are stored in the order they are formed, with the most recently deposited unit on top.
- 2 When information relevant to a referent is sought, the Work Space is the first location searched.
- 3 If no relevant information is found in the Work Space, a referent bin is identified on the basis

of a set of probe cues that specify the nature of the referent and its header is reviewed for features that may be sufficient to attain the objective at hand. If these features are found, a search of the bin is not performed. (This means that if a feature has become strongly associated with a referent, it may be used as a basis for judgment independently of any more specific information whose implications might contradict it.)

- 4 If the features of a bin header are not sufficient to attain the objective at hand, a probabilistic top-down search is performed for a representation that is potentially relevant, and a copy of this representation is retrieved. Once it has been used, the copy is returned to the top of the bin. The probability of identifying a given unit of information in the top-down search of the bin is less than one. This means that the likelihood of retrieving a particular representation in the course of searching a bin is a function of both the recency of its acquisition and use (which determines its proximity to the top of the bin) and the frequency of its use (which determines the number of copies of it that have been formed and deposited there).
- 5 As noted earlier, information is deposited in a bin that is relevant to the processing objectives that exist at the time. Thus, if information is received that John kissed Mary and a goal is to form an impression of Mary, the information would be stored in a bin pertaining to her but not to a bin pertaining to John. This means that the information is unlikely to be retrieved later if information about John is required.

These retrieval processes are theoretically invoked at all stages of processing. Thus, they govern the concepts that are retrieved for use in both interpreting new information at the time it is received and later, in the pursuit of more specific objectives (e.g., to interpret a behavior in terms of a trait concept, to infer the likelihood of an event, to report an attitude toward a person or object, etc.). Furthermore, when alternative procedures can be used to attain a certain objective, these same processes govern the goal schema that is retrieved from the goal bin and applied.

The aforementioned principles have numerous implications (Wyer and Srull,

1989), one of which may be worth mention. Note that all input information that is interpreted by the Comprehender is transmitted to the Work Space regardless of its relevance to any specific processing objective that might exist. However, only features of this information that are involved in goal-directed processing are transferred to Permanent Storage. Because the Work Space is not cleared immediately, features of the input information that are irrelevant to the goal for which the information was initially used can potentially be identified and brought to bear on other judgments and decisions that are made a short time after the information has been received. After a period of time has elapsed, however, only the results of the original goal-directed processing, which have been transmitted to Permanent Storage, can be used. This means that judgments and decisions that are made at the time information is first received can influence later judgments independently of the implications of this information. Furthermore, this influence increases over time.

Several studies support these predictions (e.g., Carlston, 1980; Lingle and Ostrom, 1979; Srull and Wyer, 1980, 1983). Carlston (1980), for example, found that people who had judged a person's honesty on the basis of a behavior that were both honest and unkind (e.g., turning in a friend for cheating on an exam) later judged him as relatively kind, whereas people who had initially judged his kindness later judged him as relatively dishonest. Moreover, this difference was greater three days after the first judgment was made than it was immediately afterwards. Research by Srull and Wyer (1980) has similar implications.

The role of procedural knowledge

The processes described in the preceding section govern the storage and retrieval of *declarative* knowledge that individuals acquire. This knowledge presumably includes

the goal schemas that are formed for the purpose of attaining a particular objective. These schemas, which presumably consist of a sequence of events similar to those that compose a script, function as mental "recipes" that people intentionally consult for information about the steps required in order to attain a goal.

However, the theory distinguishes between these knowledge representations and the routines that compose the libraries of processing units. These routines, which constitute *procedural* knowledge, are acquired through learning and, once learned, are theoretically applied without conscious deliberation. These routines can be conceptualized as "If [X], then [Y]" *productions* similar to those conceptualized by J. Anderson (1983; see also Smith, 1990), where [X] consists of a configuration of internally generated or externally impinging stimulus features and [Y] is a sequence of cognitive or motor behaviors that is activated and applied automatically when the eliciting conditions [X] are met. Thus, the configuration of features that compose [X] can include both representations of external stimuli that impinge on the cognitive system and internally generated stimuli (thoughts, proprioceptive reactions, feelings, etc.). However, the configuration is responded to as a whole. Therefore, not all of its individual features may be consciously identified in order for the behavior routines associated with them to be activated. Thus, although these routines may be activated by components of declarative knowledge, they may be performed without recourse to this knowledge.

Comprehension revisited

Although the 1989 version of the model was successful in accounting for a variety of phenomena (Wyer and Srull, 1989), subsequent research we performed made salient a number of deficiencies, most of which were traceable to problems at the early,

comprehension stage of processing. First, the Comprehender had access to only the semantic bin in Permanent Storage (see Figure 8.1). This implies that new stimulus information is first interpreted in terms of general semantic concepts and that referent-specific concepts are not applied until a later stage of processing. However, Wyer and Radvansky (1999) found that statements about well-known individuals and their behavior (e.g., "Jane Fonda did aerobics") were comprehended more quickly than equivalent statements about members of the semantic categories to which the individuals belonged (e.g., "The actress did aerobics"). Furthermore, the time required to comprehend statements about known persons was not appreciably different from the time required to verify them as true or false.

Second, the 1989 model failed to take into account the role of visual imagery. Radvansky et al. (1997) provided evidence that individuals who encounter statements about objects or events that are situationally and temporally constrained construct a mental simulation, or *situation model*, of the events that consists in part of a mental image whose features are spatially and temporally organized. Although the Comprehender was assumed to be capable of processing information in different sense modalities, the unique contribution of visual imagery in comprehending verbal information was not considered.

To account for these phenomena, the 2004 version of the model relaxed the assumption that the Comprehender had access to the semantic bin alone and allowed it to access referent-specific knowledge as well. This change not only permitted new information to be comprehended spontaneously in terms of specific concepts of its referents but also allowed for visual images to be constructed. It also allowed for the possibility that one's recognition of the validity of referent-specific information is often an inherent component of comprehension that occurs automatically rather than as a result of deliberative goal-directed cognitive activity.

Briefly, we assumed that a verbal description of a referent's observable behavior is comprehended by retrieving previously formed visual images of both the behavior and the actor and combining these images to form a new situation model that includes both. If this model is sufficiently similar to a preexisting situation model that is identified in the course of this processing, the information is spontaneously recognized as true, and if it is below a minimal threshold of similarity, it is spontaneously recognized as false. (Otherwise, the information is simply comprehended and the model of it is stored without construing its validity.)

Two implications of these processes are noteworthy. First, behaviors of a well-known person are likely to be comprehended with reference to a previously formed situation model that exemplifies them. In contrast, behaviors of an unknown person are more likely to be interpreted with reference to a more general, prototypic event representation. Colcombe and Wyer (2002) confirmed this difference.

Second, although information that is conveyed verbally may be spontaneously coded visually in the course of comprehending it (Jiang and Wyer, 2009), information that is transmitted visually is comprehended with reference to previously formed situation models but may *not* be spontaneously encoded verbally. This latter recoding may occur only if a verbal coding is necessary to attain some more specific processing objective (Adaval and Wyer, 2004; Wyer et al., 2002).

The spontaneous verification processes assumed by the 2004 extension of the model also permit the model to account for other important comprehension phenomena, such as the spontaneous reactions to statements that violate normative principles of communication (e.g., the principles that communications are typically expected to be informative and to convey the truth as the communicator sees it; Grice, 1975). That is, violations of these rules are spontaneously

recognized and stimulate higher order comprehension processes in order to construe the communication's intended meaning.

Costs and benefits

As the preceding examples indicate, the 2004 version of the model can account for several phenomena that could not be explained by the 1989 version. However, the revision has had costs as well as benefits. The costs have come as a result of the need to revise the model's assumptions about memory storage and retrieval. That is, the original model assumed that information about specific referents is retrieved from Permanent Storage by first identifying a referent bin on the basis of features contained in its header and then performing a sequential, top-down search of the bin's contents for goal-relevant information. However, the time required for this two-stage process would contradict the evidence that referent-specific information is comprehended very quickly in the absence of specific goal-directed processing. Consequently, the 2004 version of the model required a relaxation of the "bin" construct and the postulation of a memory retrieval model similar to that proposed by Ratcliff (1978), which makes few assumptions about the organization of concepts and knowledge in memory. I am reluctant to discard the bin metaphor completely, as it is a powerful tool in conceptualizing goal-directed cognitive activity at later, postcomprehension stages of processing. However, the two sets of retrieval assumptions have not been fully reconciled at this writing.

THE NATURE OF GOAL SCHEMAS: TWO EXAMPLES

The Wyer and Srull conceptualization is obviously an incomplete description of information processing in the absence of a more precise statement of the goal schemas that

govern the attainment of specific processing objectives. However, more circumscribed theoretical formulations can often be conceptualized in terms of the content of these goal schemas. Much of my work in specific areas of social information processing can be viewed as attempts to specify the general nature of these schemas.

For example, the information integration processes that were the focus of my early research in impression formation can be viewed as hypotheses concerning the routines that are stored in the Inference Maker's library and used to combine the implications of individual pieces of information to form a judgment. As such, they are not subject to conscious awareness. On the other hand, the particular routine that is applied may depend on both the type of information presented and the situational context in which the judgment is made (Wyer and Carlston, 1979). These latter contingencies may be specified in the goal schema that governs the processing of the information and, therefore, in the instructions the Executor gives to the Inference Maker along with the information to be integrated. Two other formulations we developed, each of which can be viewed as a theory of cognitive functioning in its own right, provide more specific examples of goal schemas.

Belief formation and change

My extension of McGuire's (1960) syllogistic model of belief organization, mentioned earlier, assumes that if beliefs (estimates of the likelihood that a proposition is true) are defined in units of probability, the relation between the belief in a conclusion (C) and beliefs in an antecedent (A) can be described by the equation:

$$P(C) = P(A)P(C/A) + P(\sim A)P(C/\sim A) \quad (8.1)$$

where $P(C)$ is the belief that C is true, $P(A)$ and $P(\sim A)$ are beliefs that A is and is not true,

respectively, and $P(C/A)$ and $P(C/\sim A)$ are beliefs that C is true if A is and is not true, respectively. Furthermore, if information is received that changes beliefs in A by an increment $\Delta P(A)$, its effect on the peripheral belief in C is simply,

$$\Delta P(C) = \Delta P(A)[P(C/A) - P(C/\sim A)] \quad (8.2)$$

These two equations provide remarkably accurate quantitative fits of the relations among the beliefs involved (Wyer, 1970). Furthermore, they capture the Socratic effect, or the tendency for related beliefs, once they are made salient in temporal proximity, to become more consistent over time (Rosen and Wyer, 1972). However, although this conceptualization was originally interpreted as a model of belief organization, subsequent research (Wyer and Hartwick, 1980) indicated that it was more appropriately conceptualized as a description of *inference* processes, namely, the process of inferring the likelihood of C from beliefs in a second, “informational” proposition, A , that happens to be salient at the time. For example, individuals who are asked their belief that drinking coffee is desirable (C) are likely to search memory for a second, informational proposition, A , that has implications for it and to construe the implications of this proposition. Thus, they may draw a different conclusion if the proposition they happen to retrieve is, “Drinking coffee wakes you up in the morning” than if it is, “Drinking coffee gives you insomnia” (Wyer and Hartwick, 1980). The proposition they happen to identify and use presumably depends on its accessibility in the referent bin in which it is located.

This process, which presumably is stored in the library of the Inference Maker (see Figure 8.1) is applied as a result of instructions by the Executor to make an inference about C based on information pertaining to A . To this extent, the process described by Equation (8.1) may be performed without awareness.

Impression formation

A second example of a goal schema is provided by our conceptualization of person memory and judgments (Srull and Wyer, 1989; Wyer and Srull, 1989). The conceptualization specifies the sequence of cognitive activities that occur when individuals have the goal of forming an impression of someone on the basis of a series of traits and behaviors, the judgments that are based on this representation, and the subsequent recall of the behaviors contained in it. The theory assumes that people with an impression formation objective attempt (a) to assign general personality traits to the person on the basis of the information about him, and (b) to arrive at an overall evaluation of the person as likeable or dislikeable. To this end, the following steps are performed:

- 1 *Trait encoding.* People encode the individual behavior of the person in terms of trait concepts that they exemplify. This activity establishes an association between the behavior and the trait concept it exemplifies. When several behaviors exemplify the same trait, a trait-behavior cluster is formed.
- 2 *Evaluative concept formation.* The initial information presented about the person is coded evaluatively and, if its implications are consistent, a general concept is formed of the individual as likeable or dislikeable. The person's behaviors are also encoded evaluatively and are thought about with reference to this concept, consequently becoming associated with it.
- 3 *Responses to inconsistency.* If a behavior is encountered that is evaluatively inconsistent with the general person concept that is formed in “2,” people respond in two ways. First, they think about the behavior in relation to others they have received in an attempt to reconcile its occurrence. This activity leads associations to be formed between the inconsistent behavior and others. Second, they review behaviors that are consistent with their general concept of the person in an attempt to confirm its validity. This activity strengthens the association between these behaviors and the concept.
- 4 The representations that are formed (both trait-behavior clusters and the more general person representation) are then stored in memory in a bin pertaining to the person.

Retrieval processes

Suppose that people are later asked to recall the information they have received about the person. They will retrieve one of the representations from the bin in which it is located and search its contents, beginning with the central node and progressing down one of the pathways emanating from it to a behavior. After reporting this behavior, they progress along a pathway to another behavior and so on, returning to the central node whenever they encounter a dead end. Thus, if the representation they happen to retrieve is a trait-behavior cluster, they are likely to recall the behaviors in this cluster before behaviors pertaining to other traits (for evidence, see Hamilton et al., 1980). If they retrieve and use the person representation, the first behavior they recall should be consistent with the evaluative concept of the person, as consistent behaviors are more strongly associated with the concept. In general, however, there are more pathways leading into inconsistent behaviors than consistent ones, and so inconsistent behaviors are better recalled (Hastie and Kumar, 1979; Srull, 1981).

Judgment

When people are called upon to make a judgment, they retrieve a representation whose central concept is relevant to the judgment and use its implications as a basis for the judgment without reviewing the behaviors that are associated with it. These behaviors are only consulted if the central concept has no direct implications for the judgment to be made. Thus, for example, if people are asked to evaluate the person, they should make judgments that are implied by the evaluative person concept they have formed. Yet, behaviors that are inconsistent with this concept may often be better recalled.

Qualifications

The aforementioned conceptualization is able to account for numerous memory and judgment phenomena, including the impact of stereotypes on memory (Wyer and Martin, 1986), differences in memory for individuals

and groups (Srull, 1981; Wyer et al., 1984), and the impact of information that people are told to disregard (Wyer and Budesheim, 1987). However, its applicability may nonetheless be limited to the particular paradigm that was normally used to evaluate its validity. When a person's behaviors are conveyed in the course of an informal conversation, for example, listeners tend to form an impression of the *speakers* rather than of the individuals they are describing and organize their descriptions of the person's behaviors around these impressions (Wyer et al., 1990). Thus, they think more extensively about behavior descriptions that violate conversational norms to be polite and modest (Wyer et al., 1994).

Despite these limitations, our original conceptualization of person's memory and judgment processes provides an example of a goal schema that might, in principle, be constructed and incorporated into the more general model that Wyer and Srull proposed. Furthermore, the qualifications on the original model that were identified in later research can be viewed as reflecting different goal schemas that are specific to different types of information and situational contexts.

RELATION TO OTHER THEORIES

As our examples of goal schemas indicate, a primary value of the conceptualization lies in the framework it provides for integrating the implications of more specific formulations of information processing. Although space limitations preclude a full discussion of this possibility, a few examples may illustrate its implications of other formulations and its utility in identifying further avenues for empirical investigation.

Dual processing models

Numerous conceptualizations have been proposed that fall under the heading of "dual processing models." For some reason, the

Wyer and Srull model has rarely, if ever, been formally recognized as such a model (Chaiken and Trope, 1999). As the preceding discussion indicates, however, it is actually a *multi*-process model that allows for several alternative processes at each stage of cognitive activity (comprehension, inference, etc.), some of which are automatic and others are deliberative. Furthermore, it provides a framework for conceptualizing the conditions in which different processes occur.

A model proposed by Strack and Deutsch (2004) is particularly interesting to consider in this context. Their model postulates two separate processing systems. One, a *reflective*, system comes into play in deliberative goal-directed processing and is governed by processes of which individuals are well aware. Thus, it generates judgments, decisions, and intentions that depend on the particular type of goal at hand. A second, *impulsive*, system operates automatically and is governed largely by associative processes. This system directs behavior through cognition–behavior associations that are acquired through learning. Once acquired, however, their activation is governed by general principles of knowledge accessibility (Förster and Liberman, 2007; Higgins, 1996; Wyer, 2008).

The automatic processes that compose the “impulsive” system defined by Strack and Deutsch are localized in several components of the Wyer and Srull model, including the Comprehender and the libraries of special-purpose processing units. These processes are also implied in the activities of the Executor in the course of free flow of thought. There is an important distinction between the two conceptualizations, however; that is, the reflective goal-directed processing assumed by Strack and Deutsch appears to be completely governed by goal schemas that exist as part of general knowledge. The impulsive system only operates as a default, when conscious goal-directed actions performed by the reflective system are not operating. In contrast, the Wyer and Srull model allows for

automatic (unconscious) processes to occur in the pursuit of conscious goal-directed activity. Specifically, the processes that are stored in the library of the various processing units that are activated by the model are goal-directed, but nonetheless operate automatically without consciousness of the specific cognitive operations that are involved.

Attitude formation and change

As Schwarz and Bohner (2001) point out, the conditions in which individuals consciously change their attitude or opinion in response to new information may be limited. More generally, individuals who are called upon to evaluate a stimulus may retrieve a sample of judgment-relevant information that happens to be easily accessible in memory and compute a judgment online, based on the implications of this sample. Thus, the effect of a persuasive communication may not reflect a conscious change in attitude. Rather, it would result from a difference in the information that enters into the online computation of the attitudes at different points in time.

This conceptualization is quite consistent with the Wyer and Srull model. This model assumes that an attitude toward an object is computed on the basis of a subset of the information stored in a referent bin that has most recently or frequently been used. This information can sometimes include a previously formed judgment that has been stored in the bin as a result of prior processing. To this extent, whether the attitude appears to be stable depends on whether this judgment is the *only* information retrieved or other information is retrieved as well. This depends on the recency with which other judgment-relevant information has been acquired.

Construing the implications of a persuasive message

The dominant theories of communication and persuasion have been proposed by Petty

and Cacioppo (1986) and Chaiken (1987). Petty and Cacioppo's model assumes that individuals base their judgments on different criteria, depending on their a priori assessment of the time and cognitive resources that are required to apply them. In contrast, Chaiken's (1987) model assumes that judgmental criteria are applied sequentially, with easily accessible and easy-to-apply criteria considered first and additional criteria considered only if the judgment is important and participants have little confidence in the first criteria they apply.

Of the two conceptualizations, Chaiken's is more congenial to the Wyer and Srull formulation. The sequential identification and use of alternative judgment criteria, and the threshold of confidence required to make a judgment, are presumably governed by a goal schema, and the relative accessibility of the criteria being applied would be determined by the likelihood of identifying them in a top-down search of a referent bin pertaining to the communication's referent.

Both Chaiken's conceptualization and Petty and Cacioppo's, however, leave open the question of precisely how an inference is computed. As Kruglanski et al. (1999) point out, this process could be similar regardless of whether the inference is based on the source of a message or the arguments contained in it. In fact, the syllogistic process assumed by Kruglanski et al.'s (1999) "unimodel" is very similar to the process implied by the belief inference model described earlier in this chapter as an example of a goal schema (see Wyer, 2006, for an elaboration).

Dual and implicit attitudes

Several conceptualizations postulate the existence of "dual attitudes" that exist simultaneously in memory and are called upon for use in judgments or behavioral decisions, depending on the circumstances. Greenwald and Banaji (1995), for example, distinguish

between *explicit* attitudes that individuals consciously report and *implicit* attitudes that mediate favorable thoughts or behavior toward a stimulus without necessarily any awareness of their influence or, in fact, their existence.

The existence of different attitude-based representations in memory is obviously compatible with the Wyer and Srull model, and their use as bases for judgment is presumably determined by goal schemas that are retrieved for use under the conditions at hand. For example, Brunel et al. (2004) showed that European Americans' explicit attitudes toward advertisements were similar regardless of the ethnicity of the models shown in the ads, but their implicit attitudes were more favorable when the models were white. The procedures used to assess these attitudes differed considerably. Thus, in terms of the Wyer and Srull model, these data simply suggest that the goal schema that individuals activate for use in computing a judgment depends on the type of judgment that they are asked to make.

Unconscious influences on judgments and decisions

Perhaps the most provocative body of research to emerge in recent years has shown that individuals' judgments and behavior are often governed by factors of which they are not consciously aware (cf. Bargh et al., 1996; Chartrand and Bargh, 1996; for a review, see Dijksterhuis and Bargh, 2001). These studies suggest that the perception of an event can elicit behavior automatically with little, if any, interpolated cognitive activity. A conceptualization of these effects within the framework proposed by Wyer and Srull, however, suggests the need to distinguish between (a) consciousness of the conditions that lead concepts to become accessible in memory and (b) consciousness of the connection between these conditions and the behavior that the concepts elicit.

Specifically, the processing of information by the Comprehender is not under control of the Executor and, therefore, its activities are not subject to conscious awareness. Concepts retrieved by the Comprehender can therefore become more accessible without conscious awareness of either the concepts themselves or the processes that give rise to their activation. As a consequence, the concepts are likely to be used to comprehend information that is received later (Bargh and Pietromonaco, 1982).

For concepts to come into play at later, postcomprehension stages of processing, however, they must somehow find their way into the Work Space where they can be selected as part of the set of features that the Executor compiles and either (a) uses to identify a goal schema or (b) transmits to a special-purpose processing unit, thereby influencing the particular routine that is used to attain the objective at hand. If the concepts are primed in the course of conscious goal-directed activity, they may enter the Work Space even if participants are unaware of the relation of the priming task to the behavior they perform subsequently. If concepts are primed subliminally and are not involved in postcomprehension goal-directed activity, however, this may not be the case.

The implications of this contingency are exemplified by Bargh et al.'s (1996) finding that when participants had been unobtrusively primed with concepts of the elderly, they walked more slowly to the elevator upon leaving the experiment. In this study, however, concepts were primed using a sentence-construction task that required their use. This goal-directed processing may have led the concepts to be present in the Work Space and to be fortuitously sampled by the Executor and transmitted to the Response Selector, where it influenced the selection of a goal schema that was used to generate an overt response. Thus, the effects identified by Bargh et al. might not have occurred if elderly-related concepts had been primed subliminally.

Goal-directed behavior and motivation

Kruglanski and his colleagues (Kruglanski et al., 2002; Shah et al., 2003) postulate a hierarchy of goals at different levels of specificity such that the goal at one level serves as a means of attaining goals at other levels. In support of this conceptualization, Kruglanski et al. show that subliminally activating one goal can increase the accessibility of other goals that are both higher and lower in the hierarchy. Furthermore, once goal-relevant concepts are activated, they not only facilitate goal-directed behavior to which they are relevant but also interfere with goal-directed activity to which they are not relevant (Shah and Kruglanski, 2002).

Wyer and Srull (1989) also assume that goal schemas are represented in memory at different levels of generality, and that the activation of a goal at a particular level of specificity can activate the sequence of actions that are means to its attainment. If several goal schemas can be used to attain the same objective, their accessibility in memory depends on the frequency or recency with which they have been applied.

The evidence that *subliminally* activating one component of a goal sequence can increase the accessibility of other components (Kruglanski et al., 2002) is difficult for the Wyer and Srull (1989) conceptualization to explain. The conceptualization can nevertheless account for unconscious goal-directed processing. The routines in a processing unit's library are presumably performed in the course of goal-directed activity without awareness of the general objective to which they pertain. Moreover, although goal-related concepts can be among the configuration of features that elicits a production of the sort that is contained in a processing unit's library, the configuration is responded to schematically and, therefore, might activate the production without consciousness of the specific features that compose it. Thus, goal-directed behavior can often occur without awareness

of the goal to which it is relevant (Chartrand and Bargh, 1996).

APPLICATIONS

Many aspects of the information-processing model have implications for information processing outside the laboratory. The model's implications for the spontaneous recognition of statements as true or false in the course of comprehending them (Wyer and Radvansky, 1999) and the processes that underlie humor elicitation (Wyer and Collins, 1992) are described in detail elsewhere, and space does not permit an elaboration in this chapter. My recent research has not provided rigorous tests of the conceptualization but, rather, has evaluated its more general implications for nonlaboratory phenomena. Two general areas of research may be worth noting briefly.

The role of visual imagery in responses to persuasive appeals

To reiterate, the 2004 version of the theory assumes that situation models are typically formed in the course of comprehending situation-specific events, a central component of which is a visual image. Several implications of this assumption have been investigated.

Effects of donation appeals

Appeals for donations often encourage recipients to form a mental representation (or, in terms of the 2004 theory, a "situation model") of themselves in the situation confronting the individuals in need of help. To be effective, however, the appeal must also induce the recipients to imagine themselves as a potential donor. However, it is difficult if not impossible to form mental images of oneself in two roles simultaneously. Thus, suppose recipients of an appeal have the perspective of a potential donor at the time they receive a

request for aid. In this case, an appeal that stimulates them to construct a situation model of themselves in the role of a victim may create cognitive conflict and decrease the appeal's effectiveness. Consistent with this conjecture, Iris Hung (Hung and Wyer, 2009) found that stimulating individuals to imagine a situation from the victims' perspective increased the appeal's effectiveness when a request for aid was not made until after the situation was described (as indicated by the amount of money that participants were willing to give). When participants had a disposition to imagine themselves as the donor at the outset, however, stimulating them to form a representation of the situation from the victim's perspective *decreased* the appeal's effectiveness.

Effects of advertising

Other effects of visual imagery on the impact of persuasive appeals were identified by Jiang (2008; see Wyer et al., 2008). In a representative study, participants received information about a hotel describing features that could be imagined from either the same visual perspective (i.e., inside the hotel) or different perspectives (e.g., both inside and outside). In the first case, individuals could easily form a single image-based situation model of the hotel. In the second case, however, this was impossible to do. Consequently, individuals with a disposition to form visual images evaluated the hotel more favorably in the first case than in the second. When individuals were induced to process the information semantically without forming visual images, however, this difference was eliminated.

Additional evidence that difficulty in constructing situation models can have an adverse effect on the impact of advertisements was obtained by Hung and Wyer (2008). Participants received a print ad for a product consisting of (a) a problem that the product purportedly remedied (hair loss, stained clothing, etc.) and (b) the consequences of using the product. When one component was conveyed in a picture and the

other was described verbally, recipients could construct a situation model of the sequence of events that was consistent with their general knowledge of the type of product being promoted. As a result, they evaluated the product favorably. When both components were pictured, however, flexibility in constructing a visual image of the events described was eliminated, and so a situation model of the events was more difficult. Consequently, the product was evaluated less favorably.

The role of procedural knowledge in shopping behavior

A central feature of the information processing model surrounds its ability to conceptualize the effects of a number of different processing objectives on judgments and decisions. These effects theoretically depend on how the procedures for attaining these objectives are represented in memory. Sequences of goal-directed behavior can be stored in memory either as part of semantic knowledge or as a production of the sort postulated by Anderson (1983) that is applied automatically when the preconditions for its activation are met. The impact of these different types of knowledge representations on consumer behavior has been demonstrated by Jing Xu (Xu and Wyer, 2007, 2008) and Hao Shen (Shen and Wyer, 2008), respectively.

Xu, for example, assumed that purchases are often governed by a three-stage sequence of activities involving (a) whether to make a purchase, (b) which alternative to purchase, and (c) how to implement the purchase. However, she further assumed that if situational factors activate the second stage of this sequence, consumers might perform it and proceed to the third, implemental stage without engaging in the first stage. Therefore, they might be disposed to buy one of the alternatives without considering whether they wanted to buy anything at all.

Finally, note that the process of deciding which product to buy is a special case of a

more general process of making comparative judgments. To this extent, a “which-to-buy” mindset might be activated by other exemplars of this process that have nothing to do with purchase behavior. This was in fact the case; asking persons to compare the physical attributes of wild animals, or to indicate whether one country was similar to another, was sufficient to activate a “which-to-buy” mindset and, therefore, to increase their likelihood of purchasing products that were on sale after the experiment.⁵ Further evidence that behavior in one stimulus domain can induce a mindset that affects behavior in other, quite different domains (specifically, variety seeking) was obtained by Shen and Wyer (2009).

CONCLUSION

The general information processing model described in this chapter is obviously metaphorical and, as such, must be evaluated in terms of its utility and not its validity. In my own work, I have not treated the model as a definitive theory of cognitive behavior but have viewed it as a heuristic device that calls attention to theoretical and empirical issues that might otherwise have escaped my attention. In this regard, I believe that a model should be able to defend itself and that the theorist should be its harshest critic. In this spirit, however, I want to thank the model personally for the stimulation it has provided me over the years. If the model has been able to stimulate a few others as well, it should feel doubly honored.

ACKNOWLEDGMENTS

The writing of this paper and some of the research reported was supported by grant GRF641308 from the Research Grants Council of the Hong Kong SAR, China.

NOTES

1 As implied earlier, this interest reflects my attempt to emulate W.J. McGuire, social psychology's intellectual and spiritual leader and the father of social information processing theory and research. Both his own work and the personal support and encouragement he gave to me throughout my career were an inspiration, and any success I may have had results in large part from my conscious and unconscious attempts to follow, albeit feebly, in his footsteps.

2 These meetings gradually expanded into what is now known as the "Person Memory Interest Group," initially inspired by Tom Ostrom and continued by Dave Hamilton and Eliot Smith. It now has over 80 members and is the primary organization for researchers in social cognition.

3 I earlier acknowledged my indebtedness to W.J. McGuire, whose research and theorizing obviously provided the inspiration for my own work (see footnote 1). However, his most important influence is less obvious. In 1968, almost six years after completing my PhD, I was still floundering, and wondering whether anyone had even read any of the work I had done, to say nothing of whether they cared about it. During this period of self-doubt, I submitted a paper to the *Journal of Personality and Social Psychology*, which McGuire processed as editor. Although I cannot recall his specific reactions, he undoubtedly called attention to numerous instances conceptual and expositional sloppiness, because I wrote back apologizing to him for putting him through the ordeal of evaluating it. I immediately received a response that made my day and, ultimately my career. I cannot recall his exact words. However, they were to the effect that he and others had "assumed" that I knew I was a good psychologist and that it was only in this context that they bothered to take the time to "carp" about the things he had noted. Leaving aside the fact that his comments could hardly be viewed as "carps," this was the first time that anyone, let alone someone as eminent as W.J. McGuire, had conveyed any interest whatsoever in anything I had done. His encouragement at this critical point in my career gave me the confidence to persist. Many years later, it still inspires me during times of disappointment and self-doubt. In the context of this personal history, I would like to acknowledge my indebtedness to this remarkable psychologist and equally remarkable human being.

4 On the other hand, the statement "the ball kicked the boy" cannot be automatically comprehended on the basis of routines in the Comprehender's library. Such a statement would therefore be transmitted to the Work Space along with a message signifying its status and would be comprehended

(if at all) on the basis of Executor-controlled processing.

5 An interesting speculation derived from this conceptualization is that the consumption of material goods is greater during election years, when citizens are continually being asked which of two political candidates they prefer, than in off-election years. In fact, an analysis of years between 1953 and 2000 showed that total retail-store sales in the United States were 9.4 percent higher during the three months prior to an election (August, September, and October) than during comparable months in the years before and after election years. Although this difference was not statistically significant, its consistency with our laboratory findings is provocative.

REFERENCES

Adaval, R. and Wyer, R.S. (2004). Communicating about a social interaction: Effects on memory for protagonists' statements and nonverbal behaviors. *Journal of Experimental Social Psychology*, 40, 450-465.

Anderson, J.R. (1983) *The Architecture of Cognition*. Cambridge, MA: Harvard University Press.

Anderson, N.H. (1971) Integration theory and attitude change. *Psychological Review*, 78, 171-206.

Bargh, J.A., Chen, M. and Burrows, L. (1996) Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology*, 71, 230-244.

Bargh, J.A. and Pietromonaco, P. (1982) Automatic information processing and social perception: The influence of trait information presented outside of conscious awareness on impression formation. *Journal of Personality and Social Psychology*, 43, 437-449.

Brunel, F.F., Tietje, B.C. and Greenwald, A.G. (2004) Is the Implicit Association Test a valid and valuable measure of implicit consumer social cognition? *Journal of Consumer Psychology*, 14, 385-404.

Carlston, D.E. (1980) Events, inferences and impression formation. In R. Hastie, T. Ostrom, E. Ebbesen, R. Wyer, D. Hamilton, and D. Carlston (eds), *Person Memory: The Cognitive Basis of Social Perception*, pp. 89-119. Hillsdale, NJ: Erlbaum.

Chaiken, S. (1987) The heuristic model of persuasion. In M.P. Zanna, J.M. Olson, and C.P. Herman (eds), *Social Influence: The Ontario Symposium*, 5, 3-39. Hillsdale, NJ: Erlbaum.

Chaiken, S. and Trope, Y. (1999) *Dual-process Theories in Social Psychology*. New York: Guilford Press.

Chartrand, T.L. and Bargh, J.A. (1996) Automatic activation of impression formation and memorization goals: Nonconscious goal priming reproduces effects of explicit task instructions. *Journal of Personality and Social Psychology*, 71, 464–478.

Cohen, C.E. and Ebbesen, E.B. (1979) Observational goals and schema activation: A theoretical framework for behavior perception. *Journal of Experimental Social Psychology*, 15, 305–339.

Colcombe, S.J. and Wyer, R.S. (2002) The role of prototypes in the mental representation of temporally-related events. *Cognitive Psychology*, 44, 67–103.

Dijksterhuis, A. and Bargh, J.A. (2001) The perception-behavior expressway: Automatic effects of social perception on social behavior. In M.P. Zanna (ed.), *Advances in Experimental Social Psychology*, 33, 1–40. San Diego: Academic Press.

Fishbein, M. and Hunter, R. (1964) Summation versus balance in attitude organization and change. *Journal of Abnormal and Social Psychology*, 69, 505–510.

Förster, J. and Liberman, N. (2007) Knowledge activation. In A. Kruglanski and E.T. Higgins (eds), *Social Psychology: Handbook of Basic Principles* (2nd edition), pp. 201–231. New York: Guilford.

Greenwald, A.G. and Banaji, M.R. (1995) Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, 102, 4–27.

Grice, H.P. (1975) Logic and conversation. In P. Cole and J.L. Morgan (eds), *Syntax and Semantics: Speech Acts*, pp. 41–58. New York: Academic Press.

Hamilton, D.H., Katz, L.B. and Leirer, V.O. (1980) Organizational processes in impression formation. In R. Hastie, T. Ostrom, E. Ebbesen, R. Wyer, D. Hamilton, and D. Carlston (eds) *Person Memory: The Cognitive Basis of Social Perception*, pp. 121–153. Hillsdale, NJ: Lawrence Erlbaum Associates.

Hastie, R. and Kumar, P.A. (1979) Person memory: Personality traits as organizing principles in memory for behaviors. *Journal of Personality and Social Psychology*, 37, 25–38.

Hastie, R., Ostrom, T.M., Ebbesen, E.B., Wyer, R.S., Hamilton, D.L. and Carlston, D.E. (1980) *Person Memory: The Cognitive Basis of Social Perception*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Higgins, E.T. (1996) Knowledge activation: Accessibility, applicability, and salience. In E.T. Higgins and A. Kruglanski (eds), *Social Psychology: Handbook of Basic Principles*, pp. 133–168. New York: Guilford.

Higgins, E.T., Rholes, W.S. and Jones, C.R. (1977) Category accessibility and impression formation. *Journal of Experimental Social Psychology*, 13, 141–154.

Hung, I.W.P. and Wyer, R.S. (2008) The role of implicit theories in the impact of problem-solving print advertisements. *Journal of Consumer Psychology*, 18, 223–235.

Hung, I.W. and Wyer, R.S. (2009) Differences in perspective and the influence of charitable appeals: When imagining oneself as the victim is not always beneficial. *Journal of Marketing Research*, 46, 421–434.

Jiang, Y. (2008) The role of visual imagery and visual perspective in consumer behavior. PhD dissertation, Hong Kong University of Science and Technology.

Jiang, Y. and Wyer, R. S. (2009) The role of visual perspective in information processing. *Journal of Experimental Social Psychology*, 45, 486–495.

Kruglanski, A.W., Shah, J.Y., Fishbach, A., Friedman, R., Chun, W. and Sleeth-Keppler, D. (2002) A theory of goal systems. In M.P. Zanna (ed.), *Advances in Experimental Social Psychology*, 34, 331–378. San Diego: Academic Press.

Kruglanski, A.W., Thompson, E.P. and Spiegel, S. (1999) Separate or equal? Bimodal notions of persuasion and a single process 'unimodel'. In S. Chaiken and Y. Trope (eds), *Dual-process Theories in Social Psychology*. pp. 293–313. New York: Guilford.

Lingle, J.H. and Ostrom, T.M. (1979) Retrieval selectivity in memory-based impression judgments. *Journal of Personality and Social Psychology*, 37, 180–194.

McGuire, W.J. (1960) A syllogistic analysis of cognitive relationships. In M.J. Rosenberg and C.I. Hovland (eds), *Attitude Organization and Change*, pp. 140–162. New Haven: Yale University Press.

McGuire, W.J. (1968) Personality and susceptibility to social influence. In E. Borgatta and W. Lambert (eds), *Handbook of Personality Theory and Research*. Chicago: Rand-McNally.

Nisbett, R.E. and Ross, L. (1980) *Human Inference: Strategies and Shortcomings of Social Judgments*. Englewood Cliffs, NJ: Prentice Hall.

Petty, R.E. and Cacioppo, J.T. (1986) *Communication and Persuasion: Central and Peripheral Routes to Attitude Change*. New York: Springer-Verlag.

Radvansky, G.A., Wyer, R.S., Curiel, J.M. and Lutz, M.F. (1997) Mental models and abstract relations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23, 1233–1246.

Ratcliff, R. (1978) A theory of memory retrieval. *Psychological Review*, 85, 59–108.

Rosen, N.A. and Wyer, R.S. (1972) Some further evidence for the 'Socratic effect' using a subjective probability model of cognitive organization. *Journal of Personality and Social Psychology*, 24, 420–424.

Schwarz, N. and Bohner, G. (2001) The construction of attitudes. In A. Tesser and N. Schwarz (eds), *Blackwell Handbook of Social Psychology: Intraindividual Processes*, Vol. 1. Oxford: Blackwell.

Shah, J.Y. and Kruglanski, A.W., (2002) Priming against your will: How goal pursuit is affected by accessible alternatives. *Journal of Experimental Social Psychology*, 38, 368–383.

Shah, J.Y., Kruglanski, A.W. and Friedman, R. (2003) Goal systems theory: Integrating the cognitive and motivational aspects of self-regulation. In S.J. Spencer and S. Fein (eds), *Motivated Social Perception: The Ontario Symposium*, 9, 247–275. Mahwah, NJ: Erlbaum.

Shen, H. and Wyer, R.S. (2008) Procedural priming and consumer judgments: Effects on the impact of positively and negatively valenced information. *Journal of Consumer Research*, 34, 727–737.

Shen, H. and Wyer, R.S. (2009) Cognitive and motivational influences of past behavior on variety seeking. Unpublished manuscript, Hong Kong University of Science and Technology.

Shen, H. and Wyer, R.S. (2010) The effect of past behavior on variety seeking: Automatic and deliberative influences. *Journal of Consumer Psychology*, 20, 33–42.

Smith, E.R. (1990) Content and process specificity in the effects of prior experiences. In T.K. Srull and R.S. Wyer (eds), *Advances in Social Cognition*, 3, 1–59. Hillsdale, NJ: Lawrence Erlbaum Associates.

Srull, T.K. (1981) Person memory: Some tests of associative storage and retrieval models. *Journal of Experimental Psychology: Human Learning and Memory*, 7, 440–463.

Srull, T.K. and Wyer, R.S. (1980) Category accessibility and social perception: Some implications for the study of person memory and interpersonal judgments. *Journal of Personality and Social Psychology*, 38, 841–856.

Srull, T.K. and Wyer, R.S. (1983) The role of control processes and structural constraints in models of memory and social judgment. *Journal of Experimental Social Psychology*, 19, 497–521.

Srull, T.K. and Wyer, R.S. (1989) Person memory and judgment. *Psychological Review*, 96, 58–63.

Strack, F. and Deutsch, R. (2004) Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, 8, 220–247.

Upshaw, H.S. (1965) The effect of variable perspectives on judgments of opinion statements for Thurstone scales: Equal-appearing intervals. *Journal of Abnormal and Social Psychology*, 64, 85–96.

Wyer, R.S. (1969) Effects of general response style on own attitude and the interpretation of attitude-relevant messages. *British Journal of Social and Clinical Psychology*, 8, 104–115.

Wyer, R.S. (1970) The quantitative prediction of belief and opinion change: A further test of a subjective probability model. *Journal of Personality and Social Psychology*, 16, 559–571.

Wyer, R.S. (1974) *Cognitive Organization and Change: An Information-processing Approach*. Hillsdale, NJ: Erlbaum.

Wyer, R.S. (2004) *Social Comprehension and Judgment: The Role of Situation Models, Narratives and Implicit Theories*. Mahwah, NJ: Erlbaum.

Wyer, R.S. (2006) Three models of information processing: An evaluation and conceptual integration. *Psychological Inquiry*, 17, 185–194.

Wyer, R.S. (2008) The role of knowledge accessibility in cognition and behavior: Implications for consumer information processing. In C. Haugvedt, F.R. Kardes, and P.M. Herr (eds), *Handbook of Consumer Psychology*, pp. 31–76. Mahwah, NJ: Erlbaum.

Wyer, R.S., Adaval, R. and Colcombe, S.J. (2002) Narrative-based representations of social knowledge: Their construction and use in comprehension, memory and judgment. In M.P. Zanna (ed.), *Advances in Experimental Social Psychology*, 34, 131–197. San Diego: Academic Press.

Wyer, R.S., Bodenhausen, G.V. and Srull, T.K. (1984) The cognitive representation of persons and groups and its effect on recall and recognition memory. *Journal of Experimental Social Psychology*, 20, 445–469.

Wyer, R.S. and Budesheim, T.L. (1987) Person memory and judgment: The impact of information that one is told to disregard. *Journal of Personality and Social Psychology*, 53, 14–29.

Wyer, R.S., Budesheim, T.L. and Lambert, A.J. (1990) Person memory and judgment: The cognitive representation of informal conversations. *Journal of Personality and Social Psychology*, 58, 218–238.

Wyer, R.S., Budesheim, T.L., Lambert, A.J. and Swan, S. (1994) Person memory and judgment: Pragmatic influences on impressions formed in a social context.

Journal of Personality and Social Psychology, 66, 254–267.

Wyer, R.S. and Carlston, D.E. (1979) *Social Cognition, Inference and Attribution*. Hillsdale, NJ: Erlbaum.

Wyer, R.S., Clore, G.L. and Isbell, L.M. (1999) Affect and information processing. In M.P. Zanna (ed.), *Advances in Experimental Social Psychology*, 31, 1–77. San Diego: Academic Press.

Wyer, R.S. and Collins, J.E. (1992) A theory of humor elicitation. *Psychological Review*, 99, 663–688.

Wyer, R.S. and Gruenfeld, D.H. (1995) Information processing in social contexts: Implications for social memory and judgment. In M.P. Zanna (ed.) *Advances in Experimental Social Psychology*, 27, 49–91. San Diego: Academic Press.

Wyer, R.S. and Hartwick, J. (1980) The role of information retrieval and conditional inference processes in belief formation and change. In L. Berkowitz (ed.) *Advances in Experimental Social Psychology*, 13, 241–284. New York: Academic Press.

Wyer, R.S., Hung, I.W. and Jiang, Y.W. (2008) Visual and verbal processing strategies in comprehension and judgment. *Journal of Consumer Psychology*, 18, 244–257.

Wyer, R.S. and Martin, L.L. (1986) Person memory: The role of traits, group stereotypes, and specific behaviors in the cognitive representation of persons. *Journal of Personality and Social Psychology*, 50, 661–675.

Wyer, R.S. and Radvansky, G.A. (1999) The comprehension and validation of social information. *Psychological Review*, 106, 89–118.

Wyer, R.S. and Srull, T.K. (1980) The processing of social stimulus information: A conceptual integration. In R. Hastie, T.M. Ostrom, E.B. Ebbesen, R.S. Wyer, D.L. Hamilton, and D.E. Carlston (eds), *Person Memory: The Cognitive Basis of Social Perception*. Hillsdale, NJ: Erlbaum.

Wyer, R.S. and Srull, T.K. (1986) Human cognition in its social context. *Psychological Review*, 93, 322–359.

Wyer, R.S. and Srull, T.K. (1989) *Memory and Cognition in Its Social Context*. Hillsdale, NJ: Erlbaum.

Xu, A.J. and Wyer, R.S. (2007) The effect of mindsets on consumer decision strategies. *Journal of Consumer Research*, 34, 556–566.

Xu, A.J. and Wyer, R.S. (2008) The comparative mindset: From animal comparisons to increased purchase intentions. *Psychological Science*, 19, 859–864.

Balance-Logic Theory

Chester A. Insko

ABSTRACT

Following an overview of Heider's classic statement of balance theory as a description of the perceived coherence, or lack of coherence, of interpersonal relations, this chapter reviews subsequent developments including Cartwright and Harary's (1956) application of graph theory, attitude structure, two-valued logic, self-esteem, the overlap between self-consistency and hedonism, conformity, dissonance, and consideration of the tetrahedron model for generalizing the multiplicative rule beyond two-valued distinctions.

HEIDER'S (1946, 1958) STATEMENT OF BALANCE THEORY

Elements, relations, and the concept of balance

Heider formulated balance theory as an extension of Gestalt principles to the perceived coherence, or lack of coherence, of interpersonal relations. He focuses on three types of elements and two types of relations. The elements, typically symbolized p , o ,

and x , are respectively the person whose experience is the focus of the theory, some other perceived person, and some perceived object or concept. Occasionally Heider also refers to a third person as q . Regardless as to the number of persons, it is important to recognize that these are persons as perceived by p , and also that x refers to objects or concepts that are perceived by p .

Heider classifies the relations between pairs of elements as either sentiment relations or unit relations. Sentiment relations are attitudinal, or evaluative, relations that can be either positive ("liking," "approving," "valuing") or negative ("disliking," "disapproving," "devaluing"). Positive sentiment relations are symbolized as L and negative sentiment relations as nL .

Heider (1958: 176) defines unit relations as relations that "are perceived as belonging together." The concept of unit relations involves perceived class, or category, inclusion and, like sentiment relations, can be either positive ("similar," "close," "facilitates," "belongs to") or negative ("dissimilar," "far," "interferes with," "does not belong to"). Positive unit relations are symbolized as

U and negative unit relations as *nU*. Heider relates the concept of unit relations to Wertheimer's (1923) grouping principles of similarity, proximity, and common fate (or covariation).

Cartwright and Harary (1956) point out that it is important not to confuse the negation of a positive unit relation with a negative unit relation. Thus if "owns" is a positive unit relation, the opposite, negative relation is "sells," and not "does not own." This implies, contrary to Heider, that "does not belong to" is not a negative unit relation.

Heider (1958: 180) refers to balanced relations between elements as "a harmonious state" in which elements "fit together without stress," and further elaborates the theory by discussing perceived homogeneity of another person, balance in dyads, and balance in the *p-o-x* triad.

Perceived homogeneity of another person

According to Heider (1958: 183), "if several parts, or traits, or aspects of a person are considered, the tendency exists to see them all as positive, or all as negative." This perceived homogeneity is illustrated by the well-known halo effect in the judgment of different traits in another person. Anderson (1981: 380) presents evidence for the halo effect and concludes that "the general impression of the person acts as a causal mediator in judging specific traits of that person."

Balance in dyads

Restricting attention to the *p*, *o*, and *x* elements, there are three possible dyads: *p-x*, *p-o*, and *o-x*, but Heider (1946, 1958) only discusses the first two. The *p-x* dyad is balanced if the unit relation and sentiment relation have the same signs, for example, "*p* likes the things he made; *p* wants to own the things he made; *p* values what he is

accustomed to" (1946: 108). Possible evidence for balance in the *p-x* dyad comes from the so-called secondary reinforcement effect in which a previously neutral stimulus takes on the valence of a circumstance (e.g., electric shock or food) with which the stimulus has been repeatedly paired (cf. Miller, 1951), and Zajonc's (e.g., 1968, 2001) repeated demonstrations of preference for the familiar. An intriguing illustration of preference for the familiar is Mita et al.'s (1977) finding that participants preferred a photograph of themselves with the familiar reversed image seen in a mirror to a photograph of themselves with the less familiar image seen by others.

The *p-o* dyad, like the *p-x* dyad, is balanced if the unit and sentiment relation have the same signs, for example, "*p* likes his children, people similar to him; *p* is uneasy if he has to live with people he does not like; *p* tends to imitate admired persons; *p* likes to think that loved persons are similar to him" (Heider 1946: 108). Evidence for this type of balance comes from Saegert et al.'s (1973) two experimental demonstrations in which participants expressed more liking for others when they judged tastes in the same as opposed to different booths. A manipulation of the number of trials in the same booth from 0 to 1 to 2 to 5 to 10 resulted in a corresponding increase in attraction. Insko and Wilson (1977) further found that actual interaction beyond simple contact can increase attraction. In their study, three participants, labeled *A*, *B*, and *C*, were seated in a triangular pattern facially oriented toward each other. Initially *A* and *B* were instructed to interact and get acquainted while *C* listened, and then *B* and *C* were instructed to interact and get acquainted while *A* listened. A subsequent assessment of attraction indicated that liking was greater between participants who interacted (*A* with *B*, *B* with *A*, *B* with *C*, *C* with *B*) than between participants who did not interact (*A* with *C*, *C* with *A*).

A further consideration not relevant to the *p-x* triad is that balance in the *p-o* triad requires reciprocation of the sentiment

relations. As stated by Heider (1946: 108), “a balanced state exists if pLo and oLp (or $pnLo$ and $onLp$) are true at the same time.” Support for the presence of reciprocated sentiment comes from both nonexperimental and experimental studies. For example, in an early balance study, Wiest (1965) found a positive correlation of 0.74 between schoolchildren’s reported liking for classmates and the extent to which classmates were perceived as liking them. Experimental evidence in which there is implied positive or negative evaluation of a participant by another person (e.g., Aronson and Worchel, 1966; Byrne and Griffitt, 1966; Byrne and Rhamey, 1965; Insko et al., 1973; Montoya and Insko, 2007) have consistently found evidence for reciprocated sentiment, sometimes referred to as the implied-evaluation effect.

The concept of reciprocated sentiment is, of course, not original with Heider. Shakespeare based his play *Much Ado About Nothing* on the assumption of reciprocated sentiment. Berscheid and Walster (1978) point out that this concept was the underlying assumption in Dale Carnegie’s (1937) book, *How to Win Friends and Influence People*, but that the basic idea can be traced back to a statement by the philosopher Hecato in the second century BC: “I will show you a love potion without drug or herb or any witch’s spell; if you wish to be loved, love.”

As subsequent discussion will emphasize, there are frequently additional considerations indicating that focusing on a single dyad (or triad) is an oversimplification. Two initial examples of such additional considerations occur in Heider’s discussion of the proverbs “opposites attract” and “familiarity breeds contempt.” Heider argues that if opposites do attract it is because the differences allow for the realization of some goal, and implies that for p to achieve a goal is balanced. Attraction between the sexes is an obvious example. Heider also argues that if familiarity does breed contempt it is because of too much dissimilarity. Presumably with increasing contact and familiarity there is an increasing opportunity to discover dissimilarities.

Balance in the $p-o-x$ triad

The $p-o-x$ triad involves the perceived occurrence of three relations: $p-o$, $p-x$, and $o-x$. These relations can be either sentiment or unit relations that are either positive or negative, but Heider’s theory primarily emphasizes the sign of the relation, “[I]n many cases the effects of L and U in these configurations seem to be the same” (1946: 109). On the other hand, Heider also states that “The equivalence of the L and U relations seems to be limited by the fact that often the U relation is weaker than the L relation,” (1946: 111). This issue is partially addressed by Cartwright and Harary’s (1983) above-described admonition to regard a negative unit relation, not as the negation of a positive unit relation, but as the opposite of a positive unit relation.

Restricting attention to just the sign of a relation, there are eight possible triads: $+++$, $++-$, $-+-$, $--+$, $++-$, $+--$, $-+-$, $--+$. Following an early study by Jordan (1953), the tradition has been to refer to the first sign as relating to the $p-o$ (I to other) relation, the second sign as relating to the $p-x$ (I to object) relation, and the third sign as relating to the perceived $o-x$ (other to object) relation.

Heider considers four of the eight triads ($+++$, $++-$, $-+-$, $--+$) to be balanced. These are triads in which all of the relations are positive, or two of the relations are negative and one is positive. An example of a $+++$ triad is: “ p likes o , likes the book x , and perceives that o wrote the book x .” An example of a $--+$ triad is: “ p dislikes o , dislikes the book x , and perceives that o wrote the book x .”

Heider considers three of the remaining relations in which there is a single negative sign ($+-$, $-+$, $-+$) as imbalanced, for example “ p likes the author of a disliked book.” The remaining $--$ triad, however, is considered ambiguous: “[T]he case with three negative signs does not constitute a good psychological balance, since it is too indetermined” (1946: 110). Following Cartwright and Harary (1956), subsequent researchers have universally considered the $--$ triads as imbalanced.

As pointed out above, it is frequently an oversimplification to consider just a single dyad or triad because p 's experience may include multiple dyads and triads. For example, a critic may state: "Just because I disagree with my friend about some issue it is foolish to assume that I am therefore going to dislike my friend." Indeed it may be. Note that while Byrne (1971) is well known for advocating the effect of agreement or similarity on attraction, even he does not claim an effect for the similarity of a single attitude, but rather an effect for the proportion of similar attitudes. Research has further demonstrated that the effect of the proportion of similar attitudes linearly increases as the total number of attitudes increases from 3 to 6 to 12 to 24 (Wetzel and Insko, 1974).

An additional example of the importance of considering more than one dyad or triad is the so-called "love triangle" in which two male friends are attracted to the same female. This is a + + + triad that common sense regards as unstable. Heider's discussion of this situation is succinct and to the point: " p does not want his girl friend o to fall in love with his boy friend q because oLq in this case implies $onLp$, which conflicts with pLo " (1946: 110).

Reactions to imbalance

Heider (1958) discusses reactions to imbalance by describing a simple study in which participants were asked to write how they thought a hypothetical person, Bob, would react to an imbalanced situation:

Bob thinks Jim is very stupid and a first class bore. One day Bob reads some poetry he likes so well that he takes the trouble to track down the author in order to shake his hand. He finds that Jim wrote the poems (1958: 176).

Heider coded the written reactions into five categories, and noted the frequency of each. First, 46 percent described Bob as feeling less negative about Jim; for example, "He grudgingly changes his mind about Jim"

(1958: 186). Second, 29 percent described Bob as feeling less positive about the poetry; for example, "He decides the poems are lousy" (1958: 186). Third, 5 percent described Bob as doubting that Jim wrote the poems; for example, "Bob would probably question Jim's authorship of the poems" (1958: 176-7). Fourth, 2 percent described Bob as differentiating or separating, two aspects of Jim; for example, "He thinks Jim is smart in some lines but dumb in others" (1958: 177). Fifth and finally, 18 percent did not describe Bob as resolving the imbalance but some were aware of the tension; for example, "Bob is confused and does not know what to do. He finally briefly mentions liking of the poems to Jim without much warmth" (1958: 177).

Heider interprets the fifth reaction as illustrating the important point that while imbalance will create tension, that tension will not always be resolved. In this study, of course, most of the participants did describe Bob as resolving the imbalance. The first three reactions involve simple changes in one of the relations in the $p-o-x$ triad. The fourth, or differentiation, reaction is less simple.

Differentiation involves a reconceptualization of a single element as two elements. If the element is an x , that can sometimes be easily accomplished. For example, two friends who disagree about the value of foreign aid might differentiate foreign aid into military aid, which they both dislike, from humanitarian aid, which they both like. An experiment by Stroebe et al. (1970) suggests that differentiation can also occur if the element is a person.

In the Stroebe et al. (1970) experiment participants were given information about a person, Dr. M., indicating that he was an expert or inexpert scientist and a nice or awful person. Whether Dr. M. was an expert or inexpert scientist affected evaluation of a theory with which he was associated but not of his wife. On the other hand, crosscutting information that Dr. M. was a nice or awful person affected evaluation of his wife but not of the theory. The results thus indicated that

the participants differentiated Dr. M. the scientist from Dr. M. the person.

Abelson (1959) has suggested two reactions to imbalance other than the ones listed by Heider. One of these is bolstering, or strengthening, of one of the relations producing imbalance. In terms of the above example, Bob might recall past episodes in which Jim had behaved in a boring manner. The other is transcendence, or the development of a theory accounting for the inconsistency between or among inconsistent elements. For example, Bob might develop a point of view indicating that it is human nature for poetry to be written by boring people.

Osgood and Tannenbaum (1955), Abelson and Rosenberg (1958), and McGuire (1960) have made other suggestions regarding reactions to imbalance or inconsistency. Addressing themselves to change in a relation, Osgood and Tannenbaum (1955) proposed that change is proportional to the degree of polarization in the relation. Addressing themselves to a network of relations, Abelson and Rosenberg (1958) proposed a least-effort principle according to which the change that occurs will be the one that requires the fewest number of additional changes. Finally, also relevant to a network of relations, McGuire (1960) proposed a cognitive inertia principle according to which changes in more remote relations will require more time than changes in relations closest to an encountered inconsistency.

CARTWRIGHT AND HARARY'S MULTIPLICATIVE RULE

Cartwright and Harary (1956) address balance theory's elements and relations with a graph theory representation of points and lines. Positive relations are indicated with solid lines and negative relations with dashed lines. Such an approach allows for the easy representation of sets of relations involving more elements, or points, than dyads or triads. A closed loop of points and lines is

referred to as a *cycle*. In a situation in which it becomes important to distinguish the direction of a relation, for example the difference between p to o and o to p , the lines are given arrowheads and the interconnected points are referred to as a *semicycle*. Thus the p to o and o to p dyad is a semicycle, as is the p to o , p to x , o to x triad.

A cycle or semicycle is balanced only if the product of the signs is positive. Thus, for example, p — o — x triads, or semicycles, that contain three positive signs or one positive sign and two negative signs are balanced, and *semicycles* that contain one negative sign and two positive signs or three negative signs are imbalanced. Note that this multiplicative rule applies equally well to Heider's p — x dyad, or semicycle, that is balanced if the sentiment relation and unit relation have the same sign, and also to the p — o dyad, or semicycle, that has the additional consideration that the p to o relation and the o to p relation have the same sign. If a set of positive and negative lines, referred to as a signed graph or an s-graph, contains more than one cycle, or semicycle, that graph is balanced if and only if all cycles, or semicycles, are balanced.

Cartwright and Harary state a theorem regarding balance in an s-graph: "An s-graph is balanced if and only if its points can be separated into two mutually exclusive subsets such that each positive line joins two points of the same subset and each negative line joins two points from different subsets" (1956: 286). This theorem describes what is sometimes referred to as a "black and white" attitude in which, for example, p perceives positive sentiment relations among the ingroup members and negative sentiment relations with the similarly perceived out-group members. The theorem also describes one interpretation of the Judeo-Christian conception of heaven and hell.

Cartwright and Harary's consideration of multiple linkages points to new and important topics beyond dyads and triads, for example, attitude structures and conformity effects discussed below. Cartwright and Harary's discussion also raises an issue

regarding the utility of consistent thought. It is clear, as Cialdini (2009) points out, that without a tendency to think consistently “our lives would be difficult, erratic, and disjointed” (2009: 53). However, Cialdini also correctly indicates that the tendency to think consistently can create a “foolish fortress” (2009: 54) in which premature closure leads to disregard of important facts and considerations. The mixed utility of consistency is captured in Ralph Waldo Emerson’s frequently quoted statement: “A foolish consistency is the hobgoblin of little minds, adored by little statesmen and philosophers and divines.” Consider, for example, consistent application of the “Reaganomics” principle that “government is the problem.”

ATTITUDE STRUCTURES AND COGNITIVE RESPONSES

While Heider developed balance theory primarily in the context of interpersonal relations, Peak (1955), Rosenberg (1956, 1960), and Fishbein took a similar, consistency approach to attitudes that were related with linking beliefs. Rosenberg refers to a single attitude object with linking beliefs to other attitude objects, or values, as having a structure. He illustrates the basic idea by describing a doctor who negatively evaluates socialized medicine. The doctor believes that socialized medicine interferes with high medical standards and positively evaluates high medical standards. Further, the doctor also believes that socialized medicine facilitates government control and negatively evaluates government control. Each of these two examples (socialized medicine interferes with high medical standards, and socialized medicine facilitates government control) is referred to as a cognitive band. Rosenberg sees an attitude structure as consisting of a series of such cognitive bands. The linking beliefs in each cognitive band of particular interest to Rosenberg were instrumental, or facilitates-interferes-with, cognitions.

In Heider’s language these are positive or negative unit relations.

Some readers may question what happened to p . The person, p , is the doctor. Thus a cognitive band could be stated as a $p-x-y$ triad; the doctor, p , dislikes socialized medicine ($-$), dislikes government control ($-$), and believes that socialized medicine facilitates government control ($+$).

Rosenberg assumes that attitude structures tend to maintain a stable homeostatic, or consistent, state. One procedure that Rosenberg followed to test this theory involved a number of steps. First, participants responded to a $+3$ (strongly facilitates) to -3 (strongly interferes with) scale relating a central attitude object to a number of other attitude objects, or values. Second, participants responded to a $+3$ (very favorable) to -3 (very unfavorable) scale for each of the linked values. Third, the belief and evaluative ratings within each cognitive band were multiplied. Fourth, the products of the belief by evaluative ratings within each cognitive band were summed across bands to give an index expected to predict attitude toward the central attitude object.

Rosenberg (1956, 1960) found that the sum of products across cognitive bands significantly predicted attitude toward free speech for communists, and segregated African-American housing, two topical issues of the time. Other investigators who used variations of Rosenberg’s procedure found similar results. The earliest of these studies by Woodruff and DiVesta (1948) was concerned with attitude toward the abolishment of fraternities and sororities. A subsequent study by Fishbein (1963) was concerned with attitude toward African Americans, and a study by Insko et al. (1970) was concerned with attitude toward the use of birth control. These studies found evidence for consistent attitude structures.

One approach to developing an attitude-structure questionnaire involves the conducting of prior interviews with a sample of respondents who are probed for beliefs related to the attitude issue in question.

An alternative approach involves the use of a thought-listing technique originally developed as a measure of cognitive responses to a persuasive message (Brock, 1967; Greenwald 1967, 1968). The technique requires that participants list their thoughts, one thought per line. Brock had judges code the listed thoughts for counterarguments to the point of view advocated in a communication, and then created an index that was the simple sum of the number of such counterarguments. Greenwald had the participants themselves categorize their thoughts as favorable or unfavorable to the point of view advocated in the communication, and then created an index that was the difference between the number of favorable and unfavorable thoughts. Both Brock and Greenwald found that their indices were related to the participants' postcommunication attitude. Greenwald, for example, found a positive correlation of 0.65 between his index and postcommunication attitude.

Even though the possible relevance of listed thoughts to cognitive bands was pointed out (Insko, 1981), the relevance was generally not recognized. Research with the listed-thought technique has revealed that some issue-relevant thoughts are explicit cognitive bands; for example, an explicit counterargument, and some others can be easily interpreted as implicit cognitive bands. For example, if following a communication advocating the building of nuclear power-generating plants, a respondent lists the thought "problem of radioactive waste," there is an implicit reference to a consistent cognitive band: "Nuclear power generating plants are bad (-) because they generate (+) dangerous radioactive waste (-)."

TWO-VALUED BALANCE AND TWO-VALUED LOGIC

It is clear that Heider regarded balance theory as an application of Gestalt principles rather than logic. For example, in his posthumously

published notes (Benesh-Weiner, 1988), Heider writes: "For balance: there is a gestalt quality (pleasantness, fittingness) correlated to a 'structure.' The structure is a relation between the relations, between the parts ..." (1988: 52).

In an influential article Abelson and Rosenberg (1958) distinguished between logic and psycho-logic (their term for balance). However, two logicians, Runkel and Peizer (1968) pointed out that if relations are restricted to two-valued, plus or minus, distinctions, the multiplicative rule is perfectly mapped by two-valued logic. Quoting directly: "Once it is realized that at most two categories are available for assigning perceptually associated elements, the practical distinction between psycho-logic and ordinary logic becomes superfluous from the point of view of the behavioral scientist" (1968: 61).¹

As illustrated by the familiar syllogism in which Socrates either is or is not mortal, it is important to recognize that deductive logic makes only two-valued distinctions, and thus in order for logic to map the implications of balance theory, balance-imbalance must also be restricted to two-valued distinctions. Working through the eight possible *p-o-x* triads makes this mapping fairly obvious. For example, restricting the similar-dissimilar dimension to the two values of same and different, it follows logically that if *a* is the same as *b*, and *b* is different from *c*, then *a* is different from *c*, consistent with the multiplicative rule (+ - -).

Seemingly deductive logic's two-valued assumption is not always understood. A well-known social psychologist once commented: " $5 \neq 6$ and $6 \neq 7$ does not imply that $5 = 7$." The problem here is that several degrees of inequality were assumed. Note that if only the values of equal and unequal are allowed, $a \neq b$ and $b \neq c$ does imply that $a = c$.

Abelson appears not to have been convinced by Runkel and Peizer (1968) because at a later time he critiqued balance theory as due to its historical dependence on a "weak ... linkage drawn between configurations of stimulus elements and configurations of

cognitive elements" (1983: 411). Harary (1983) replied to Abelson that the basis of balance theory is really Boolean algebra – the logical calculus of truth values that results in the same conclusions as the semantic approach to logic.

There appears to have been a general lack of appreciation of the logical basis of the multiplicative rule. Newcomb (1968), for example, developed a version of balance theory in which balance was restricted to triads, or semicycles, in which the *p* to *o* relation is positive. This, of course, departs from the multiplicative rule and logic. But are participants responsive to balance in semicycles in which the *p* to *o* relation is negative? Aronson and Cope (1968) found that just as participants inferred that my friend's friend is my friend, they also inferred that my enemy's enemy is my enemy (see also Gawronski and Walther, 2008, Experiment 2).

HENLE ON LOGICAL THINKING

The fact that the multiplicative rule is mapped by logic raises a question regarding whether thinking is logical. Some social psychologists have assumed that thinking is not logical. For example, Bruner et al. wrote: "Much of human reasoning is supported by a kind of thematic process rather than by an abstract logic" (1956: 104).

Henle (1962) points out that whether or not thought is logical has been an issue of dispute among philosophers. An older generation of philosophers, for example, John Stuart Mill (1874), regarded logic as descriptive of the laws of thought. However, more recent philosophers – for example, M.R. Cohen – have argued otherwise: "That the laws of logic are not the universal laws according to which we do actually think is conclusively shown not only by the most elementary observation of introspection, but by the very existence of fallacies" (Cohen 1944: 2–3).

Henle notes that this change in philosophical opinion appears to have been based more

on a cultural shift than on the discovery of new evidence, and that the older philosophers were also aware of error. For example, Mill wrote that a person "... has it almost always in his power to make the syllogism good by introducing a false premise; and hence it is scarcely ever possible decidedly to affirm that any argument involves a bad syllogism" (1874: 560).

The fact that the philosophical disagreement related to the interpretation of error provided Henle with a justification for collecting open-ended data on participants' reports of their thoughts. Psychology graduate students were presented with a series of syllogisms in the context of everyday discussions, and then asked to state whether the conclusion of each syllogism followed and to describe their reasoning. As expected errors did occur, but the stated reasons for the false judgments could be categorized as: (1) failure to accept the requested logical task; (2) alteration of the meaning of a premise; (3) omission of one or more premises; and (4) adding a premise.

Henle gives two interesting examples of the prevalence and importance of syllogisms in everyday life, syllogisms that Aristotle referred to as "practical syllogisms":

It is too far to walk to the Public Library; I must take a subway or bus. The fifth Avenue bus passes the Public Library. I do not want to wear the same dress two days in succession. I wore this dress yesterday; so I do not want to wear this dress today (1945: 374).

Henle argues that without reliance on the practical syllogism it is not obvious that people could cope with the "ordinary tasks of life ... understand each other, follow one another's thinking, reach common decisions, and work together" (1945: 374).

Despite such examples of explicit reasoning, Henle quotes, with apparent agreement, Aristotle's assertion that the reasoning in the practical syllogism may be implicit:

The mind does not stop and consider at all one of the two premises, namely, the obvious one; for example, if walking is good for a man, one does

not waste time over the premise "I am myself a man." Hence such things as we do without calculation, we do quickly" (1945: 701).

Although Henle does not discuss balance theory, it is arguable that syllogisms can be recast as semicycles. For Henle's two above examples we have balanced $+ - -$ semicycles: "Taking the fifth avenue bus (+) avoids (-) having to walk too far to the Public Library (-)," and "Wearing a different dress today (+) avoids (-) wearing the same dress two days in a row (-)." Quite possibly human thought more typically takes the form of semicycles than syllogisms. Perhaps this is why Bruner et al. (1956) made the above statement that human reasoning is more "thematic" than logical.

THE ASSUMPTION OF HIGH SELF-REGARD

One of the least developed aspects of balance theory relates to the self. Heider (1946) does indicate that his discussion of dyads and triads is based on an assumption of high self-regard:

High self regard of p can be expressed by pLp , low self regard by $pnlp$ (although the two p 's in this expression are not strictly equivalent). All of the examples so far considered presupposed pLp (1946: 111).

In addition to an unexplored assumption of high self-regard, Heider also seems to assume the existence of two related, but different, selves. Wiest (1965) makes a similar assumption when he argues that if p has a positive unit relation with the self-concept, or s , in order for balance to occur there should also be a positive sentiment relation with the self-concept, pLs .

Although not explicitly stated, both Heider and Wiest appear to agree with James' (1890) classic distinction between the self as "I" and the self as "Me," or the self as subject and the self as object. The self as subject is the self that is assumed to direct attention to perceptions

and thoughts and to multitask with a resulting increase in knowledge. The self as object, sometimes referred to as the self-concept, is the knowledge that is acquired regarding the person. Leary and Tangney (2003) relate the distinction between the two self-processes to the distinction between attention and cognition.

There is evidence that people tend to have high self-esteem, or high self-regard. This evidence appears to relate more obviously to the self as object than to the self as subject, but given Wiest's (1965) assumption of a positive unit relation between the two selves, they should be similarly evaluated. Support for unrealistically high self-regard has been documented by Taylor and Brown (1988), by Gilovich (1991), and by Alicke and associate's evidence for the "better than average" effect (e.g., Alicke, 1985). Alicke's evidence initially came from a demonstrated tendency for undergraduates to rate desirable traits as more characteristic of themselves than the average college student, and undesirable traits as less characteristic of themselves than the average college student – particularly if the traits were considered controllable.

Baumeister et al. (1989) obtained further evidence for the prevalence of high self-esteem. They found that for 20 different studies and 12 different self-esteem scales the scores were clustered toward the upper half of the possible range with the mean or median well above the mathematical midpoint. Baumeister et al. (1989) indicated that "low" self-esteem participants do not characterize themselves as possessing strongly negative traits, but tend to select the neutral-middle of most negative to positive dimensions with "somewhat" or "sometimes" or "average" responses.

Consistent with the above evidence, a cross-national study by Schmitt and Allik (2005) obtained striking support for the prevalence of high self-esteem. In this study the widely used Rosenberg (1979) Self-Esteem Scale (RSES) was translated into 28 different languages and administered to participants in 53 different nations. The results

indicated that, although there was cross-national variation, “All nations scored above the theoretical midpoint of the RSES, indicating generally positive self-evaluation may be culturally universal” (Schmitt and Allik, 2005: 623).²

The assumption of high self-esteem allows for balance interpretations of various effects, such as effects in the conformity and dissonance literatures discussed below. The following discussion will be limited to a single example relating to the so-called similarity-attraction effect.

Heider (1946, 1958) clearly assumes that if *p* and *o* are similar *p* should be attracted to *o*, and there is evidence (Byrne, 1971) that in many situations this generalization holds for attitudes and values. However, for traits the situation is less clear. Thus, for example, an introvert may or may not be more attracted to introverts than to extroverts. How can this apparent difference between attitudes and traits be explained? There is evidence that attraction is more closely related to similarity to ideal self than to similarity to actual self (LaPrell et al., 1990, 1991; Wetzel and Insko, 1974). Note that since in most instances attitudes are more readily changeable than are traits, actual and ideal attitudes are more likely confounded than are actual and ideal traits. Thus it follows that attraction should be more closely related to similar attitudes than to similar traits. This interpretation is consistent with Alicke’s (1985) above described evidence that the “better than average” effect held more obviously for personal traits that were controllable.

LOGIC AND HEDONISM

There is a conceptually interesting overlap between logic and a hedonistic, or reward-cost, perspective that is captured by the common reference to “rational economic decisions.” The implicit assumption of this statement is that some decisions are both hedonistic and rational in that they allow for the maximization

of outcomes by following the logical implication of behaving consistently with a positive self. Note that, given high self-regard, it is both balanced and hedonic for a positive self (+) to receive (+) a reward (+) or for a positive self (+) to avoid (−) a cost (−), and that it is both imbalanced and antihedonic for a positive self (+) to avoid (−) a reward (+) or for a positive self (+) to receive (+) a cost. The overlap between logic and hedonism stands in contrast to the frequent assumption, for example McGuire (1960), that wishful thinking and logical thinking are fundamentally different. The overlap between logic and hedonism suggests two issues that will be briefly considered. One of these relates to testability and the other to the evolutionary priority of logic and hedonism.

Testability

A classic issue associated with hedonism relates to its testability. To use a contemporary example, is the suicidal behavior of terrorists inconsistent with hedonism or rather an indication that terrorists are concerned with increasing the rewards of an afterlife by harming infidels? Because the hedonistic perspective does not generally specify for every circumstance what will be regarded as a reward or cost, the theory is in a general sense not testable. This issue, however, is easily misunderstood. Even though the general theory is not testable, what is testable are specific predictions regarding the rewards or costs that will control behavior for particular people in particular situations.

Given the conceptual overlap between logic and hedonism, it follows that the testability issue applies to the behavioral implications of both perspectives. Recall the above-described argument of Henle’s (1962) that the apparent occurrence of logical errors could be due to a variety of causes, such as the omission of a premise, and her quote of John Stuart Mill’s assertion that “...it is scarcely ever possible decidedly to affirm that any argument involves a bad syllogism” (Mill 1874: 560).

Evolutionary priority

Given the overlap between consistency with a positive self and hedonism, there is a fundamental question regarding which of these processes played a crucial role in evolutionary development? Clearly an organism (or its genes) that did not seek rewards (such as food) or avoid costs (such as predators) would be selected against. But does such behavior flow from consistency with a positive self, or hedonism, or something more primitive than either? Actually, it is difficult to imagine anything more primitive than a simple, consistency algorithm.

Note, however, that the postulation of such an algorithm seemingly requires a positive self, and it is implausible that simple one-celled organisms possess a self. Elsewhere (Insko, 1999: 133) I have speculated that the problem could be solved by assuming the emergence of a primal-positive sign that made survival logical. While the postulation of such a primitive process is highly speculative, it is arguable that without the emergence of such a primal sign, survival would not have been possible. Note also that the postulation of a primal positive sign is no more “off the wall” than an assumption that one-celled organisms are motivated by a desire to receive rewards and avoid costs.

If the survival of primitive life required a primal-positive sign, there is the further interesting possibility that through additional evolutionary development the primal-positive sign was elaborated into a more complex self. Although they do not postulate a primal-positive sign, such a possibility is consistent with Sedikides and Skowronski’s (1997) argument that rudimentary forms of the self exist in simpler, nonhuman animals.

STUDIES OF HYPOTHETICAL *p*–*o*–*x* SEMICYCLES

Beginning with Jordan (1953), a student of Heider’s, numerous studies were conducted in which participants rated the pleasantness

of each of the eight possible semicycles generated by positive or negative *p* to *o*, *p* to *x*, and *o* to *x* relations (+ + +, + – –, – + –, – – +, + + –, + – +, – + +, – – –). Researchers have generally followed Jordan’s precedent of letting the first sign refer to the *p* to *o* relation, the second to the *p* to *x* relation, and the third to *o* to *x* relation. As presented to participants “I” was substituted for “*p*” as in “I like *O*, I like *X*, and *O* likes *X*.”

Jordan found that mean pleasantness-unpleasantness ratings on a scale in which low numbers indicated pleasantness were as follows: 26.2 (+ + +), 39.5 (+ – –), 55.3 (– + –), 62.4 (– – +), 57.0 (+ + –), 58.2 (+ – +), 54.8 (– + +), 58.4 (– – –). These data indicate an approximate tendency for the first two semicycles (+ + + and + – –) to differ from the remaining six. The first two semicycles are semicycles in which *p* likes *o* and agrees with *o* by liking or disliking *x*. According to the multiplicative rule these two semicycles are balanced. However, the two further balanced semicycles, in which *p* dislikes *o* and disagrees with *o* (– + – and – – +), were not rated as particularly pleasant.

Zajonc (1968) summarized the literature on hypothetical triads by calculating attraction, agreement, and balance indices for each study. Each index was a ratio. For example, the attraction index was the ratio of the four semicycles in which *p* likes *o* to the four semicycles in which *p* dislikes *o*. A more familiar approach would have been to view the semicycles as the eight cells generated by a three-factor analysis of variance design in which the factors were *p* to *o*, *p* to *x*, and *o* to *x*. From this perspective the attraction index, or effect, is a main effect for the *p* to *o* factor such that semicycles in which *p* likes *o* were rated as more pleasant than semicycles in which *p* dislikes *o*. The agreement effect is a double interaction of the *p* to *x* and *o* to *x* factors such that the semicycles in which *p* and *o* agree by both liking or both disliking *x* were rated as more pleasant than the semicycles in which *p* and *o* disagree in their liking or disliking of *x*. And finally, the balance effect is the triple interaction among all three

factors such that the balanced semicycles in which p likes and agrees with o or dislikes and disagrees with o were rated as more pleasant than the imbalanced semicycles in which p dislikes and agrees with o or likes and disagrees with o .

Across all of the reviewed studies, Zajonc (1968) found that “the results are in favor of the agreement variable, with balance holding a close second, and attraction a decided third place” (1968: 347). Note that the tendency of the agreement effect to be descriptively larger than the attraction effect indicates that over all studies Jordan’s (1953) two versus six pattern did not exactly replicate because in the four cells in which attraction and agreement were in opposition (+ + -, + - +, - + +, - - -), agreement was rated more pleasant than attraction.

The results for Zajonc’s three indices led him to conclude that: “This rough summary is damaging to the balanced principle” (1968: 347). Indeed it is the case that the manipulated p - o - x semicycle does not adequately account for the results. However, as was observed above, it is frequently an oversimplification to assume that the only relevant semicycle is the semicycle that is observed or manipulated. The argument here parallels Henle’s (1962) above-described argument that the occurrence of an apparent error in the judgment of a syllogism does not necessarily indicate that illogical thinking has occurred because the participant may, for example, add one or more additional premises. In the present context the argument is that participants may assume additional relations that create one or more additional semicycles. Beginning with a suggestion by Aderman (1969) regarding contact, these possibilities are reviewed below.³

The assumption of p - o contact

Aderman (1969) suggested that when the typical participant considers a hypothetical p - o - x semicycle, he or she tends to assume contact between p and o . Although Aderman

did not explicitly develop the implications of this insight, a positive unit relation between p and o would create two additional semicycles in addition to the manipulated semicycle. These are an attraction-contact semicycle that is balanced when p likes and has contact with o and an agreement-contact semicycle that is balanced when p agrees with and has contact with o . Consideration of all three semicycles indicates that the only cells in which all semicycles are balanced are the two + + + and + - - cells that Jordan (1953) found to be more pleasant than the remaining six cells. In each of the remaining six cells only one of the three semicycles is balanced. For example in the - + + cell the agreement semicycle is balanced but the attraction semicycle and the manipulated semicycle are imbalanced.

An initial question regarding the contact hypothesis is whether participants who rate the hypothetical situations do, in fact, assume contact between themselves and the other person. Research with postexperimental questionnaires (Insko, 1984) has found that 72 to 91 percent of the participants indicated that they did assume such contact when making their ratings. Further experimental studies were conducted in which ratings were made on five scales (pleasantness, harmony, expectancy, consistency, stability) and participants were instructed to make certain assumptions regarding contact or, in a standard condition, were given no instructions regarding contact. The results of these studies were at least somewhat supportive of the contact hypotheses. Most notably, assumed breaking p and o contact reversed the direction of the agreement effect in three of four studies and reversed the direction of the attraction effect in all four studies.

While the results, particularly the breaking-contact results, imply that assumed contact does play a role in accounting for attraction and agreement effects, clearly more than contact is involved. First, and perhaps most obvious, even when participants were instructed to assume no contact between p and o , the agreement effect was significant in

three of the four experiments and the attraction effect was significant in one of the four experiments. Second, the contact interpretation cannot account for the fact that the agreement effect has been routinely found to be descriptively larger than the attraction effect. Third, a comparison among the five different rating scales indicated that the attraction and agreement effects were larger on the affective scales (pleasantness, harmony) than on the more cognitive scales (consistency, expectancy, stability). There is no obvious way in which the attraction-contact and agreement-contact semicycles can account for this difference.

Assumed o to p sentiment and the concern with being liked

The typical study of $p-o-x$ semicycles manipulated the p to o sentiment relation but not the o to p sentiment relation. However, there should be a balance tendency to assume reciprocal o to p sentiment in the $p-o$ dyad, or semicycle, and also to assume agreement-consistent o to p sentiment in the o to p , o to x , p to x triad, or semicycle. Thus the balanced-implied assumption of o to p sentiment creates additional semicycles in addition to the manipulated semicycle. Furthermore, research has indicated that when participants are asked to rate o to p liking they report more such liking when p likes o than when p dislikes o and when p agrees with o than when p disagrees with o (Insko, 1984). However, an additional semicycle is still implied.

The additional semicycle is a semicycle that includes the self as an element and that assumes the transformation of the o to p sentiment relation into an element, being liked or being disliked. Note that it is balanced for a positive self (+) to be (+) liked (+), and that it is imbalanced for a positive self (+) to be (+) disliked (-), and that this difference suggests a concern with being liked.

The postulated concern with being liked agrees with the assumption of sociometry

theory that the perception that one is liked by others leads to more positive self-evaluation. Although not typically interpreted as a consistency effect, there is an abundance of experimental evidence (e.g., Leary and Baumeister, 2000; Leary et al., 2003), and also nonexperimental evidence (e.g., Srivastava and Beer, 2005) in agreement with this assumption.

In the context of research with hypothetical situations, the self-relevant semicycle implying a concern with being liked is particularly important because it suggests a reason why attraction and agreement effects should be larger with ratings on affective scales than with ratings on cognitive scales. Note that while consistency may or may not relate to the self, affect exists only in relation to a self or organism. Thus an affective rating should be particularly sensitive to semicycles including the self as an element.

The research literature (summarized in Insko, 1984) includes two experiments testing the above reasoning. One experiment included a condition in which o to p liking or disliking was consistent with agreement or disagreement, and the other included a condition in which o to p liking or disliking was consistent with p to o liking or disliking. The first experiment found that the explicit statement of o to p liking consistent with agreement increased the agreement effect relative to a standard condition, and did so primarily on the affective scales. The second experiment found that the explicit statement of o to p liking consistent with p to o liking increased the attraction effect and also did so primarily on the affective scales. Thus the assumption of o to p sentiment provides a possible explanation of the agreement and attraction effects and also that such effects are larger on affective than cognitive scales.

Assumption that o is right and the concern with being right

Although the concern with being liked provides a possible explanation as to why

attraction and agreement effects should be larger on affective than cognitive scales, it does not account for the fact that agreement effects have repeatedly been found to be larger than attraction effects. However, social-comparison theory (Festinger, 1950, 1954a, 1954b) suggests a consideration that could create an additional semicycle relating just to agreement, and thus accounting for the larger size of the agreement effect.

According to Festinger (1950, 1954a, 1954b), people have a drive to hold correct opinions. The drive can sometimes be satisfied through simple observation (i.e., by referring to physical reality), but when this cannot be done, as for example with political, religious, or moral opinions, the drive is satisfied by agreeing with others (i.e., by referring to social reality).

As applied to hypothetical semicycles, the implication is that o is right about x and that by agreeing with o , p is also right: o is (+) right, p agrees with (+) o , p is (+) right. But when considering hypothetical $p-o-x$ situations, do participants assume that the other person has some basis for his or her evaluation of x ? A single study found that one-third did not, but that two-thirds did (see Insko, 1984).

For participants who do assume that the other person has a basis for his or her evaluation of x , there is the possibility of a self-relevant semicycle. Note that it is balanced for a positive self (+) to be (+) right (+), and that it is imbalanced for a positive self (+) to be (+) wrong (-), and that this difference suggests a concern with being right.

A single study that manipulated o 's knowledge of x altered the magnitude of the agreement effect and thus provided an explanation as to why agreement effects are larger than attraction effects. However, the increased magnitude of the agreement effect was not larger on the affective than the cognitive scales and thus suggests that if the semicycle is hypothetical there is little or no salience of the self-relevant concern with being right. The following section will review some evidence that in nonhypothetical

situations the concern with being right can play a role.

CONFORMITY

Normative and informational social influence

Deutsch and Gerard (1955) are widely referenced for interpreting conformity as due to two different factors: normative social influence and informational social influence. They define normative influence somewhat abstractly as “influence to conform to the positive expectations of another,” and they define informational influence as “influence to accept information obtained from another as *evidence* about reality” (1955: 629).

The concerns with being liked and being right

Although Deutsch and Gerard (1955) did not explicitly relate normative and informational social influence to the self, there is an obvious similarity between normative social influence and self-relevant concern with being liked (or not disliked), and between informational social influence and self-relevant concern with being right (or not being wrong).

To examine the role of these two concerns in a nonhypothetical context Insko et al. (1983); and Insko et al. (1985) conducted two different experiments. Both experiments followed the Asch (1952) procedure of having a series of confederates and a single participant seated in a row of chairs, with the actual participant seated next to last. Unlike in Asch's research the judgments were not of line lengths but of colors, for example, whether the blue-green in the middle was more like the blue on the left or the green on the right. Confederates responded incorrectly on a subset of the trials.

Both experiments included two factors. One factor related to whether the participants

responded publicly or privately. In the private condition the participants wrote their judgments on booklets that were to be dropped in a box with other unsigned booklets and "stirred" by the experimenter. The other factor related to whether the correct judgments could be determined, or not determined, by the readings of a color spectrometer.

In the determined condition, but not the undetermined condition, it was explained that the spectrometer readings would be available at the end of the experiment. If the possibility of being proven right or wrong created a greater concern with being right and if, as social-comparison theory implies, group judgments were regarded as a source of information, it follows that there should have been more conformity in the determined than the undetermined condition.

The results revealed significant main effects for both factors. The public-private main effect revealed greater conformity in the public condition than in the private condition. The determined versus undetermined main effect revealed greater conformity in the determined condition than in the undetermined condition. These main effects are consistent with both the concern with being liked and the concern with being right.

Baumeister (1982) has argued that many social phenomena, including conformity, can be explained as a manifestation of concern with self-presentation. He pointed out that "the most common procedure for testing for self-presentational motives is by comparing two situations that are identical in all respects except that some circumstance is public in one situation but private in the other" (1982: 4), and he emphasized that in order for a circumstance to be truly private, the participant must believe that the experimenter is unaware of how he or she performed.

In the private condition of the two Insko et al. (1983, 1985) experiments, participants anticipated that their booklets containing unsigned color judgments were to be dropped in a box containing other unsigned booklets and then "stirred" by the experimenter.

In view of the private nature of the private condition it is important to note that the interaction between the determined versus undetermined factor and the private versus public factor was not significant in either experiment. If the effect for the determined versus undetermined factor had been due to a concern with self-presentation, there should have been an interaction with the private versus public factor such that the difference between the determined and undetermined means was greater in the public condition. The nonsignificance of the interaction does not support an interpretation that the concern with being right was associated with, or dependent on, a self-presentational concern.

Although the evidence is inconsistent with the possibility that the concern with being right was dependent on a self-presentational concern, it is still possible that the public-private main effect was due, not just to an intrinsic concern with one's self-worth, but to appearing worthy to others. Stated differently, the greater conformity in public than private situations may have been due to a concern with "looking good," to a concern with "being good," or to both.

A possible role for contact

The above-described studies of hypothetical *p-o-x* semicycles provided some evidence that assumed *p* to *o* contact played at least a minor role in producing agreement effects. Circumstantial evidence that contact may also be important in nonhypothetical situations comes from the fact that in the two Insko et al. (1983, 1985) experiments the number of conformity-related errors in the private-undetermined cell, while relatively low, was nonetheless greater than the number that occurred in a control-no-influence condition. More definitive evidence for the role of contact would require the comparison with a condition in which the group judgments were, for example, presented on videotape.

Self-relevant individualistic and collectivistic values

Although contact and the twin concerns with being liked and being right may partially account for conformity effects, there is convincing evidence that these variables do not provide a complete explanation. Bond and Smith (1996) conducted a meta-analysis of cross-cultural line-judgment studies similar to those used by Asch (1952, 1956) and found that conformity was related significantly to Hofstede's (1980) index of individualism-collectivism. Conformity was higher in collectivistic cultures.

What makes the Bond and Smith evidence particularly relevant to the present argument is the possibility that individualistic and collectivistic values are related to the self. Markus and Kitayama (1991) argue that members of individualistic cultures base self-esteem on the "ability to express self, validate internal attributes" (1991: 230, Table 1), while members of collectivistic cultures base self-esteem on the "ability to adjust, restrain self, maintain harmony with social context" (1991: 230, Table 1). Consistent with Markus and Kitayama's interpretation, Sedikides et al. (2003) report evidence relating the differential association of self-positivity to individualistic and collectivistic behaviors and traits. They compared the self-positivity of US college students of Japanese or non-Japanese backgrounds on reported hypothetical individualistic or collectivistic behaviors (e.g., "Put yourself before your group" or "Defend your group's decisions," [1991: 64, Table 1]) and traits (e.g., "independent" or "loyal," [1991: 64, Table 2]). Following Alicke's (1985) finding of a "better than average" effect, self-reports compared own behavior or traits with the behavior or traits of the supposed typical other group member. The results indicated that students of Japanese backgrounds saw themselves as relatively more positive than others to a greater extent on collectivistic than individualistic behaviors and traits while students of non-Japanese backgrounds saw themselves

as relatively more positive than others to a greater extent on individualistic behaviors and traits than collectivistic behaviors and traits.

The conformity dilemma

Both the concern with being liked and the concern with being right could be interpreted as flowing from consistency with positive self-esteem and thus as implying that self-esteem should be positively correlated with conformity. In view of the persuasive evidence supporting sociometry theory, this possibility is particularly compelling to the extent that conformity flows from the concern with being liked.

Even though the evidence consistent with sociometry theory could be interpreted as implying that self-esteem should be positively associated with conformity, the literature on individualism and collectivism implies that for people with individualistic values self-esteem should be negatively associated with conformity. If both hypothesized processes occur, one would expect that for people with individualistic values the dilemma posed by conformity pressure would be particularly acute. The dilemma is between maintaining high self-esteem by giving in to conformity pressure and obtaining the assumed approval of the group or maintaining high self-esteem by resisting conformity pressure and behaving consistently with individualistic values. On the other hand, for individuals with collectivistic values, the dilemma posed by conformity pressures should be much less acute.⁴

There is indirect evidence that participants in Western conformity experiments do recognize the dilemma. Allen (1975: 18) reports some unpublished data indicating that perceived deviation from group consensus in the Johnny Rocco case (Schachter, 1951) led to the anticipation of being voted out of the group, and Sabini et al. (2001) cite Jahoda (1959) as reporting that both conforming and nonconforming participants in

the Asch (1952, 1956) studies were upset by the experience.

IMBALANCE RESOLUTION AND DISSONANCE REDUCTION

Aronson (1968), Steele (1988), Steele and Liu (1983, and Schlenker (1982), have all seen the self as related to dissonance. However, none of these social psychologists extended their interpretations to include imbalance in self-related semicycles. The relevance of imbalance in self-related semicycles to dissonance can be illustrated with two examples from the literature on so-called free-choice situations – situations in which the participants chooses between two desirable alternatives.⁵

The first investigation of free-choice situations was a study by Brehm (1956). Consistent with the theoretical argument that a choice between desirable alternatives creates dissonance, Brehm reported that the chosen alternative increased in rated desirability and the rejected alternatives decreased in rated desirability. Such so-called spreading of the alternatives was interpreted as evidence for dissonance reduction. However, later research by Shultz et al. (1999) found that the spreading was due, not to the chosen alternative increasing in desirability, but to the rejected alternative decreasing in desirability. Such a pattern is consistent with a balance interpretation.

Note that for a positive self (+) to choose (+) a desirable alternative (+) is consistent and thus does not provide a two-valued basis for change. However, for a positive self (+) to reject (–) a desirable alternative (+) is imbalanced and does provide a basis for change. The balance-achieving change could involve lowering self-esteem but that would create inconsistency in the self-concept. Thus the simplest route to imbalance resolution would involve decreasing the perceived desirability of the rejected alternative.⁶

Further evidence consistent with a balance interpretation of alternative spreading comes from Brock's (1963) finding that spreading of the alternatives was greater when the alternatives were objectively dissimilar than when the alternatives were objectively similar. Note quite simply that objective similarity is a positive unit relation and that when the alternatives are perceived as associated by a positive sign different evaluation of the alternatives would create an imbalanced semicycle.

BEYOND TWO-VALUED TO MANY-VALUED RELATIONS – THE TETRAHEDRON MODEL

Although logic is two-valued, human thought is obviously capable of many-valued distinctions. How then can logic's two-valued distinctions be extended to many-valued distinctions? One possibility is by combining logic and probability theory as in the probabilogical models of McGuire (1960) and Wyer (1974). A potential objection to these models is that they relied on traditional formulations of logic when, as argued above, it is possible that the multiplicative rule provides a better description of human thought than does the syllogism.

An approach that relies on a corollary of the multiplicative rule is Rosenberg's (1956, 1960), above-described technique of multiplying the bipolar ratings of the value and instrumental cognition within a cognitive band. This approach was successful in demonstrating a correlation between the sum of such multiplications across cognitive bands and evaluation of the central attitude in the structure. However, insofar as the concern is with predicting a specific evaluation of the central attitude there is an obvious problem in that the multiplication will, in some instances, result in a product that goes off the scale. Fortunately, this problem does not occur with the approach of the tetrahedron model, a model first suggested by

Wiest (1965) but given mathematical form by Wellens and Thistlethwaite (1971a, 1971b).

Consider each of the three relations in the p - o - x semicycle, not just as having either plus or minus values, but as having more than two values on a bipolar scale, for example $-2, -1, 0, +1, +2$. Wiest associated each of these relations with one of the left-right, down-up, and near-far dimensions of a cube. Any cube, of course, has eight corners, four on the "top" and four on the "bottom." Wiest assumed that the corners of the cube represented the pure, two-valued possibilities and thus that four of the corners ($+++$, $++-$, $-+-$, $--+$) were balanced and the remaining four were imbalanced. Connecting the balanced corners inside the cube creates a tetrahedron – a spatial form that can be thought of as a three-sided pyramid. Wiest generalized beyond the two-valued corners by assuming that any point on the surface of the tetrahedron or inside the tetrahedron was balanced, and thus suggested an elegantly simple solution to the perplexing problem as to how to generalize beyond two-valued possibilities.

It is, of course, arbitrary which relation is associated with which spatial dimension, but suppose that the p - o relation is associated with the down-up dimension with negative values down and positive values up. Connecting any two values of the p - x and o - x dimensions on the top of the cube creates a point that can be projected with a straight line into the cube. This line will intersect the tetrahedron at its upper boundary, proceed through the tetrahedron, and emerge from the tetrahedron at its lower boundary. Wellens and Thistlethwaite (1971a, 1971b) wrote one equation that, given any paired values for the p - x and o - x relations, predicts the upper-boundary intersection and a second equation that predicts the lower-boundary intersection.

Because Wiest assumed that any point on the surface of the tetrahedron or inside the tetrahedron was balanced, it follows that any predicted value for p - o from the upper

boundary down to and including the lower boundary would be balanced. Since, however, for paired p - x and o - x values that are not highly polarized the specified area inside the tetrahedron is large, the range of balanced possibilities is also large. A simple solution to this problem is to assume that the most probable predicted value is midway between the upper and lower boundaries, or is a simple average of the upper- and lower-boundary predictions. This approach to combining the upper- and lower-boundary predictions creates Wellens and Thistlethwaite's equal weights model.

Other models are, of course, possible, but Wellens and Thistlethwaite present only one, a so-called unequal weights model in which the upper-boundary prediction is weighted 0.75 and the lower-boundary prediction is weighted 0.25. How do the predictions for the equal and unequal weights models differ? If, as assumed above, the upper end of the p - o dimension represents the positive values, the predicted p - o values for less polarized p - x and o - x pairings are more positive or less negative for the unequal weights model than for the equal weights model. The unequal-weights model predicts less p to o disliking as a function of p and o disagreement than does the equal weights model.

Research by Wellens and Thistlethwaite (1971a, 1971b), and Tashakkori and Insko (1979, 1981) in which the p - o relation (attraction) was the dependent variable has supported the unequal weights model. Wellens and Thistlethwaite used a procedure in which participants played the role of a p who held each of five evaluations ($-2, -1, 0, +1, +2$) regarding a student demonstration, and considered an o who also held each of the five evaluations. Tashakkori and Insko used a variation of the Byrne (1971) anonymous stranger technique. In this latter research participants initially spent some time at a computer responding to a large number of attitude items that had been pilot tested to represent the different levels of the p - x variable. After random assignment of the participants to one cell of the p to x and o to

x design the computer selected five x 's from each participant's prior responses, and asked the participants to rate their liking of the other person whose evaluations of these x 's were given. As indicated above, all of this research supported the unequal weights model.

Why with a p - o , or attraction, dependent variable should the unequal weights model have produced the superior fit? A possible explanation parallels the above postulated concern with being liked (or not being disliked). Recall that in the case of p and o disagreement the unequal weights model predicts less disliking than does the equal weights model. Note that p to o disliking should result in reciprocal o to p disliking in the p - o , o - p semicycle, and that transforming the o to p disliking relation into an element, being disliked, would produce imbalance in a self-relevant semicycle, I (+) am (+) disliked (−). Stated more simply, it is the reluctance to entertain the possibility of negative self-evaluation that results in the reluctance to assume extreme dislike of o .⁷

The problem with modeling the p - o - x semicycle is familiar to examining a single semicycle in isolation from other semicycles that may be implicitly, or explicitly, involved. In research in which the p to o relation is the dependent variable the argument is that it is the reluctance to assume being disliked that shifts p to o sentiment toward the more positive, or less negative, upper boundary and thus provides a superior fit for the unequal weights model. The argument here parallels the above argument that the concern with being liked (or not disliked) partially accounts for conformity effects and for agreement effects in hypothetical semicycles.

Tashakkori and Insko (1981) reasoned that they could block the tendency to avoid implied dislike by manipulating o to p and testing for o to x . A single experiment found, as expected, that the results fit the equal weights model.

Although these results are preliminary and research with other semicycles should be

conducted, the findings do suggest that the tetrahedron model has promise. Understanding how to extend the multiplicative rule to other than two-valued distinctions and understanding how to conceptualize low self-regard are the two outstanding problems confronting further theoretical development.

NOTES

1 Philosophers use the term "fuzzy logic" to refer to a controversial form of many-valued logic (cf. Haack, 1978: 165–167; Haack, 1996: 229–258). Haack (1978) characterizes fuzzy logic as partially related to fuzzy set theory according to which membership in a set is not restricted to present or absent but a matter of degree represented by a real number between 0 and 1. For example, if person A belongs to the degree 0.2 to the set of tall people, it follows in fuzzy logic that the proposition "A is tall" has the value 0.2, or has a low degree of truth. Haack (1996: 230) stated that 'I remain convinced, first ... that truth does not come in degrees, and, second that fuzzy logic is not a viable competitor of classical logic.'

2 Despite the fact that low self-esteem is only low in a relative sense, and may be more characteristic of a short-term "state" than a long-term "trait" (Heatherton and Polivy, 1984), low self-esteem is nonetheless, as Baumeister (1995) suggests, a "puzzle." This puzzle is amply illustrated by Swann et al.'s (1987) finding that while participants low in social self-esteem consistently judged unfavorable feedback regarding their body language as more valid than favorable feedback, they reacted to such unfavorable feedback with relatively more negative affect. Such inconsistency between affect and cognition is indeed a puzzle and merits further study. Does the negative affect relate to the implied difficulty of bolstering long-term, trait self-esteem?

3 In view of the fact that responses to hypothetical situations may differ from responses to nonhypothetical situations, some readers may wonder why suggestions regarding hypothetical p - o - x semicycles merit investigation. This question has at least two answers. First, in view of the fact that the influential Zajonc critique (1968) was based on results from the study of hypothetical p - o - x semicycles, it is theoretically important to thoroughly explore those semicycles. Second, although nonhypothetical situations may be of more practical significance than hypothetical situations, balance theory is a theory regarding thought processes regardless as to whether those processes relate to practical-nonhypothetical situations or impractical-hypothetical situations.

4 Since everyone has learned to rely on the evidence of their senses, the dilemma, even though less acute for individuals with collectivistic values, should still be present.

5 For evidence relating to the insufficient-reward effect (Festinger and Carlsmith, 1959) see Insko et al. (1975).

6 The inconsistency with positive self-interpretation has an interesting fit with Baumeister et al.'s (2001: 323) extensive documentation that "bad is stronger than good" and that self-protective motivation is stronger than self-enhancement motivation.

7 The above interpretation of the superior fit of the unequal weights model does require that a relation be transformed into an element, but since ordinary speech routinely transforms verbs into nouns this requirement does not pose any particular difficulty. Note, further, that the interpretation does not require that the participant go through an explicit reasoning process. Although the reasoning process could be explicit, it is more likely that reasoning regarding any self-relevant semicycle is implicit and automatic. The argument parallels Aristotle's (1945) above described assumption that in deducing the implications of a practical syllogism we do not bother with an explicit consideration of self-relevant premises.

REFERENCES

Abelson, R.P. (1959) Modes of resolution of belief dilemmas. *Journal of Conflict Resolution*, 3, 343–352.

Abelson, R.P. (1983) Whatever became of consistency theory. *Personality and Social Psychology Bulletin*, 9, 37–54.

Abelson, R.P. and Rosenberg, M.J. (1958) Symbolic psycho-logic: A model of attitudinal cognition. *Behavioral Science*, 62, 990–1008.

Aderman, D. (1969) Effects of anticipating future interaction on the preference for balanced states. *Journal of Personality and Social Psychology*, 11, 214–219.

Allen, V.L. (1975) Social support for nonconformity. In L. Berkowitz (ed.), *Advances in Experimental Social Psychology*, 8, 1–43. New York: Academic Press.

Alicke, M.D. (1985) Global self-evaluation as determined by the desirability and controllability of trait adjectives. *Journal of Personality and Social Psychology*, 49, 1621–1630.

Anderson, N.H. (1981) Integration theory applied to cognitive responses and attitudes. In R.E. Petty, T.M. Ostrom and T.C. Brock (eds), *Cognitive Responses in Persuasion*. Hillsdale, NJ: Erlbaum.

Aristotle (1945) *Movement of Animals* (E.S. Forster, Trans.). Cambridge: Harvard University Press.

Aronson, E. (1968) Dissonance theory: Progress and problems. In R.P. Abelson, E. Aronson, W.J. McGuire, T.M. Newcomb, M.J. Rosenberg and P.H. Tannenbaum (eds), *Theories of Cognitive Consistency: A Sourcebook*, pp. 5–27. Chicago: Rand McNally.

Aronson, E. and Cope, V. (1968) My enemy's enemy is my friend. *Journal of Personality and Social Psychology*, 8, 8–12.

Aronson, E. and Worchel, P. (1966) Similarity versus liking as determinants of interpersonal attractiveness. *Psychonomic Science*, 5, 157–158.

Asch, S.E. (1952) *Social Psychology*. Englewood Cliffs, NJ: Prentice-Hall.

Asch, S.E. (1956) Studies of independence and conformity: A minority of one against a unanimous majority. *Psychological Monographs*, 70(46).

Baumeister, R.F. (1982) A self-presentational view of social phenomena. *Psychological Bulletin*, 91, 3–26.

Baumeister, R.F. (2011) *Self Esteem: The Puzzle of Low Self-regard*. New York: Plenum Press.

Baumeister, R.F., Bratslavsky, E., Finkenauer, C. and Vohs, K.D. (2001) Bad is stronger than good. *Review of General Psychology*, 5, 323–370.

Baumeister, R.F., Tice, D.M. and Hutton, D.G. (1989) Self-presentation motivations and personality differences in self-esteem. *Journal of Personality*, 56, 571–598.

Benesh-Weiner, M. (ed.) (1988) *Fritz Heider: The Notebooks: Vol. 4: Balance Theory*. Weinheim: Springer-Verlag.

Berscheid, E. and Walster, P. (1978) *Interpersonal Attraction*. Menlo Park, CA: Addison-Wesley.

Bond, R. and Smith, P.M. (1996) Culture and conformity: A meta-analysis using Asch's (1952b, 1956) line judgment task. *Psychological Bulletin*, 119, 111–137.

Brehm, J. (1956) Postdecision changes in the desirability of alternatives. *Journal of Abnormal and Social Psychology*, 60, 379–383.

Brock, T.C. (1963) Effects of prior dishonesty on post-decision dissonance. *Journal of Abnormal and Social Psychology*, 66, 325–331.

Brock, T.C. (1967) Communication discrepancy and intent to persuade as determinants of counterargument production. *Journal of Experimental Social Psychology*, 3, 296–309.

Bruner, G., Goodnow, J.J. and Austin, G. (1956) *A Study of Thinking*. New York: Wiley.

Byrne, D. (1971) *The Attraction Paradigm*. New York: Academic Press.

Byrne, D. and Griffitt, W. (1966) Similarity versus liking: A clarification. *Psychonomic Science*, 6, 295–296.

Byrne, D. and Rhamey, R. (1965) Magnitude of positive and negative reinforcements as a determinant of attraction. *Journal of Personality and Social Psychology*, 2, 884–889.

Carnegie, D. (1937) *How to Win Friends and Influence People*. New York: Simon and Schuster.

Cartwright, D. and Harary, F. (1956) Structural balance: A generalization of Heider's theory. *Psychological Review*, 63, 277–293.

Cialdini, R.B. (2009) *Influence: Science and Practice*. New York: Allyn and Bacon.

Cohen, M.R. (1944) *A Preface to Logic*. New York: Holt.

Deutsch, M. and Gerard, H.B. (1955) A study of normative and informational social influences upon individual judgment. *Journal of Abnormal and Social Psychology*, 51, 629–636.

Festinger, L. (1950) Informal social communication. *Psychological Review*, 57, 271–282.

Festinger, L. (1954a) A theory of social comparison processes. *Human Relations*, 7, 117–140.

Festinger, L. (1954b) Motivation leading to social behavior. In M.R. Jones (ed.), *Nebraska Symposium on Motivation*, 2, 121–218. Lincoln, NE: University of Nebraska Press.

Festinger, L. and Carlsmith, J.M. (1959) Cognitive consequences of forced compliance. *Journal of Abnormal and Social Psychology*, 58, 203–211.

Fishbein, M. (1963) An investigation of the relationships between beliefs about an object and the attitude toward that object. *Human Relations*, 16, 233–240.

Gawronski, B. and Walther, E. (2008) The TAR effect: When the ones who dislike become the ones who are disliked. *Personality and Social Psychology Bulletin*, 34, 1276–1289.

Gilovich, T. (1991) *How We Know What Isn't So*. New York: Free Press.

Greenwald, A.G. (1967) An amended learning model of persuasion. Paper presented at the meeting of the American Psychological Association, Washington, DC.

Greenwald, A.G. (1968) Cognitive learning, cognitive responses to persuasion, and attitude change. In A.G. Greenwald, T.C. Brock and A.T.M. Ostrom (eds), *Psychological Foundations of Attitudes*. New York: Academic Press.

Haack, S. (1978) *Philosophy of Logics*. Cambridge: Cambridge University Press.

Haack, S. (1996) *Deviant Logic, Fuzzy Logic: Beyond the Formalism*. Chicago: University of Chicago Press.

Harary, R. (1983) Consistency theory is alive and well. *Personality and Social Psychology Bulletin*, 9, 60–64.

Heider, F. (1946) Attitudes and cognitive organization. *Journal of Psychology*, 21, 107–112.

Heider, F. (1958) *The Psychology of Interpersonal Relations*. New York: Wiley.

Henle, M. (1962) On the relation between logic and thinking. *Psychological Review*, 69, 366–378.

Hofstede, G. (1980) *Cultural Consequences: International Differences in Work-related Values*. Beverly Hills, CA: Sage.

Insko, C.A. (1981) Balance theory and phenomenology. In R. Petty, T. Ostrom and T. Brock (eds), *Cognitive Responses in Persuasion*, pp. 309–338. New Jersey: Erlbaum.

Insko, C.A. (1984) Balance theory, the Jordan paradigm, and the Wiest tetrahedron. In L. Berkowitz (ed.), *Advances in Experimental Social Psychology*, 18, 89–140. New York: Academic Press.

Insko, C.A. (1999) A balance-logic perspective on Kruglanski and Thompson's single-route approach to persuasion. *Psychological Inquiry*, 10, 127–137.

Insko, C.A., Blake, R.R., Cialdini, R.B. and Muliak, S.A. (1970) Attitude toward birth control and cognitive consistency: Theoretical and practical implications of survey data. *Journal of Personality and Social Psychology*, 16, 228–237.

Insko, C.A., Drenan, S., Solomon, M.R., Smith, R.H. and Wade, T. (1983) Conformity as a function of the consistency of positive self-evaluation with being liked and being right. *Journal of Experimental Social Psychology*, 19, 341–358.

Insko, C.A., Smith, R.H., Alicke, M.D., Wade, J. and Taylor, S. (1985) Conformity and group size: The concern with being right and the concern with being liked. *Personality and Social Psychology Bulletin*, 11, 41–50.

Insko, C.A., Thompson, V.D., Wtroeb, W., Shaud, K.F., Pinner, B.E. and Layton, B.D. (1973) Implied evaluation and the similarity-attraction effect. *Journal of Personality and Social Psychology*, 25, 297–308.

Insko, C.A. and Wilson, M. (1977) Interpersonal attraction as a function of social interaction. *Journal of Personality and Social Psychology*, 35, 903–911.

Insko, C.A., Worchel, S., Folger, R. and Kutkus, A. (1975) A balance theory interpretation of dissonance. *Psychological Review*, 82, 169–183.

James, W. (1890) *Principles of Psychology*. New York: Dover.

Jahoda, M. (1959) Conformity and obedience. *Human Relations*, 12, 99–120.

Jordan, N. (1953) Behavioral forces that are a function of attitude and of cognitive organization. *Human Relations*, 6, 273–287.

LaPrelle, J., Hoyle, R.H., Insko, C.A. and Berenthal, P. (1990) Interpersonal attraction and descriptions of the traits of others: Ideal similarity, self similarity, and liking. *Journal of Research in Personality*, 24, 216–240.

LaPrelle, J., Insko, C.A., Cooksey, L. and Graetz, K. (1991) Self similarity, ideal similarity, and undesired self similarity as indicators of attraction. *Representative Research in Personality*, 19, 65–74.

Leary, M.R. and Baumeister, R.F. (2000) The nature and function of self-esteem: sociometer theory. In M.P. Zanna (ed.), *Advances in Experimental Social Psychology*, 32, 1–62. New York: Academic Press.

Leary, M.R., Gallagher, B., Fors, E., Buttermore, N., Baldwin, E., Kennedy, K. et al. (2003) The invalidity of disclaimers about the effects of social feedback on self-esteem. *Personality and Social Psychology Bulletin*, 5, 623–636.

Leary, M.R. and Tangney, J.P. (2003) The self as an organizing construct in the behavioral and social sciences. In M.R. Leary and J.P. Tangney (eds), *Handbook of Self and Identity*, pp. 3–14. New York: Guilford.

Markus, H.R. and Kitayama, S. (1991) Culture and the self: Implications for cognition, emotion and motivation. *Psychological Review*, 98, 224–253.

McGuire, M.J. (1960) A syllogistic analysis of cognitive relationships. In C.I. Hovland and M.J. Rosenberg (eds), *Attitude Organization and Change: An Analysis of Consistency Among Attitude Components*, pp. 65–111. New Haven, CT: Yale University Press.

Miller, N.E. (1951) Learnable drives and rewards. In S.S. Stevens (ed.), *Handbook of Experimental Psychology*, pp. 435–472. New York: Wiley.

Mill, J.S. (1874) *A System of Logic*. New York: Harper.

Mita, T.H., Dermer, M. and Knight, J. (1977) Reversed facial images and the mere exposure hypothesis. *Journal of Personality and Social Psychology*, 35, 597–601.

Montoya, M. and Insko, C.A. (2007) Toward a more complete understanding of the reciprocity of liking effect. *European Journal of Social Psychology*, 38, 477–498.

Newcomb, T.M. (1968) Interpersonal balance. In R.P. Abelson, E. Aronson, W.J. McGuire, T.M. Newcomb, M.J. Rosenberg and P.H. Tannenbaum (eds), *Theories of Cognitive Consistency: A Sourcebook*. Chicago: Rand McNally.

Osgood, C.E. and Tannenbaum, P.H. (1955) The principle of congruity in the prediction of attitude change. *Psychological Review*, 62, 42–55.

Peak, R. (1955) Attitude and motivation. In M.R. Jones (ed.), *Nebraska Symposium on Motivation*, 3, 149–188. Lincoln: University of Nebraska Press.

Rosenberg, M. (1979) *Conceiving the Self*. New York: Basic Books.

Rosenberg, M.J. (1956) Cognitive structure and attitudinal affect. *Journal of Abnormal and Social Psychology*, 53, 367–372.

Rosenberg, M.J. (1960) An analysis of affective-cognitive consistency. In C.I. Hovland (ed.), *Attitude Organization and Change*, 15–64. New Haven: Yale University Press.

Runkel, P.J. and Peizer, D.B. (1968) The two-valued orientation of current equilibrium theory. *Behavioral Science*, 13, 56–65.

Saegert, S., Swap, W. and Zajonc, R.B. (1973) Exposure, context, and interpersonal attraction. *Journal of Personality and Social Psychology*, 25, 234–242.

Sabini, J., Siepmann, M. and Stein, J. (2001) The really fundamental attribution error in social psychological research. *Psychological Inquiry*, 12, 1–15.

Schachter, S. (1951) Deviation, rejection, and communication. *Journal of Abnormal and Social Psychology*, 46, 190–207.

Schlenker, B.R. (1982) Translating actions into attitudes: An identity-analytic approach to the explanation of social conduct. In L. Berkowitz (ed.), *Advances in Experimental Social Psychology*, 15, 193–247. San Diego: Academic Press.

Schmitt, D.P. and Allik, J. (2005) Simultaneous administration of the Rosenberg self-esteem scale in 53 nations: Exploring the universal and culture-specific features of global self-esteem. *Journal of Personality and Social Psychology*, 89, 623–642.

Sedikides, C. (1993) Assessment, enhancement, and verification determinants of the self-evaluation process. *Journal of Personality and Social Psychology*, 65, 317–338.

Sedikides, C., Gaertner, L. and Toguchi, Y. (2003) Pancultural self-enhancement. *Journal of Personality and Social Psychology*, 84, 60–79.

Sedikides, C. and Skowronski, J.J. (1997) The symbolic self in evolutionary development. *Personality and Social Psychology Review*, 1, 80–102.

Srivastava, S. and Beer, J. (2005) How self-evaluations relate to being liked by others: Integrating sociometer and attachment perspectives. *Journal of Personality and Social Psychology*, 89, 966–977.

Shultz, T.R., Leveille, E. and Lepper, M.R. (1999) Free choice and cognitive dissonance revisited: Choosing

'lesser evils' versus 'greater good'. *Personality and Social Psychology Bulletin*, 25, 40–48.

Steele, C.M. (1988) The psychology of self-affirmation: Sustaining the integrity of the self. In L. Berkowitz (ed.), *Advances in Experimental Social Psychology*, 21, 261–302. San Diego: Academic Press.

Steele, C.M. and Liu, T.J. (1983) Dissonance processes as self-affirmation. *Journal of Personality and Social Psychology*, 45, 5–19.

Stroebe, W., Thompson, V.D., Insko, C.A. and Reisman, S.R. (1970) Balance and differentiation in the evaluation of linked attitude objects. *Journal of Personality and Social Psychology*, 16, 48–54.

Swann Jr., W.B., Griffin, J.J., Predmore, S.C. and Gaines, B. (1987) The cognitive-affective crossfire: When self-consistency confronts self-enhancement. *Journal of Personality and Social Psychology*, 52, 881–889.

Tashakkori, A. and Insko, C.A. (1979) Interpersonal attraction and the polarity of similar attitudes: A test of three balance models. *Journal of Personality and Social Psychology*, 37, 2262–2277.

Tashakkori, A. and Insko, C.A. (1981) Interpersonal attraction and person perception: Two tests of three balance models. *Journal of Experimental Social Psychology*, 17, 266–285.

Taylor, S.E. and Brown, J.D. (1988) Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, 103, 193–210.

Wellens, A.R. and Thistlthwaite, D.L. (1971a) An analysis of two quantitative theories of cognitive balance. *Psychological Review*, 78, 141–150.

Wellens, A.R. and Thistlthwaite, D.L. (1971b) Comparison of three theories of cognitive balance. *Journal of Personality and Social Psychology*, 20, 89–92.

Wertheimer, M. (1923) Untersuchungen zur Lehre von der Gestalt. *Psychologische Forschung*, 4, 301–350.

Wetzel, C. and Insko, C.A. (1974) Preacquaintance attraction as an interactive function of the proportion and number of similar attitudes. *Representative Research in Social Psychology*, 1974, 27–34.

Wiest, W.M. (1965) A quantitative extension of Heider's theory of cognitive balance applied to interpersonal perception and self-esteem. *Psychological Monographs*, 79 (14, Whole No. 607).

Woodruff, H. and DiVesta, F. (1948) The relationship between values, concepts, and attitudes. *Educational and Psychological Measurement*, 8, 645–660.

Wyer, R.S. (1974) *Cognitive Organization and Change: An Information Processing Approach*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Zajonc, R.B. (1968) Cognitive theories in social psychology. In G. Lindzey and E. Aronson (eds), *Handbook of Social Psychology*, 1, 320–411. Reading, MA: Addison-Wesley.

Zajonc, R.B. (2001) Mere exposure: A gateway to the subliminal. *Current Directions in Psychological Science*, 10, 224–227.

Lay Epistemic Theory

Arie W. Kruglanski

ABSTRACT

This chapter traces the development of the theory of lay epistemics as it has unfolded over the last three decades. The chapter describes the serendipitous beginnings of the theory, its roots in twentieth-century philosophy and sociology of science, and its reaction to social cognitive theory on attribution and consistency processes. The general theoretical depiction of the knowledge formation process gave rise to three streams of lay epistemic research described herein. The respective research programs revolve around (1) the evidential base of knowledge formation giving rise to research on the *unimodel*, (2) the unique sociopsychological nature of evidence giving rise to research on *epistemic authority*, and (3) the motivational basis of knowledge formation giving rise to research on the *need for cognitive closure*. The evolution of these research programs is described in detail as are the factors affecting the theoretical decisions that lend direction to the lay epistemic framework. Because subjective knowledge forms the basis of a wide variety of social behavior the lay epistemic theory is relevant to a considerable range of real world concerns, a sample of which is described here.

INTRODUCTION

Thinking back to the origins of my lay epistemic theory, I can discern three major sources of influence on its formation and development. One source was my PhD advisor Harold Kelley who, in a passing chat with me in 1968, mentioned an interest in investigating something he labeled as “lay epistemology.” I didn’t think about it much at the time but the term stuck in my mind, apparently. This constituted the first source of influence. Years later, as a senior lecturer at Tel Aviv University, I became strongly interested in philosophy of science and issues of the scientific method (owing partly to the influence of another mentor, Barry Collins, a UCLA professor and a devoted disciple of Donald Campbell). Leaving for a sabbatical at the University of North Carolina, I took along a bunch of philosophy of science books and was reading them voraciously in my free moments. I was particularly struck by the ideas of Karl Popper and his notion

that the construction of scientific knowledge is similar in essence to the construction of lay knowledge, and hence that science is “common sense writ large” as he succinctly put it (Popper, 1959: 22). That notion resonated with my latent interest in pursuing the study of lay epistemology, stimulated by Kelley’s casual remark,¹ and I was particularly taken by Popper’s portrayal of the scientific process as a possible model of the way lay persons go about constructing their knowledge.

The third source of influence stemmed from a conviction, also indebted to Popper’s impact, that an aspirational ideal in theory construction is the maxim of generality. In Popper’s terms, scientific theories ought to be bold and comprehensive; that is, as rich as possible in empirical contents creating a correspondingly large set of opportunities for falsification. This precept that I found compelling² evoked my concern that the work in which I was engaged at the time just wasn’t general enough and that it ultimately targeted surface structures rather than deep structures, eluding the underlying principles that govern manifest phenomena. An explanation of how this transpired requires a slight detour in which that line of work is briefly described.

Internal–external or endogenous–exogenous attributions?

My earlier line of work had to do with my critique of the internal–external distinction in attribution theory and my proposal to replace it with the endogenous–exogenous distinction (Kruglanski, 1975). Whereas the internal–external distinction juxtaposed the *person* and the *situation* as the essential attributional categories, the endogenous–exogenous distinction juxtaposed actions performed for their own sake, constituting *ends in themselves*, and actions performed for ulterior ends, serving as *means* to further goals. I argued that both the interest in the task and desire for the external reward are

equally *internal* to the actor, whereas both the task as such and the external reward are equally *external*. It made more sense to me that when participants perform a task for its interest value and appeal, they treat the task as an end in itself. I called their reasons in such case *endogenous* to the task. Conversely, when participants perform the task because they expect a reward or fear a punishment they treat the task as a means to an end. I called their reasons in such case as *exogenous* to the task.

The endogenous–exogenous distinction explained considerable amount of attributional findings and offered numerous possibilities for further inquiry, described in a comprehensive *Psychological Review* article (Kruglanski, 1975) and related empirical papers (Kruglanski et al., 1972, 1975a, 1975b). Indeed, for a while it seemed that the endogenous–exogenous theory would become the launching pad for my research program for years to come, and that I would continue exploring its implications and highlighting its insights into a wide range of social behaviors. This wasn’t meant to pass, however, for soon after the publication of the endogenous–exogenous paper, I became troubled by the following thought: that besides the endogenous–exogenous and internal–external attributional distinctions, one can propose further and further distinctions that could possibly mediate human behavior in various domains. Thus, Weiner (1979, 1985) distinguished between attributions of performance outcomes to “effort,” “ability,” “luck,” and “task difficulty,” and creatively explored the implications of these attributional categories for achievement behavior. But the “scary” thought for me was that one could actually make attributions to an open-ended host of potential categories having to do with individuals’ traits (e.g., aggressiveness, intelligence, honesty, friendliness), motivational and emotional states (e.g., fear, fatigue, greed, love, anxiety), environmental conditions (e.g., temperature, crowding, availability of resources), gender, ethnicity, and so on. All these might serve as

useful attributional categories for some people in some circumstances, but somehow their exploration didn't appear to contribute much to the illumination of the *ways* in which people make causal assignments – which is what I thought attribution theory was all about!

This realization led me to the distinction between the *process* of causal attribution and the attributional *contents*. The two were insufficiently distinguished in prior attributional writings (e.g., Kelley, 1967), probably because most attributional analyses centered around a limited number of content categories (e.g., between internal and external attributions) that meshed easily with process elements. But once it was considered that attributional contents are innumerable, whereas the attributional process is unitary by definition (constituting the processes whereby causality generally is ascribed), the process–content distinction came into high relief (for me, that is). Because of a commitment to general principles, I quickly centered my efforts on delineating the “pure” attribution *process* stripped from the attributional *contents* attached to it in former publications.

Presumably such process was contained in the “attributional criteria” described in Kelley's (1967) famous “ANOVA cube” consisting of the “consensus,” “consistency,” and “distinctiveness” rules. Specifically, according to Kelley, an attribution to an “entity” was assumed to be made if the effect (to be attributed) was distinctive to the entity and absent for other entities (the distinctiveness criterion), was consistently present across times and modalities when the entity was present (the consistency criterion), and was similarly produced by all relevant actors (the consensus criterion). But wait just a moment! Do these criteria depict a pure, content-free, attribution *process*? A moment's thought reveals that they do not. Instead, they pertain to a situation wherein the attributor is deciding among specific potential content categories of causality: the *entity*, the *person*, the *time*, and the *modality*. Thus, an entity

attribution is warranted where the effect *covaries* with the entity, and it *does not covary* with time, modality and person. Assuming that these four content categories need not be of universal interest to individuals and could be readily replaced by alternative causal categories about which one might be concerned, what is left then as an invariant process element is the principle of covariation. Indeed, in a subsequent publication Kelley (1971) identified this particular principle as the main vehicle of causal assignment, abandoning (so it would seem) the attributional criteria central to his prior exposition.

But why should covariation be evidence for causality? Very simply, it is part of the common definition of what a cause is. A cause (in common understanding) is something that covaries with the effect. Put differently, most people probably subscribe to the belief that “*if* something is a cause of an effect *then* it covaries with it.” This premise affords the derivation that “*if* something does not covary with the effect it is NOT its cause.” It is precisely that kind of reasoning that allowed Kelley (1967) to rule out the *time*, *modality*, and *person* categories as causes of the effect, assigning the causal status to the only candidate category that exhibited covariation.

The astute reader would notice that the foregoing is an example of syllogistic reasoning. One starts with a major premise (concerning in this case covariation, a concept derived from the construct of causality), combines it with a minor premise (say, that a given entity exhibited covariation with some effect of interest) to deduce a conclusion, that the entity is (probable) cause of the effect. This framing immediately begged the question whether the covariation principle, apparently a major vehicle in the rendition of causal inferences isn't a special case of a broader principle involved in all possible inferences.

The syllogistic structure of reasoning (underlying the deduction of causality from covariation evidence) recalled Popper's

assertion that all scientific knowledge is arrived deductively. Combined with Popper's notion that lay knowledge is arrived in the same manner as scientific knowledge, this notion of syllogistic reasoning, or what psychological researchers (e.g., Sloman, 1996) referred to as "rule following," represents the essential principle whereby humans arrive at all their knowledge. This realization posed the challenge of transcending the realm of attribution processes, and to refocusing my research program on the broader issue of knowledge construction, marking the start of my theorizing about lay epistemics.³

Specifically, my analysis of causal inference was generalized to a major assumption of lay epistemic theory: that all knowledge⁴ is positively derived from evidence. In this respect, my reasoning differed from Popper's who insisted that knowledge can only be *falsified* by evidence. Let me explain. Popper (1959) assumed that the process of hypothesis testing is represented by the premise: "If H (hypothesis) is true then E (evidence) should be the case," which implies that one can only disconfirm a hypothesis via a logical *modus tollens* (if E is false, H too must be false), but not verify it by evidence as I am suggesting. But, it is plausible that the knower may depart from a different (and stronger) assumption; specifically, the assumption that *if and only if* hypothesis H were true evidence E would be observed. The *if and only if* framing implies a bidirectional implication (in the terminology of logic known as the *equivalence* relation) asserting that *both* "if H then E" and "if E then H" are true. For instance, one may feel confident (to the point of trusting it with one's life!) that the road is clear if the evidence of her or his eyes indicated this. The knower's premise in this instance may be expressed as "if and only if my eyesight indicates that the road is clear then it is indeed clear".

In other words, given the *if and only if* premise (that may "feel" justifiable under specific circumstances) an individual could logically deduce the hypothesis from the evidence in a *modus ponens* fashion, whereby

if E is affirmed (the evidence is observed) then H is inferred (the hypothesis is supported). Of course, the *if and only if* assumption might need to be modified on the basis of subsequent information which would cast doubt on the original conclusion that H was proven to be valid. For instance, if a rival alternative hypothesis H₁ appeared plausible and the need arose to distinguish it from the original H, one would have to formulate a subsequent inference rule whereby *if and only if H but not H₁* were true then E₁ would obtain, and so on.⁵

Terminating the epistemic sequence

Philosophers of knowledge have long maintained that the sequence of hypothesis generation and testing (whether in science or lay belief formation) has no unique, or objective point of termination. In principle, it could continue interminably as one keeps engendering further and further alternative hypotheses without ever crystallizing firm knowledge on a topic. Obviously, such epistemic "obsession" would be highly dysfunctional and paralyzing. Most of the people, most of the time are quite capable of forming judgments based on available evidence, and of self-regulating adaptively on the basis of those judgments.

The theory of lay epistemics identifies two mechanisms that effect the cessation of the hypothesis generation sequence and the crystallization of confident knowledge; one of these mechanisms is *motivational*, the other is *cognitive*. The motivational mechanism is based on the need for cognitive closure (Kruglanski, 2004). Two types of such need have been distinguished, referred to respectively as the needs for *nonspecific* and *specific* closure. The need for nonspecific closure denotes a desire for a firm answer to a question, any firm answer as compared to confusion and ambiguity. The need for a specific closure denotes a specific desirable answer to a question; for example, an esteem

enhancing answer, or an optimistic answer. Each of these needs is assumed to vary in degree and lie on a continuum ranging from a low to a high motivational magnitude. Thus, one may desire nonspecific closure strongly, mildly or not at all, actually craving to avoid it. Similarly, one may desire to reach a particular conclusion (or specific closure) with varying degrees of strength. Both types of need determine the length of the epistemic sequence of hypothesis generation and testing. The higher the need for nonspecific closure the shorter the sequence and the stronger the tendency to “seize and freeze” on accessible, closure-affording evidence. The higher the need for a specific closure, the stronger the tendency to terminate the sequence when the available evidence appears to yield the desired conclusion, or to keep the sequence going until such conclusion seems implied by the evidence (Ditto and Lopez, 1992).

The cognitive mechanism relates to the “authority” of the evidence, or what we have referred to as *epistemic authority*. (Kruglanski et al., 2005). Specifically, the theory assumes that some kinds of evidence are “incontrovertible,” because their source is deemed indubitable and beyond reproach. For instance, one might trust the evidence yielded by one’s senses (e.g., seeing something with one’s own eyes), or evidence provided by a trusted expert (one’s physician, an auto mechanic, a parent, an admired leader priest or cleric), or evidence derived from “common knowledge,” that is from one’s group’s shared reality (Hardin and Higgins, 1996).

EXPLORATIONS IN LAY EPISTEMICS

The transition of my research interests from attribution processes to those of knowledge formation was described in a *Psychological Review* article (Kruglanski, 1980) that followed (by five years) the endogenous–exogenous theory described earlier. What ensued was three decades of research on various facets of lay epistemics organized

around three major research programs. These addressed, respectively: (1) the *unimodel* of social judgment, (2) the construct of *epistemic authority*, and (3) the *need for cognitive closure*. Let me now describe the major theoretical and empirical developments in each of those domains.

The unimodel of human judgment: all you need is one! Lay epistemic theory and the dual process models

As already noted, the lay epistemic theory assumes that all knowledge derives from evidence, broadly conceived. To reiterate, in constructing new knowledge, or in forming a new judgment, the individual is assumed to use an inference rule whereby if a given evidence E obtains, conclusion C follows (or hypothesis H is supported).

The (lay epistemic) notion that all knowledge is derived via unitary, syllogistic, process seemed inconsistent with the popular claim that judgments can be formed via two qualitatively different modes, referred to as *central* and *peripheral* modes (Petty and Cacioppo, 1986) or *heuristic* and *systematic* modes (Chaiken et al., 1989). On close examination, I became convinced that the difference between these “modes” lies in *contents* of the inferential rules, and that judgments reached via both modes are unexceptionally syllogistic in fact. Based on that assumption, I proceeded to examine large bodies of research interpreted as supportive of the dual mode distinction, and to offer an alternative framing of their findings cast from the lay epistemic perspective.

Persuasion research

A pervasive finding in dual-mode inspired persuasion research has been that “peripheral” or “heuristic” cues exert judgmental impact (i.e., effect change in recipients’ attitudes or opinions) under conditions of low processing resources (e.g., where recipients’ interest in the task is low, when they are

cognitively busy or distracted, when their need for cognition is low, etc.). By contrast, “message arguments” typically exerted their effects under high processing resources (e.g., high interest in the task, or ample cognitive capacity). Though the dual mode analyses never stated explicitly *how* message arguments or cues differentially mediate judgments, it seems plausible that in both cases the inferential mechanism is the unitary rule-based process described earlier. In case of “peripheral” or “heuristic” reasoning, the rule might be “if expert then correct,” or “if supported by majority then correct,” whereas in case of “central” or “systematic” reasoning the rule might be “if involving a tuition hike then unacceptable,” or “if promoting individual rights then deserving of support,” for example. In other words, the syllogistic rule-following process seemed the same in both presumptively distinct modes, and they only seemed to differ in contents of the rules employed.

But this fundamental commonality did little to explain the pervasively found differences in reliance on cues/heuristics or messages under different motivational or cognitive resource conditions. If the lay epistemic theory were to offer a truly comprehensive account of knowledge formation, this puzzle needed to be resolved. The way it was finally resolved rested on an empirical “discovery,” involving a previously unidentified *confounding* present in many dual-mode studies. It consisted of an inadvertent linkage of low processing resources with the use of easy-to-process rules and of high processing resources with the use of relatively difficult-to-process rules (for reviews see Erb et al., 2003; Kruglanski and Gigerenzer, 2011; Kruglanski and Thompson, 1999a, 1999b; Kruglanski et al., 1999, 2007; Pierro et al., 2005).

Specifically, because in dual mode persuasion studies the message arguments were typically lengthier, more complex and placed later in the informational sequence, their processing may have imposed higher processing demands than the processing of

“cues” that were invariably brief, simple, and presented upfront. When these confoundings were experimentally removed, the previously found differences between conditions under which the “cues” versus the “message arguments” (or vice versa) exerted their persuasive effects were eliminated. Once the message arguments were made shorter and simpler, they were found to mediate judgments under limited resource conditions just as did cues and heuristics in prior research. Similarly, once the cues were presented in a lengthier and more complex form (e.g., an expert’s qualifications portrayed via a lengthy resume), they mediated judgments only under ample resource conditions, as well as led to more persistent attitudes that were more predictive of behavior (Pierro et al., 2008).

Dispositional attributions

Inattention to a potential confounding between task difficulty and informational contents afford a lay epistemic reinterpretation of findings from alternative research domains as well, unrelated to persuasion. Thus, a major attributional question concerned the process whereby an actor’s behavior is causally ascribed to the *situation*, or to the actor’s *disposition*. In an important paper devoted to this topic, Trope and Alfieri (1997) found that ambiguous behavior tends to be disambiguated by assimilation to the context in which it is taking place. For instance, an ambiguous facial expression is likely to be perceived as sad if the context was sad as well (e.g., a funeral), and as happy if the context was happy (e.g., a party). Once the behavior had been identified, however, and the question of its causal origin was pondered, the context plays a subtractive (rather an assimilative) role in determining the behavior’s attribution. Specifically, the role of the context is subtracted to determine the role of the actor’s disposition in producing the behavior. For instance, if the context was sad, an individual’s sad expression would tend to not be attributed to the actor’s dispositional sadness because other persons in the same situation would probably be sad as well.

Of present interest, Trope and Alfieri (1997) found that the assimilative process of behavior identification was independent of cognitive load, whereas the subtractive process of dispositional attribution was undermined by load. These investigators also found that invalidating the information on which the behavior identification process was based, by stating that the actor was unaware of the potential situational demands on their behavior, did not alter these identifications, whereas invalidating that same information did alter the dispositional judgments. Two alternative explanations may account for these results: (1) that the two processes are qualitatively distinct, or (2) that for some reason the behavior identification task in Trope and Alfieri's (1997) work was less demanding than the dispositional attribution task. Consistent with the latter interpretation, Trope and Gaunt (2000) discovered that when demands associated with the dispositional attribution task were lowered (e.g., by increasing the salience of the information given), the subtraction of context from dispositional attributions was no longer affected by load. Furthermore, Chun et al. (2002) found that when the behavior identification task was made more difficult (e.g., by decreasing the salience of the information given) it was also undermined by load. Under those conditions too, invalidating the information on which the behavioral identifications were based did alter these identifications. These findings are consistent with the lay epistemic notion that, when a given judgmental task (e.g., "behavior identification" or "dispositional attribution") is sufficiently demanding, it is exigent of cognitive resources and can be undermined by load. As Chun et al. (2002) point out, *behavior identification* and *dispositional attribution* appear to be deduced in the same (syllogistic) way from relevant evidence (Kruglanski and Gigerenzer, 2011). Furthermore, each inferential task may be more or less difficult depending on the circumstances, requiring respectively more or less resources.

Base-rate neglect

Inattention to the psychological difficulty inherent in using different types of evidence applies to a classic phenomenon in judgment and decision making, the neglect of base rates in favor of heuristic evidence concerning representativeness. In the original demonstrations of this effect (Kahneman and Tversky, 1973) base rates of some pertinent categories (the proportion of lawyers and engineers) in a given sample was presented to participants along with a vignette describing an individual randomly drawn from that sample. A careful analysis of typical studies on this topic carried out by Chun and Kruglanski (2006) revealed that the base rate information in many of those studies was typically presented briefly, via a single sentence, and upfront. By contrast, the individuating ("representativeness") information was presented subsequently via a relatively lengthy vignette. If one assumes that participants in such studies had sufficient motivation and cognitive capacity to process the entire informational "package" with which they were presented – they might have been challenged to fully process the later, lengthier, and hence more demanding vignette information and to have given it considerable weight in the ultimate judgment. This is analogous to the finding in persuasion studies that the lengthier, later-appearing message argument information but not the brief, upfront-appearing "cue" information has impact under ample processing resources (e.g., of high processing motivation and cognitive capacity). If the above is true, we should be able to "move" base-rate neglect around by reversing the relative length and ordinal position in the informational sequence of the base-rate and the individuating ("representativeness") information. A series of studies by Chun and Kruglanski (2006) accomplished just that, and showed that when the representativeness information is rendered easier to process than the base rate information (by making the former brief and simple and the latter lengthy and complex), the representativeness information

is relied upon more when the participants' resources are depleted by cognitive load. Of even greater interest, when the base rate information is presented briefly and simply and the representativeness information is more difficult to process as in the standard paradigm, it tends to be utilized more under depleted (versus intact) resource conditions. In short, *base rates* and *representativeness* information seem to constitute different contents of evidence for likelihood estimates (e.g., that a randomly drawn individual from a population is a member of a given profession), but they afford the relevant judgments via the same inferential process. As with the persuasion and attribution tasks discussed earlier, these types of evidence may be presented in different formats varying in their difficulty of application hence requiring corresponding amounts of processing resources (Kruglanski and Gigerenzer, 2010).

Relative relevance, task demands, and processing resources

Let us switch now to a very different confounding, namely of information type and subjective relevance of information. Often, the different types of information presented to research participants have (inadvertently) differed in their *subjective relevance* to these persons. For instance, in the domain of persuasion Pierro et al. (2004) carried out an extensive content analysis of experimental materials in persuasion studies to conclude that, typically, the "cues" presented to participants were judged as less relevant to the judgmental (attitudinal) topic than the "message arguments." Recall that in much persuasion research the "cues" but not the "message arguments" exerted judgmental impact under low processing resources, whereas the "message arguments" did so under high processing resources. From the present perspective, it is possible to generalize these findings in terms of the following derivations: (1) given ample processing resources, the *more relevant* information (e.g., the "message arguments" in much persuasion research) would have a greater

judgmental impact than the *less relevant* information; however (2) given limited processing resources (relative to the task demands), the *easier-to-process* information (of above threshold relevance) would have a greater judgmental impact than the *difficult-to-process* information (Kruglanski and Gigerenzer, 2010).

Pierro and colleagues (2004) tested these notions in three experimental studies and found that when later, and hence the more difficult-to-process information (whether "heuristic or message argument-based") was more subjectively relevant to the judgmental topic than early information, it exerted judgmental (persuasive impact) only under high motivation conditions but not under low motivation. By contrast, the early, less-relevant information exerted its effect only under low motivation but not under high motivation. A very different pattern obtained where the early information was more relevant than the latter information. Here, the impact of the early information invariably overrode that of the latter information: Under low processing motivation this may have been because the earlier information was easier to process than the latter information, and under high processing motivation because the early information was in fact more relevant than the latter information.

To summarize, a growing body of evidence from a variety of domains (persuasion, attribution, judgment under uncertainty, person perception) is consistent with the lay epistemic notion that the inferential mechanism whereby judgments are made is the same across different content domains of judgment. The often observed differences in the use of different kinds of information (e.g., "central" versus "peripheral" information, base rate versus heuristic information, etc) seem explicable in terms of (locally determined) difficulty or ease with which the different types of information were presented to research participants, requiring correspondingly different amounts of processing resources. Because difficulty or ease of information processing lies on a continuum

(rather than representing a dichotomy), I concluded that the notion of two qualitatively distinct modes of making judgments is unnecessary. Instead, it seemed important to highlight the universal inferential mechanism (the syllogistic “if then” reasoning structure elaborated earlier) involved in all judgments, whose application may be more or less difficult across different situations requiring different amounts of processing (cognitive and motivational) resources.

From dual modes to dual systems

The *dual mode* theories considered earlier (e.g., the dichotomous persuasion, or attribution models) didn’t elaborate much on the process whereby each of the modes is assumed to mediate judgments, focusing instead on informational contents (e.g., of base rates versus representativeness information). In recent years, however, a different family of dualistic models was proposed in which an attempt was made to be more explicit about processes. This category of theories, generally known as *dual systems* models distinguishes between *associative* and *rule-following* processes (e.g., Kahneman, 2003; Sloman, 1996; Strack and Deutsch, 2004). Associative processes were often characterized as automatic, unconscious, and independent of resources; rule-following processes, in contrast, are typically thought of as deliberative, conscious, and resource intensive. Because of the centrality accorded in the lay epistemic theory to inferential (if-then) rules it was generally classified in the rule-following category (cf. Strack and Deutsch, 2004). If valid, the dual systems distinction would pose a challenge to the presumed universality of the knowledge formation process depicted in lay epistemic theory. Thus, the next phase of lay epistemic theory consisted in a careful examination of the dual systems claims.

Associative processes are rule-based

Consider classical conditioning. Though it has been viewed as a prototypic case of an associative process, compelling evidence

exists (Holyoak et al., 1989; Rescorla, 1985; Rescorla and Holland, 1982; Rescorla and Wagner, 1972) that, in fact, it is fundamentally *rule-based*. In this vein, Holyoak and colleagues (1989) explicitly stated that “representations of the environment take the form of ... (if then) rules ... the rat’s knowledge about the relation between tones and shocks might be informally represented by a rule ‘if a tone sounds in the chamber then a shock will occur,’” (Holyoak et al., 1989: 320). In present terms, the rule represents a major premise, that when combined with a minor premise “a tone is sounding,” warrants the deduction “a shock is coming” and elicits the warranted crouching behavior.

Whereas the work reviewed by Holyoak et al. (1989) addressed classical conditioning, a recent review of *evaluative conditioning* attests that it too is “propositional,” hence *rule-following* (Mitchell et al., in press). In evaluative conditioning a neutral CS (e.g., a book) is presented concomitantly with an affectively laden UCS (e.g., a smiling, or a pouting face); subsequently, the CS acquires the affective valence of the UCS. Though evaluative conditioning differs in a number of respects from classical conditioning (for discussions see Baeyens et al., 1988; Walther et al., 2005) the rule-following nature of the conditioning process appears common to both (Mitchell et al., in press). In short, the two paradigmatic examples of associative processes (classical and evaluative conditioning) appear to be rule-driven. This realization effectively undermines the distinction proposed by some authors between *rule-following* and *associative* processes (e.g., Gawronski and Bodenhausen, 2006; Sloman, 1996).

Automaticity of rule-following

Rules involved in conditioning may be applied with considerable ease and rapidity. The notion that “automatic” phenomena in the realm of skill acquisition involve a routinization of “if... then” sequences has been central to Anderson’s (1983) ACT* model that Smith (1984, 1989; Smith and Branscombe,

1988; Smith et al., 1988) generalized to social judgment. That research has demonstrated that social judgments represent a special case of procedural learning based on practice that strengthens the “if... then” components resulting in increased efficiency (or “automaticity”).

Awareness

Efficiency implies, in turn, a lowered need to commit attentional resources to the execution of social judgments. In William James’ (1890: 496) eloquent turn of phrase, “consciousness deserts all processes when it can no longer be of use.” According to James’ *parsimony principle* of consciousness, routinization removes the need for conscious control, rendering awareness of the process superfluous. In a related vein, Logan (1992) suggested that automatization of certain skills effects an attentional shift to higher organizational levels.

It is in this sense, then, that some judgmental phenomena, mediated by well routinized “if...then” rules, may take place outside conscious awareness. Already Helmholtz (1910/2000) discussed the notion of unconscious *inference* in the realm of perception. More recently, social cognitive work on spontaneous trait inferences (Newman and Uleman, 1989; Uleman, 1987) suggests that lawful (i.e., rule-following) inferences presumably can occur without explicit inferential intentions, and without conscious awareness of making an inference. The spontaneous trait inference that John is “clumsy” on basis of the information that he “stepped on Stephanie’s foot while dancing” (Newman and Uleman, 1989: 156), surely requires the inference rule, “if stepping on a partner’s foot, then clumsy” or some variant thereof. A person who rejects that premise would be unlikely to reach that conclusion.

In summary, a variety of evidence and theoretical considerations converge on the lay epistemic notion that judgments are rule-based and in this sense, derived from “evidence.” To make a judgment is to go beyond the “information given” (Bartlett, 1932; Bruner, 1973),

by using it as testimony for a conclusion in accordance with an “if... then” statement to which the individual subscribes. Such implicational structure appears fundamental to explicit human inferences (Anderson, 1983), implicit conclusion-drawing (Newman and Uleman, 1989), conditioning responses in animal learning studies (Holyoak et al., 1989; Rescorla and Wagner, 1972), and perceptual judgments of everyday objects (Gregory, 1997; Pizlo, 2001; Rock, 1983). The elementary “if... then” form appears essential to all such inferences, whether conscious or non-conscious, instantaneous or delayed, innate or learned. It is a fundamental building block from which all knowledge appears to be constructed.

Parametric determinants of informational impact

Subjective relevance The lay epistemic analysis of the dual mode and dual systems theories allowed the identification of major factors involved in determining whether the “information given,” (i.e., present in the environment or provided by a source) will influence an individual’s judgments. To highlight their continuous nature, I labeled these factors *judgmental parameters*. The first parameter refers to the degree to which an individual subscribes (whether consciously or unconsciously) to a given inferential rule or major premise (“if X then Y”). The stronger the subjective implication of Y by X, the greater the evidential relevance of X to Y, and all else being equal, the stronger the tendency to reach the judgment Y upon confronting X.

It is useful to conceptually distinguish between *potential relevance* of X to Y reflecting the degree to which the “If X then Y” inferential rule has been generally learned and believed in, and *contextual or perceived relevance* reflecting the degree to which X is recognized as relevant to Y *in a given situation*. Beyond one’s degree of belief in the rule, or availability of the rule in a person’s memory, *perceived relevance* is affected by the rule’s accessibility, difficulty of

identifying the X, and individual's motivational and cognitive resources available for overcoming the difficulty (for discussion see Kruglanski et al., in press).

Difficulty of the inferential task Retrieval of the inferential rule, or the major premise, from a person's memory could be more or less difficult depending on the rule's history of activation as defined by its frequency and recency (Higgins, 1996). Similarly, situational discernment of the minor premise; that is, of information that instantiates the antecedent (X) of the rule, may be more or less difficult in different task environments. Jointly, challenges posed by activities of *retrieval* and *discernment* define the (continuous) parameter of task difficulty on which different inferential contexts may vary.⁶

Availability of cognitive resources In approaching an inferential task individuals may come to it with different reserves of cognitive resources. This may depend on their prior engagement in resource-depleting activities, or the presence of concomitant attention-demanding tasks. In general, individuals' ability to cope with difficult judgmental tasks depends on the resources at their disposal. For instance, if the discernment of relevant information is difficult to accomplish, only individuals with sufficient resources would base their judgments on that information. Similarly, if individuals' resources are depleted, they may tend to base their judgments on easy-to-process (e.g., simple, brief, and/or salient) information.

Motivation Besides depleting alternative activities, availability of attentional resources may depend on individuals' motivational states. For instance, under high (versus low) motivational involvement in a judgmental topic, individuals may mobilize greater attentional resources for an inferential task, and hence be capable of coping with relatively difficult information-processing requirements, (posed, for example, by lengthy

and complex information) (for reviews see Chaiken et al., 1989; Kruglanski et al., 2007; Petty and Cacioppo, 1986).

In summary, application of the lay epistemic theory to the realm of dual mode and dual systems models resulted in a reanalysis of research findings adduced in their support, and a parametric reinterpretation focused on the interaction between the subjective relevance of evidence, the difficulty of its discernment in given circumstances, and the cognitive and motivational resources that individuals may bring to bear on the judgmental task at hand.

Epistemic authority: source as evidence

According to lay epistemic theory, all evidence types function in the knowledge construction process in the same, syllogistic way. Because people's concerns typically extend beyond their domains of expertise, they often rely on other people as knowledge providers. Indeed, an essential aspect of people's social nature is their informational interdependence (Kelley and Thibaut, 1969), and the fact that they share their view of reality with significant others (Hardin and Higgins, 1996). In pondering this issue, it appeared to me that the role of *trusted sources of information* merits its own consideration in a theory of knowledge formation. Accordingly, I introduced the concept of *epistemic authority* to denote the pervasive and often determinative function of others' opinions on individuals' own judgments. Specifically, individuals may subscribe to general "if X then Y" rules in which the antecedent X denotes a given epistemic authority; for example, of an expert ("If Expert says so then it is correct"), the group ("If the Group believes so, then it is correct"), or the self ("If I believe X, then X is correct").

The concept of "epistemic authority" is akin to the notion of *source credibility* (encompassing a conjunction of perceived

expertise and *trustworthiness*), and it addresses the extent to which an individual is inclined to treat a source's opinions as incontrovertible evidence for her or his own judgments. The ascribed epistemic authority of various sources in the individuals' social environments may vary, and the authority of a given source may vary across domains as well as across individuals' lifespan phases.

In the social psychological literature, source characteristics (such as expertise) were often implied to offer a somewhat *inferior* counsel as to correct judgments, and were treated as suboptimal heuristics employed under limited resource conditions and when one's "sufficiency threshold" of required confidence was low (Chaiken et al., 1989). Yet, as common observations suggest, the compelling power of source authority can be considerable, (e.g., authority accorded to a religious prophet, a parent, a political leader, or the printed word) and carry substantial weight in instilling confidence in judgments. Indeed, it can often override other types of information and exert a determinative influence on individuals' opinions and corresponding behaviors.

Whereas prior treatments of source credibility affects exclusively addressed sources *external* to self (cf. Chaiken et al., 1989; Hovland et al., 1953; Kruglanski and Thompson, 1999a, 1999b; Petty and Cacioppo, 1986), the lay epistemic theory considers the *self* as a particularly important target of epistemic authority assignments. Research summarized by Kruglanski et al. (2005) has revealed: (1) developmental trends involving a decline in authority assigned to the primary caregivers, coupled with an increase in epistemic authority attributed to the *self*, and involving an increase in differentiation and specificity of epistemic authorities across domains; (2) stable individual differences in epistemic authority effects; (3) a hierarchical structure and operation of epistemic authorities; and (4) the relative role of the self and external sources as perceived epistemic authorities.

Developmental trends

Raviv et al. (1990) found that from age four to ten, (a) the perception of parents as epistemic authorities remains relatively stable, with decreases in a few knowledge areas, (b) the perception of the teacher as an epistemic authority remains stable with an increase in the area of science, and (c) the perceived epistemic authority of friends increases in the social domain. Raviv et al. (1990) also found that across age groups the perception of teachers and friends varied more as a function of knowledge areas than the perception of parents. Specifically, parents tended to be perceived as *overall authorities*, possibly due to continued emotional and material dependence on them inducing a motivation to view them as all powerful and knowledgeable.

Individual differences in the distribution of epistemic authority assignments across sources

Individuals differ systematically in their distributional profiles of epistemic authority across sources. These differences, in turn, affect individuals' search for, and use of, information. Specifically, Bar (1983, 1999) found that individuals turn first to information provided by sources whom they regard as highest in epistemic authority, that they process such information more extensively, that they derive from it greater confidence, and that they tend more to act in accordance with its perceived implications.

Self-ascribed epistemic authority and learning from experience

The concept of "experience" has long been privileged in psychological theory. The use of experiential learning in training and education has been inspired by John Dewey's (1916, 1958) instructional philosophy, Carl Rogers' (1951, 1967) person-centered approach to therapy and humanistic psychology more generally (e.g., Shafer, 1978). In social psychology, Fazio and Zanna (1981) suggested that attitudes acquired via direct experience with the attitude object are the

strongest and most tightly related to behavior. Yet, these authors also hinted at *moderators* that qualify the power of experience in shaping attitudes. As they put it, “An attitude formed by indirect means could conceivably also be held with extreme confidence. For example, a child’s attitude may be held with great confidence, even though formed indirectly because of his or her parents’ extreme credibility” (Fazio and Zanna, 1981: 184).

Whether an individual would treat her or his personal experience as a reliable knowledge source may depend on this person’s self-ascribed epistemic authority in a domain. In a study designed to investigate these notions, Ellis and Kruglanski (1992) assessed their participants’ self-ascribed epistemic authority in mathematics via a questionnaire specifically designed for this purpose. Participants also responded to the numerical aptitude test (Cattell and Epstein, 1975) to serve as a control measure for their actual math ability, and they filled out a postexperimental questionnaire designed to assess their perceptions of their own and the instructor’s epistemic authority in mathematics.

In the *experiential* condition, participants were given self-instruction booklets with exercises related to the five arithmetic rules. In the *instructional-principles* condition, the experimenter was introduced as a PhD in mathematics, and he explicitly articulated the relevant mathematical principles. In the intermediate, *instructional-examples* condition, the instructor solved the problems on the board and stated the arithmetic principle underlying each solution. The results indicated that method of instruction significantly interacted with participants’ self-ascribed epistemic authority (SAEA). Controlling for participants’ actual mathematical ability, in the *experiential* condition, participants with a high SAEA did significantly better than participants with a low SAEA. By contrast, in the *instructional principles* condition, low SAEA participants did better than their high SAEA counterparts, and in the intermediate, *instructional-examples* condition the high

and low SAEA participants did not differ in their performance.

The foregoing findings identify a significant boundary condition on the efficacy of *experience* as a basis of *learning*. It appears that in order to be able to learn from their own experience; that is, without instruction from others, individuals need to believe in their ability to draw inferences from the experience; that is, possess high self-ascribed epistemic authority in a domain. Crucially, self-ascribed epistemic authority can be empirically distinct from actual ability in a domain. In the Ellis and Kruglanski (1992) study, the correlation between the two, though significant, was relatively low ($r = 0.36$), and the interaction between SAEA and method of instruction remained significant, even after controlling for actual mathematical ability.

Summary

Even though all evidence may function in the same (syllogistic) manner, the evidence category subsumed under the notion of epistemic authority is special. It embodies the fundamental notion that human knowledge is socially constructed and that it is heavily influenced by the opinions of significant others whose judgments one holds in high regard.

Need for closure research: on epistemic dynamics

The lay epistemic treatment of the concept of *evidence* (addressed in research programs on the *unimodel* and on *epistemic authority*) touches on the mechanism of knowledge formation, or the *how* of individuals construction of their judgments and opinions. But any knowledge is potentially insecure. New facts may become revealed, trusted epistemic sources may express views at odds with one’s prior opinions, and novel hypotheses may be generated to account for familiar phenomena. Philosophers and historians of science (like Karl Popper or Thomas Kuhn), whose writings have fascinated me, highlight

in their analyses the dynamic aspect of scientific knowledge represented by scientific revolutions (Popper, 1959), or the replacement of ruling paradigms by alternative frameworks (Kuhn, 1962).

In considering the potentially shifting aspect of human knowledge, it seemed that a major issue in this regard concerns people's open-mindedness and their inclination to continue the hypothesis generation and testing process that results in knowledge. Because the quest for new information and the generation of new hypotheses is potentially interminable, it seemed to me that the determinants of open and closed mindedness must be *subjectively driven*, and are likely to be *motivational*. In this vein, I introduced the concept of need for closure⁷ as a major psychological determinant of epistemic stability and change, inspiring the most extensive research program to date carried out within the lay epistemic framework.

I defined the need for cognitive closure as the desire for certainty on a topic. I further postulated that the magnitude of this need is determined by the perceived benefits of closure, and by the costs of lacking closure. In those terms, the need for closure was assumed to be elevated where action was required because the launching of intelligible action requires prior closure. Additionally, the need for closure was assumed to be elevated in circumstances where the possession of closure would obviate costly or laborious information processing, as may occur under time pressure, in the presence of ambient noise, or when a person is fatigued or intoxicated (see Kruglanski, 2004, for a review). In addition to the transient situational determinants of the need for closure, this motivation was also assumed to represent a dimension of individual differences, and we constructed a scale to assess it (Webster and Kruglanski, 1994). To date, this scale has been translated into numerous languages and has been shown to yield similar results with situational manipulations of the need for closure;⁸ an improved version of the scale was recently published by Roets and van Hiel (2007), and

a shortened yet unpublished version of the scale has been effectively used as well (Pierro and Kruglanski, 2008).

Intrapersonal phenomena: seizing and freezing

The dynamic aspect of knowledge involves what Kurt Lewin labeled as the "freezing," "unfreezing," and "refreezing" of individual's beliefs. The need for cognitive closure is seen to underlie these fundamental epistemic phenomena. Specifically, a heightened need for cognitive closure may induce in individuals the tendency to "seize" on early, closure-affording evidence and "freeze" upon the judgments it suggests. These tendencies were studied in reference to several classic effects in social cognition and perception, including primacy and recency effects in impression formation, the use of stereotypes, and anchoring effects among others (for reviews see Kruglanski and Webster, 1996; Kruglanski, 2004). In all these cases, a heightened need for closure induced a freezing on judgments afforded by recently or frequently activated information (e.g., early information about a social target, or a prevalent social stereotype).

In a recent demonstration of need for closure's impact on the use of recently activated information, Pierro and Kruglanski (2008) conducted a study on the *transference effect* in social judgment. The Freudian concept of transference refers to the process by which a psychotherapeutic patient superimposes onto the therapist her or his childhood fantasies about a significant childhood figure (typically a parent). Andersen and her colleagues (e.g., Andersen and Cole, 1990; Andersen et al., 1995) showed, however, that the transference effect could be part and parcel of normal social-cognitive functioning in which a significant other's schema is mistakenly applied to a new target that resembles the significant other in some respects. In a first session of Pierro and Kruglanski's (2008) experiment, participants completed the revised 14-item need for closure scale (Pierro and Kruglanski, 2005) and were asked to

visualize and describe a significant other. In a second session, participants were presented with information about a target person with whom they expected to interact. The target person was either described in similar terms as their significant other, or was depicted as dissimilar from that person. After having studied this information, participants were presented with a recognition test of their memory for the target. Items about the target person that were not presented in the description were included in the recognition test. The degree of transference was operationally defined as the proportion of statements falsely recognized as having been included in the description of the target person that were consistent with the representation of the significant other provided in the first session. The results indicated that participants high on the need for closure exhibited a more pronounced transference effect, as indicated by higher false alarm rates, in the similar (versus dissimilar) condition than did participants low on the need for closure.

Interpersonal phenomena

Beyond its effects on intrapersonal phenomena in the domain of social judgment, need for closure was shown to exert a variety of interpersonal phenomena in realms of linguistic expression, communication and persuasion, empathy, and negotiation behavior.

Linguistic expression Several studies looked at need for closure effects on language abstractness in interpersonal communications. Abstract language indicates a *permanence* of judgments across situations, and hence a greater stability of closure. For instance, characterizing an individual's behavior in a given situation as reflecting this person's aggressiveness (an abstract depiction) implies that he or she may be expected to behave aggressively in other contexts as well. By contrast, depicting the same behavior as a "push" (i.e., a concretely situated occurrence) carries fewer trans-situational implications. Accordingly, it is possible to predict that individuals under high (versus low) need for

closure would generally tend to employ abstract terms in their communications. Evidence consistent with this prediction was obtained by Boudreau et al. (1992), Rubini and Kruglanski (1997) and Webster et al. (1997).

Persuasion Research by Kruglanski et al. (1993) explored the conditions under which need for closure may increase or decrease the susceptibility to persuasion. To do this, participants were presented with information about a legal case, allowed time to process the information, and then later interacted with a partner (a fellow "juror") in order to reach a joint verdict in the case. When participants were given complete information about the case, including a (fictitious) legal analysis suggesting the appropriate verdict, individuals high (versus low) on the need for closure were less likely to be persuaded by their fellow juror (who argued for the opposite verdict). However, when high-need-for-closure individuals were given incomplete information lacking the legal analysis, they were more likely to be persuaded by their fellow juror than their low-need-for-closure counterparts. In short, individuals high (versus low) on the need for closure tend to resist persuasion attempts when they have formed a crystallized opinion about a topic, but tend to change their attitudes when presented with persuasive appeals when they lack an opinion about the topic.

Empathy Because high-need-for-closure individuals tend to "freeze" on their own perspective, they are less able to empathize with their interaction partners, especially when those are dissimilar from themselves (Webster-Nelson et al., 2003). In related work, Shytenberg et al. (2008) found that high (versus low) scorers on the need for closure scale were less sensitive to injustice done to their teammate by the experimenter (perceiving the experimenter as less unfair). Richter and Kruglanski (1999) found that individuals with high (versus low) dispositional need for closure tended less to "tune" their messages

to their audience's perspective;⁹ as a consequence, their communications were less effectively decoded by their recipients.

Negotiation behavior Need for closure was found to exert significant effects on negotiation behavior. Thus, DeDreu et al. (1999) measured participants dispositional need for closure and then (after a 30 minute delay) had them engage in a task in which they acted as sellers and interacted with presumed buyers (actually simulated by pre-programmed responses). It was found that individuals with high (versus low) dispositional need for closure tended more to adhere to cut-off values (of minimal acceptable profits) given them by the experimenter; they also made smaller concessions to their negotiation partners and engaged in less systematic information processing. In another study on negotiation, De Dreu and Koole (1997) found that, under conditions where a majority suggested a competitive strategy, lowering participants' need for closure decreased their tendency to behave competitively and to reach an impasse.

Group centrism

Some people are more group-oriented than others, and most people are more group-oriented in some situations than in other situations. Kruglanski et al. (2006) defined the concept of "group centrism" by the degree to which individuals strive to enhance the "groupness" of their collectivity. Groupness, in turn, has been defined by a firm, consensually supported "shared reality" (Hardin and Higgins, 1996), unperturbed by dissents and disagreements. While reality sharing has been regarded as the defining essence of groupness (e.g., Bar-Tal, 1990, 2000), its attainment may be facilitated by several aspects of group interaction, all essentially enhanced by the need for closure. At the initial phases of group formation, this can involve members' attempts to arrive at a speedy consensus, by exerting uniformity pressures on each other (DeGrada et al., 1999).

The positive relation between need for closure and autocracy (De Grada et al., 1999; Pierro et al., 2003) is consistent with Gelfand's (2008) cross-cultural research in 35 countries across the globe in which she finds a significant relationship between the country's degree of autocracy and situational constraints and the inhabitants' need for closure. Though these results may reflect the notion that high need for closure individuals tend to construct autocratic societies, they may also mean that life in tight autocratic societies tends to engender members with a high need for closure. These two are not necessarily incompatible. Their existence and interrelation could be profitably probed in further research.

In addition to influencing group structure, intensified quest for uniformity under heightened need for closure was found to lead to an intolerance of diversity (Kruglanski et al., 2002; Shah et al., 1998). Diversity may impede the arrival at consensus, thereby reducing the group's ability to reach closure. In this vein, heightened need for closure, through the implementation of time pressure and ambient noise, has been shown to lead to a rejection of opinion deviates in a working group (Kruglanski and Webster, 1991). Elevated need for closure was also found to foster favoritism toward one's ingroup, in direct proportion to its degree of homogeneity and opinion uniformity. Finally, need for closure was found to foster outgroup derogation (Kruglanski et al., 2002; Shah et al., 1998), whose degree was *inversely* related to the outgroup's homogeneity and opinion uniformity (Kruglanski et al., 2002). These findings are consistent with the notion that high need for closure individuals are attracted to groups (whether ingroups or outgroups) that promise to afford firm shared realities to their members, providing stable cognitive closure.¹⁰

The quest for stable shared reality on part of individuals with high need for closure may express itself in conservatism and the upholding of group norms and traditions. Indeed, both political conservatism

(Jost et al., 2003a, 2003b) and the tendency to maintain stable group norms across generational cycles (Livi, 2003) were found to be related to a heightened need for closure.

Kosic et al. (2004) found evidence that need for closure augments loyalty to one's ingroup and instills a reluctance to abandon it and 'defect' to alternative collectivities. Such loyalty persists to the extent that one's ingroup is salient in the individuals' social environment. If, however, an alternative group's views became overridingly salient, high need for closure may prompt members instead to switch groups. In this vein, Croat and Polish immigrants to Italy who were high (versus low) on need for closure tended to assimilate less to the Italian culture (i.e., they maintained loyalty to their culture of origin) if their social environment at entry consisted of their coethnics. However, if it consisted of members of the host culture (i.e., of Italians) high (versus low) need for closure immigrants tended more to "defect" and assimilate to the Italian culture.

In summary, considerable research findings attest to the considerable role that need for cognitive closure plays in intrapersonal, interpersonal, and group phenomena. At the individual level, these processes affect the formation of social judgments, attitudes and impressions. At the interpersonal level they enter into communication and persuasion, empathy, and negotiation behavior, and at the group level, into the formation of consensus and the forging of stable social realities for the members. In all these domains, and at all these levels of analysis, the need for closure has been shown to constitute a variable with implications for major classes of social psychological phenomena.

REAL-WORLD IMPLICATIONS

The lay epistemic theory has a variety of real world implications. The present emphasis on the concept of evidence and its subjective nature suggests that effective persuasion

should be tailored to recipients' idiosyncratic inference rules, and include the provision of evidence that fits those rules. This suggests the importance of identifying the contents of those inference rules as these may differ across individuals, groups, and cultures. For instance, in recent work with detainees suspected of jihadist terrorism, an attempt has been under way to get a handle on the evidential basis of their ideology in order to counter it effectively in various national deradicalization programs (Bin Kader, 2009; Kruglanski and Gelfand, 2009). Insofar as jihadist suspects are dispersed across detention centers around the globe, and belong to diverse national and cultural groups, their beliefs and rules of inference might vary widely. Hence it is of great importance to divine what these beliefs are in order for them to be appropriately addressed. Research on these issues is currently underway in our labs.

A related topic concerns revered epistemic authorities as a basic evidential category for one's beliefs. Bar (1983, 1999) demonstrated how individuals differ in their hierarchy of epistemic authorities, and how such a hierarchy determines their readiness to accept the statements of various authorities as evidence for their own beliefs. Determining the relative epistemic authority of various sources is imperative when it comes to real-world persuasion efforts in various domains. In the deradicalization programs mentioned earlier, it would be well to ascertain the epistemic authority of different clerics for different detainees, in order to maximize the efficacy of the religious dialogues aimed at undermining the detainees' jihadist ideology.¹¹

Similarly, determination of therapists' epistemic authority for clients may serve to improve the outcomes of clinical treatments (Kruglanski and Jaffe, 1988; Abramson et al., 1990). The same should hold for determining the epistemic authority of teachers for students in educational settings (Ellis and Kruglanski, 1992). Though some initial research exists in these domains, additional applications are well worthy of further investigation.

The need-for-closure concept also has numerous real-world relevancies. In clinical therapy contexts, lowering the need for closure could promote the “unfreezing” of maladaptive (“irrational”) beliefs (Beck et al., 1987; Ellis and Yeager, 1989) and refreezing more adaptive ones in their stead (Kruglanski and Jaffe, 1988). The “unfreezing” and “refreezing” processes may also be applied to alternative belief systems; for example, those in the political domain. Recently, Gayer et al. (2009) carried out three studies in which they attempted to “unfreeze” Israelis’ attitudes toward the Israeli–Palestinian conflict by pointing to the potential costs of failing to reach agreement on a two-state solution with the Palestinians, likely inducing a “fear of invalidity” that would lower the need for closure on this topic (Kruglanski, 2004). “Unfreezing” in this research was assessed by research participants’ readiness to expose themselves to a wide array of information sources (a broad variety of press, television, and radio programs). It was found that participants’ general perception of losses (in Study 1) was significantly correlated with their “unfreezing” tendency, and that priming evidence about losses (Studies 2 and 3) augmented “unfreezing.”

Whereas lowering the need for closure may induce “unfreezing,” heightening this need may augment “freezing.” This might have several important real-world consequences as well. One of these, already commented on, involves the adoption of conservative political attitudes (Jost et al., 2003a, 2003b). The reason for this is that political conservatism includes resistance to change and traditionalism as its central features, hence “freezing” on one’s prior opinions and perspectives. Though need for closure might also induce a “freezing” on prior liberal ideologies, hence inducing a “dogmatism of the left,” the contents of liberal ideologies are less closure-promoting, which makes “dogmatism of the right” more likely than “dogmatism of the left” (Jost et al., 2003a).

Because need for closure augments one’s attraction to ingroups and the derogation of

outgroups, it induces a particularly aggressive response in cases of intergroup conflict. In this vein, Orehk et al. (2008) recently found in a series of studies that need for closure is correlated with support for tough antiterrorism measures (including torture of suspected terrorists, and their imprisonment in foreign lands). Need for closure was also found to be positively correlated with support for a decisive and rigid presidential candidate who was likely to provide closure and to be negatively correlated with a flexible and open-minded candidate who was unlikely to do so.

In conclusion, because the processes of knowledge formation and change are related to almost all conceivable domains of human endeavors, the lay epistemic theory may enable insights to a particularly broad range of real-world problems. The full extent of this potential has yet to be explored.

NOTES

1 Years later, I discovered that what Kelley meant by “lay epistemology” was a far cry from what I took it to mean. Whereas I took it to mean the *process* of knowledge construction, Kelley intended by it the *contents* of people’s beliefs and their ways of representing an array of social situations and personal relationships. Kelley’s theoretical work on this topic culminated in his publication with a number of colleagues of the *Atlas of Interpersonal Situations* (Kelley et al., 2003), a volume that has hardly to do with knowledge construction issues that formed the body of my own “lay epistemics.”

2 Not necessarily because of the increased *falsification opportunity* (as explained later), but because science essentially aims at reducing the observed variability in phenomena by identifying them as special instances of general principles that they share in common. In that sense, the more general the principle, the greater the theory’s explanatory power, and the greater the theory’s correspondence with the quintessential aim of science, which is explanation (Popper, 1959).

3 This line of reasoning also suggests that covariation shouldn’t be regarded as the exclusive evidence for causality. Thus, temporal precedence of cause over effect is another implication of the cause concept and could represent an important category of evidence for causality. Causality could also be inferred

from statements of a general epistemic authority (see Kruglanski et al., 2005) that something is the cause of something else representing another category of evidence for causality (and other things) that could lead to confident causal attributions.

4 Popper and other philosophers of science (e.g., Paul Feyerabend or Imre Lakatos) have noted that whereas knowledge formation is guided by the ideal of truth, one can never be certain that this ideal has been realized. This implies that the concept of "knowledge" is best understood in its subjective sense, as a *belief*. This hardly implies that knowledge must be solipsistic or idiosyncratic. To the contrary, knowledge typically is socially shared, and knowledge construction (whether scientific or lay) is accomplished via a communal process (Hardin and Higgins, 1996).

5 The logic of confirming a hypothesis on the basis of evidence seems immanent in the rationale of a research design in which all alternative hypotheses are controlled for that to the researcher's mind seem capable of producing a given effect, hence (subjectively) warranting the premise that (insofar as all other possible factors are controlled for) *if and only if* the focal research hypothesis is true, a given data pattern will be observed.

6 The degree to which a rule is activable in given circumstances may depend on its degree of routinization. Some rules can be routinized to the point of "automatization" which may render their use relatively independent of processing resources. In other words, the distinction between "automatic" and "deliberative" processes may represent the (quantitative) parameter of processing difficulty rather than a qualitatively different mode of judgment.

7 In early publications (e.g., Kruglanski, 1980; Kruglanski and Freund, 1983) I used the term 'need for structure' to denote the motivational stopping mechanism of knowledge formation. However, my close friend and colleague Tory Higgins convinced me that what primarily mattered to me was the closed-mindedness induced by the motivational force I was attempting to identify rather than the structured character of the resulting knowledge. Thus, the term "need for closure" was born.

8 In a recent paper, Roets et al. (2008) argued that in addition to exerting a direct motivational effect similar to that of dispositional NFCC, situational manipulations of need for closure (via time pressure or noise) exert an effect on cognitive capacity as well as manifesting itself in deteriorated task performance.

9 This may depend also on whether the object is evaluatively ambiguous or not. In this vein, Higgins and McCann (1984) found that when the object is evaluatively ambiguous or unclear there is even more tuning in to the audience, in particular if the latter has high epistemic authority.

10 This may depend on whether the individual feels that she/he could join the outgroup or not. If she/he couldn't, a highly cohesive and consensual outgroup might elicit even greater derogation.

11 In the extensive deradicalization program in Iraq headed up by US Major General Douglas Stone, it was clear that American sources would have, if anything, a negative epistemic authority for the detainees. Consequently, this successful program (on which basis no less than 18,000 suspected terrorists have been released) relied exclusively on Iraqi sources (whether religious or secular) in work with the detainees.

REFERENCES

Abramson, L.Y., Alloy, L.B. and Dykman, B.M. (1990) Cognitive theory and therapy of emotional disorders: A lay epistemic reconstruction? *Psychological Inquiry*, 1, 198–219.

Andersen, S.M. and Cole, S.W. (1990) 'Do I know you?' The role of significant others in general social perception. *Journal of Personality and Social Psychology*, 59, 384–399.

Andersen, S.M., Glassman, N.S., Chen, S. and Cole, S.W. (1995). Transference in social perception: The role of chronic accessibility in significant-other representations. *Journal of Personality and Social Psychology*, 69, 41–57.

Anderson, J.R. (1983) *The Structure of Cognition*. Cambridge, MA: Harvard University Press.

Baeyens, F., Crombez, G., Van den Bergh, O. and Eelen, P. (1988) Once in contact always in contact: Evaluative conditioning in resistance to extinction. *Advances in Behavior Research and Therapy*, 10, 179–199.

Bar, R. (1983) Hierarchy of epistemic authority test. Unpublished MSc thesis, Tel-Aviv University.

Bar, R. (1999) The impact of epistemic needs and authorities on judgment and decision making. Unpublished PhD dissertation, Tel Aviv University.

Bar-Tal, D. (1990) *Group Beliefs: A Conception for Analyzing Group Structure, Processes, and Behavior*. New York: Springer-Verlag.

Bar-Tal, D. (2000) *Shared Beliefs in a Society: A Social Psychological Analysis*. Thousand Oaks, CA: Sage Publications.

Bartlett, F.C. (1932) *Remembering: A Study in Experimental and Social Psychology*. New York: Cambridge University Press.

Beck, A.T., Shaw, B.F., Rush, A.J. and Emory, G. (1987) *Cognitive Theory of Depression*. New York: The Guilford Press.

Bem, D.J. (1972) Constructing cross-situational consistencies in behavior: Some thoughts on Aiker's critique of Mischel. *Journal of Personality*, 40, 17–26.

Bin Kader, A.H. (Ed.) (2009) *Countering Radicalism: The Next Generation and Challenges Ahead*. Singapore: Taman Bacaan Pemuda pemudi Melayu Singapura.

Boudreau, L.A., Baron, R.M. and Oliver, P.V. (1992) Effects of expected communication target expertise and timing of set on trait use in person description. *Journal of Personality and Social Psychology*, 63, 447–451.

Bruner, J.S. (1973) *Beyond the Information Given: Studies in the Psychology of Knowing*. Oxford: Norton.

Cattell, R.B. and Epstein, A.R. (1975) *Comprehensive Ability Battery*. Champaign, IL: Institute for Personality and Ability Testing.

Chaiken, S., Liberman, A. and Eagly, A.H. (1989) Heuristic and systematic information processing within and beyond the persuasion context. In J.S. Uleman and J.A. Bargh (eds), *Unintended Thought*. New York: Guilford Press.

Chirumbolo, A. (2002) The relationship between need for cognitive closure and political orientation: The mediating role of authoritarianism. *Personality and Individual Differences*, 32, 603–610.

Chun, W.Y. and Kruglanski, A.W. (2006) The role of task demands and processing resources in the use of base-rate and individuating information. *Journal of Personality and Social Psychology*, 91, 205–217.

Chun, W.Y., Spiegel, S. and Kruglanski, A.W. (2002) Assimilative behavior Identification can also be resource dependent: A unimodel perspective on personal-attribution phases. *Journal of Personality and Social Psychology*, 83, 542–555.

Dechesne, M., Schultz, J.M., Kruglanski, A.W., Orehek, E. and Fishman, S. (2008) A psychology of borders: need for closure and the allure of group impermeability. Unpublished manuscript, University of Maryland.

DeDreu, C.K.V. and Koole, S.L. (1997) Motivated use of heuristics in negotiation. Unpublished raw data.

DeDreu, C.K.W., Koole, S.L. and Oldersma, F.L. (1999) On the seizing and freezing of negotiator inferences: Need for cognitive closure moderates the use of heuristics in negotiation. *Personality and Social Psychology Bulletin*, 25, 348–362.

De Grada, E., Kruglanski, A.W., Mannetti, L. and Pierro, A. (1999) Motivated cognition and group interaction: Need for closure affects the contents and processes of collective negotiations. *Journal of Experimental Social Psychology*, 35, 346–365.

Dewey, J. (1916) *Democracy and Education*. New York: Free Press.

Dewey, J. (1958) *Experience and Nature*. New York: Collier Macmillan.

Ditto, P.H. and Lopez, D.F. (1992) Motivated skepticism: Use of differential decision criteria for preferred and non preferred conclusions. *Journal of Personality and Social Psychology*, 63, 568–584.

Erb, H.P., Kruglanski, A.W., Chun, W.Y., Pierro, A., Mannetti, L. and Spiegel, S. (2003) Searching for commonalities in human judgment: The parametric unimodel and its dual mode alternatives. In W. Stroebe and M. Hewstone (eds), *European Review of Social Psychology*, 14, 1–47.

Ellis, S. and Kruglanski, A.W. (1992) Self as epistemic authority: Effects on experiential and instructional learning. *Social Cognition*, 10, 357–375.

Ellis, A. and Yeager, R.J. (1989) *Why Some Therapies Don't Work: The Dangers of Transpersonal Psychology*. Amherst: Prometheus Press.

Fazio, R.H. and Zanna, M.P. (1981) Direct experience and attitude-behavior consistency. In L. Berkowitz (ed.), *Advances in Experimental Social Psychology*, 14, 161–203. San Diego, CA: Academic Press.

Festinger, L. and Carlsmith, M.E. (1959) Cognitive consequences of forced compliance. *Journal of Abnormal Psychology*, 58, 203–210.

Gawronski, B. and Bodenhausen, G.V. (2006) Associative and propositional processes in Evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin*, 132, 692–731.

Gayer, C., Landman, S., Halperin, E. and Bar-Tal, D. (2009) Overcoming psychological barriers to peaceful conflict resolution: The role of arguments about losses. Unpublished manuscript, Tel Aviv University.

Gelfand, M.J. (2008) Situated culture: A multilevel analysis of situational constraint across 35 nations. Paper presented at Social Psychology Winter Conference, Park City, Utah.

Gregory, R. L. (1997) *Eye and Brain: The Psychology of Seeing* (5th edition). Princeton: Princeton University Press.

Hardin, C.D. and Higgins, E.T. (1996) Shared reality: How social verification makes the subjective objective. In R.M. Sorrentino and E.T. Higgins (eds), *Handbook of Motivation and Cognition*, 3, 28–84. New York: Guilford Press.

Helmholtz, H. v. (1910/2000) *Helmholtz's Treatise on Physiological Optics*. Bristol: Thoemmes Press.

Higgins, E.T. (1996) Knowledge activation: Accessibility, applicability and salience. In E.T. Higgins and A.W. Kruglanski (eds), *Social Psychology: Handbook of*

Basic Principles, 133–168. New York: Guilford Press.

Holyoak, K.J., Kohl, K. and Nisbett, R.E. (1989) A theory of conditional reasoning: Inductive learning within rule-based hierarchies. *Psychological Review*, 96, 315–340.

Hovland, C.I., Janis, I.L. and Kelley, H.H. (1953) *Communication and Persuasion*. Newhaven, CT: Yale University Press.

James, W. (1890) *The Principles of Psychology*. New York: Henry Holt and Co.

Jost, J.T., Glaser, J., Kruglanski, A.W. and Sullaway, F.J. (2003a) Exceptions that prove the rule: using a theory of motivated social cognition to account for ideological incongruities and political anomalies. *Psychological Bulletin*, 129, 383–393.

Jost, J.T., Glaser, J., Kruglanski, A.W. and Sullaway, F.J. (2003b) Political conservatism as motivated social cognition. *Psychological Bulletin*, 129, 339–375.

Kahneman, D. (2003) A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, 58, 697–720.

Kahneman, D. and Tversky, A. (1973) On the psychology of prediction. *Psychological Review*, 80, 237–251.

Kelley, H.H. (1967) Attribution theory in social psychology. *Nebraska Symposium on Motivation*, 15, 192–238.

Kelley, H.H. (1971) Attribution in social interaction. In E.E. Jones, D.E. Kanouse, H.H. Kelley, R.E. Nisbett, S. Valins and B. Weiner (eds), *Attribution: Perceiving the Causes of Behavior*. Morristown, NJ: General Learning Press.

Kelley, H.H., Holmes, J.G., Kerr, N.L., Reis, H.T., Rusbult, C.E. and Van Lange, P.A.M. (2003) *An Atlas of Interpersonal Situations*. Cambridge: Cambridge University Press.

Kelley, H.H. and Thibaut, J.W. (1969) Group problem solving. In G. Lindzey and E. Aronson (eds), *The Handbook of Social Psychology*. Reading, MA: Addison-Wesley.

Kosic, A., Kruglanski, A.W., Pierro, A. and Mannetti, L. (2004) Social cognition of immigrants acculturation: effects of the need for closure and the reference group at entry. *Journal of Personality and Social Psychology*, 86, 796–813.

Kruglanski, A.W. (1975) The endogenous–exogenous partition in attribution theory. *Psychological Review*, 82, 87–406.

Kruglanski, A.W. (1980) Lay epistemologic process and contents: Another look at attribution theory. *Psychological Review*, 87, 70–87.

Kruglanski, A.W. (2004) *The Psychology of Closed Mindedness*. New York: Psychology Press.

Kruglanski, A.W., Alon, S. and Lewis, T. (1972) Retrospective misattribution and task enjoyment. *Journal of Experimental Social Psychology*, 8, 493–501.

Kruglanski, A.W. and Gigerenzer, G. (2011) Intuitive and deliberative judgments are based on the same principles. *Psychological Review*.

Kruglanski, A.W., Dechesne, Orehek, E. and Pierro, A. (in press) Three decades of lay epistemics: The why, how and who of knowledge formation. *European Review of Social Psychology*.

Kruglanski, A. and Freund, T. (1983) The freezing and un-freezing of lay-inferences: Effects on impressional primacy, ethnic stereotyping and numerical anchoring. *Journal of Experimental Social Psychology*, 19, 448–468.

Kruglanski, A.W. and Gelfand, M.J. (2009) Deradicalization. Paper presented at ?

Kruglanski, A.W. and Jaffe, Y. (1988) Curing by knowing: The epistemic approach to cognitive therapy. In L. Abramson (ed.), *Social Cognition and Clinical Psychology*. New York: The Guilford Press.

Kruglanski, A.W., Pierro, A., Mannetti, L. and DeGrada E. (2006) Groups as epistemic providers: Need for closure and the unfolding of group-centrism. *Psychological Review*, 113, 84–100.

Kruglanski, A.W., Pierro, A., Mannetti, L., Erb, H-P. and Chun, W.Y. (2007) On the parameters of social judgment. In M.P. Zanna (ed.), *Advances in Experimental Social Psychology*, 39, 345–392.

Kruglanski, A.W., Riter, A., Amitai, A., Margolin, R., Shabtai, L. and Zaksh, D. (1975a) Can money enhance intrinsic motivation: A test of the content-consequence hypothesis. *Journal of Personality and Social Psychology*, 31, 744–750.

Kruglanski, A.W., Riter, A., Arazi, D., Agassi, R., Monteqio, J., Peri, I. and Peretz, M. (1975b) The effects of task intrinsic rewards on extrinsic and intrinsic motivation. *Journal of Personality and Social Psychology*, 21, 699–706.

Kruglanski, A.W., Raviv, A., Bar-Tel, D., Raviv, A., Sharvit, K., Ellis, S., Bar, R., Pierro, A. and Mannetti, L. (2005) Says who? Epistemic authority effects in social judgment. In M.P. Zanna (ed.), *Advances in Experimental Social Psychology*, 37, 345–392.

Kruglanski, A.W., Shah, J.Y., Pierro, A. and Mannetti, L. (2002) When similarity breeds content: Need for closure and the allure of homogeneous and self-resembling groups. *Journal of Personality and Social Psychology*, 83, 648–662.

Kruglanski, A.W. and Thomson, E.P. (1999a) Persuasion by a single route: A view from the unimodel. *Psychological Inquiry*, 10, 83–109.

Kruglanski, A.W. and Thomson, E.P. (1999b) The illusory second mode or, the cue is in the message. *Psychological Inquiry*, 10, 182–193.

Kruglanski, A.W., Thomson, E.P. and Spiegel, S. (1999) Separate or equal? Bimodal notions of persuasion and a single process 'unimodel'. In S. Chiken and Y. Trope (eds.), *Dual-Process Theories in Social Psychology*, pp. 293–313. New York: Guilford Press.

Kruglanski, A.W. and Webster, D.M. (1991) Group members' reactions to opinion deviates and conformists at varying degrees of proximity to decision deadline and of environmental noise. *Journal of Personality and Social Psychology*, 61, 212–225.

Kruglanski, A.W., Webster, D.M. and Klem, A. (1993) Motivated resistance and openness to persuasion in the presence or absence of prior information. *Journal of Personality and Social Psychology*, 65, 861–876.

Kuhn, T.S. (1962) *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Livi, S. (2003) *Il bisogno di chiusura cognitiva e la trasmissione delle norme nei piccoli gruppi* [The need for cognitive closure and normtransmission in small groups]. Unpublished PhD dissertation, University of Rome 'La Sapienza'.

Logan, G.D. (1992) Attention and preattention in theories of automaticity. *American Journal of Psychology*, 105, 317–339.

Mitchell, C.J., De Houwer, J. and Lovibond, P.F. (in press) The propositional nature of human associative learning. *Behavioral and Brain Sciences*.

Newman, L.S. and Uleman, J.S. (1989) Spontaneous trait inference. In J.S. Uleman and J.A. Bargh (eds), *Unintended Thought*. New York: Guilford Press.

Orehek, E., Fishman, S., Dechesne, M., Doosje, B., Kruglanski, A.W., Cole, A.P., et al. (2008) Certainty quest and the social response to terrorism. Unpublished manuscript, University of Maryland.

Petty, R.E. and Cacioppo, J.T. (1986) The elaboration likelihood model of persuasion. In L. Berkowitz (ed.), *Advances in Experimental Social Psychology*, 19, 123–205.

Pierro, A. and Kruglanski, A.W. (2005) Revised need for closure scale. Unpublished data, University of Rome 'La Sapienza'.

Pierro, A. and Kruglanski, A.W. (2008) 'Seizing and freezing' on a significant-person schema: Need for closure and the transference effect in social judgment. *Personality and Social Psychology Bulletin*, 34, 1492–1503.

Pierro, A. and Mannetti, L. (2004) Motivated consumer search behavior: The effects of epistemic authority. Unpublished manuscript, University of Rome 'La Sapienza'.

Pierro, A., Mannetti, L., DeGrada, E., Livi, S. and Kruglanski A.W. (2003) Autocracy bias in groups under need for closure. *Personality and Social Psychology Bulletin*, 29, 405–417.

Pierro, A., Mannetti, L., Erb, H-P., Spiegel, S. and Kruglanski, A.W. (2005) Informational length and order of presentation as determinants of persuasion. *Journal of Experimental Social Psychology*, 41, 458–469.

Pierro, A., Mannetti, L., Kruglanski, A.W. and Sleeth-Keppler, D. (2004) Relevance override: On the reduced impact of 'cues' under high motivation conditions or persuasion studies. *Journal of Personality and Social Psychology*, 86, 251–264.

Pierro, A., Mannetti, L., Orehek, E. and Kruglanski, A.W. (2008) Persistence of attitude change and attitude-behavior correspondence based on extensive processing of source information. Unpublished manuscript, University of Rome 'La Sapienza'.

Pizlo, Z. (2001) Perception viewed as an inverse problem. *Vision Research*, 41, 3145–3161.

Popper, K.R. (1959) *The Logic of Scientific Discovery*. New York: Basic Books.

Raviv, A., Bar-Tal, D., Raviv, A. and Houminer, D (1990) Development in children's perception of epistemic authorities. *British Journal of Developmental Psychology*, 8, 157–169.

Rescorla, R.A. (1985) Conditioned inhibition and facilitation. In R.R. Miller and N.E. Spear (eds), *Information Processing as Animals: Conditioned Inhibition*, pp. 299–326. Hillsdale, NJ: Erlbaum.

Rescorla, R.A. and Holland, P.C. (1982) Behavioral studies of social learning in animals. *Annual Review of Psychology*, 33, 265–308.

Rescorla, R.A. and Wagner, A.R. (1972) A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A.H. Black and W.F. Prokasy (eds), *Classical Conditioning II: Current Research and Theory*, pp. 64–99. New York: Appleton-Century-Crofts.

Richter, L. and Kruglanski, A.W. (1999) Motivated search for common ground: Need for closure effects on audience design in interpersonal communication. *Personality and Social Psychology Bulletin*, 25, 1101–1114.

Rock, I. (1983) *The Logic of Perception*. Cambridge, MA: MIT Press.

Roets, A. and van Hiel, A. (2007) Separating ability from need: Clarifying the dimensional structure of

the Need for Closure Scale. *Personality and Social Psychology Bulletin*, 33, 266–280.

Roets, A., van Hiel, A., Cornelis, I. and Soetens, B., (2008) Determinants of task performance and invested effort: A need for closure by relative cognitive capacity interaction analysis. *Personality and Social Bulletin*, 34, 779–792.

Rogers, C.R. (1951) *Client-Centered Therapy: Its Current Practice, Implications, and Theory*. Oxford: Houghton Mifflin.

Rogers, C.R. (1967) *The Therapeutic Relationship and Its Impact: A Study of Psychotherapy with Schizophrenics*. Oxford: Wisconsin Press.

Rubini, M. and Kruglanski, A.W. (1997) Brief encounters ending in estrangement: Motivated language-use and interpersonal rapport. *Journal of Personality and Social Psychology*, 12, 1047–1060.

Shytenberg, G., Gelfand, M.J., Imai, L., Bell, C.B. and Mayer, D. (2008) When others' injustices matters: The role of collectivism and epistemic needs. *Journal of Personality and Social Psychology*.

Shafer, J.B.P. (1978) *Humanistic Psychology*. Englewood Cliffs, NJ: Prentice Hall.

Shah, J.Y., Kruglanski, A.W. and Thompson, E.P. (1998) Membership has its (epistemic) rewards: Need for closure effects on ingroup bias. *Journal of Personality and Social Psychology*, 75, 383–393.

Sloman, S.A. (1996) The empirical case for two systems of reasoning. *Psychological Bulletin*, 119, 3–22.

Smith, E.R. (1984) Model of social inference processes. *Psychological Review*, 91, 392–413.

Smith, E.R. (1989) Procedural efficiency: general and specific components and effects on social judgment. *Journal of Experimental Social Psychology*, 25, 500–523.

Smith, E.R. and Branscombe, N.R. (1988) Category accessibility as implicit memory. *Journal of Experimental Social Psychology*, 24, 490–504.

Smith, E.R., Branscombe, N.R. and Bormann, C. (1988) Generality of the effects of practice on social judgment tasks. *Journal of Personality and Social Psychology*, 54, 385–395.

Strack, F. and Deutsch, R. (2004) Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, 8, 220–247.

Trope, Y. and Alfieri, T. (1997) Effortfulness and flexibility of dispositional judgment processes. *Journal of Personality and Social Psychology*, 73, 662–674.

Uleman, J.S. (1987) Consciousness and control: The case of spontaneous trait inferences. *Personality and Social Psychology Bulletin*, 13, 337–354.

Walther, E., Nagengast, B. and Trasselli, C. (2005). Evaluative conditioning in social psychology: Facts and speculations. *Cognition and Emotion*, 19, 175–196.

Webster, D.M. and Kruglanski, A.W. (1994) Individual differences in need for cognitive closure. *Journal of Personality and Social Psychology*, 67, 1049–1062.

Webster, D.M., Kruglanski, A.W. and Pattison, D.A. (1997) Motivated language use in intergroup contexts: Need for closure effects on the linguistic intergroup bias. *Journal of Personality and Social Psychology*, 72, 1122–1131.

Webster-Nelson, D., Klein, C.T. and Irvin, J. E. (2003) Motivational antecedents of empathy: Inhibiting effects of fatigue. *Basic and Applied Social Psychology*, 25, 37–50.

Weiner, B. (1979) A theory of motivation for some classroom experiences. *Journal of Educational Psychology*, 71–3–25.

Weiner, B. (1985) An attributional theory of achievement motivation and emotion. *Psychological Review*, 92, 548–573.

The Elaboration Likelihood Model

Richard E. Petty and Pablo Briñol

ABSTRACT

This chapter traces the development of the elaboration likelihood model (ELM) across three decades of research. The ELM began as a theory about the processes responsible for attitude change and the strength of the attitudes that result from those processes. It has now been applied to a wide variety of judgmental change phenomena. By focusing on the core mechanisms of change, the ELM has served to organize the many different theories, outcomes, and variables relevant to persuasion and influence. This review describes four fundamental ideas from the ELM and six phases of ELM research. A key strength of the ELM is that it provides a useful framework from which to understand the moderation and mediation of changes in attitudes as well as other judgmental outcomes from reducing prejudice to the impact of classic heuristics that influence choice and decision making.

INTRODUCTION

The study of attitudes and persuasion has one of the longest histories in social psychology (see Briñol & Petty, *in press*). At one point,

the study of attitudes was considered to be the single most indispensable topic in the discipline (Allport, 1935). Empirical studies on persuasion were among the first in the field and Carl Hovland's massive program of research on attitude change during and after World War II set the core topics and provided the research agenda for decades afterwards (see McGuire, 1968). The study of attitude change became so popular that by the 1970s, there were hundreds of studies and many conceptual analyses. Indeed, so much research and so many specific theories had accumulated that this area of inquiry was in danger of collapse from the weight of competing theories and conflicting findings.

One problem was that seemingly simple variables such as the credibility of the message source that were expected to have a relatively straightforward impact on persuasion according to the persuasion theories of the time, instead produced a mystifying diversity of findings. The accumulated research results just did not support the widespread simple main effect assumptions that accepted theories had for the persuasion

outcomes of many variables (Petty, 1997). For instance, associating a message with an expert source, though usually good for persuasion, sometimes led to reduced influence. Another critical problem was that the core concept of attitudes was under attack largely because in some studies attitudes appeared to be consequential (e.g., guiding behavior) but more often, it seemed, they were not (Wicker, 1969). The surprising complexity of research findings caused most reviewers of the attitudes literature in the 1970s to be quite pessimistic about the usefulness of additional research (e.g., see Fishbein and Ajzen, 1972; Kiesler and Munson, 1975).

It was against this backdrop that the elaboration likelihood model (ELM) was born as a collaborative effort between Richard Petty and John Cacioppo while they were graduate students at Ohio State University in the mid-1970s. For his dissertation, Petty decided to tackle the problem of why some attitude changes persisted over time whereas others were very ephemeral. Drawing on the available literature and personal experience, Petty speculated that when attitude change was produced thoughtfully (such as after listening to strong arguments presented by John Cacioppo), the new judgments were relatively persistent whereas when attitude change occurred with relatively little thinking (such as when deciding you liked someone based on a first impression), the resulting judgment was more transitory. When Tim Brock, advisor for the dissertation, first learned about the planned studies on the persistence of persuasion, he challenged Petty to be more grandiose and propose a more general theory of attitude change rather than focusing on the more narrow attitude persistence hypothesis alone.

Intrigued by the challenge, Petty drew his friend and roommate, Cacioppo, into a long series of late-night (sometimes heated) conversations about the formation and change of attitudes that served as the foundation for the theory that was to come. The core two routes to persuasion idea (i.e., relatively thoughtful or not) was first presented in the

final chapter of Petty's dissertation following empirical studies focusing on memory for one's own issue-relevant thoughts as a determinant of the persistence of attitude change (see Petty, 1977). The dissertation also benefited greatly by the presence of Tony Greenwald on the Ph.D. committee. Greenwald (1968) had earlier proposed a "cognitive response" approach to attitude change which focused on a high elaboration-mechanism by which persuasion occurred or was resisted (i.e., actively generating favorable or unfavorable thoughts to the message arguments). The addition of a low thinking route to persuasion built on Greenwald's earlier approach.

The two routes to persuasion theory did not receive the *elaboration likelihood model* (ELM) name until it was first used in a textbook on persuasion that Petty and Cacioppo (1981) wrote in their first few years out of graduate school. The name was developed after John Harvey, editor of the series in which the book was to appear, advised that a formal name was *essential* if the idea was to stick. In hindsight, it was clear that he was right! The title of the theory was selected to convey the core idea that the high versus low thought processes of persuasion formed a continuum rather than a discrete pair.

Interestingly, at about the same time, Shelly Chaiken, a graduate student at the University of Massachusetts working on her dissertation with Alice Eagly, was also developing the idea that persuasion was sometimes the result of effortful thinking but was sometimes the result of a lower effort reliance on simple heuristics such as "experts are correct" (see Chaiken, 1978). Without awareness of each other's dissertation work, both Petty and Chaiken entered the job market in the same year and even competed for the same positions at several universities, probably (to the bewilderment of the audience) giving similar job talks. Over time, they became good-natured rivals and friends. Chaiken's theory was first called the heuristic model (Chaiken, 1987) to emphasize this unique low effort mechanism

of persuasion and eventually the heuristic-systematic model (HSM) in order to highlight the low versus high effort processes involved (see Chaiken et al., 1989). Although the ELM and HSM stem from different conceptual traditions (i.e., cognitive response theory versus message learning theory) and use different language and terminology, the theories have far more in common than they have points of divergence (see Petty and Wegener, 1998, 1999). Most importantly, the joint appearance of these theories and the research inspired by them did much to foster a more general interest in what became an explosion of dual process (see Chaiken and Trope, 1999) and dual system (e.g., Deutsch and Strack, 2006) approaches to judgment.

In any case, by the mid-1980s a good number of studies had emerged testing various ELM ideas and Petty and Cacioppo (1986a) summarized the accumulated research in a monograph in which the ELM was first presented as a series of seven formal postulates (see also Petty and Cacioppo, 1986b). In the years since then, as more work on the ELM was published, various new summaries of research guided by the ELM have appeared (e.g., Petty and Wegener, 1999) of which this chapter is the most current. From its inception, the ELM was developed to account for the complicated, contradictory, and even perplexing results obtained in the accumulated persuasion literature. It also aimed to provide an integrative framework from which past research findings could be understood as well as new predictions generated in the attitudes domain and beyond. In describing the development of the ELM over time, we will also highlight some of the key people who played important roles.

FOUR CORE ELM IDEAS

The ELM has been presented both schematically (e.g., Petty, 1977; Petty and Cacioppo, 1981, 1986a, 1986b; see Figure 11.1) and as

a series of formal propositions (Petty and Cacioppo, 1986a, 1986b; Petty and Wegener, 1999). Stripped to its bare bones, however, the ELM does four essential things. First, it highlights the fact that modifying people's attitudes or other judgments can be done with a high degree of thought or a relatively low degree of thought. That is, the "elaboration continuum" ranges from low to high.

Second, the ELM holds that there are numerous specific processes of change that operate along this continuum (e.g., classical conditioning and mere exposure require relatively little thought and operate at the low end of the continuum, but expectancy-value and cognitive response models require high degrees of thought and operate along the upper end of the continuum). When the operation of processes at the low end of the continuum determines attitudes, persuasion is said to follow the *peripheral route* whereas when the operation of processes at the high end of the continuum determines attitudes, persuasion is said to follow the *central route*. Of course, much of the time, persuasion is determined by a mixture of these processes.

The third thing the ELM does is to postulate that it matters whether persuasion occurs as the result of relatively high or low amounts of thought. This is because the degree of thought behind a judgment determines how consequential that judgment is. Specifically, the more a judgment is based on thinking about the merits of an issue, the more it tends to persist over time, resist attempts at change, and has consequences for other judgments and behavior (Petty et al., 1995).

The fourth and arguably most useful thing that the ELM does is to organize the many specific processes by which variables can affect attitudes into a finite set that operate at different points along the elaboration continuum. For example, the ELM postulates that one of the things that variables such as the attractiveness of the source of a message or the incidental emotion a person is experiencing can do is to affect how much thinking a person is doing – placing them somewhere

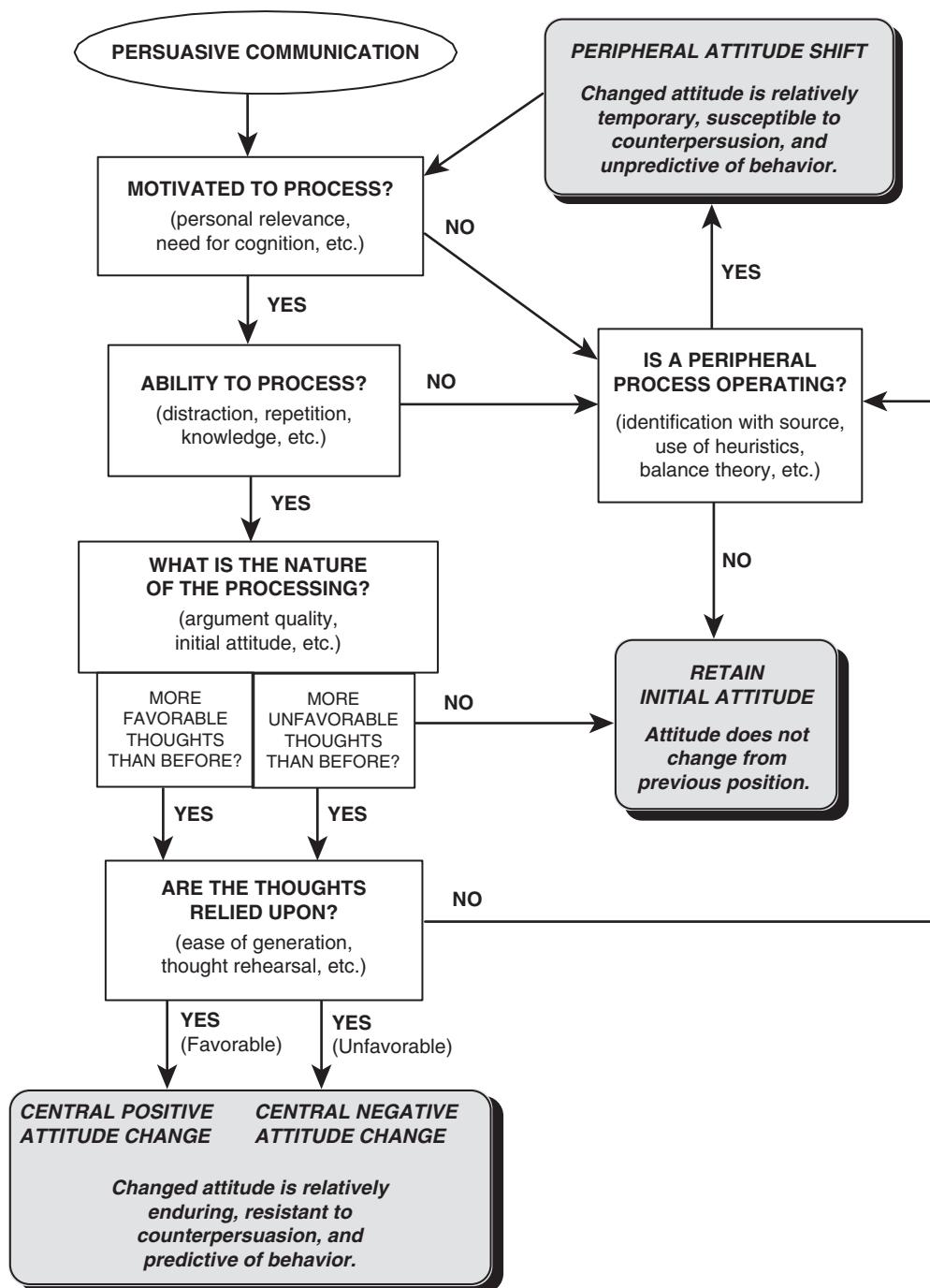


Figure 11.1 Schematic depiction of the Elaboration Likelihood Model

along the elaboration continuum. However, if circumstances have already conspired to place the person at the low end of the thinking continuum, then variables can serve as simple cues, affecting attitudes in a direction that is consistent with their valence (e.g., an attractive source or a positive emotion would lead to positive persuasion outcomes). If the person is at the high end of the elaboration continuum, then there are three other ways in which the variable can affect judgments. Specifically, the variable (1) can be examined as an argument (does the fact that the source is attractive or that the person feels good provide some relevant evidence as to the true merit of what is being advocated?), (2) can affect the valence of the thoughts that come to mind (e.g., exposure to an attractive source or being in a good mood can make positive thoughts more likely to come to mind), and/or (3) can affect a structural feature of the thoughts generated (e.g., an attractive source or feeling happy could make one's thoughts be held with greater confidence). These roles are described in more detail shortly.

SIX PHASES OF ELM RESEARCH

Given a theory with the ambitions and complexity of the ELM, it could not, of course, be tested in a single study, or two or ten. Instead, research on the theory proceeded in a series of stages, and our review will follow these phases in a roughly chronological order. The first stages of work, conducted mostly by Petty and Cacioppo and their various peer and student collaborators, focused on the four core ideas just outlined. Thus, the first phase focused on simply establishing that there was a thinking continuum and that this continuum was consequential for persuasion. The second stage focused on providing evidence for the idea that the mechanism of persuasion could be different under high and low thinking conditions. A third phase examined the consequences of attitudes

changed by high versus low thinking conditions. A fourth phase provided evidence for the so-called “multiple roles” postulate – the idea that any one persuasion variable could affect attitudes in different ways depending on the likelihood of thinking.

Once the four core ELM ideas were supported in the first phases of the research program, a fifth phase of research focused on extending the ELM principles to other judgmental areas beyond the persuasion domain. Although work on each of these phases continues, the most active current phase of research focuses on exploration of a particular role that variables can assume in modifying attitudes or other judgments. Whereas prior research focused on primary cognition – the original association of an attitude object with some attribute – current work is examining the role of secondary cognition (i.e., metacognition). In particular, this work focuses on how and when people assess the validity of their thoughts and what the consequences of this are. We next review the six phases of research on the ELM and present a study that illustrates each.

Phase 1: Exploring the elaboration continuum

In contrast to the earliest attitude change theories that focused on just one process of change (e.g., classical conditioning; Staats and Staats, 1958), the ELM allows for multiple processes that can involve different degrees of thinking. Because different processes of change occur along the thinking continuum, it was important early on to determine the situational and individual difference variables that place people along this continuum. Points along the continuum are determined by how motivated and able people are to assess the fundamental (central) merits of a person, issue, or position (i.e., the attitude object). The ELM assumes that when making an evaluative judgment, the default goal is to determine how good or bad the object truly is. That is, people want to have

attitudes that are subjectively correct. However, people neither have the desire nor the ability to attain equal confidence in every attitude. Thus, motivational and ability factors will determine how much thinking they do in any given situation. For example, it is not worthwhile to exert considerable mental effort to achieve correctness in all situations and people do not always have the requisite knowledge, time, or opportunity to thoughtfully assess the merits of a proposal.

Amount of thinking

In early research relevant to the ELM, it was useful to show that differences in the underlying extent of thinking (elaboration) could provide an explanation for the persuasive effects of variables that had been accounted for in different ways by prior theories.¹ The idea that variables could affect the extent of thinking was also important in explaining how any one variable could both increase and decrease persuasion. As an example, consider a variable like external distraction. Prior research guided by a message-learning approach (e.g., Hovland et al., 1953) suggested that distraction should be bad for persuasion because it would disrupt learning of the message arguments. Prior research guided by dissonance theory, however, suggested that distraction could be good for persuasion because people would have to justify the extra effort they put into processing the message (Baron et al., 1973). Another possibility, suggested by the elaboration continuum idea, was that distraction would affect how much thinking people did about a message.

Imagine a person who is exposed to a message containing eight cogent arguments. The high elaboration processor might think of two or three favorable implications of each of the arguments, whereas the low elaboration processor might think of only one favorable implication (because he or she is not thinking as much). The effect of this is that the high elaboration processor will likely have more favorable attitudes toward the issue than the

low elaboration processor because he or she will have generated more favorable implications of the strong arguments presented and use these thoughts as a guide to the correct attitude. Thus, if distraction reduces processing, when the message contains strong arguments, distraction will be associated with reduced persuasion because fewer favorable thoughts will be generated. This result would be consistent with both learning theory and the ELM. However, what if the message contains weak rather than strong arguments? In this case the high elaboration processor might think of many unfavorable implications of the arguments (i.e., counterarguments), whereas the low elaboration processor might think of only a few. This time the effect is that the high elaboration processor will have less favorable attitudes toward the issue than the low elaboration processor because he or she will have generated more unfavorable implications of the specious arguments presented. When this hypothesis was tested in an empirical study in which the extent of distraction and argument quality were jointly manipulated, the interaction pattern on the attitude data that was expected by the elaboration hypothesis was obtained (Petty et al., 1976, see Figure 11.2, top panel). That is, distraction reduced persuasion when the arguments were strong but increased persuasion when the arguments were weak.²

Following this study on distraction – the first to use a manipulation of argument quality to examine how a variable affects thinking – many other investigations have also used this paradigm. Today, there is a long list of variables that have been shown to affect the extent of thinking and thereby influence attitudes. These variables include message repetition, accountability, and emotion, to name just a few (see Petty and Wegener, 1998, for a review). Perhaps the most studied variable in this regard is the personal relevance of the communication. Linking the message to virtually any aspect of the self appears to increase motivation to think about it (Petty and Cacioppo, 1990). For example, in one study (Petty and

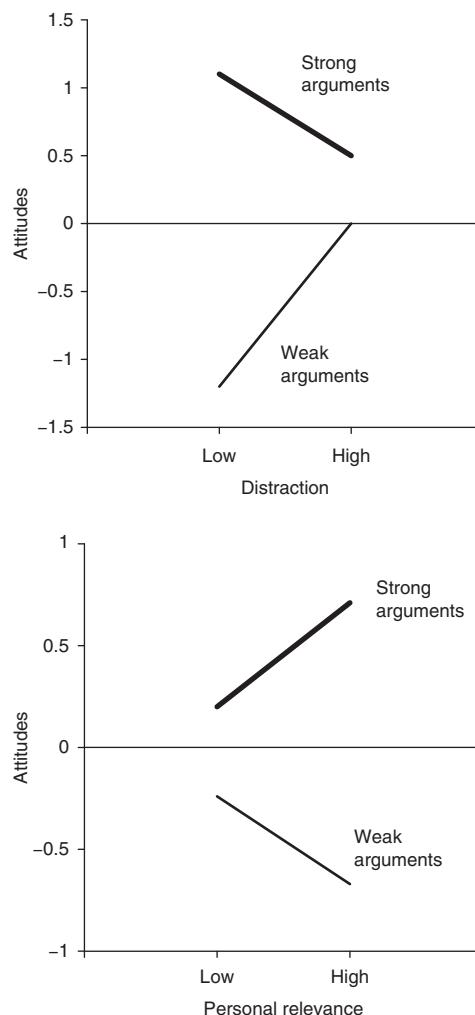


Figure 11.2 **Interactions of variables with argument quality.** Top panel depicts distraction (data from Petty et al., 1976). Bottom panel depicts personal relevance (data from Petty and Cacioppo, 1979b)

Cacioppo, 1979b), undergraduate students were told that a new comprehensive exam policy was going into effect next year or not until ten years in the future. The proposal for requiring students to pass an exam in their major area of concentration as a prerequisite to graduation was supported with either strong or weak arguments. When the policy was said to affect the students personally, argument quality effects were larger than when the policy had no personal relevance

(see bottom panel of Figure 11.2). Or viewed differently, increasing personal relevance tended to increase persuasion when the arguments were strong but to decrease persuasion when the arguments were weak.

In the previous examples, people were more motivated to be thoughtful if the message was linked to the self and they were more able to attain a thoughtful assessment of the arguments if the persuasion context was free of distraction. Although a motive to

be accurate is assumed to be the default goal and underlies the effect on elaboration of variables such as personal relevance, it is not the only motive that affects the extent of information processing. For example, putting people in a positive mood gets them to think more about pleasant messages, not because positive moods or pleasant messages increase the desire to be correct, but because thinking about a pleasant message is hedonically rewarding, and people in positive moods are especially attentive to the hedonic consequences of their actions (Wegener and Petty, 1994; Wegener et al., 1995). In a similar vein, some individuals generally take greater pleasure in thinking than others and thus these individuals (i.e., those high in *need for cognition*; Cacioppo and Petty, 1982) tend to engage in effortful thought because of its intrinsic enjoyment without respect to the importance of the issue or the need to be correct (see Cacioppo et al., 1996; Petty et al., 2009b for reviews). People also generally think more when something makes them feel doubt rather than certainty in their attitudes because doubt is generally less comfortable than certainty and people can try to reduce that discomfort by reassessing their attitudes (see Briñol et al., 2006).³

Biases in thinking

It is important to note that just because a person is thinking intently about a message, the ELM does not assume that the thinking will be totally objective. Rather, the ELM holds that variables can affect not only how much thinking a person is doing, but also whether the thoughts are relatively objective or biased. Consider the personal relevance of the message. We have seen that the more the message connects to the self, the more thinking the message elicits. But is that thinking relatively objective or biased? As outlined by Petty and Cacioppo (1990), this depends on a number of factors. For example, does the message threaten or support one's current views? If a message is relevant (versus irrelevant) to one's outcomes, values, identities, possessions, and so forth, it will engage more

processing. If the message takes a position that is consistent and supportive of one's outcomes, values, and so forth, it will lead to positively biased processing. However, if it takes a position that is counter to or threatening to one's outcomes, values, and so forth, it will lead to negatively biased processing (see Petty et al., 1992).

A number of motivational and ability variables have been shown to bias processing (i.e., affect the valence of the thoughts that come to mind). For example, if a message induces psychological reactance (see Brehm, 1966) by placing undue pressure on an individual to change his or her mind, the person will be motivated to resist and therefore counterargue the message (see Petty and Cacioppo, 1979a). If balance motives (Heider, 1958) are operating, people would prefer to adopt the position of a liked source but distance themselves from a disliked source. If impression management motives (Tedeschi et al., 1971) are in ascendance, people would prefer to hold whatever position they think would be ingratiating and avoid those that would make them look bad. If self-affirmation motives (Steele, 1988) are high, people prefer the position that would make them feel best about themselves, and so forth. Clearly, there are a host of motives that can produce biases in information processing (see Briñol and Petty, 2005). Or, in the absence of any motivational forces, certain factors can uniquely enable positive or negative thoughts (e.g., positive emotions can make positive thoughts more accessible; Petty et al., 1993). In sum, the ELM holds that two of the ways in which a variable can affect attitudes are to (1) affect how much thinking takes place (amount of thinking), and (2) determine whether the thinking is relatively favorable or unfavorable (bias in thinking).

Phase 2: Central and peripheral routes to persuasion

Variables such as distraction and personal relevance can determine where a person falls

along the elaboration continuum. At the high end of the continuum, people assess object-relevant information in relation to knowledge that they already possess, and arrive at a reasoned (though not necessarily unbiased) attitude that is well articulated and bolstered by supporting information (the “central route” to judgment where the focus is on assessing information central to the merits of the attitude object). When people are thinking intently, whether the thoughts are favorable or unfavorable are the key determinants of influence and there are many factors that can motivate or enable favorable or unfavorable thoughts. At the low end of the elaboration continuum, information scrutiny is reduced. Nevertheless, attitude change can still result from a low-effort scrutiny of the information available (e.g., examining less information than when elaboration is high or examining the same information less carefully).

Furthermore, if people are generating few thoughts relevant to the merits of the issue, the ELM holds that there are additional change mechanisms that can come into play to influence attitudes. These mechanisms require relatively little in the way of cognitive resources and include processes such as classical conditioning (Staats and Staats, 1958), self-perception (Bem, 1972), and the use of heuristics (Chaiken, 1987). In one early demonstration of different persuasion mechanisms under high and low thinking conditions, Petty et al. (1981) manipulated the personal relevance of a message along with argument quality just as in the study mentioned earlier (Petty and Cacioppo, 1979b). In addition, however, the expertise of the message source was varied (i.e., whether the message on an educational issue was said to come from a Princeton University Professor or a local high school student). Under high relevance conditions, attitudes were determined by the quality of the arguments, just as in the earlier study. The expertise of the source mattered little when thinking was very high. The new result was what happened under low relevance conditions.

Here, argument quality made little difference and attitudes were only affected by source expertise with more attitude change to the high than the low expert source. This study suggested that attitudes were determined by a high thought process – evaluation of the arguments presented – when motivation to think was high, but by a low thought process – reliance on an expertise cue – when motivation to think was low.

In a critique of the ELM, the HSM, and other dual process theories more generally, Kruglanski and Thompson (1999) correctly noted that many of the early studies on dual processes of persuasion (such as the study just described) compared the impact of relatively simple cues (e.g., expertise) described briefly with more complex verbal arguments (e.g., nine consequences of adopting a recommendation each presented in a separate paragraph). This fact led them to suggest that perhaps there was only one mechanism of persuasion that operated and it only appeared as if there were two separate processes because two separate kinds of content were available to process. The problem, as they saw it, was that evidence for dual processes came from studies in which the central route (or high effort processing) resulted from the impact of complex message factors, and the peripheral route (or low effort processing) resulted from the impact of simple source and other nonmessage factors such as one’s mood.

However, it is not the case that all dual process studies suffer from this confound. At the conceptual level, in the ELM, content (e.g., source versus message variables; simple versus complex presentation) and process (e.g., effortful scrutiny, reliance on cues) are orthogonal. That is, one can engage in effortful scrutiny for merit of message *and* source factors, and these features of the persuasion context can also serve as simple cues to persuasion if thinking is low. Similarly, one can process simple or complexly presented material with relatively high or low amounts of effort. Thus, although some ELM research has manipulated simple source

versus complex message variables to study high versus low effort attitude change as Kruglanski and Thompson noted (e.g., Petty et al., 1981), other ELM research has presented only complex message information to show how it could be processed differently depending on whether motivation to think was relatively high or low (e.g., Petty and Cacioppo, 1984b). Furthermore, some research has manipulated very simple to process source factors (e.g., attractiveness) and pointed to its evaluation as an argument under high thinking conditions but as a peripheral cue under low thinking conditions (Petty and Cacioppo, 1984a).

The point is that when a person's goal in scrutinizing all of the information is to determine the true merits of the proposal, the person will use whatever information seems useful in reaching that goal. Thus, if providing a message recipient with extensive information about the source convinces the person more of the validity of the position when the source information is scrutinized, the impact of the source information could be even larger under high than under low thinking conditions. Conversely, if the source information proves irrelevant to the merits of the attitude object when scrutinized (e.g., an attractive source arguing for a new tax law), then its impact will be reduced under high thinking conditions. Illustrating that different processes can be applied to the *same* information under low and high thinking conditions requires a study in which the information serving as a cue and as a substantive argument is exactly the same (i.e., there are no differences in length, complexity of information, placement, etc.). To demonstrate dual processes, all that should vary is how individuals process and use the *same* information that is presented.

In the relevant conditions of one study, Petty and Cacioppo (1984b) compared how people would respond to a message with three strong arguments versus one with three strong plus three weak arguments. If people are carefully processing the arguments, there should be no more persuasion when

three weak reasons are added to three strong ones. Indeed, the extent of persuasion could even be reduced as negative issue-relevant thoughts are combined with positive thoughts. If people are not processing the messages carefully, however, then evaluation might occur by a different, less effortful process. People might simply *count* the arguments and reason that six arguments are better than three, leading to more persuasion.

To examine whether the same information could be processed differently leading to different persuasion outcomes under conditions fostering relatively high versus low motivation to think, Petty and Cacioppo (1984b) varied the personal relevance of the message topic along with the message type. When relevance was high, adding weak arguments to strong ones did not enhance persuasion but when relevance was low, adding weak arguments to strong ones led to a significant increase in agreement. This study reveals that even though high and low self-relevance individuals were exposed to the exact same information (three strong plus three weak arguments versus three strong only), they used a different evaluation strategy (i.e., processed the information differently) under high and low relevance conditions leading to very different persuasion outcomes. Research such as this demonstrates that the same information can be processed in qualitatively different ways depending on a person's overall motivation and ability to think (see Petty and Briñol, 2006; and Petty et al., 1999, for additional discussion of multi-versus single-process models of persuasion).

Phase 3: Elaboration affects attitude strength

According to the ELM, attitudes that are changed with relatively high versus low amounts of issue-relevant thought are postulated to be stronger than attitudes that are changed to the same extent as a result of minimal object-relevant thought. By stronger,

we mean that the attitudes are more likely to persist over time, resist change, and have an impact on other judgments and behavior (see Krosnick and Petty, 1995). This is true regardless of whether the enhanced thinking taking place is relatively objective or biased. There are several reasons for this. First, as thinking increases during attitude change, people should acquire more support for their attitudes (knowledge) and their attitudes should become more accessible and internally consistent. Furthermore, as a result of thinking, people should become more confident in their views. Each of these factors would increase the likelihood that attitudes would be consequential (see Petty and Krosnick, 1995).

The available evidence supports the idea the elaboration enhances attitude strength. For example, in one set of studies, individuals who engaged in greater thinking during attitude formation showed greater persistence over time and more resistance to change when their newly formed attitudes were challenged immediately compared to individuals who formed similar initial attitudes but with less thinking (Haugtvedt and Petty, 1992). However, it is important to note that persistence over time and resistance to change can be independent such as when multiple pairings of an attitude object with positive cues lead it to persist over time, but do not help it resist attack (Haugtvedt et al., 1994). This is because pairing an attitude object with positive cues can make the favorable attitude memorable, but these cues will not help the attitude resist an attacking message that relies on argumentation (see Wegener et al., 2004, for a review).

Once a person's attitude has changed, behavior change requires that the person's new attitude rather than the old attitude or previous habits guide action. If a new attitude is based on high thought, it is likely to be highly accessible and comes to mind automatically in the presence of the attitude object. Therefore, it will be available to guide behavior even if people do not think much before acting (Fazio, 1990). However, even if

people do engage in some thought prior to action, research suggests that attitudes based on high thinking are still more likely to guide behavior because these attitudes are held with more certainty and people are more willing to act on attitudes in which they have confidence. So strong is the inferential link between thinking and confidence that people do not have to actually engage in more thinking to attain confidence – they only have to believe they have engaged in more thinking (see Barden and Petty, 2008).

Phase 4: Multiple roles for persuasion variables

We have now seen that there is a continuum of thinking that underlies persuasion and that attitudes can be changed by both high and low thought processes with the former attitudes tending to be more consequential than the latter. In outlining these ideas, we have already highlighted several of the roles that a variable can play in producing persuasion. We have seen that variables can serve as cues or as arguments, or they can affect the extent (amount) or direction (bias) in thinking. A fifth role that variables can play when thinking is high is affecting what people think about their thoughts. Since this is the most recent role for variables that has been studied, we discuss it in more depth in a subsequent section (phase 6). But first, it is important to illustrate the ELM principle that any one variable can affect attitudes in multiple ways.

In describing the roles for variables so far, we have mostly used different variables to illustrate each role. Thus, we have seen how distraction can affect the amount of thinking or that source expertise can serve as a simple cue. However, the ELM holds that any *one* variable can serve in each of these roles depending on a number of other factors. In fact, earlier in this chapter we briefly described how an attractive source or a person's good mood could affect attitudes by different processes in different situations.

Empirical research has supported this “multiple roles” view. For example, in one study (Petty et al., 1993) placing an advertisement for a pen in the context of a comedy or bland documentary affected attitudes differently depending on whether people were motivated or not to think about the ad. When motivation to think was high, the pleasant feelings from the positive program led people to have more positive thoughts about the product and these thoughts led to more favorable attitudes. When motivation to think was low, however, the good feelings from the program induced more favorable attitudes toward the product without enhancing the favorability of the thoughts generated (i.e., good feelings served as a simple cue). The low thinking results are what would be expected from relatively low effort theories of attitude change such as classical conditioning (Staats and Staats, 1958) or the use of an ‘affect heuristic’ (Chaiken, 1987; Slovic et al., 2002). Under high thinking conditions, however, the indirect influence observed is what would be expected from relatively high effort theories of the use of affect such as the “affect infusion” hypothesis (Forgas, 1995) in which emotions can make retrieval and generation of affectively congruent cognitive material more likely (see Petty et al., 2003, for a review of emotions and persuasion).

According to the ELM, however, these are just two of the roles that variables can play in persuasion settings. When thinking is high, not only should emotions bias the thoughts that come to mind, but also the emotion itself can be evaluated as an argument. The “mood as input” model of emotions was designed to account for just such situations where people scrutinize their emotions as evidence (see Martin, 2000). There is one more process by which emotions can operate when thinking is high – affecting confidence in thoughts (Briñol et al., 2007), and we discuss this role in the sixth phase of ELM research.

Finally, when the likelihood of thinking is not constrained to be high or low by other variables, emotions can affect the extent of thinking. The “mood as information” theory

of emotions is one of several theories that makes this prediction. The idea is that negative emotions signal that the world is unsafe or problematic and thus information processing is needed. Positive emotions signal the opposite – that the world is safe and thus thinking is not necessary (Schwarz et al., 1991). If sadness, for instance, leads to more thinking than happiness, then people would actually be more persuaded when sad than happy if the message arguments are strong, but less persuaded when sad than happy if the arguments are weak (Bless et al., 1990).

Although different theories of emotion and judgment have developed around each of the specific roles for variables that the ELM holds to be possible, and some theories of emotions have even considered more than one role (e.g., seeForgas, 2005), no other theory incorporates all of these processes. Perhaps more importantly, unlike the specific theories of emotion, the ELM holds that these same fundamental processes can be applied to a host of other variables such as source attractiveness or recipient power that have nothing to do with emotion.

Phase 5: Extending beyond the persuasion context

As described earlier, the ELM was originally proposed as a theory of persuasion (attitude change), but Petty and Cacioppo (1986a) noted that the same principles could be applied to virtually any judgment. Over time, the ELM was used as a framework to study a diversity of persuasive messages on all sorts of topics and in a variety of domains (e.g., health communications, consumer advertisements, legal appeals). A pioneer in moving the ELM beyond persuasion studies was Duane Wegener. Petty met Wegener when the latter came to Ohio State for graduate study in the early 1990s. Wegener was notorious for keeping his advisor (Petty) at work late into the evening with “just one more idea” that he wanted to discuss. Following his PhD, Wegener became a faculty member at Yale,

then Purdue, and he ultimately returned to his alma mater as a faculty member in 2010. Although Wegener developed several influential lines of research that did not involve the ELM (e.g., see Wegener and Petty, 1997), an important ELM contribution was to show that the four core ideas of the ELM outlined above have broader applicability than in the traditional attitude change arena. For example, in the domain of stereotyping, Wegener et al. (2006) showed that a person's existing stereotypes can serve in multiple roles when forming attitudes about a particular member of the stereotyped group. Prior research on stereotyping had focused either on how stereotypes can bias information processing (a high effort process; for example, Kunda and Sherman-Williams, 1993) or on how stereotypes can serve as simple heuristics to judgment (a low effort process; for example, Bodenhausen, 1990). Wegener noted that according to the ELM, however, both roles for stereotypes should be possible depending on the likelihood of thinking.

In one study demonstrating high and low thought roles for stereotypes, Wegener and colleagues (2006) had college students watch a videotape of a child working on some intelligence test questions in which they could observe the answers the child provided. Prior to the videotape, the students learned that that child came from either a high or a low socioeconomic status (SES) background. When not under cognitive load, higher SES led the students to give higher estimates of the child's intelligence and this was mediated by the thoughts listed about the child consistent with the idea that SES could bias processing of the information observed about the child. However, when under cognitive load, the SES information was also associated with greater estimates of intelligence, but this effect was not mediated by thoughts consistent with the use of SES as a heuristic. In a second study, Wegener et al. (2006) showed differential strength consequences for these judgments. That is, the initial impressions of the child that were influenced by the SES stereotype were more resistant to

change by subsequent contradictory information when the initial impressions were formed under high rather than low thought conditions.

Although the studies just described did not use a typical persuasion paradigm, they did involve making evaluative judgments about a target's intelligence. Thus, the ELM might reasonably be expected to operate. What if the judgment requested had nothing to do with evaluation? For a second example of the applicability of ELM principles beyond the persuasion context we turn to another series of studies conducted by Wegener and colleagues, this time on numerical anchoring.

The *anchoring effect* occurs when exposure to a seemingly high (versus low) random number influences participants' numeric responses to a question (Tversky and Kahneman, 1974). For example, if participants are asked to write the last four digits of their social security number on a piece of paper before estimating the age George Washington was when he died, those with high SSNs estimate a higher age than those with low SSNs. Some theories of anchoring assume that the effect occurs by a relatively high effort process conceptually similar to biased processing (e.g., see Mussweiler and Strack, 1999). That is, the anchor biases thoughts in an anchor consistent direction. Other theories, however, assume that anchors work by a less cognitively effortful route. For example, the anchor could provide a simple hint that the answer is large or small (Schwarz, 1994) or prime a general feeling of high or low quantity which is used to infer the answer (Oppenheimer et al., 2008).

As should be clear by now, the ELM suggests that both high and low effort anchoring processes are possible but would operate at different points along the elaboration continuum. To examine this idea, in one study, Blankenship et al. (2008), asked students whether the answer to a particular question (e.g., the age of Neil Armstrong when he walked on the moon) was higher or lower than a presumably randomly generated high or low number. For some participants, during

the anchoring questions (four with high anchors and four with low anchors), they were given a secondary task to perform that would disrupt the anchor from biasing thinking. Other respondents were not distracted during the anchoring task. Finally, all participants responded to the questions both in the initial session and then one week later. At the delayed questioning, no distraction was present. The results of the study revealed that there was a similar anchoring effect initially for both individuals under high and low cognitive load. However, when asked again one week later, the individuals who had presumably used the anchor thoughtfully (low cognitive load) showed greater persistence of the anchoring bias consistent with the idea that when elaboration is involved, it can enhance the strength of any judgment. In another study, the anchoring effect was also shown to be more resistant to counter influence when it was challenged immediately. Thus, the work by Wegener and colleagues shows that the ELM strength postulate appears to hold beyond the prototypical attitude change domain.

Phase 6: A new role for variables – self-validation

As we have seen, in the original formulation of the ELM, under the central route to persuasion, much attention was paid to the number and the valence of thoughts people generated to a persuasive message. Other aspects of thoughts, though mentioned briefly in original treatments of the ELM, received scant research attention. However, in the past decade a particular aspect of thoughts has proven to be very important – the overall confidence people have in the thoughts that they generate. Thought confidence is a metacognition that refers to a sense of how valid one's thoughts seem. Thought confidence is consequential because the extent of thought confidence affects whether people use their thoughts in forming their judgments. This idea is referred to as the *self-validation*

hypothesis (Petty et al., 2002) and is compatible with the lay epistemic notion (Kruglanski, 1990) that people not only generate ideas, but also seek to determine their correctness.

Research on self-validation might not have occurred had Petty not met Pablo Briñol at a two-day conference on “two roads to persuasion” hosted by the University of Salamanca (Spain) in November of 1998. Briñol was a graduate student in social psychology at the Universidad Autónoma de Madrid (UAM) when he decided to attend the conference to learn more about behavioral factors in persuasion, the intended topic of his dissertation. Briñol approached Petty after his talk to ask some questions and ended up serving as translator for Petty for the remaining talks – all given in Spanish. During the session breaks, the pair planned some studies that were aimed at pinning down the mechanism by which the effects of an earlier behavioral manipulation – head nodding (Wells and Petty, 1980) – affected attitudes. When the results of the planned studies subsequently turned out in a surprising way, the self-validation hypothesis was developed.

Specifically, the research on head nodding, which became Briñol's dissertation under the supervision of Petty and Alberto Becerra, showed that head nodding (moving one's head up or down or side to side during exposure to a message) interacted with argument quality to affect attitudes. This interaction result normally would be interpreted as evidence that head nodding affected the extent of thinking about the message, but there was no evidence that this pattern resulted from differences in the number or nature of the thoughts produced. Rather, it appeared that vertical head movements validated the thoughts that people had, magnifying their impact on attitudes. The argument was that nodding (vs. shaking) one's own head served to validate one's own thoughts similar to how other people nodding (vs. shaking) their heads in response to an individual speaking would validate (or invalidate) what the individual was saying via social consensus (Festinger, 1954). When this research was

written for publication, reviewers found the explanation to be a little odd and unconvincing, so Briñol and Petty, along with a new Ohio State graduate student, Zakary Tormala, decided to conduct some more direct tests of the self-validation idea. Following his dissertation defense, Briñol joined the faculty at UAM and for every year since he has spent each fall as a visiting scholar at Ohio State. During this period, much progress on the self-validation hypothesis was made.

In the first direct test of the self-validation notion (Petty et al., 2002: Study 1), Ohio State students were asked to list their thoughts on the issue of a new campus proposal and then rate the confidence they had in their thoughts as well as their attitudes on the topic. A key result of this study was that not only were attitudes affected by the number and valence of thoughts listed (as many prior studies had shown), but also by thought confidence. People were more likely to use thoughts in forming their attitudes when confidence in those thoughts was high rather than low.⁴

Once it was clear that thought confidence was an important factor in translating thoughts into attitudes, it suggested that influencing thought confidence would be one more way in which variables can impact attitudes. Demonstrated ways to affect thought confidence now include head nodding (Briñol and Petty, 2003) and many other variables. As one additional example, consider the well-studied variable of source credibility. We have already noted several roles that credibility could play in producing persuasion (e.g., serving as a simple cue when thinking is low, biasing the thoughts message recipients have when thinking is high, etc.). It is now clear that under certain conditions, source credibility can also affect thought confidence.

In one study (Tormala et al., 2006), information about source credibility was presented *after* participants had processed a message containing either strong or weak arguments. The key idea was that people would reason that if the information presented by the

source was valid (or invalid as inferred from source credibility), their own thoughts in response to the message would also be valid (or invalid). Consistent with this notion, when the message presented strong arguments and thoughts were mostly favorable, increased source credibility was associated with more persuasion because people relied on their positive thoughts. However, when the message presented weak arguments and thoughts were mostly unfavorable, increased source credibility was associated with less persuasion because people relied on their negative thoughts. In other research examining source credibility effects under high thinking conditions, source credibility biased thinking when it preceded the message but affected thought confidence when it came after processing was completed (Tormala et al., 2007).

This work suggests that research on persuasion can benefit from considering the timing of the key manipulations as placement of the independent variable in the sequence of persuasion stimuli can have an impact on the mechanism by which it operates. In accord with the ELM multiple roles idea, the self-validation mechanism operates at the high end of the elaboration continuum and occurs when the sense of confidence experienced is most naturally attributed to one's own thoughts, such as when the feeling of confidence is concurrent with or follows thought generation (see Briñol and Petty, 2009, for a review of the many variables that have now been shown to influence thought confidence).

ADVANTAGES OF THE ELM

The ELM is a multi-faceted theory. It points to different attitude change processes that operate in different circumstances. It suggests that any one variable can work in multiple ways and sometimes produce opposite outcomes (e.g., high source credibility leading to more persuasion when it serves as

a cue but to less persuasion when it enhances thinking about weak arguments). It further indicates that the same persuasion outcome can be produced by different processes (e.g., source credibility leading to more persuasion both when it serves as a cue and when it enhances thinking about strong arguments, validates one's favorable thoughts, or biases thoughts). And, it postulates that not all judgmental outcomes that look the same on the surface really are the same (e.g., the same judgments induced by high versus low thinking processes are differentially persistent over time). In the remainder of this chapter we summarize some of the key benefits of such a multifaceted theory for the field of persuasion and beyond.

Coherence in the field of persuasion

Integration of empirical outcomes

In our view, the ELM has brought some coherence to an attitude change literature that had gotten quite messy. As noted earlier, in the 1970s, numerous scholars complained about the bewildering array of seemingly inconsistent findings in the field and bemoaned the fact that even simple variables could sometimes increase persuasion but at other times reduce it. The ELM explains how and when these different outcomes can occur. It was also confusing that sometimes changed attitudes appeared to be consequential but at other times changed attitudes were not meaningful. The ELM also explains how and when each effect is likely.

In addition to addressing these longstanding puzzles, the ELM has been useful for understanding some current controversies. As one example, consider recent research on implicit measures of attitudes. Contemporary implicit measures aim to assess evaluations that come to mind automatically with little thinking whereas deliberative measures allow some time for reflection (see Petty et al., 2009c). Although the ELM has focused on

how the extent of thinking during attitude formation affects whether attitudes are based on central or peripheral processes, it is possible to apply the elaboration continuum idea to the extent of thinking that occurs during attitude expression. Paralleling previous ELM findings, current research is consistent with the idea that simple cues that do not affect attitudes that are reported on deliberative measures often still have an impact on attitudes that are assessed with measures allowing for little thinking (see Petty and Briñol, 2010, for further review).

Integration of different theories of persuasion

Our discussion of the ELM so far has focused on the ELM as a primary theory of judgment. However, the ELM was also intended as a metatheory (theory about theories) in that it specified what the domain of operation of different theories was. As an early example, Petty and Cacioppo (1986a) noted that the ELM could be used to understand differences between the competing dissonance (Festinger, 1957) and self-perception (Bem, 1972) theories. From the vantage point of the ELM, each of these theories attempted to account for many of the same phenomena (e.g., why people changed their attitudes more when advocating something for a small rather than a large incentive), but did so by very different mechanisms in different situations. Most importantly, self-perception theory relied on a simple inference process and thus it should be more likely to operate on the low end of the elaboration continuum, whereas dissonance theory relied on extensive cognitive activity and thus should be more likely to operate when motivation and ability to think were high. Similarly, we earlier noted how separately developed theories of the impact of emotion on judgment could be organized according to the ELM processes.

Indeed, according to the ELM framework, most of the major theories of attitude change are not necessarily competitive or contradictory, but rather operate in different circumstances. Some theories (e.g., cognitive

response, cognitive dissonance, mood as input) refer to processes that require diligent and effortful information-processing activity, whereas others (e.g., classical conditioning, self-perception, affect heuristic) postulate processes that proceed with considerably less mental effort (see Petty and Cacioppo, 1986b; Petty and Wegener, 1998). The ELM does not diminish the importance of the individual theories. Rather, these theories can be viewed as specifying in more detail the specific process involved under relatively high and low thought conditions. That is, whereas the ELM lumps all kinds of simple cue processes together and all kinds of biased processing theories together, the more specific theories are useful for fleshing out the mechanistic details.

The ELM lumps theories into broad process categories based on the common mechanisms involved, the situations in which they operate, and the consequences observed. For example, cue theories have in common that attitude change moves in the direction of the valence (positive or negative) of the cue, occurs with relatively little thinking, and results in a judgment that is less consequential than a judgment rendered with higher thought. But, the specific way in which this occurs (e.g., conditioned association, use of a heuristic) is also worthy of study. The ELM is designed to be a general approach that can explain the effects of a wide array of variables that have been examined separately under the rubric of different theories.

Integration of source, message, recipient, and context variables

Because of the ELM postulate that any one variable can produce persuasion in multiple ways, the classic source, message, recipient, and context variables that affect attitudes can be examined from a common perspective. That is, one can see how very different variables such as source credibility and a person's emotions operate to influence attitudes by the very same fundamental mechanisms. Furthermore, the ELM provides a useful framework for approaching completely novel

variables. For example, if one wondered how the color of the paper on which a message was printed would influence attitudes, one would look for simple cue effects when thinking was low (e.g., the most liked color would produce the most favorable attitudes), but would look for other effects (e.g., affecting thinking, biasing thinking, validating thoughts) as the elaboration likelihood was increased.

Furthermore, the ELM can shed new light in looking at traditional variables that the literature appears to have relegated to just one role. Consider the operation of self-relevance. Much research has shown that when the self-relevance of a message is made salient prior to a communication, it influences the amount of thinking (Petty and Cacioppo, 1979b). However, when self-relevance is induced after the message, it affects thought confidence (Petty and Briñol, 2011). Although in this case the two processes lead to a similar result (i.e., a greater argument quality effect under high vs. low self-relevance), the underlying mechanism is quite different.

Real-world applications of the ELM

A discussion and review of the many areas of application of the ELM is well beyond the scope of this chapter. Thus, we just briefly note that although much ELM research has been conducted in the laboratory, there is considerable work that has been conducted in field settings as well (e.g., Bakker, 1999). The ELM has proven especially useful in the domains of marketing and advertising (Haugtvedt and Kasmer, 2008; Rucker et al., 2007) and health communication (Briñol and Petty, 2006; Petty et al., 2009a), though there are also applications in the legal, environmental, political, and educational fields as well. Indeed, the ELM has provided practical guidelines for developing effective communications on a wide variety of topics. Tutorials are available to illustrate the actual steps

policymakers and others might take in improving their persuasive appeals using ELM principles (e.g., Briñol and Petty, 2006; Rucker and Petty, 2006).

One of the reasons the ELM has been so widely applied is because persuasion is everywhere, playing an essential role in politics, religion, psychotherapy, education, and day-to-day social interactions. Given that people attempt to persuade others and are also the targets of persuasion, they often wonder about questions such as: are attractive people particularly persuasive? Are experts more persuasive than nonexperts in convincing a jury? Is fear a good emotional tool or is it counterproductive in order to stop people from engaging in risky behaviors? Humans have a longstanding curiosity about such questions and contemporary scholars continue to study these issues as well. The ELM provides answers based on experimental research to many of these questions or suggests ways to initiate new investigations.

We have already noted several of the benefits of focusing on the basic processes underlying effective persuasion. First, identifying the processes by which variables impact attitudes is essential for determining which outcome (increased or decreased persuasion) will occur. Second, we have seen that the process by which an attitude is formed or changed has considerable consequences for the strength of the attitude. Even though both high and low effort processes can sometimes result in the same extent of influence, the attitudes induced by low thinking mechanisms tend to be less stable and predictive of behavior than the ones produced by higher thinking mechanisms. Thus, understanding process is important because it informs us about both immediate and long-term consequences.

As a final illustration of this point, consider our recent research examining whether the principles of the ELM can be applied to the reduction of prejudiced attitudes. Consistent with the ELM, Martin et al. (2011) found that changing attitudes toward

stigmatized groups can be affected by both simple processes that require little thinking and also by traditional elaborative forms of persuasion. Importantly, even when the obtained attitude change was equivalent for processes requiring a low versus a high degree of thinking, there were important benefits of high elaboration prejudice reduction. That is, although both high and low thinking processes were associated with a reduction in the extremity of prejudiced attitudes, the reductions in prejudice produced by high thinking processes were more persistent and resistant to subsequent attacks than equivalent changes produced by less thoughtful mechanisms. As illustrated by this example, the ELM can serve as a basis for, and shed light on, a variety of phenomenon not only relevant to attitude change but also to numerous other judgments, ranging from reducing prejudice to the operation of various heuristics and biases that influence choice and decision making.

NOTES

1 The term "elaboration" is used in the theory to connote that people thoughtfully add something to the information provided externally rather than simply mentally rehearsing the original information. In this sense, the term is more restrictive than "cognitive response" (Greenwald, 1968) which would include the former as well as mere restatements of the message.

2 The arguments are developed in pretesting so that strong arguments elicit primarily favorable thoughts when people are instructed to think about them but weak arguments elicit primarily unfavorable thoughts with the same instructions. All arguments are presented as supporting the advocacy but the strong arguments do so in a more compelling way (e.g., pointing to consequences that are more desirable and likely if the advocacy is adopted; see Petty and Cacioppo, 1986a, for an extended discussion).

3 One exception to this is when people feel certain in an ambivalent attitude. In this case, people engage in greater information processing than if they are uncertain of the ambivalent attitude (Tormala et al., 2008). Similarly, if people feel certainty in a doubted attitude, they could engage in

greater information processing than if they felt uncertainty in a doubted attitude (see Wichman et al., 2010).

4 Thought confidence also predicted attitudes above and beyond other aspects of the thoughts listed such as the likelihood and desirability of the consequences inherent in the thoughts.

REFERENCES

Allport, G.W. (1935) Attitudes. In C. Murchison (ed.), *Handbook of Social Psychology*, 798–884. Worcester, MA: Clark University Press.

Bakker, A.B. (1999) Persuasive communication about AIDS prevention: Need for cognition determines the impact of message format. *AIDS Education and Prevention*, 11, 150–162.

Barden, J. and Petty, R.E. (2008) The mere perception of elaboration creates attitude certainty: Exploring the thoughtfulness heuristic. *Journal of Personality and Social Psychology*, 95, 489–509.

Baron, R.S., Baron, P.H. and Miller, N. (1973) The relation between distraction and persuasion. *Psychological Bulletin*, 80, 310–323.

Bem, D.J. (1972) Self-perception theory. In L. Berkowitz (ed.), *Advances in Experimental Social Psychology*, 6, 1–62. New York: Academic Press.

Blankenship, K.L., Wegener, D.T., Petty, R.E., Detweiler-Bedell, B. and Macy, C.L. (2008) Elaboration and consequences of anchored estimates: An attitudinal perspective on numerical anchoring. *Journal of Experimental Social Psychology*, 44, 1465–1476.

Bless, H., Bohner, G., Schwarz, N. and Strack, F. (1990) Mood and persuasion: A cognitive response analysis. *Personality and Social Psychology Bulletin*, 16, 331–345.

Bodenhausen, G.V. (1990) Stereotypes as judgmental heuristics: Evidence of circadian variation in discrimination. *Psychological Science*, 1, 319–322.

Brehm, J.W. (1966) *A Theory of Psychological Reactance*. New York: Academic Press.

Briñol, P. and Petty, R.E. (2003) Overt head movements and persuasion: A self-validation analysis. *Journal of Personality and Social Psychology*, 84, 1123–1139.

Briñol, P. and Petty, R.E. (2005) Individual differences in persuasion. In D. Albaracín, B.T. Johnson and M.P. Zanna (eds), *Handbook of Attitudes and Attitude Change*, pp. 575–616. Hillsdale, NJ: Erlbaum.

Briñol, P. and Petty, R.E. (2006) Fundamental processes leading to attitude change: Implications for cancer prevention communications. *Journal of Communication*, 56, 81–104.

Briñol, P. and Petty, R.E. (2009) Persuasion: Insights from the self-validation hypothesis. In M.P. Zanna (ed.), *Advances in Experimental Social Psychology*, 41, 69–118. New York: Elsevier.

Briñol, P. and Petty, R.E. (in press) The history of attitudes and persuasion research. In A. Kruglanski and W. Stroebe (eds), *Handbook of the History of Social Psychology*. New York: Psychology Press.

Briñol, P., Petty, R.E. and Barden, J. (2007) Happiness versus sadness as a determinant of thought confidence in persuasion: A self-validation analysis. *Journal of Personality and Social Psychology*, 93, 711–727.

Briñol, P., Petty, R.E. and Wheeler, S.C. (2006) Discrepancies between explicit and implicit self-concepts: Consequences for information processing. *Journal of Personality and Social Psychology*, 91, 154–170.

Cacioppo, J.T. and Petty, R.E. (1982) The need for cognition. *Journal of Personality and Social Psychology*, 42, 116–131.

Cacioppo, J.T., Petty, R.E., Feinstein, J. and Jarvis, W.B.G. (1996) Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin*, 119, 197–253.

Chaiken, S. (1978) The use of source versus message cues in persuasion: An information processing analysis. Unpublished PhD dissertation, University of Massachusetts.

Chaiken, S. (1987) The heuristic model of persuasion. In M.P. Zanna, J. Olson and C.P. Herman (eds), *Social Influence: The Ontario Symposium*, 5, 3–39. Hillsdale, NJ: Lawrence Erlbaum Associates.

Chaiken, S., Liberman, A. and Eagly, A.H. (1989) Heuristic and systematic processing within and beyond the persuasion context. In J.S. Uleman and J.A. Bargh (eds), *Unintended Thought*, pp. 212–252. New York: Guilford Press.

Chaiken, S. and Trope, Y. (eds.) (1999) *Dual Process Theories in Social Psychology*. New York: Guilford Press.

Deutsch, R. and Strack, F. (2006) Duality models in social psychology: From dual processes to interacting systems. *Psychological Inquiry*, 17, 166–172.

Fazio, R.H. (1990) Multiple processes by which attitudes guide behavior: The MODE model as an integrative framework. In M. Zanna (ed.), *Advances in Experimental Social Psychology*, 23, 75–109. San Diego, CA: Academic Press.

Festinger, L. (1954) A theory of social comparison processes. *Human Relations*, 7, 117–140.

Festinger, L. (1957) *A Theory of Cognitive Dissonance*. Stanford, CA: Stanford University Press.

Fishbein, M. and Ajzen, I. (1972) Attitudes and opinions. *Annual Review of Psychology*, 23, 487–544.

Forgas, J.P. (1995) Mood and judgment: The Affect Infusion Model (AIM). *Psychological Bulletin*, 117, 39–66.

Greenwald, A.G. (1968) Cognitive learning, cognitive response to persuasion, and attitude change. In A. Greenwald, T. Brock and T. Ostrom (eds), *Psychological Foundations of Attitudes*, pp. 148–170. New York: Academic Press.

Haugtvedt, C.P. and Kasmer, J. A. (2008) Attitude change and persuasion. In C.P. Haugtvedt, P.M. Herr and F.R. Kardes (eds), *Handbook of Consumer Psychology*, pp. 419–436. New York: Erlbaum.

Haugtvedt, C.P. and Petty, R.E. (1992) Personality and persuasion: Need for cognition moderates the persistence and resistance of attitude changes. *Journal of Personality and Social Psychology*, 63, 308–319.

Haugtvedt, C.P., Schumann, D.W., Schneier, W. and Warren, W. (1994) Advertising repetition and variation strategies: Implications for understanding attitude strength. *Journal of Consumer Research*, 21, 176–189.

Heider, F. (1958) *The Psychology of Interpersonal Relations*. New York: Wiley.

Kiesler, C.A. and Munson, P.A. (1975) Attitudes and opinions. *Annual Review of Psychology*, 26, 415–456.

Krosnick, J.A. and Petty, R.E. (1995) Attitude strength: An overview. In R.E. Petty and J.A. Krosnick (eds), *Attitude Strength: Antecedents and Consequences*, 1–24. Mahwah, NJ: Erlbaum.

Kruglanski, A.W. (1990) Lay epistemic theory in social-cognitive psychology. *Psychological Inquiry*, 1, 181–197.

Kruglanski, A.W. and Thompson, E.P. (1999) Persuasion by a single route: A view from the unimodel. *Psychological Inquiry*, 10, 83–110.

Kunda, Z. and Sherman-Williams, B. (1993) Stereotypes and the construal of individuating information. *Personality and Social Psychology Bulletin*, 19, 90–99.

Martin, L.L. (2000) Moods do not convey information: Moods in context do. In J.P. Forgas (ed.), *Feeling and Thinking: The Role of Affect in Social Cognition*, pp. 153–177. Cambridge, MA: Cambridge University Press.

Martin, M.A., Briñol, P., Horcajo, J. and Petty, R.E. (2011) Changing prejudiced attitudes by thinking about persuasive messages: An attitude strength perspective. Unpublished manuscript.

McGuire, W.J. (1968) Personality and attitude change: An information-processing theory. In A.G. Greenwald, T.C. Brock and T.M. Ostrom (eds), *Psychological Foundations of Attitudes*, pp. 171–196. New York: Academic Press.

Mussweiler, T. and Strack, F. (1999) Hipótesis/consistent testing and semantic priming in the anchoring paradigm: A selective accessibility model. *Journal of Experimental Social Psychology*, 35, 136–164.

Oppenheimer, D.M., LeBoeuf, R.A. and Brewer, N.T. (2008) Anchors aweigh: A demonstration of cross-modality anchoring and magnitude priming. *Cognition*, 106, 13–26.

Petty, R.E., Wells, G.L. and Brock, T.C. (1976) Distraction can enhance or reduce yielding to propaganda: Thought disruption versus effort justification. *Journal of Personality and Social Psychology*, 34, 874–884.

Petty, R.E. (1977) A cognitive response analysis of the temporal persistence of attitude changes induced by persuasive communications. Unpublished PhD dissertation, Ohio State University.

Petty, R.E. and Cacioppo, J.T. (1979a) Effects of forewarning of persuasive intent and involvement on cognitive responses and persuasion. *Personality and Social Psychology Bulletin*, 5, 173–176.

Petty, R.E. and Cacioppo, J.T. (1979b) Issue involvement can increase or decrease persuasion by enhancing message-relevant cognitive responses. *Journal of Personality and Social Psychology*, 37, 1915–1926.

Petty, R.E., Cacioppo, J.T. and Goldman, R. (1981) Personal involvement as a determinant of argument-based persuasion. *Journal of Personality and Social Psychology*, 41, 847–855.

Petty, R.E. and Cacioppo, J.T. (1981) *Attitudes and Persuasion: Classic and Contemporary Approaches*. Dubuque, IA: Wm. C. Brown.

Petty, R.E., Ostrom, T.M. and Brock, T.C. (1981) *Cognitive Responses in Persuasion*. Hillsdale, NJ: Erlbaum.

Petty, R.E. and Cacioppo, J.T. (1984a) Source factors and the elaboration likelihood model of persuasion. *Advances in Consumer Research*, 11, 668–672.

Petty, R.E. and Cacioppo, J.T. (1984b) The effects of involvement on responses to argument quantity and quality: Central and peripheral routes to persuasion. *Journal of Personality and Social Psychology*, 46, 69–81.

Petty, R.E. and Cacioppo, J.T. (1986a) *Communication and Persuasion: Central and Peripheral Routes to Attitude Change*. New York: Springer-Verlag.

Petty, R.E. and Cacioppo, J.T. (1986b) The Elaboration Likelihood Model of persuasion. In L. Berkowitz (ed.), *Advances in Experimental Social Psychology*, 19, 123–205. New York: Academic Press.

Petty, R.E. and Cacioppo, J.T. (1990) Involvement and persuasion: Tradition versus integration. *Psychological Bulletin*, 107, 367–374.

Petty, R.E., Cacioppo, J.T. and Haugvedt, C. (1992) Involvement and persuasion: An appreciative look at the Sherifs' contribution to the study of self-relevance and attitude change. In D. Granberg and G. Sarup (eds), *Social Judgment and Intergroup Relations: Essays in Honor of Muzafer Sherif*, pp. 147–174. New York: Springer/Verlag.

Petty, R.E., Schumann, D.W., Richman, S.A. and Strathman, A.J. (1993) Positive mood and persuasion: Different roles for affect under high and low elaboration conditions. *Journal of Personality and Social Psychology*, 64, 5–20.

Petty, R.E., Haugvedt, C. and Smith, S.M. (1995) Elaboration as a determinant of attitude strength: Creating attitudes that are persistent, resistant, and predictive of behavior. In R.E. Petty and J.A. Krosnick (eds), *Attitude Strength: Antecedents and Consequences*, pp. 93–130. Mahwah, NJ: Erlbaum.

Petty, R.E. and Krosnick, J.A. (1995) (eds) *Attitude Strength: Antecedents and Consequences*. Mahwah, NJ: Erlbaum Associates.

Petty, R.E. (1997) The evolution of theory and research in social psychology: From single to multiple effect and process models. In C. McGarty and S.A. Haslam (eds), *The Message of Social Psychology: Perspectives on Mind in Society*, pp. 268–290. Oxford, UK: Blackwell Publishers.

Petty, R.E. and Wegener, D.T. (1998) Attitude change: Multiple roles for persuasion variables. In D. Gilbert, S. Fiske and G. Lindzey (eds), *The Handbook of Social Psychology*, Vol. 1 (4th edition), pp. 323–390. New York: McGraw-Hill.

Petty, R.E. and Wegener, D.T. (1999) The Elaboration Likelihood Model: Current status and controversies. In S. Chaiken and Y. Trope (eds), *Dual Process Theories in Social Psychology*, pp. 41–72. New York: Guilford Press.

Petty, R.E., Wheeler, S.C. and Bizer, G.Y. (1999) Is there one persuasion process or more? Lumping versus splitting in attitude change theories. *Psychological Inquiry*, 10, 156–163.

Petty, R.E., Briñol, P. and Tormala, Z.L. (2002) Thought confidence as a determinant of persuasion: The self-validation hypothesis. *Journal of Personality and Social Psychology*, 82, 722–741.

Petty, R.E., Fabrigar, L.R. and Wegener, D.T., (2003) Emotional factors in attitudes and persuasion. In R.J. Davidson, K.R. Scherer and H.H. Goldsmith (eds), *Handbook of Affective Sciences*, pp. 752–772. Oxford, UK: Oxford University Press.

Petty, R.E. and Briñol, P. (2006) Understanding social judgment: Multiple systems and processes. *Psychological Inquiry*, 17, 217–223.

Petty, R.E., Barden, J. and Wheeler, S.C. (2009a) The elaboration likelihood model of persuasion: Developing health promotions to produce sustained behavior change. In R.J. DiClemente, R.A. Crosby and M. Kegler (eds), *Emerging Theories in Health Promotion Practice and Research* (2nd edition), pp. 185–214. San Francisco, CA: Jossey-Bass.

Petty, R.E., Briñol, P., Loersch, C. and McCaslin, M.J. (2009b) The need for cognition. In M.R. Leary and R.H. Hoyle (eds), *Handbook of Individual Differences in Social Behavior*, pp. 318–329. New York: Guilford Press.

Petty, R.E., Fazio, R.H. and Briñol, P. (eds) (2009c) *Attitudes: Insights From the New Implicit Measures*. New York: Psychology Press.

Petty, R.E. and Briñol, P. (2010) Attitude structure and change: Implications for implicit measures. In B. Gawronski and B.K. Payne (eds), *Handbook of Implicit Social Cognition: Measurement, Theory, and Applications*, pp. 335–352. New York: Guilford Press.

Petty, R.E. and Briñol, P. (2011) Effects of self-relevance on persuasion: A self-validation analysis. Unpublished manuscript.

Rucker, D.D. and Petty, R.E. (2006) Increasing the effectiveness of communications to consumers: Recommendations based on the Elaboration Likelihood and attitude certainty perspectives. *Journal of Public Policy and Marketing*, 25, 39–52.

Rucker, D.D., Petty, R.E. and Priester, J.R. (2007) Understanding advertising effectiveness from a psychological perspective: The importance of attitudes and attitude strength. In G.J. Tellis and T. Ambler (eds), *The Handbook of Advertising*, pp. 73–88. Thousand Oaks, CA: Sage.

Schwarz, N. (1994) Judgments in a social context: Biases, shortcomings, and the logic of conversation. In M. Zanna (ed.), *Advances in Experimental Social Psychology*, 26, 125–162. San Diego, CA: Academic Press.

Schwarz, N., Bless, H. and Bohner, G. (1991) Mood and persuasion: Affective status influence the processing of persuasive communications. In M. Zanna (ed.), *Advances in Experimental Social Psychology*, 24, 161–197. San Diego, CA: Academic Press.

Slovic, P., Finucane, M.L., Peters, E. and Macgregor, D.G. (2002) The affect heuristic. In T. Gilovich, D. Griffin, and D. Kahneman (eds), *Heuristics and Biases: The Psychology of Intuitive Judgment*, pp. 397–420. New York: Cambridge University Press.

Staats, A.W. and Staats, C. (1958) Attitudes established by classical conditioning. *Journal of Abnormal and Social Psychology*, 67, 159–167.

Steele, C.M. (1988) The psychology of self-affirmation: Sustaining the integrity of the self. In L. Berkowitz (ed.), *Advances in Experimental Social Psychology*, 1, 261–302. New York: Academic Press.

Tedeschi, J.T., Schenker, B.R. and Bonoma, T. (1971) Cognitive dissonance: Private ratiocination or public spectacle? *American Psychologist*, 26, 685–695.

Tormala, Z.L., Briñol, P. and Petty, R.E. (2006) When credibility attacks: The reverse impact of source credibility on persuasion. *Journal of Experimental Social Psychology*, 42, 684–691.

Tormala, Z.L., Briñol, P. and Petty, R.E. (2007) Multiple roles for source credibility under high elaboration: It's all in the timing. *Social Cognition*, 25, 536–552.

Tormala, Z.L., Rucker, D.D. and Seger, C.R. (2008) When increased confidence yields increased thought: A confidence-matching hypothesis. *Journal of Experimental Social Psychology*, 44, 141–147.

Tversky, A. and Kahneman, D. (1974) Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124–1130.

Wegener, D.T., Clark, J. K. and Petty, R.E. (2006) Not all stereotyping is created equal: Differential consequences of thoughtful versus non-thoughtful stereotyping. *Journal of Personality and Social Psychology*, 90, 42–59.

Wegener, D.T. and Petty, R.E. (1994) Mood management across affective states: The hedonic contingency hypothesis. *Journal of Personality and Social Psychology*, 66, 1034–1048.

Wegener, D.T. and Petty, R.E. (1997) The flexible correction model: The role of naive theories of bias in bias correction. In M.P. Zanna (ed.), *Advances in Experimental Social Psychology*, 29, 141–208. San Diego, CA: Academic Press.

Wegener, D.T., Petty, R.E. and Smith, S. M. (1995) Positive mood can increase or decrease message scrutiny: The hedonic contingency view of mood and message processing. *Journal of Personality and Social Psychology*, 69, 5–15.

Wegener, D.T., Petty, R.E., Smoak, N.D. and Fabrigar, L.R. (2004) Multiple routes to resisting attitude change. In E.S. Knowles and J.A. Linn (eds), *Resistance and Persuasion*, pp. 13–38. Mahwah, NJ: Erlbaum.

Wells, G.L. and Petty, R.E. (1980) The effects of overt head movements on persuasion: Compatibility and incompatibility of responses. *Basic and Applied Social Psychology*, 1, 219–230.

Wichman, A.L., Briñol, P., Petty, R.E., Rucker, D.D., Tormala, Z.L. and Weary, G. (2010) Doubting one's doubt: A formula for confidence? *Journal of Experimental Social Psychology*, 46, 350–355.

Wicker, A.W. (1969) Attitudes versus actions: The relationship of verbal and overt behavioral responses to attitude objects. *Journal of Social Issues*, 25, 41–78.

A Theory of Heuristic and Systematic Information Processing

Shelly Chaiken and Alison Ledgerwood

ABSTRACT

The heuristic-systematic model proposes two distinct modes of thinking about information. Systematic processing involves attempts to thoroughly understand any available information through careful attention, deep thinking, and intensive reasoning, whereas heuristic processing involves focusing on salient and easily comprehended cues that activate well-learned judgmental shortcuts. Heuristic processing is a more efficient and relatively automatic mode of processing, but more often than not confers less judgmental confidence. Systematic processing confers more confidence but is relatively effortful and time-consuming. Thus, individuals tend to engage in heuristic processing unless they are both motivated and able to think carefully about information, in which case the two modes of processing can have additive, attenuating, or interactive effects. Furthermore, both modes of processing can be relatively open-minded, driven by accuracy concerns, or relatively biased, driven by defense or impression concerns. This chapter situates the heuristic-systematic model within its intellectual and personal history, and highlights key empirical findings that support the model's central tenets.

INTRODUCTION

Attitudes have been a primary focus of theory and research in social psychology since the 1920s. Nine decades of research have produced a sizeable and complex body of literature that speaks to questions of how people's attitudes are formed, maintained, and changed, and provides an ever-growing foundation upon which new questions arise and new answers unfold. In 1980, a foundation of process-oriented models that explained attitude change based on how people understand and evaluate persuasive argumentation set the stage for one question in particular: Was careful argument scrutiny the only kind of process by which attitude change could occur? Or might we sometimes change our minds in more efficient, but less effortful, ways?

The heuristic-systematic model of persuasion (Chaiken, 1980, 1987; Chaiken et al., 1989, 1996; Chen and Chaiken, 1999)

answers this question by proposing two distinct modes of information processing. The first mode, *systematic processing*, involves attempts to thoroughly understand any and all available information through careful attention, deep thinking, and intensive reasoning (e.g., thinking carefully about the arguments presented, the person arguing, and the causes of the person's behavior). This information is combined and used to guide subsequent attitudes, judgments, and behaviors. For instance, a systematic approach to thinking about a proposed economic policy might involve reading as many magazine and newspaper reports as possible to learn and develop an opinion about the "best" course of action for the economy. The heuristic-systematic model suggests that such systematic thinking entails a relatively high degree of mental effort, and thus requires that a person (1) *can* devote a certain amount of attention to thinking about the issue, and (2) *wants* to devote this attention. Thus, systematic processing is unlikely to occur unless a person is both *able* and *motivated* to do so.

Heuristic processing is much less demanding in terms of the mental work required and much less dependent on having the ability (e.g., enough knowledge and enough time) to think carefully about information. In fact, heuristic processing can be viewed as relatively automatic because it can occur even when people are not motivated and able to deliberately think about a topic. Heuristic processing involves focusing on easily noticed and easily understood cues, such as a communicator's credentials (e.g., expert versus nonexpert), the group membership of the communicator (e.g., Democrat or Republican), the number of arguments presented (many or few), or audience reactions (positive or negative). These cues are linked to well-learned, everyday decision rules known as *heuristics*. Like other knowledge structures (e.g., stereotypes), heuristics can vary in their availability and accessibility, as well as in their perceived reliability (i.e., the extent to which a particular person perceives

a heuristic to be a valid guide for judgment in a given situation; see Chen and Chaiken, 1999; Darke et al., 1998). Moreover, they can be used self-consciously or non-self-consciously: People may consciously decide to invoke a heuristic in order to inform a subsequent judgment, but heuristics can also influence judgments without intention or self-awareness.

Examples of heuristics include "experts know best," "my own group can be trusted," "argument length equals argument strength," and "consensus implies correctness." These simple, intuitive rules allow people to form judgments, attitudes, and intentions quickly and efficiently, simply on the basis of the easily noticed cues, and with little critical thinking. A heuristic approach to a proposed economic plan might involve simply adopting the opinion of a noted economist. In other words, heuristic thinking is what we do when we do not have much ability or time to think about something and want to make a reasonable decision as quickly as possible.

The theory further proposed that two principles act in conjunction to determine the mode and extent of information processing that occurs in any given context (Chaiken, 1980, 1987; Chaiken et al., 1989). The model's *least effort principle* reflects the assumption that individuals try to arrive at attitudinal decisions as efficiently as possible (see also Allport, 1954). Thus, all else equal, people should tend to prefer a less effortful mode of processing (i.e., heuristic processing) to one that requires more time and cognitive resources (i.e., systematic processing).

Meanwhile, however, the *sufficiency principle* asserts that individuals are sometimes motivated to exert additional cognitive effort in order to reach a certain level of judgmental confidence. They must therefore balance their preference for maximizing cognitive efficiency with the desire to satisfy their motivational concerns, such as the goal to reach an accurate conclusion (Chaiken et al., 1989; see also Simon, 1976). The heuristic-systematic model suggests that this balance point is determined by a *sufficiency*

threshold, defined as the degree of confidence to which an individual aspires in a given judgmental situation (Chaiken et al., 1989; Eagly and Chaiken, 1993). The sufficiency threshold can be conceptualized as a point located on a continuum of judgmental confidence. The extent of information processing is determined by the size of the discrepancy that exists between an individual's actual level of confidence in their judgment and the sufficiency threshold (i.e., their desired confidence). Thus, effortful information processing should only occur when

actual confidence falls below the sufficiency threshold, and should continue (when capacity allows) until this confidence gap is closed. Extent of information processing will therefore depend on both a particular person's actual level of judgmental confidence in a given persuasion setting, as well as their desired level of confidence in that setting (see Figure 12.1).

Together, the least effort and sufficiency principles suggest that—assuming adequate cognitive capacity—individuals will engage in systematic processing insofar as the less

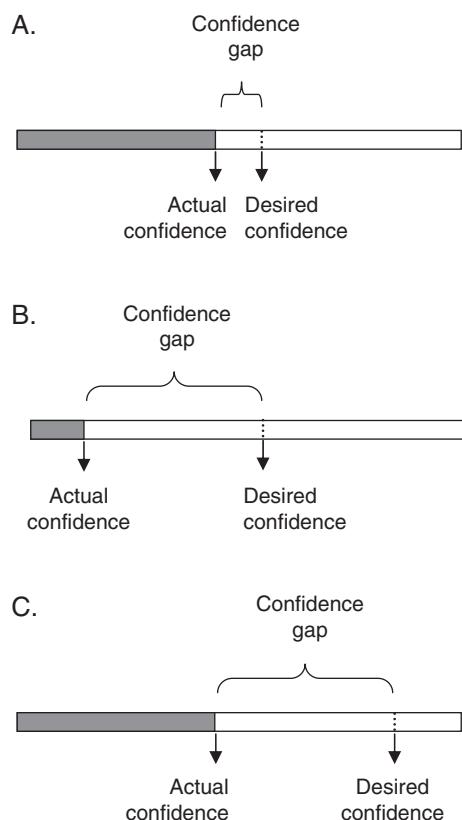


Figure 12.1 A person with a small gap between actual and desired confidence might be able to reach their desired level of confidence (the sufficiency threshold) using only heuristic processing (Panel A). If the confidence gap is larger, either due to a low level of actual confidence (Panel B) or a high level of desired confidence (Panel C), it is less likely that people can reach their desired level of judgmental confidence using only heuristic processing. When people cannot attain their desired level of confidence using only heuristic processing, they will engage in systematic processing in an effort to finish closing the confidence gap, assuming they have the ability to do so

effortful heuristic mode does not yield sufficient judgmental confidence (either because heuristic processing cannot occur, as in situations that do not offer easily processed heuristic cues, or because it is insufficient to close the confidence gap). Systematic processing will therefore be increased by factors that either decrease actual confidence, increase desired confidence, or both.

THE MULTIPLE-MOTIVE HEURISTIC-SYSTEMATIC MODEL

Although the heuristic-systematic model was initially developed to apply to individuals motivated by accuracy concerns to seek valid judgments, later work expanded the model to incorporate two other broad motivations that can lead to selective information processing geared toward arriving at a particular attitudinal position (Chaiken et al., 1989, 1996; Chen and Chaiken, 1999). The first of these, *defense motivation*, was intended to reflect the impact of such self-focused variables as ego-involvement and personal commitment (see, e.g., Kiesler, 1971; Sherif and Cantril, 1947). According to the multiple-motive model of heuristic-systematic processing, these factors arouse a desire to confirm and defend the validity of preferred attitudinal positions (like one's pre-existing opinion), while challenging the validity of nonpreferred positions. *Impression motivation*, on the other hand, reflects the impact of other-focused variables such as impression-relevant involvement, communication goals, and affiliative concerns (e.g., Higgins and McCann, 1984; Johnson and Eagly, 1989; Smith et al., 1956), which arouse a desire to express attitudes that are socially acceptable.

Like accuracy motivation, defense and impression motivations can engender heuristic and/or systematic processing. However, unlike accuracy motivation, these directional motives tend to lead people to process information selectively, rather than open-mindedly. The biases engendered by these

directional motives largely occur outside of awareness; people usually operate under the illusion that they are thinking in an open-minded fashion. In the case of defense-motivated processing, for instance, individuals may selectively choose heuristics that help to confirm a preferred position. A defense-motivated person might therefore invoke the heuristic "experts know best" if the position of an expert source reinforces her cherished values and social identity, but might choose a different heuristic (e.g., "outgroup sources can't be trusted") if the position threatens her social identity. Likewise, impression-motivated heuristic processing entails selective application of heuristics that ensure a smooth interaction with specific others. For example, when interacting with a person or group whose views on an issue are unknown or vague, a perceiver might invoke the heuristic "moderate judgment minimizes disagreement." On the other hand, when others' views are known, a "go along to get along" heuristic might better serve the same goal.

With sufficient cognitive capacity and higher levels of motivation, defense- or impression-motivated people will also process systematically, but they will again do so selectively. Thus, a defense-motivated perceiver will attend to, elaborate on, and recall information that serves to bolster his preferred, self-protective position, while an impression-motivated perceiver will systematically process information in such a way as to convey a desired impression to (real or imagined) others.

The multiple-motive model thus proposed three general categories of motives that give rise to three distinct processing goals, any of which can engender heuristic and/or systematic processing. Expanding the theory in this way broadened its applicability to a much wider range of persuasion and social influence situations. In essence, the multiple-motive heuristic-systematic model allowed a rapidly increasing laundry list of persuasion-relevant variables to be understood in terms of their effects on a few key factors—processing goal, cognitive capacity, actual

confidence, and desired confidence—which could in turn suggest a reliable prediction about the extent of attitude change that should occur in a given setting. The strength of this basic dual-process model to organize and generate predictions in the persuasion literature led to its application across a wide range of settings (Chen and Chaiken, 1999; Ledgerwood et al., in press; Mackie, 1987; Stroebe and Diehl, 1988; see Eagly and Chaiken, 1993; Ledgerwood et al., 2006, for reviews). More broadly, it was one of several theories that helped to precipitate a flowering of dual-process models across multiple areas of social psychology, as researchers began to see similar basic principles at work in a number of different domains including stereotyping, impression formation, and decision making (see Chaiken and Trope, 1999).

PERSONAL HISTORY OF THE THEORY: AN AUTHOR BY LITERATURE INTERACTION

In 1972, the first author entered graduate school at the University of Massachusetts Amherst armed with a math major, a psychology minor, and a vague interest in social influence. I began working with Alice Eagly, who was at the time examining the importance of message comprehensibility within the context of Bill McGuire's information processing paradigm. In fact, my masters research project involved testing an idea about comprehensibility that McGuire had tucked away in the depths of a handbook chapter; namely, that the importance of comprehensibility in determining the effectiveness of persuasive appeals would depend on the modality of the communication (Chaiken and Eagly, 1976). Looking back, I can trace part of the development of the heuristic-systematic model to this project. The idea was that message comprehensibility should matter more when the message is in written form rather than audio or video, partly because there is more flexibility to carefully

scrutinize a message when reading it than when hearing it spoken. Later, we returned to this idea to examine whether a different type of persuasion variable—source cues—might become increasingly influential as one moved from written to audio and visual modalities (Chaiken and Eagly, 1983). This research suggested that different persuasion variables might be more or less influential depending on how a message was presented.

Here then was one seed for the heuristic-systematic model: different types of persuasion variables had more or less impact depending on a recipient's ability to carefully scrutinize a message. Other seeds were in the recent and current literature at that time: articles on correspondent inference theory, Kelley's covariation theory, and self-perception theory populated the reading lists for my coursework, and I was intrigued both by attribution models and by the simplicity of self-perception (Bem, 1972; Jones and Davis, 1965; Kelley, 1972, 1973). With Alice, I helped develop an attribution model of source characteristics based on Kelley's (1973) cube model, particularly his notions of discounting and augmentation. Yet compared to self-perception theory, analyzing the covariances of even a somewhat superficial persuasion variable like source characteristics seemed effortful and deliberative. Could people really be so careful, so thoughtful, all the time?

The simple if-then's of self-perception theory appealed to me—why engage in some arduous analysis of your own thoughts and behaviors when you could simply reason: if I'm yelling, I must be angry? Years before, when Kennedy and Nixon had been running for president, I remember listening to my parents consider the intricacies of the various political issues at stake. Meanwhile I (with a young child's preference for the simple that I still haven't seemed to grow out of completely) *knew* that Kennedy was the man to vote for; after all, he looked better. And it wasn't just me; in graduate school, I read about data showing that although those who heard the first Kennedy–Nixon debate on the

radio believed Nixon had won, those who watched it on television were convinced that Kennedy had in fact prevailed. If the average person was really what McGuire (1969) had dubbed “the lazy organism,” might a simple if-then suffice for most of us, much of the time?

Then, in 1975, I came across Shelley Taylor’s recently published dissertation, which shed some additional light on self-perception processes. Female participants were shown pictures of three different men who varied in attractiveness. Some participants were given false physiological feedback suggesting that they showed a strong preference for one of the men pictured. Participants were also led to believe that they would have the opportunity to meet one of these men in a few weeks (high consequences condition), or were not led to expect a future meeting (no consequences). They then rated each of the three men on attractiveness. The results suggested that participants in the no consequences condition engaged in self-perception: they based their ratings of attractiveness on the physiological feedback provided. Most interestingly to me, however, participants in the high consequences condition were not affected by the feedback manipulation. Instead, there was some evidence to suggest they were thinking more carefully and critically about the three candidates: they spent more time looking at the pictures, and content analyses suggested that they made more critical comments.

I remember thinking to myself that surely this could apply to persuasion. High and low consequences provided a variable that could perhaps predict when a lazy organism would opt for a simple “if-then” versus a more complicated analysis of available information. I built my main dissertation experiment around this idea, testing whether high versus low consequences would moderate the persuasive impact of source cues (the most frequently studied noncontent variable at the time) and content (extent of strong persuasive argumentation). I reasoned that source cues such as likeability can be processed

quite easily and efficiently by a lazy organism unmotivated by future consequences. When future consequences were present, however, participants should be motivated to process information more carefully, and extent of strong argumentation should play a greater role in persuasion.

And (amazingly, to me at the time), the study worked. I started calling the more deliberative mode of thinking systematic, but was unsure what to call the other one until Icek Ajzen, another important mentor for me in graduate school, suggested the name “heuristic.” As I continued the line of research, the notion of consequences became abstracted into motivation to process information. Like many other theories at that time, the default motivation was implicitly assumed to be accuracy; I began to explicitly label the motivation “accuracy motivation” only later in order to emphasize that both modes of thinking served the same motive (rather than one being rational and the other irrational). Drawing on my earlier modality research, I also added capacity as a second variable that seemed necessary for deliberative processing to occur.

Over the years, I tried to expand the model to other kinds of cues, and to test its assumptions in various ways. Perhaps inevitably, given that my intellectual genes were steeped in classic functional theories of attitudes (Alice Eagly had worked with Herb Kelman), it occurred to me that accuracy was not the only motive in town, and I began to try to group the major attitudinal functions I had learned about in graduate school into a few broad categories of motives. Over time, we developed and tested predictions deriving from a multiple-motive heuristic-systematic model that included not just accuracy motivation, but also impression and defense motives (see Chaiken et al., 1996; Chen and Chaiken, 1999, for reviews).

Very gradually, then, the theory expanded—first under the influence of functional theories, and then following new developments in social cognition. I had always thought of heuristics as simple decision rules

that were likely represented in memory, and such a conceptualization lent itself to new theorizing about availability and accessibility in social cognition. By now I was at New York University, where hallway conversations with Tory Higgins and John Bargh inevitably turned toward basic principles of social cognition. I began to think that heuristics ought to vary in their availability, accessibility, and reliability, and that this would have important consequences for when a given heuristic would be applied. Furthermore, heuristics seemed to me to be relatively automatic, in at least some senses of the term (see Bargh, 1994). I always thought of them as a kind of shortcut; thus, at the very least they were automatic in the sense of being efficient. It also seemed likely that they often (but not always) operated outside of awareness. Over time, the results of accumulating studies provided support for this social-cognitive side of the model as well (see Chen and Chaiken, 1999, for a review).

Conceptualizing heuristics as a form of automatic social cognition highlights one way in which the basic processes underlying the heuristic-systematic model extend beyond the persuasion context to other domains. It became apparent early on that a dual-process perspective was not restricted to a persuasion context; that it would be fruitful to look across different domains to understand the common mechanisms at work in all of them. And indeed, the heuristic-systematic model was just one of a growing family of dual-process models that began to populate social psychology in the 1980s and 1990s, as researchers across different domains converged on a similar set of mechanisms to explain information processing in a variety of settings (see Chaiken and Trope, 1999).

INTELLECTUAL HISTORY OF THE THEORY

Like many of the models in this family, the heuristic-systematic model suggests that

individuals can think about information not only via a bottom-up, data-driven process but also via a more top-down process that depends on the pre-existing knowledge structures they bring to a particular context. Although this represented a radical reorientation in the field of persuasion at the time, the notion that we can rely on learned associations to structure understanding emerged as early as 1930 in Kohler's discussion of sensation and perception, in which he suggested that our perceptions are shaped as much by a top-down application of knowledge derived from past experiences as by bottom-up, sensory experience (see also Moskowitz et al., 1999; Yates, 1985). For instance, upon sensing a pattern of colors and lines with our eyes, we can draw on our past experiences and associations with this pattern to label it a "chair" and infer its form and function. Subsequently, Bruner's "new look" emphasized the notion that our perceptions are substantially shaped by expectation and motivation (Bruner, 1957). Research on mental schemas developed this idea to suggest that we can quickly organize and "fill in the blanks" about our world using generalized mental structures built from our past experiences (e.g., Anderson and Pichert, 1978; Brewer and Treyens, 1981; see Fiske and Linville, 1980; Taylor and Crocker, 1981; and Fiske and Taylor, 2008, for reviews). Together, these literatures highlight a relatively quick, efficient, top-down method of understanding the world that capitalizes on past experience to structure current understanding, and suggests that these mental shortcuts may be applied to a range of different domains (see also Tversky and Kahneman, 1974).

Within this historical context, the heuristic-systematic model proposed that individuals might sometimes rely on quick, efficient, cognitive shortcuts to make judgments about the validity of information they encounter. Thus, rather than carefully scrutinizing any and all available information, people might instead draw on simple if-then associations learned through repeated experience to

inform their attitude judgments. For instance, given that experts tend to be correct, individuals might develop a learned association between experts and correctness that allows them to easily and efficiently infer that a subsequently encountered expert is likely to be right (“if expert, then correct”).

As noted earlier, because the model assumed that heuristics are like other knowledge structures, it invited connections to social-cognitive research on the principles governing the activation and use of stored knowledge (Chaiken et al., 1989, Chen and Chaiken, 1999). In other words, heuristics should be subject to the same principles of availability, accessibility, and applicability that underlie the use of stored knowledge in other domains (e.g., Higgins, 1989; Higgins et al., 1982). Considerable research supports this claim (see Chen and Chaiken, 1999, for a review). For instance, in order to be used to inform attitudes in a given setting, a heuristic must be (1) accessible (e.g., because it has been situationally primed), and (2) applicable (e.g., because an individual believes it to be a reliable, or usable, guide for judgment; Chaiken et al., 1992).

Heuristic processing thus represented a very different mode of thinking from the more systematic, comprehensive mode that had occupied the center stage of persuasion theory and research for some time. Furthermore, the heuristic-systematic model suggested that these modes of processing involved a tradeoff between optimal judgments (maximized by systematic processing) and efficient judgments (maximized by heuristic processing). The model’s original formulation proposed that heuristic or systematic processing would predominate depending on the relative importance of accuracy or economic concerns for a given person in a given context (Chaiken, 1980). Subsequently, this notion was refined to emphasize a continuum of judgmental confidence, along which two critical points can be located: a person’s actual confidence, and their desired confidence or *sufficiency threshold* (Eagly and Chaiken, 1993). As lazy

organisms (McGuire, 1969), people first attempt to close this gap in confidence via heuristic processing. Only when this easier strategy fails to confer sufficient judgmental confidence will people exert the cognitive effort required by systematic processing, assuming they are able to do so.

Considerable research supports this central claim that individuals will process information heuristically unless they are both motivated and able to engage in more effortful systematic processing. Heuristic cues alone tend to guide judgments when ability is low (such as when participants possess little knowledge about the topic, when they are under time pressure, or when situational constraints diminish cognitive capacity) and when motivation is low (such as under conditions of low task importance or personal relevance; Giner-Sorolla et al., 2002; Petty et al., 1976; Ratneshwar and Chaiken, 1991; Wood et al., 1985). As ability and motivation increase, systematic processing plays an increasing role in influencing attitudes (e.g., Chaiken, 1980; Martin et al., 2007; Petty and Cacioppo, 1984; see Eagly and Chaiken, 1993, for a review). Importantly, the processing modes are by no means mutually exclusive: given adequate levels of ability and motivation, heuristic and systematic processing often co-occur (Chaiken, 1980, 1987; Eagly and Chaiken, 1993). We return to this assumption of *concurrent processing* later in the section.

Bridging beyond the persuasion context

Although the heuristic-systematic model was initially developed within the context of the paradigmatic persuasion experiment, in which a source conveys a message to a target with some effect, it quickly became clear that the fundamental processes at work within this context were mirrored in other domains. At its heart, the persuasion paradigm involves individuals making judgments in light of information, as they do in many

other domains. Because it focuses on the basic processes underlying persuasion effects, the heuristic-systematic model provided a natural bridge from persuasion to many other, conceptually similar, areas. Across various domains, individuals can make judgments based on quick shortcuts or more effortful, extensive processing, and motivation and ability play a key role in guiding the extent to which effortful processing occurs.

Indeed, as noted earlier, the heuristic-systematic model was among several early dual-process models in social psychology. Together, these paved the way for a rapid proliferation of information-processing theories in a variety of domains that distinguished between a relatively automatic, fast, reflexive mode of thinking based on well-learned associations, and a more controlled, analytic, effortful mode based on systematic reasoning (e.g., Brewer, 1988; Devine, 1989; Fazio and Towles-Schwen, 1999; Fiske et al., 1999; Gawronski and Bodenhausen, 2006; Gilbert, 1989; see Chaiken and Trope, 1999; Smith and DeCoster, 2000, for reviews). In their 1999 volume, Chaiken and Trope brought together a variety of dual-process models from diverse fields that converged in their basic distinction between these two types of processes, illustrating that these perspectives are really a family of theories with a common core.

EMPIRICAL FINDINGS

The first experiment designed to test the heuristic-systematic model examined whether involvement would moderate the extent to which a heuristic cue (communicator likeability) versus message content (extent of supportive argumentation) affected people's attitudes (Chaiken, 1980: Study 1). Undergraduate participants read a transcript of an interview with a university administrator who in the course of the interview either praised undergraduates (likeable source condition) or disparaged them (unlikeable

source condition). Later, the administrator stated his opinion on an issue (e.g., changing from a semester to a trimester system) and provided either a weak message (containing only two arguments) or a strong message (containing six different arguments) in support of his opinion.

To test whether participants' level of motivation would determine the extent to which they relied on the heuristic cue or engaged in more effortful processing of message content, the experiment also manipulated participants' involvement by leading them to expect that they would discuss either the same issue or a different issue at a subsequent experimental session. Participants who expected to discuss the same issue should be more motivated to reach an accurate conclusion about whether the administrator's position was valid, compared to those who expected to discuss a different issue, and should therefore engage in more systematic processing. Consistent with the study's hypotheses, high involvement participants showed greater attitude change in response to a strong (versus weak) message, but were unaffected by communicator likeability. In contrast, low involvement participants showed greater attitude change in response to the likeable (versus unlikeable) communicator, but were unaffected by message content. Furthermore, substantiating the notion that attitude change was mediated via systematic processing in the high involvement condition, these participants showed greater recall of arguments and reported more issue-relevant thoughts, compared with those low in involvement. Thus, which factors produced persuasion—and *how* they produced persuasion—depended critically on participants' level of motivation.

Importantly, by delineating the dual processes underlying people's thinking about persuasive appeals, the heuristic-systematic model was able to shed light on the role played by motivational variables, as well as source cues and message content, in influencing attitudes. For instance, previous research had reported conflicting findings

regarding the impact of involvement on persuasion (e.g., Pallak et al., 1972; Sherif and Hovland, 1961). Our results (Chaiken, 1980) suggested that involvement could either increase or decrease attitude change in response to a persuasive message, depending on the valence of available heuristic cues and the strength of the message content. Similarly, Axsom et al. (1987) showed that whereas involvement increased the impact of argument quality on persuasion, it decreased the impact of the heuristic cue of audience response (i.e., whether an overheard message audience sounded enthusiastic or unenthusiastic). The heuristic-systematic model thus provided a theoretical framework within which to organize a large number of persuasion-related factors in a literature that had often produced contradictory results.

The concurrent processing assumption

It was in large part the prevalence of such contradictory results that motivated the development of the heuristic-systematic model. Looking back, the historical assumptions discussed earlier, combined with the current climate in the persuasion literature, created a unique context within which the logic of a dual-process perspective was perhaps more likely to be discovered. And, in fact, two dual-process models of persuasion independently emerged from this context: the heuristic-systematic model and the elaboration-likelihood model (ELM; Petty and Wegener, 1999). Both provided an organizing framework for understanding the impact of various persuasion variables by suggesting two routes to persuasion: the heuristic or “peripheral” route, and the systematic or “central” route. However, they differed in some important ways. For instance, whereas the ELM assumed that the peripheral and central routes to persuasion were mutually exclusive, the heuristic-systematic model suggested that they could co-occur and even interact.

Thus, although many of the initial dual-process studies of persuasion suggested that heuristic cues do not impact attitudes when people are motivated and able to process systematically (e.g., Axsom et al., 1987; Chaiken, 1980; Petty et al., 1981; Wood et al., 1985), the heuristic-systematic model suggested that this pattern was only one possible outcome of the two modes of information processing. Specifically, these results seemed to represent cases in which systematic processing *attenuated* the judgmental impact of heuristic processing because it took into account information that contradicted the valence of the available heuristic cues. If systematic processing instead yielded information that was congruent with heuristic processing, the heuristic-systematic model suggested an *additivity hypothesis* whereby heuristic processing could exert a direct effect on judgment over and above the impact of systematic processing. Supporting this hypothesis, Maheswaran and Chaiken (1991; see also Maheswaran et al., 1992) found that when heuristic cues and message content were congruent, attitude change was mediated by both heuristic and systematic processing.

Importantly, however, the heuristic-systematic model proposed that the two processes could not only co-occur, but could also interact to exert interdependent effects on judgment. Specifically, heuristic processing could *bias* systematic processing by influencing people’s expectations about the validity of arguments presented in a persuasive appeal (Chaiken et al., 1989). To test this notion, Chaiken and Maheswaran (1994) presented participants with a novel attitude object (a new telephone answering machine called the “XT-100”) and assigned them to one cell of a 2 (accuracy motivation: low versus high) by 2 (heuristic cue: valid versus invalid) by 3 (argument quality: strong versus ambiguous versus weak) design. This study manipulated accuracy motivation by varying the importance and personal relevance of participants’ decisions regarding this new product. Whereas participants in the high

importance condition learned that they were part of a small list of respondents, that their input would be heavily weighted, and that the product would be distributed in their geographical area, participants in the low importance condition learned that they were part of a large group of respondents, that individual opinions were unimportant, and that the product would be distributed in a different geographical area.

Participants next received a positive message about the product that contained a heuristic cue conveying either high or low validity. Specifically, they learned that the product description in the message was taken from *Consumer Reports*, a credible source, or from a promotional Kmart pamphlet, a noncredible source. The product description contained either strong arguments, weak arguments, or an ambiguous mixture of the two. Participants then reported their attitudes toward the XT-100 and listed their thoughts about the product description.

As in previous studies (e.g., Chaiken, 1980), the relatively unmotivated participants in the low importance condition expressed attitudes that reflected the source credibility cues, but not the quality of the arguments presented in the product description. Thus, participants were more favorable toward the XT-100 when they had read a positive message from a credible (versus noncredible) source, regardless of actual message content. Moreover, this effect of source cue on attitudes was direct, rather than mediated by cognitive elaboration, consistent with the notion that participants were directly inferring the validity of the message from the source's credibility (i.e., processing heuristically by using a well-learned association between credibility and correctness).

Meanwhile, the results for participants in the high importance condition who read an unambiguous message also replicated past research: highly motivated participants who read a strong (versus weak) persuasive message expressed more positive attitudes toward the XT-100, and this effect was mediated by participants' cognitive elaborations about the

product. Additional analyses revealed that when source cue and message content were contradictory in their implications for message validity (i.e., a credible source paired with weak arguments, or a noncredible source and strong arguments), systematic processing alone determined attitudes. This is consistent with the *attenuation hypothesis* suggesting that systematic processing can override the effects of heuristic processing. However, when source credibility and message content were congruent (i.e., a credible source and strong arguments, or a noncredible source and weak arguments), there was both a direct effect of the heuristic source cue on attitudes and an effect of message content mediated by systematic processing. Thus, when the information provided by heuristic and systematic processing were congruent, the results supported the *additivity hypothesis* suggesting that both modes of processing can independently influence attitudes.

Finally, highly motivated participants who read an ambiguous message were influenced both by the source cue and by systematic processing of the high (versus low) quality arguments. Supporting the *bias hypothesis*, these participants' cognitive elaborations about the attitude object were influenced by the validity information provided by the source cue, such that the high credibility source biased systematic processing in a positive direction, whereas the low credibility source biased systematic processing in a negative direction. In addition, attitudes in this condition were also directly influenced by the heuristic cue.

In other research examining the bias hypothesis, Darke et al. (1998) studied the impact of consensus information presented in the absence of persuasive argumentation on college students' support for comprehensive exams. Accuracy motivation was manipulated via personal relevance. Participants in the high relevance condition were led to believe that the exam policy would have direct personal consequences (i.e., it would take effect the following academic year, and

would thus apply to current students), whereas those in the low relevance condition were led to believe that there would be no personal consequences (i.e., the policy would take effect in ten years, and therefore have no impact on current students). Participants then learned that 80 percent of students either supported or opposed instituting comprehensive exams, based on either a small poll (a sample size of ten students) or a large poll (a sample size of 1,000 students). Consistent with the bias hypothesis, participants in the high personal relevance condition generated thoughts that were biased in the direction of the available consensus cue, and these thoughts then influenced their attitudes. In contrast, the consensus information exerted a direct, heuristic influence on participants' attitudes in the low personal relevance condition. Interestingly, highly motivated participants also discriminated between the more and less reliable heuristic cues: participants in the high relevance condition were more persuaded by the consensus information when the poll was based on a large versus small sample of students, whereas participants in the low relevance condition were persuaded by consensus information regardless of the poll's size.

Together, then, these studies highlight the complex interplay between heuristic and systematic processing (see also Chen et al., 1996; Erb et al., 1998; Ziegler et al., 2005). Importantly, they demonstrate that the two modes of processing can influence attitudes both independently and interactively, suggesting that they may best be conceptualized as two interdependent and potentially co-occurring ways of thinking (see Eagly and Chaiken, 1993: Chapter 7, for further discussion).

Multiple motives

Another unique feature of the heuristic-systematic model is that it jointly considers the influence of multiple modes of processing on the one hand and multiple motives on

the other. The tripartite analysis of motives in the heuristic-systematic model has its historical roots in the literature on attitude function, although it should be noted that similar classes of motives that center on understanding, protecting the self, and affiliating with others are echoed across multiple domains (e.g., Allport, 1954; Baumeister and Leary, 1995; Deutsch and Gerard, 1955; Fiske, 2002; Heider, 1958; Lerner and Tetlock, 1999; Tesser and Campbell, 1983). The notion that individuals are often motivated to form and hold attitudes that square with relevant facts built on Katz's (1960) knowledge function and Smith et al.'s (1956) object appraisal function of attitudes, which emphasized the role often played by attitudes in organizing experience and guiding action with respect to an individual's ongoing concerns. The heuristic-systematic model was thus initially designed to apply to persuasion contexts in which the message recipient is concerned with assessing the validity of a persuasive appeal (Chaiken, 1980, 1987; Chaiken et al., 1996). We subsequently extended the model beyond validity-seeking persuasion contexts, adding impression and defense motives to encapsulate two other broad classes of attitude functions in the literature (Chaiken et al., 1989). The concept of impression motivation was designed to capture other-oriented, affiliative functions such as Smith et al.'s (1956) social adjustment function, which emphasized the role that attitudes can play in helping people establish and maintain relationships with other individuals or groups (see also McGuire, 1969). Meanwhile, defense motivation encapsulated self-oriented defensive functions such as Katz's (1960) ego-defensive function and Smith et al.'s (1956) externalization function, which suggested that some attitudes serve to protect individuals' self-image against internal or external threats.

Considerable evidence supports the notion that impression motivation can guide heuristic and systematic processing (see Chaiken et al., 1996, for a review). For example, Chen

et al. (1996: Study 2) led participants to anticipate a discussion about a social issue with a partner who ostensibly held either a favorable or an unfavorable opinion on the issue. Before this discussion, participants read a series of fictitious scenarios designed to prime either the accuracy goal of determining a valid opinion, or the impression goal of getting along with other people. After this task, participants familiarized themselves with the discussion issue by reading an evaluatively balanced essay concerning the issue (in this case, whether election returns should be broadcast while polls are still open). Participants then listed the thoughts that had occurred to them as they read the essay and indicated their own attitudes toward the issue.

Impression-motivated participants expressed attitudes that were much more congruent with their alleged partners' attitudes than did accuracy-motivated participants: when the partner favored one side of the issue, they favored the same side, whereas when the partner opposed it, they opposed it. Interestingly, accuracy-motivated and impression-motivated participants exhibited the same amount of systematic processing (as measured by the number of issue-relevant thoughts that were listed). However, whereas accuracy-motivated participants' systematic processing was open-minded and unbiased by their partners' attitudes, impression-motivated participants exhibited systematic processing that was biased toward their partners' attitudes. For example, when the partner favored allowing broadcasts of election returns while the polls were still open, impression-motivated participants listed thoughts that revealed much more favorable thinking about arguments supporting the broadcasting of returns and more unfavorable thinking about arguments opposing it.

Like impression motivation, defense motivation can also guide heuristic and systematic processing in a directional fashion, as individuals attempt to close the gap between actual and desired confidence that a judgment will protect their cherished beliefs

and self-views (e.g., Ditto and Lopez, 1992; Giner-Sorolla and Chaiken, 1997; Liberman and Chaiken, 1992; Lord et al., 1979). For instance, Giner-Sorolla and Chaiken (1997) found that participants' vested interest in a campus issue biased their judgments of a consensus cue's reliability, when additional information that would permit systematic processing was unavailable. Specifically, participants rated the consensus information (an opinion poll of their fellow students) as more reliable, and criticized it less, when the poll results supported rather than opposed their vested interests. When additional information was available, participants also displayed a defensive bias in their systematic processing, cognitively elaborating the arguments presented in a selective manner that reflected their vested interests. Interestingly, when both types of information were available, exposure to a hostile consensus cue appeared to undermine judgmental confidence and increase systematic processing of the arguments presented: In these conditions, the influence of vested interests on participants' subsequent attitudes was mediated by their cognitive elaborations about the issue. In contrast, exposure to a congenial cue appeared to close the confidence gap, such that participants simply used their vested interests to directly inform their subsequent attitudes, rather than engaging in additional heuristic or systematic processing. Thus, as with accuracy and impression motives, both heuristic and systematic processing can be used to serve self-protective processing goals.

IMPLICATIONS FOR SOCIAL ISSUES

Because it focuses on the basic mechanisms by which persuasion can occur, the heuristic-systematic model can predict how a wide range of variables will influence attitudes and judgments in various situations. It is therefore a particularly powerful tool for understanding and influencing information

processing in ways that can help effect positive social change, and has been applied to diverse issues such as increasing individuals' acceptance of potentially threatening health information, improving the design of product warning labels, identifying and decreasing bias in jury decision-making, increasing recycling behavior, and developing more effective programs for preventing substance abuse among teens (e.g., Brewer and Hupfeld, 2004; ForsterLee et al., 2006; Harris and Napper, 2005; Howard et al., 2006; Jepson and Chaiken, 1990; Liberman and Chaiken, 1992; Scott, 1996; Werner et al., 2002; Zuckerman and Chaiken, 1998). Here, we discuss the implications of the heuristic-systematic model for two areas that we find particularly interesting: negotiation and political decision-making.

Negotiation and conflict resolution

Research exploring heuristic and systematic processing in simulated negotiations has confirmed the utility of a dual-process perspective for understanding information processing in conflict settings (see Ledgerwood et al., 2006, for a review). Specifically, when negotiators have modest levels of motivation (or low cognitive capacity), they often rely on heuristics such as fixed-pie assumptions (the perception that a negotiation is a zero-sum game), initial anchor values (e.g., first offers, or information about the typical outcome of similar negotiations), and stereotypes about an opponent's group membership (De Dreu et al., 1999; Thompson and Hastie, 1990; see De Dreu, 2004, for a review). In contrast, when motivation and capacity are relatively high, sole reliance on these heuristics tends to decrease as systematic processing increases.

Researchers have identified several factors that influence the extent to which people process information in negotiations (see De Dreu, 2004). These factors include both

stable individual differences and temporary elements of a given situation that influence motivation and/or capacity. For instance, negotiators who are high in the dispositional need for cognitive closure—that is, the desire to reach a judgment quickly and avoid ambiguity (Webster and Kruglanski, 1994)—are more likely to rely solely on heuristics than are those who have a low need for closure (De Dreu et al., 1999).

Temporary, situation-specific factors such as the presence of a highly involving task or process accountability (the need to justify the way in which a decision is made) tend to increase the extent of systematic processing, whereas time pressure and capacity-degrading conditions (e.g., noise) tend to decrease such processing (e.g., De Dreu, 2003; Tetlock et al., 1989; see Ledgerwood et al., 2006, for a review). For example, De Dreu (2003) examined the effect of time pressure on fixed-pie perceptions. Business students were placed into pairs and asked to play the role of a buyer or seller in a negotiation over the purchase of a car. The negotiation task was designed to hold integrative potential: the different issues varied in importance to the two negotiators, so that an integrative solution that capitalized on this variation in priorities would be more beneficial to both negotiators than a 50:50 split based on a fixed-pie assumption. Participants were led to believe that they had either plenty of time to complete the negotiation (low time pressure condition), or relatively little time (high time pressure condition). Participants were more likely to revise their fixed-pie assumptions, which led to higher joint outcomes, under low rather than high time pressure. These results suggest that time pressure reduces systematic processing, heightening reliance on heuristic cues such as fixed-pie perceptions and preventing negotiators from capitalizing on integrative potential.

In contrast, when an individual expects to discuss an issue with, justify a decision to, or be evaluated by an unknown audience, he or she tends to engage in pre-emptive

self-criticism, displaying a heightened motivation to arrive at an accurate conclusion (Lerner and Tetlock, 1999; Tetlock et al., 1989). In terms of the heuristic-systematic model, holding a person accountable to an audience whose views are unknown can increase desired confidence for a correct judgment and thereby stimulate accuracy-motivated systematic processing. Confirming this idea, De Dreu et al. (2000) randomly assigned business student participants to high-accountability and low-accountability conditions before asking them to engage in a mock negotiation over the purchase of a car. In the high-accountability condition, participants expected that their negotiation strategies and decisions would be reviewed and evaluated several days later by an experienced negotiator and a psychologist. In the low-accountability condition, participants did not receive this information. The results showed that under high accountability, participants were more likely to revise their fixed-pie assumptions and tended to obtain higher joint outcomes. Together, these studies suggest that negotiation outcomes can be improved by reducing the impact of variables that decrease accuracy motivation and capacity (like time pressure), as well as by facilitating factors that increase accuracy motivation (like accountability to an impartial expert).

Political attitudes

The heuristic-systematic model can also be used to shed light on political decision-making and voting behavior (e.g., Forehand et al., 2004; Marcus et al., 2000; Mondak, 1993; Newman and Perloff, 2004), and suggests that the impact of various factors on political judgments and intentions will depend on a voter's ability and motivation to think about available information. When people are motivated and able to process political information, they will tend to weigh the quality of the arguments put forth regarding an issue or candidate. In contrast,

when people are low in motivation to process information about political issues or candidates (e.g., involvement and personal relevance are low), or when they lack the ability to process systematically (e.g., they are stressed or under time pressure), they may tend to rely on heuristics such as party labels, expert or celebrity endorsements, and source cues such as attractiveness or group membership. For example, a low-motivation or low-capacity voter might oppose a state ballot initiative because Oprah opposes it, support a senator because the letter ("D" or "R") next to the name matches the voter's typical political preferences, or vote for a presidential candidate because their facial features convey an air of competence (see Hall et al., 2009; Todorov et al., 2005).

Political psychologists have identified five broad categories of heuristics that can influence voting behavior: party affiliation, ideological affiliation, endorsements, polls (i.e., consensus information), and candidate appearance (Lau and Redlawsk, 2001). Although in an ideal world, citizens participating in a democratic process would usually think carefully and critically about political information before arriving at conclusion, heuristic processing is thought to guide a substantial portion of political decision-making. For instance, echoing Converse's (1964) observation that the majority of Americans display relatively low levels of political sophistication and knowledge, Mondak (1993) suggested that most voters face a range of pressing everyday concerns that tend to take precedence over political matters, increasing the likelihood that voters will rely on heuristics when processing political information (see also Ledgerwood and Chaiken, 2007). Consistent with this notion, Lau and Redlawsk (2001) found a high rate of heuristic use among individuals participating in a mock presidential election. Using a process-tracing methodology, these researchers were able to track the extent to which participants accessed different kinds of information about the candidates on a computer: They provided participants with a

list of available types of information (e.g., “Issue Stance,” “Past Experience,” “Endorsements”), each of which could be opened with a mouse click to display the relevant information, and then recorded which kinds of information participants chose to access. Information from each of Lau and Redlawsk’s five political heuristic categories was accessed by over 90 percent of participants. Interestingly, different participants also appeared to prefer different types of heuristics: those higher in political expertise were more likely to use ideology and endorsement heuristics, whereas those lower in expertise were more likely to use candidate appearance heuristics. As the heuristic-systematic model would predict, participants were more likely to use heuristics when their ability to engage in more effortful processing was limited (i.e., when the information environment was made more complex by having the information labels actively scroll past participants on the computer screen rather than remain static).

Lau and Redlawsk’s (2001) study suggests that all five categories of heuristics are likely to play a role in a given election; however, some types have been studied more frequently than others. For example, given the prolific use of endorsements for a wide variety of political attitude objects (including everything from local ballot initiatives to presidential candidates), and from a wide variety of endorsers (ranging from political organizations to celebrities), political scientists have been particularly interested in how endorsement heuristics influence political opinions and voting behavior. Using data from a California poll regarding an upcoming election for members of the State Supreme Court, Mondak (1993) showed that endorsements increased voters’ willingness to express an opinion and influenced the direction of that opinion when they had relatively little information about the issue. Specifically, respondents were more likely to say a Supreme Court justice should be retained or recalled (rather than choosing “not sure”) when told which governor had appointed the justice, and they used their evaluation of the

governor to guide their evaluation of the justice in question. In other words, they used the governor’s endorsement as a heuristic in forming an attitude toward the associated Supreme Court justice. Consistent with the heuristic-systematic model, this was more likely to occur when respondents had been previously exposed to relatively little media information regarding the justice (thereby limiting their ability to engage in systematic processing) and for respondents scoring higher on a need for cognitive efficiency measure (designed to tap both motivation and ability to carefully process information).

In Mondak’s (1993) study, the heuristic implication of an endorsement from a politician depended on a voter’s attitude toward that politician. However, the impact of an endorsement could also depend on the perceived *reliability* of the heuristic for a particular judgmental task; that is, the extent to which a perceiver deems a heuristic to be a valid guide for judgment in a given situation (see Chen and Chaiken, 1999). For example, when considering an environmental issue, a voter might feel that an endorsement from Greenpeace affords a sizeable increase in judgmental confidence, whereas an endorsement from the National Basketball Association does not, despite equivalent evaluations of the two organizations. Indeed, Forehand et al. (2004) found that participants expressed more favorable attitudes toward a hypothetical initiative when it was endorsed by a well-known and issue-relevant source rather than a fictional or issue-irrelevant source. Supporting the heuristic-systematic model’s sufficiency principle, this difference emerged in a low motivation context (in which participants expected to justify their preferences about an unimportant and unrelated issue, ballot formatting) but not a high motivation context (in which participants expected to be held accountable for their position on the initiative itself).

Group endorsements can also act to bias systematic information processing about an issue or a candidate. Individuals may be motivated by defense or impression concerns

to agree with an ingroup and disagree with an outgroup, and may therefore process information selectively to arrive at these preferred judgments (Fleming and Petty, 2000). For example, Cohen (2003: Study 4) asked liberal undergraduate students to evaluate a (stereotypically liberal) proposal for a generous federally funded job-training program. Half the participants learned that Democrats (their political ingroup) opposed and Republicans (their political outgroup) supported the program, while half received no information about group endorsement. On average, participants in the no-endorsement condition supported the program, in keeping with their ideological beliefs. However, when participants were told that their ingroup opposed the program, they showed biased processing of the information presented in the proposal, selectively interpreting ambiguous information and selectively attending to unambiguous information to support the ingroup position. As a result, participants in the ingroup-opposed condition were more likely to oppose the program themselves, compared to participants in the no-information condition. Moreover, the Democratic participants believed that group endorsement influenced the attitudes of other Democrats and (even more strongly) Republicans, but perceived themselves to be relatively unaffected by this information. Thus, consistent with the notion that heuristic processing need not involve intentionality and self-awareness (see Chaiken et al., 1989; Chen and Chaiken, 1999), it seems likely that people are unaware of the extent to which group endorsements bias their thinking about an issue. This may tend to exacerbate political conflicts: whereas Democrats and Republicans might agree on a policy in the absence of endorsement information, merely attaching a party label to a proposal can distort information processing and lead partisans to adopt divergent positions. Interestingly, then, bipartisan proposals may be particularly likely to gain public support not only because their actual content may better address the political goals of both groups, but also because the absence

of a link to a particular party may help to promote more open-minded information processing.

CONCLUSION

Looking back, we see the heuristic-systematic model as very much a product of its historical context, building on theories both within the attitudes domain and outside of it, and developing beyond the study of basic social psychological processes to shed light on important and relevant social issues. To us, this illustrates the benefit of working in an area with such a long and cumulative history that both influences and draws from other psychological and related social-science disciplines. In coming years, we hope that the field continues to develop the heuristic-systematic model in concert with other dual-process theories, drawing from the research that has already been done to influence that which is yet to come.

REFERENCES

Allport, G. (1954) *The Nature of Prejudice*. Reading, MA: Addison-Wesley.

Anderson, R.C. and Pichert, J.W. (1978) Recall of previously unrecallable information following a shift in perspective. *Journal of Verbal Learning and Verbal Behavior*, 17, 1–12.

Axsom, D., Yates, S.M. and Chaiken, S. (1987) Audience response as a heuristic cue in persuasion. *Journal of Personality and Social Psychology*, 53, 30–40.

Bargh, J.A. (1994) The Four Horsemen of automaticity: Awareness, efficiency, intention, and control in social cognition. In R.S. Wyer, Jr. and T.K. Srull (eds), *Handbook of Social Cognition*, 2nd Edition, pp. 1–40. Hillsdale, NJ: Erlbaum.

Baumeister, R.F. and Leary, M.R. (1995) The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, 117, 497–529.

Bem, D.J. (1972) Self-perception theory. In L. Berkowitz (ed.), *Advances in Experimental Social Psychology*, 6, 1–62. New York: Academic Press.

Brewer, M.B. (1988) A dual-process model of impression formation. In T. Srull and R. Wyer (eds), *Advances in Social Cognition*, 1, 1–36. Hillsdale, NJ: Erlbaum.

Brewer, N. and Hupfeld, R.M. (2004) Effects of testimonial inconsistencies and witness group identity on mock-juror judgments. *Journal of Applied Social Psychology*, 34, 493–513.

Brewer, W.F. and Treyens, J.C. (1981) Role of schemata in memory for places. *Cognitive Psychology*, 13, 207–230.

Bruner, J.S. (1957) On perceptual readiness. *Psychological Review*, 64, 123–152.

Chaiken, S. (1980) Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology*, 39, 752–766.

Chaiken, S. (1987) The heuristic model of persuasion. In M.P. Zanna, J.M. Olson and C.P. Herman (eds), *Social Influence: The Ontario Symposium*, 5, 3–39. Hillsdale, NJ: Erlbaum.

Chaiken, S., Axsom, D., Liberman, A. and Wilson, D. (1992) Heuristic processing of persuasive messages: Chronic and temporary sources of rule accessibility. Unpublished manuscript, New York University.

Chaiken, S. and Eagly, A.H. (1976) Communication modality as a determinant of message persuasiveness and message comprehension. *Journal of Personality and Social Psychology*, 34, 605–614.

Chaiken, S. and Eagly, A.H. (1983) Communication modality as a determinant of persuasion: The role of communicator salience. *Journal of Personality and Social Psychology*, 45, 241–256.

Chaiken, S., Giner-Sorolla, R. and Chen, S. (1996) Beyond accuracy: multiple motives in heuristic and systematic processing. In P.M. Gollwitzer and J.A. Bargh (eds), *The Psychology of Action: Linking Motivation and Cognition to Action*, pp. 553–578. New York: Guilford Press.

Chaiken, S., Liberman, A. and Eagly, A.H. (1989) Heuristic and systematic information processing within and beyond the persuasion context. In J.S. Uleman and J.A. Bargh (eds), *Unintended Thought*, pp. 212–252. New York: Guilford Press.

Chaiken, S. and Maheswaran, D. (1994) Heuristic processing can bias systematic processing: Effects of source credibility, argument ambiguity, and task importance on attitude judgment. *Journal of Personality and Social Psychology*, 66, 460–473.

Chaiken, S. and Trope, Y. (eds) (1999) *Dual-Process Theories in Social Psychology*. New York: Guilford Press.

Chen, S. and Chaiken, S. (1999) The heuristic-systematic model in its broader context. In S. Chaiken and Y. Trope (eds), *Dual-process Theories in Social Psychology*, pp. 73–96. New York: Guilford Press.

Chen, S., Shechter, D. and Chaiken, S. (1996) Getting at the truth or getting along: Accuracy- versus impression-motivated heuristic and systematic processing. *Journal of Personality and Social Psychology*, 71, 262–275.

Cohen, G.L. (2003) Party over policy: The dominating impact of group influence on political beliefs. *Journal of Personality and Social Psychology*, 85, 808–822.

Converse, P.E. (1964) The nature of belief systems in mass publics. In D.E. Apter (ed.), *Ideology and Discontent*, pp. 206–261. New York: Free Press.

Darke, P.R., Chaiken, S., Bohner, G., Einwiller, S., Erb, H.-P. and Hazlewood, J.D. (1998) Accuracy motivation, consensus information, and the law of large numbers: Effects on attitude judgment in the absence of argumentation. *Personality and Social Psychology Bulletin*, 24, 1205–1215.

De Dreu, C.K.W. (2003) Time pressure and closing of the mind in negotiation. *Organizational Behavior and Human Decision Processes*, 91, 280–295.

De Dreu, C.K.W. (2004) Motivation in negotiation: A social psychological analysis. In M. Gelfand and J.M. Brett (eds), *Handbook of Negotiation and Culture*, pp. 114–135. Palo Alto, CA: Stanford University Press.

De Dreu, C.K.W., Koole, S. and Oldersma, F.L. (1999) On the seizing and freezing of negotiator inferences: Need for cognitive closure moderates the use of heuristics in negotiation. *Personality and Social Psychology Bulletin*, 25, 348–362.

De Dreu, C.K.W., Koole, S.L. and Steinle, S. (2000) Unfixing the fixed pie: A motivated information-processing approach to integrative negotiation. *Journal of Personality and Social Psychology*, 79, 975–987.

Deutsch, M. and Gerard, H.B. (1955) A study of normative and informational social influences upon individual judgment. *Journal of Abnormal Social Psychology*, 51, 629–636.

Devine, P.G. (1989) Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, 56, 5–18.

Ditto, P.H. and Lopez, D.F. (1992) Motivated skepticism: Use of differential decision criteria for preferred and nonpreferred conclusions. *Journal of Personality and Social Psychology*, 63, 568–584.

Eagly, A.H. and Chaiken, S. (1993) *The Psychology of Attitudes*. Belmont, CA: Wadsworth.

Erb, H.P., Bohner, G., Schmalzle, K. and Rank, S. (1998) Beyond conflict and discrepancy: cognitive bias in minority and majority influence. *Personality and Social Psychology Bulletin*, 24, 620–633.

Fazio, R.H. and Towles-Schwen, T. (1999) The MODE model of attitude-behavior processes. In S. Chaiken and Y. Trope (eds), *Dual-process Theories in Social Psychology*, pp. 97–116. New York: Guilford.

Fiske, S.T. (2002) Five core social motives, plus or minus five. In S.J. Spencer, S. Fein, M.P. Zanna and J. Olson (eds), *Motivated Social Perception: The Ontario Symposium*, 9, 233–246. Mahwah, NJ: Erlbaum.

Fiske, S.T., Lin, M.H. and Neuberg, S.L. (1999) The Continuum Model: Ten years later. In S. Chaiken and Y. Trope (eds), *Dual-process Theories in Social Psychology*, pp. 231–254. New York: Guilford Press.

Fiske, S.T. and Linville, P.W. (1980) What does the schema concept buy us? *Personality and Social Psychology Bulletin*, 6, 543–557.

Fiske, S.T. and Taylor, S.E. (2008) *Social Cognition: From Brains to Culture*. New York: McGraw-Hill.

Fleming, M.A. and Petty, R.E. (2000) Identity and persuasion: An elaboration likelihood approach. In D.J. Terry and M.A. Hogg (eds), *Attitudes, Behavior, and Social Context: The Role of Norms and Group Membership*, pp. 171–199. Mahwah, NJ: Lawrence Erlbaum Associates.

Forehand, M., Gastil, J. and Smith, M.A. (2004) Endorsements as voting cues: Heuristic and systematic processing in initiative elections. *Journal of Applied Social Psychology*, 34, 2215–2233.

ForsterLee, R., ForsterLee, L., Horowitz, I.A. and King, E. (2006) The effects of defendant race, victim race, and juror gender on evidence processing in a murder trial. *Behavioral Sciences and the Law*, 24, 179–198.

Gawronski, B. and Bodenhausen, G.V. (2006) Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin*, 132, 692–731.

Gilbert, D.T. (1989) Thinking lightly about others: Automatic components of the social inference process. In J.S. Uleman and J.A. Bargh (eds), *Unintended Thought*, pp. 189–211. New York: Guilford.

Giner-Sorolla, R. and Chaiken, S. (1997) Selective use of heuristic and systematic processing under defense motivation. *Personality and Social Psychology Bulletin*, 23, 84–97.

Giner-Sorolla, R., Chaiken, S. and Lutz, S. (2002) Validity beliefs and ideology can influence legal case judgments differently. *Law and Human Behavior*, 26, 507–526.

Hall, C.C., Goren, A., Chaiken, S. and Todorov, A. (2009) Shallow cues with deep effects: Trait judgments from faces and voting decisions. In E. Borgida, J.L. Sullivan, and C.M. Federico (eds), *The Political Psychology of Democratic Citizenship*, pp. 73–99. Oxford: Oxford University Press.

Harris, P.R. and Napper, L. (2005) Self-affirmation and the biased processing of health-risk information. *Personality and Social Psychology Bulletin*, 31, 1250–1263.

Heider, F. (1958) *The Psychology of Interpersonal Relations*. Hoboken, NJ: Wiley.

Higgins, E.T. (1989) Knowledge accessibility and activation: Subjectivity and suffering from unconscious sources. In J.S. Uleman and J.A. Bargh (eds), *Unintended Thought*, pp. 75–115. New York: Guilford Press.

Higgins, E.T., King, G.A. and Mavin, G.H. (1982) Individual construct accessibility and subjective impressions and recall. *Journal of Personality & Social Psychology*, 43, 35–47.

Higgins, E.T. and McCann, C.D. (1984) Social encoding and subsequent attitudes, impressions, and memory: 'Context-driven' and motivational aspects of processing. *Journal of Personality and Social Psychology*, 47, 26–39.

Howard, M.V.A., Brewer, N. and Williams, K.D. (2006) How processing resources shape the influence of stealing thunder on mock-juror verdicts. *Psychiatry, Psychology and Law*, 13, 60–66.

Jepson, C. and Chaiken, S. (1990) Chronic issue-specific fear inhibits systematic processing of persuasive communications. *Journal of Social Behavior and Personality*, 5, 61–84.

Johnson, B.T. and Eagly, A.H. (1989) The effects of involvement on persuasion: A meta-analysis. *Psychological Bulletin*, 106, 290–314.

Jones, E.E. and Davis, K.E. (1965) From acts to dispositions: The attribution process in social psychology. In L. Berkowitz (ed.), *Advances in Experimental Social Psychology*, 2, 219–266. New York: Academic Press.

Katz, D. (1960) The functional approach to the study of attitudes. *Public Opinion Quarterly*, 24, 163–204.

Kelley, H.H. (1972) *Causal Schemata and the Attribution Process*. New York: General Learning Press.

Kelley, H.H. (1973) The processes of causal attribution. *American Psychologist*, 28, 107–128.

Kiesler, C.A. (1971) *The Psychology of Commitment: Experiments Linking Behavior to Belief*. San Diego, CA: Academic Press.

Kohler, W. (1930) *Gestalt Psychology*. London: Bell.

Lau, R.R. and Redlawsk, D.P. (2001) An experimental study of information search, memory, and decision making during a political campaign. In J. Kuklinski (ed.), *Citizens and Politics: Perspectives from Political Psychology*, pp. 136–159. New York: Cambridge University Press.

Ledgerwood, A. and Chaiken, S. (2007) Priming us and them: Automatic assimilation and contrast in group attitudes. *Journal of Personality and Social Psychology*, 93, 940–956.

Ledgerwood, A., Chaiken, S., Gruenfeld, D.H. and Judd, C.M. (2006) Changing minds: Persuasion in negotiation and conflict resolution. In M. Deutsch, P.T. Coleman, and E.C. Marcus (eds), *The Handbook of Conflict Resolution: Theory and Practice*, 2nd Edition, pp. 455–485. San Francisco: Jossey-Bass.

Ledgerwood, A., Mandisodza, A.N., Jost, J.T. and Pohl, M.J. (in press) Working for the system: Motivated defense of meritocratic beliefs. *Social Cognition*.

Lerner, J.S. and Tetlock, P.E. (1999) Accounting for the effects of accountability. *Psychological Bulletin*, 125, 255–275.

Liberman, A. and Chaiken, S. (1992) Defensive processing of personally relevant health messages. *Personality and Social Psychology Bulletin*, 18, 669–679.

Lord, C.G., Ross, L. and Lepper, M.R. (1979) Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, 37, 2098–2109.

Mackie, D.M. (1987) Systematic and nonsystematic processing of majority and minority persuasive communications. *Journal of Personality and Social Psychology*, 53, 41–52.

Maheswaran, D. and Chaiken, S. (1991) Promoting systematic processing in low motivation settings: Effect of incongruent information on processing and judgment. *Journal of Personality and Social Psychology*, 61, 13–25.

Maheswaran, D., Mackie, D.M. and Chaiken, S. (1992) Brand name as a heuristic cue: The effects of task importance and expectancy confirmation on consumer judgments. *Journal of Consumer Psychology*, 1, 317–336.

Marcus, G.E., Neumann, W.R. and MacKuen, M.B. (2000) *Affective Intelligence and Political Judgment*. Chicago: University of Chicago Press.

Martin, R., Hewstone, M. and Martin, P.Y. (2007) Systematic and heuristic processing of majority and minority endorsed messages: The effects of varying 'levels of orientation' and outcome relevance on attitude and message processing. *Personality and Social Psychology Bulletin*, 33, 43–56.

McGuire, W.J. (1969) The nature of attitudes and attitude change. In G. Lindzey and E. Aronson (eds), *Handbook of Social Psychology*, 2nd Edition, 3, 136–314. Reading, MA: Addison-Wesley.

Mondak, J.J. (1993) Public opinion and heuristic processing of source cues. *Political Behavior*, 15, 167–192.

Moskowitz, G.B., Skurnik, L and Galinsky, A.D. (1999) The history of dual-process notions, and the future of pre-conscious control. In S. Chaiken and Y. Trope (eds), *Dual-process Theories in Social Psychology*, pp. 12–36. New York: Guilford Press.

Newman, B.I. and Perloff, R.M. (2004) Political marketing: theory, research, and applications. In L.L. Kaid (ed.), *Handbook of Political Communication Research*, pp. 17–43. Mahwah, NJ: Erlbaum.

Pallak, M.S., Mueller, M., Dollar, K. and Pallak, J. (1972) Effect of commitment on responsiveness to an extreme consonant communication. *Journal of Personality and Social Psychology*, 23, 429–436.

Petty, R.E. and Cacioppo, J.T. (1984) Source factors and the elaboration likelihood model of persuasion. *Advances in Consumer Research*, 11, 668–672.

Petty, R.E., Cacioppo, J.T. and Goldman, R. (1981) Personal involvement as a determinant of argument-based persuasion. *Journal of Personality and Social Psychology*, 41, 847–855.

Petty, R.E. and Wegener, D.T. (1999) The elaboration likelihood model: Current status and controversies. In S. Chaiken and Y. Trope (eds), *Dual-process Theories in Social Psychology*, pp. 37–72. New York: Guilford.

Petty, R.E., Wells, G.L. and Brock, T.C. (1976) Distraction can enhance or reduce yielding to propaganda: Thought disruption versus effort justification. *Journal of Personality and Social Psychology*, 34, 874–884.

Ratneshwar, S. and Chaiken, S. (1991) Comprehension's role in persuasion: The case of its moderating effect on the persuasive impact of source cues. *Journal of Consumer Research*, 18, 52–62.

Scott, C.G. (1996) Understanding attitude change in developing effective substance abuse prevention

programs for adolescents. *The School Counselor*, 43, 187–195.

Sherif, M. and Cantril, H. (1947) *The Psychology of Ego-involvements: Social Attitudes and Identifications*. New York: Wiley.

Sherif, M. and Hovland, C.I. (1961) *Social Judgment: Assimilation and Contrast Effects in Communication and Attitude Change*. New Haven, CT: Yale University Press.

Simon, H.A. (1976). *Administrative Behavior*, 3rd Edition. New York: Free Press.

Smith, E.R. and DeCoster, J. (2000) Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. *Personality and Social Psychology Review*, 4, 108–131.

Smith, M.B., Bruner, J.S. and White, R.W. (1956) *Opinions and Personality*. New York: Wiley.

Stroebe, W. and Diehl, M. (1988) When social support fails: Supporter characteristics in compliance-induced attitude change. *Personality and Social Psychology Bulletin*, 14, 136–144.

Taylor, S.E. (1975) On inferring one's attitudes from one's behavior: Some delimiting conditions. *Journal of Personality and Social Psychology*, 31, 126–131.

Taylor, S.E. and Crocker, J. (1981) Schematic bases of social information processing. In E.T. Higgins, C.P. Herman, and M.P. Zanna (eds), *Social Cognition: The Ontario Symposium*, 1, pp. 89–134. Hillsdale, NJ: Erlbaum.

Tesser, A. and Campbell, J. (1983) Self-definition and self-evaluation maintenance. In J. Suls and A. Greenwald (eds), *Social Psychological Perspectives on the Self*, 2, 1–31. Hillsdale, NJ: Erlbaum.

Tetlock, P.E. (1989) Structure and function in political belief systems. In A.R. Pratkanis, S.J. Breckler, and A.G. Greenwald (eds), *Attitude Structure and Function*, pp. 129–151. Hillsdale, NJ: Erlbaum.

Tetlock, P.E., Skitka, L. and Boettger, R. (1989) Social and cognitive strategies for coping with accountability: Conformity, complexity, and bolstering. *Journal of Personality and Social Psychology*, 57, 632–640.

Thompson, L.L. and Hastie, R. (1990) Social perception in negotiation. *Organizational Behavior and Human Decision Processes*, 47, 98–123.

Todorov, A., Mandisodza, A.N., Goren, A. and Hall, C.C. (2005) Inferences of competence from faces predict election outcomes. *Science*, 308, 1623–1626.

Tversky, A. and Kahneman, D. (1974) Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124–1131.

Webster, D.M. and Kruglanski, A.W. (1994) Individual differences in need for cognitive closure. *Journal of Personality and Social Psychology*, 67, 1049–1062.

Werner, C.M., Stroll, R., Birch, P. and White, P.H. (2002) Clinical validation and cognitive elaboration: Signs that encourage sustained recycling. *Basic and Applied Social Psychology*, 24, 185–203.

Wood, W., Kallgren, C.A. and Preisler, R.M. (1985) Access to attitude-relevant information in memory as a determinant of persuasion: The role of message attributes. *Journal of Experimental Social Psychology*, 21, 73–85.

Yates, J. (1985) The content of awareness is a model of the world. *Psychological Review*, 92, 249–284.

Ziegler, R., von Schwischow, A. and Diehl, M. (2005) Matching the message source to attitude functions: Implications for biased processing. *Journal of Experimental Social Psychology*, 41, 645–653.

Zuckerman, A. and Chaiken, S. (1998) A heuristic-systematic processing analysis of the effectiveness of product warning labels. *Psychology and Marketing*, 15, 621–664.

The Continuum Model and the Stereotype Content Model

Susan T. Fiske

ABSTRACT

Two models address respectively processes and contents of social categories in impression formation. In the continuum model (CM), people default to category-based processes to form impressions of others, so they often respond rapidly, captured by schema-triggered affect, stereotypic associations, and discriminatory responses. According to the CM, adequate information and appropriate motivation can override categories and encourage more individuating processes. What are the default categories that capture social cognition? In the stereotype content model (SCM), two fundamental dimensions describe clusters of apparently universal categories for interpersonal and intergroup perception. The warmth dimension divides people into friends and foes, while the competence dimension divides people by status. Warmth-by-competence combinations identify (a) the ingroup plus its allies, seen as both warm and competent, eliciting pride (e.g., middle-class people), contrasting with three outgroup types; (b) extreme, disgust-inducing groups, allegedly neither warm nor competent (e.g., homeless people); (c) ambivalent, pitied groups, allegedly warm but incompetent (e.g., older people); and (d) ambivalent, envied groups, allegedly competent but not warm (e.g., rich people). Distinct

active and passive discrimination target each quadrant, along with preliminary evidence of neural correlates. Not all categories are equivalent; CM and SCM principles describe both usage and patterns of content with policy implications.

CATEGORICALLY CAPTURED: THE CONTINUUM MODEL AND THE STEREOTYPE CONTENT MODEL

Puzzle #1: At my high-school best-friend's party, over a pounding Motown beat, one random dance partner asked me, "What sign are you?" Prior to that, he knew me only as a white girl about his age, probably attending the local private school. I knew him only as a black guy about my age, presumably from the neighborhood, though not from my school. When I answered, between beats, "Leo," he turned and walked away, in the middle of the song. I was less hurt than merely puzzled as to what that category meant for him.

People must perceive with the aid of categories; as Allport noted, “[O]lderly living depends upon it” (1954: 19). Consequently, our social categories capture our impressions. We cannot help noticing a person’s gender, race, and age within a fraction of a second (see Fiske and Taylor, 2008, for a review). We also instantly perceive cues to social class, status, and some predispositions such as extraversion (though astrological signs do not appear among the most immediately visible cues). These accessible expectations shape our responses from the first moments. Two of my theories, developed with a little help from my friends, describe, respectively, the processes and the contents of our social categories. And they solve some puzzles along the way, including how people use categories such as “Leo.”

A PERSONAL TALE OF TWO THEORIES

I admit to being obsessed by categories. Attending Hyde Park’s University of Chicago Laboratory Schools – later to enroll Sasha and Malia Obama – meant even then encountering an abundantly multiracial and economic diversity of students, from nursery school onward. We played at each other’s houses and knew each other’s families. We collaborated and competed in class, clubs, and sports. We danced and dated. Not that we failed to notice race, class, or gender, but we had additional impressions to guide our play, our projects, and our schoolwork. Nevertheless came Puzzle #2: In the clique-infested halls of middle school, former elementary-school friends suddenly separated into identity groups based on income, race, and ethnicity, as well as intellectual and athletic aspirations. The cafeteria seating clusters would be familiar to anyone attending middle or high school. However, being a small community, we had known each other too long to view each other as aliens. How did we live in these two worlds of identity politics and deep familiarity?

Graduation later scattered us, and my own college experience of Boston shocked me into realizing that Chicago’s Hyde Park possessed a distinct microclimate. The rest of the world seemed to operate along stricter racial and economic divides than I knew. Puzzle #3: What made the difference?

Transition to graduate school in the 1970s, when issues of race, class, gender, and group conflict dominated late-night discussions in dorms, apartments, and – yes – my communal house. One debated question was why a local doorman kept confusing me with Kathleen Kennedy. Apart from gender, age, race, and class, why would he mix us up? (Puzzle #4). With perfect timing, my graduate advisor, Shelley Taylor, invited me on board some experiments studying social categorization; I leapt at the opportunity. As some of our first studies indicated (Taylor et al., 1978), people code each other by gender and race, so much so that they confuse other people within category (say, blacks with other blacks, and whites with other whites) more than between categories (rarely confusing blacks with whites or vice versa).

In one of our studies, Harvard undergraduates watched a prerecorded brainstorming conversation, in which each person offered suggestions about how to publicize a play. Afterward, in a surprise recognition task, participants were more likely to make within-category errors than between-category ones, as predicted. That is, they misremembered some of a black person’s suggestions as coming from a different black person. Ditto for whites with other whites, as well as men with other men and women with other women. This was some of the first evidence for the role of basic social categories in social cognition (later much replicated). Puzzle #4 again: Why wouldn’t people pay more attention, beyond race and gender?

And now for something completely different: in graduate school, as is typical, I worked as a teaching assistant for several semesters. I began to observe a phenomenon regarding my impressions of my students, namely that even when I could not remember someone’s

full identity, I always knew how I felt about the person. How did these impressions form separately, the slippery facts and the reliable likability? This Puzzle #5 fit another research project, completely different from the category-oriented one, so it requires a little background.

At this time, I was also working with Reid Hastie on impression formation according to Norman Anderson's (1981) information integration model of evaluative impressions (elaborated in the next section). His theory predicted how perceivers combine individual cues to form an overall evaluation. Our study examined whether such evaluations and the relevant memory follow the same or different cognitive processes. For Elizabeth Drebén's senior thesis (Drebén et al., 1979), Harvard undergraduates read a series of positive, negative, and neutral sentences describing a person's behavior (e.g., "Alan bought groceries for an elderly lady next door who was ill"; "Alan pressed the button and waited for the elevator to come"). People apparently formed and updated their impressions online, as they received each cue, but their final impressions did not depend on memory for the behaviors. This study provided some of the first evidence for the independence of impressions and memory in the encoding of individual cues (Hastie and Park, 1986). At each moment, as they receive each piece of information, people immediately know how they feel about other people, continually updating that impression, but later they cannot necessarily produce the relevant evidence. Ever known exactly how you feel about somebody, without recalling any of the reason for those feelings? Puzzle #5 again. Now turn to a more formal statement of our models that ultimately help explain these puzzles.

INTELLECTUAL HISTORY OF THE MODELS

Why were the Taylor and Hastie studies precedent-setting? Why was it so exciting to

be present at their inception? First, some intellectual context explains why, and then the next sections describe two Fiske et al. models that emerged to address some of social cognition's continuing puzzles.

Intellectual Context: Birth of Social Cognition

Both these graduate-school collaborations joined the opening salvos of the social cognitive revolution in psychological sciences. The first half of the twentieth century had seen psychology dominated by learning theories that valorized stimulus and response; that is, directly observable conditions and behaviors (Boring, 1950). The mind – intervening mental processes of organisms (from lab rats to undergraduates) – had been off-limits, judged scientifically irrelevant, unimportant, or even suspect. For example, Guthrie famously made fun of Tolman's maze-rat "buried in thought" and never choosing (1935: 172).

At mid-century, this tyranny was overthrown by the cognitive revolution, advocating the study of humans as information processors, with minds oriented toward encoding, representation, retrieval, and response. The computer metaphor informed studies of attention, memory, and problem solving. Cognitive psychology flourished. Before and during the cognitive revolution, meanwhile, offices on psychology departments' other floors had long opposed the field's prior domination by learning-theory diktats: social psychologists had always maintained their focus on indirectly observable mental constructs, such as attitudes and stereotypes (for an overview, see Fiske and Taylor, 2008). With the cognitive revolution, some social psychologists leapt into familiar territory (impressions) seen from a new angle. The new ideas turned out to have exciting implications for how people make sense, not just of their inanimate world, but also of each other.

Into this heady intellectual era, my research launched two separate trajectories.

My graduate work with Taylor on social categories continued in studies of category salience (Taylor and Fiske, 1978). Among them were studies of how solo (or token) members of under-represented groups suffer exaggerated, category-constrained judgments. The explanatory potential of categories appeared crucial.

At the same time, my dissertation picked up the Anderson information integration theory. With its focus on the isolated impact of each separate social cue, I hoped eventually to design studies that would quantify the impact of race, gender, and other impermissible categories on impression formation, relative to the other information available. My dissertation applied the information integration formulas to individual behaviors because they seemed to constitute a more elementary unit than personality traits, let alone social categories. In my rash cross-country phone consult, Norman Anderson firmly advised me not to use the rich but messy behavioral stimuli proposed.

Undaunted, my dissertation project did examine behaviors, namely the disproportionate weight of negative and extreme behaviors, relative to positive and moderate ones (Fiske, 1980). In a painstaking 16×16 Latin square, participants viewed images of 16 separate young men, each depicted at one of four levels of sociability (from isolation to full involvement at a picnic) and at four levels of civic responsibility (from rejection to promotion of a petition against child pornography). Participants' looking-time recorded for each slide, they rated each person's likability in turn. As predicted, the theoretically derived impression weights for each behavior closely mirrored people's online visual attention to each behavior. The study linked these computed impression weights to observable behavior, showing that both negative and extreme behaviors independently elicit both weight and attention. The study almost won the Society of Experimental Social Psychology (SESP) dissertation award, and helped me land a

job at that information-processing Mecca, Carnegie-Mellon University.

Just as my two separate lines of research seemed to be launched, nearby Ohio State University invited me for a colloquium in my first year as an assistant professor. An astute but now-nameless graduate student listened to the account of my dissertation research, knowing about my other work on social categories, and asked the life-changing question: "how can you do both?" The apparent contradiction struck me all of a heap: on one hand, the elemental, piecemeal, information-integration approach famously denied any interaction among the elements of an impression. Mathematically, it proposes a weighted averaging model, wherein each cue possesses: s_i , a set value on a likability scale (e.g., on a 100-point scale, how unlikable it is to reject friends at a picnic), and w_i , a weight reflecting its relative importance in the overall impression (e.g., social rejection might matter more than civic responsibility to one's overall likability). Regardless of its other algebraic assumptions, the model stated clearly that the elements would not affect each other's evaluations (e.g., the sociability of an irresponsible person is just as likable as the sociability of a responsible person). From a mathematical perspective, this meant that the individual elements could not interact or multiply. Hence, the terms *elemental* or *piecemeal* both describe this theory of impression formation.

On the other hand, category theories (or schema theories, as they were often known) had an intrinsically interactive flavor. For example, two landmark studies in my professional cohort established the importance of such processes. Nancy Cantor (e.g., Cantor and Mischel, 1977) showed that category prototypes intrude on related trait judgments, so that conceptually related but not presented items bias recognition memory. Hazel Markus (e.g., 1977) showed that the most central parts of people's self-concepts operate like memory schemas organizing their own traits. These and other early social cognition studies built on older Gestalt theories that argued

for configural features of perception. Applied to person perception, the *configural* approach implies that the meaning of any given cue (e.g., intelligent) depends on the context of its surrounding traits. In Solomon Asch's famous (1946) demonstration, a warm intelligent person is wise, whereas a cold intelligent person is sly.

The Gestalt meaning-change versus Andersonian fixed-scale-value perspectives inspired heated debates (Hamilton and Zanna, 1974). Because one of the protagonists taught at Ohio State (Ostrom, 1977), perhaps the student's pivotal question reflected local preoccupations. The question certainly catalyzed mine. That puzzle (#6, but who's counting) began to resolve some of the earlier ones, as a result of motivating the continuum model.

A category process model: the continuum model of impression formation

The wordy subtitle to our eventual main theoretical statement provides a précis of the continuum model (CM): from category-based to individuating processes, influence of information and motivation on attention and interpretation (Fiske and Neuberg, 1990). The CM resulted directly from the Ohio State graduate student pointing out the apparent contradiction between my two lines of research, examining both elemental and categorical processes. On reflection, choosing one process, to the exclusion of the other, made no sense. Too much research already demonstrated the viability of each type of process using different paradigms. My hunch was that people would do both, but under different circumstances that not coincidentally reflected the respective paradigms used to establish the distinct processes. The theoretical task was to demonstrate how people do which one when.

Our team's initial efforts were more empirical than explicitly theoretical, anticipated in a 1982 grant proposal, whose overall goal

was to contrast category-based processing, characteristic of stereotyping, with individuating processes, characteristic of piecemeal impressions. Nevertheless, the studies test the model that eventually evolved out of that work, so I will start here with the subsequently official CM hypotheses (Fiske and Neuberg, 1990):

- Category-based processes take priority over attribute-based process, both because people start with categories and because they stop at categories, if the categories work well enough. Categories are instantly available, convenient, and linked to ready responses – stereotypes, prejudices, and discriminatory tendencies. If the target individual adequately fits the category prototype, then perceivers should not bother collecting and analyzing details about the target.
- Perceivers progress along the continuum, from category-based processes, to recategorization, to more attribute-based processes, depending on adequate information fit. Although captured by categories, perceivers are no fools, and if the category obviously does not fit (Obama does not appear to be a stereotypic black man), people first recategorize by picking another category (maybe he is a Muslim), subtyping (maybe he is a black professional), making a new category (biracial, cosmo-ethnic), or relating him to self (he's a citizen of the world, like me). If none of those compromises fit, then people construct a more deliberate, attribute-by-attribute, piecemeal impression (Hawaiian birth, Kansas upbringing, anthropologist mother, Kenyan father, Indonesian elementary school, Columbia, Harvard, Chicago community organizer, Illinois legislator...).
- Attention mediates the use of attribute information, beyond the initial category. Each step requires progressively more attentional resources, and focus of attention determines information use. For example, attentional focus on category-consistent information tends to reconfirm the category, whereas attention to inconsistent information allows category disconfirmation.
- Motivation, especially interdependence, moderates progress along the continuum. People accept initial, default categories when they have no motivation to expend resources examining further. When accuracy especially matters (because they depend on the other person, for

example, or because they personally prioritize accurate impressions), people are motivated to try harder.

- Attention also mediates the effects of motivation on impressions. Accuracy-motivated people focus on apparent inconsistencies, in an effort to resolve them by changing their meaning to fit the category or revising their impression. People also can be motivated to confirm their stereotypes, and in that case, they might focus on category-confirming attributes.

Antecedents of each hypothesis came from prior theorists. A version of the first hypothesis had long ago appeared in Asch's (1946) paper that favored Gestalt configurational processes over more elemental (algebraic) processes, although we were among the first applying modern information-processing theories to these ideas. Information fit, as a moderator, built on person-memory research that contrasted congruent and incongruent information (e.g., Hastie et al., 1980). Motivation, as a moderator, built on dissatisfactions with the cold-cognition approaches (Clark and Fiske, 1982). And the mediator in both cases, attention, built on my dissertation and prior graduate work on salience (Taylor and Fiske, 1978).

First, our lab tried to tackle information use as revealing each process, and soon afterward we turned to motivational moderators with more consistent success. Our earliest information-analytic work appeared only in book chapters (Fiske, 1982; Fiske and Pavelchak, 1986), as we struggled to find paradigms that would adequately test the CM informational predictions. Several studies manipulated category fit, showing that matches (i.e., category-consistent attributes) elicited evaluations consistent with the overall category, in what we termed *schema-triggered affect*. Mismatches (i.e., a category label but inconsistent attributes) elicited evaluations more in line with the attributes. One early study tested an old-flame hypothesis, whereby we tailored stimulus people to match (or not) each person's past or current partner on traits, appearance, or both. However appealing the idea and promising

the preliminary results, we had confounded positivity with match. Susan Andersen later examined significant-other transference much more elegantly (for a summary and theory, see Andersen and Chen, 2002). Another converging effort, to manipulate politician schemas (mostly negative), suffered related confounds.

Yet another, but more convincing, study contrasted local stereotypes of engineers and artists, respectively "vegetables" and "fruits," on that campus. By using each stereotype's content as the other's control, we finally unconfounded valence from category fit, by comparing stereotypic engineers to: other engineers with artistic traits, artists with engineer traits, and stereotypic artists. Thus, holding information constant, schema matches elicited negative responses in line with these negative stereotypes, more than did mismatches. The astute reader will note that, although we had unconfounded valence from consistency, the inconsistent stimulus people still might have been more interesting or salient.

Luckily for the CM, the first peer-reviewed evidence for dual processes of impression formation (Fiske et al., 1987) was an eventual pair of experiments designed to demonstrate category-based processes using two converging operationalizations, versus attribute-based processes, again using two (other) converging operationalizations. The main CM-related prediction was: affective responses will reflect the category when categorization is easy, but they will reflect the attributes when categorization is difficult.

The two easily categorized conditions provided either (1) a category label, followed by stereotype-consistent attributes (e.g., a loan shark who is opportunistic, shady, greedy, shrewd, and heartless); or (2) a category label, followed by uninformative attributes (e.g., an artist who is adult, medium height, employed, television-viewer, brown-haired). Two difficult-categorization conditions included either: (3) a category label, followed by category-inconsistent attributes (although consistent with another stereotype

in the stimulus set; that is, a doctor who is bored, obedient, unenterprising, uneducated, and efficient does not fit a doctor stereotype, but does fit a maid stereotype); or (4) no category label, followed by attributes consistent with another stereotype in the stimulus set (e.g., a person who is practical, educated, scientific, skilled, and observant, attributes stereotypic for a doctor, but not immediately cuing the label). As a manipulation check, typicality ratings confirmed the predicted ease of categorization in each condition.

In a pretest, as Table 13.1 indicates, participants had provided separate ratings of the category labels and the trait sets. This enabled us to calculate precise correlations between each person's likability rating for a particular stimulus combination and that same person's own prior separate ratings of the component category and attribute set, a technique originally crafted in my advisee Mark Pavelchak's dissertation (1989, see below). The critical correlations revealed

that, in the two easily categorized conditions, a stimulus person's overall likability ratings correlated with the category labels (as well as the redundant or uninformative attributes). But in the two difficult-to-categorize conditions, likability ratings mainly correlated with the attributes but not the provided category label (which did not fit the attributes). What's more, in a second experiment, participants thought aloud into a tape recorder as they formed impressions. Many of the CM's hypothesized processes occurred: spontaneous comments about typicality, relying on the category when the attributes matched, recategorizing when they did not, and simply combining the attributes when no category fit.

Simultaneous with this information-based mechanism for manipulating category-based versus attribute-based processes, Pavelchak (1989) developed the idea of manipulating the two processing modes by instruction. First, he developed trait lists that fit campus

Table 13.1 Stimulus materials for Experiments 1 and 2 in Fiske et al. (1987): categories, attributes, and their respective likabilities

Experiment 1				
<i>Loan shark</i>	<i>Artist</i>	<i>Doctor</i>	<i>Hotel maid</i>	<i>Person</i>
2.40	5.91	6.09	4.99	6.03
Opportunistic	Nonconforming	Practical	Bored	Adult
Shady	Creative	Educated	Obedient	Medium height
Greedy	Eccentric	Scientific	Unenterprising	Employed
Shrewd	Idealistic	Skilled	Uneducated	Television viewer
Heartless	Fashionable	Observant	Efficient	Brown-haired
2.05	6.23	6.46	3.91	5.38
Experiment 2				
<i>Professor</i>	<i>Reporter</i>	<i>Construction Worker</i>	<i>Politician</i>	<i>Person</i>
5.7	4.7	5.2	4.7	5.3
Intellectual	Curious	Strong	Selfish	Ordinary
Productive	Energetic	Loud	Power-hungry	Normal
Preoccupied	Perceptive	Rowdy	Pragmatic	Nice
Self-regulated	Aggressive	Red-neck	Opinionated	Usual
Hard-working	Liberal	Closed-minded	Smiley	Unremarkable
6.7	6.9	2.0	2.9	5.4

Note. All ratings are on 9-point scales, where 1 = not at all likable, 9 = extremely likable. Number of judges for Experiment 1 stimuli is 173; number of judges for Experiment 2 stimuli is 10, so fewer decimal places are given.

stereotypes for various academic majors (but did not automatically cue them). In an initial session, participants rated the likability of 35 majors and 50 relevant traits. In the second session, participants saw six stimulus people consisting only of four traits each. Half the participants first guessed the person's academic major (thereby activating a category) and then evaluated the person; the other half evaluated each person and then afterward guessed the academic majors. In the category-first condition, evaluations correlated with each participant's own prior evaluation of the category, more than with the average of the relevant attributes. In the no-category (piecemeal) condition, evaluations correlated with the average of the attributes, rather than the category that the participant guessed afterward. These (and some other Fiske-lab studies) laid the ground for informational conditions that moderate category and attribute-based processes.

At the same time, motivation shared our research agenda and soon dominated our work, for several reasons. First, my empirical talents emerged as stronger in interpersonal experimental social psychology than in heavily cognitive social psychology, as scores of old, unpublished manuscript files would prove. Second, on a more intellectual note, early on, Bill Swann and I met at Nag's Head workshops on social cognition, discovering that we shared a suspicion about the relentless person-memory findings that people's free-recall favors expectancy-inconsistent information (e.g., Srull, 1981). This inconsistency bias simply could not hold in the hurly-burly of the real world, where people habitually rely on stereotypes and expectancies to bolster their memories; both our work attested to the power of category-based consistencies (Snyder and Swann, 1978; Taylor et al., 1978), and as later meta-analyses indeed indicated (Stangor and McMillan, 1992). How could the (stereotype) consistency and the (memory) inconsistency biases both be true? Maybe in a person-memory experiment with massive numbers of stimulus people (as in my own earlier work;

Dreben et al., 1979; Fiske, 1980), perceivers do engage piecemeal processing that would favor inconsistencies, but otherwise, most of the time, in keeping with the CM, they should use their easily-at-hand category-based expectancies.

Swann and I wondered if it would matter whether people really cared about the other person; although we intended to run and later combine parallel collaborative studies, his interests shifted away. As it happens, my lab did design a study that showed no motivational effects on memory, but instead strong motivational effects on attention, attributions, and inferences. That study and its replications, heavily vetted by my local colleagues (as the acknowledgments to them and Swann attest), was our best paper to date, and probably my only empirical paper ever accepted outright, without revision (Erber and Fiske, 1984).

Graduate student Ralph Erber and I designed a situation in which people would expect to interact with another person, but in some cases would care, and in other cases not care, to go beyond superficial (category-based) impressions. We predicted that under ordinary circumstances, people would focus on expectancy-consistent information, but that if people cared about the person as a unique individual, they would go beyond their expectancies to focus on expectancy-inconsistent information. Although designed with stereotyping in mind, the experiment started with a simple positive-negative valenced expectancy.

Volunteers showed up to work as the novice in an expert-novice collaboration with a teacher-in-training (the expert). Together, they would design educational games for children, using silly wind-up toys. Undergraduates met a friendly but bland female confederate, who then exited, while they separately completed some introductory questionnaires. The partner's questionnaire communicated the valenced expectancy: she predicted that she might be really good at this task, or that she thought her ability would be low. Apparently at our request, she had also

brought with her some teaching comments allegedly written by her peers, each one a single sentence in different handwriting on a separate index card (this was the low-tech early 1980s). The comments pretested as positive were written in blue ink, the negative ones in black ink (or vice versa, across conditions). With a stopwatch in each jacket pocket (again, primitive but serviceable), the experimenter could surreptitiously time how long the participant examined each type of comment (blind to the match between color and valence). After reading the comments at their own speed to form a preliminary impression of their partner, participants rated her before an expected interaction that never came.

The crucial manipulation was motivation in the form of interdependence: how much they depended on their partner. Everyone expected to brainstorm ideas alone at first and then to work together with the partner. Half the participants expected to be eligible for a prize based on their work alone in the first half, while the other half expected to be eligible for the prize based on their joint work. We reasoned that the outcome-dependent participants would be motivated to pay more attention to their partner, especially to the most diagnostic information, the parts that disputed their expectancy, because they cared having an accurate understanding, to work together effectively. Indeed, outcome dependency uniquely predicted attention to inconsistent information (regardless of valence). Apparently in the service of deeper understanding, outcome-dependency also predicted more dispositional attributions, characterized as individuating the partner, beyond the simplistic expectancy. As the CM predicts, motivation moderated category- and attribute-based impression processes.

Conceptual replications of these studies bolstered their support for the CM. Actual stereotypes (of a recently discharged hospital patient, who was either an anxiety-provoking schizophrenic or a harmless heart patient) showed patterns similar to merely valenced

expectancies, as well as effects on category-based versus attribute-based evaluations (Neuberg and Fiske, 1987). The crucial motivation appeared as an effort toward accuracy, consistent with the idea that people are trying to control their own outcomes by having a sense of accuracy about their partners who influence their potential success. Consistent with the accuracy mediator, people's motivations to use category-based versus individuating processes respond to direct instruction ("try to be accurate"; Neuberg, 1989).

Interdependence presupposes correlated outcomes. If one's contingency on another person constitutes the crucial motivation, then – besides collaborative interdependence (positively correlated outcomes) – negative interdependence (competition) should show the same patterns. Interpersonal competition does exactly this (Ruscher and Fiske, 1990). More broadly, people respond to situational motivators in general. For example, people use the two processes according to perceived norms, for those who care about norms, and according to self-concept, for those who care about that (Fiske and Von Hendy, 1992). We were encouraged in all these studies by our colleagues, by then at University of Massachusetts, Amherst, and by National Science Foundation and National Institute of Mental Health support.

Having explored various moderators (motivation, instructions, norms, self-concept) and mediators (attention) predicted by the CM, the research program's next step was to explore boundary conditions. All the studies to date had concerned interpersonal interactions, where the correlated outcomes occurred within dyads; intergroup interdependence, having more complex contingencies, should not show the same effects (Ruscher et al., 1991). People prioritized individuating their teammates, with whom they had positive interdependence, but neglected the opposing team. Presumably, knowing the teammates is a necessary condition for understanding the group-on-group competition. Besides, the opponents can be assumed to have a simple shared purpose,

namely besting one's own team. Learning their individual weaknesses can wait.

We further pursued the intergroup dynamics because an aggregate can operate as a monolithic, impenetrable group or as a bunch of knowable individuals. If one's outcomes depend on a hostile and homogenous group, presumably less open to influence, one might stay with categorical processes, but if the group were heterogeneous, presumably less tightly knit and therefore more open to influence, one might be motivated to use individuating processes. As predicted, the interpersonal differed from the intergroup context, apparently because of differences in the potential for control (Dépret and Fiske, 1999).

The perceiver's possibility of control began to appear more and more crucial. In an earlier study (Ruscher and Fiske, 1990), unconfident participants, who did not expect to do well at the joint task, simply gave up, not showing the same selective attention and dispositional inferences to the most diagnostic (inconsistent) information about their partners. The more we thought about interdependence as motivating attempts at accuracy, to control one's outcomes, the more explanatory value the concept of control seemed to have. A moderating role for control had also appeared in perceivers being subject to higher degrees of a homogenous outgroup's power (Dépret and Fiske, 1999), being a minority (Guinote et al., 2006), or being on powerholders' territory (Guinote and Fiske, 2003).

In other studies, perceived control also seemed to characterize our standard task-related outcome-dependency conditions, in which working harder or smarter at a specific task could presumably yield some control over one's outcomes. In contrast, however, under evaluation-related outcome dependency conditions, an absent powerholder would judge one's performance, without any interaction or personal contact, as is all-too-often the case. In this context, participants apparently felt they could not exert much control over their fate; instead they showed a positivity bias, hoping for the best from their evaluator (Stevens and

Fiske, 2000). Similarly, in another setting that entailed an evaluation of oneself as an overall person – romantic outcome-dependency – participants again undermined accuracy motivation, settling for wishful thinking (Goodwin et al., 2002). When unable to control their fate, people apparently hope for the best.

During the course of these studies that established for us the importance of control, another variable also yielded a more precise understanding as we operationalized it in different ways. Interdependence implies that two people each depend on the other, as in one-on-one cooperation or competition. So far, we had shown that mutual interdependence increases motivations to be accurate and control one's outcomes, resulting in attention to diagnostic information, dispositional inferences that individuate the other person, and relatively attribute-based impressions. Theoretically, though, we had a confound. The interdependence could cause people either (1) to worry about their own, contingent outcomes (our hypothesis) or (2) to feel responsible for the other person; either case could motivate people to form more attribute-based, careful impressions. Two lines of work resulted from this insight.

First, a series of studies isolated the participant's own outcome dependency on (versus independence from) the partner. For example, people who depend on their partner, either symmetrically or asymmetrically, respond in similar ways. Compared with an independent control group, asymmetrically outcome-dependent participants – that is, those who depend on their partner but the partner does not depend on them – still strive for accuracy and control (Stevens and Fiske, 2000).

Second, we began to take power seriously. So far, we had examined outcome dependency from the bottom up. What about the top-down perspective? Defined as control over another's outcomes, power should make people neglect the other person as a unique individual, according to the CM logic, because the powerful are less motivated to go

beyond their default categorical impressions. A theory of power as control predicted: The powerful “need not, cannot, and want not” to attend to their subordinates because of, respectively, their own outcome control, the sheer numbers of subordinates, and potential personality correlates of seeking power (Fiske, 1993: 621). Laying out these predictions did not immediately recruit graduate students to collaborate on the project. American students seemed to react as if power were a rude, taboo topic. Who wants to study Machiavellian people who manipulate other people? A French advisee, however, was the first enthusiast, and power as asymmetrical control informed our studies of power from the bottom-up (Dépret and Fiske, 1999, as noted).

Then, finally, some other graduate students wanted to examine the powerholders themselves; that is, people vulnerable to stereotyping others whom they do not need for anything. Powerholders with more control indeed engage in more stereotyping than powerholders with less control, and trait dominance has similar effects (Goodwin et al., 2000). Individually dominant powerholders do not necessarily attune to the task-relevant competence of their subordinates, whereas low-dominant powerholders do (Operario and Fiske, 2001). Consistent with being clueless, the dominant powerholders are entertained and pleased by sociable subordinates, but they do not retain them for the task at hand.

Nearly a decade after the original published theory (Fiske and Neuberg, 1990) and almost two decades after the first published chapter (Fiske, 1982), we re-evaluated the CM (Fiske et al., 1999). Evidence for the model suggested that its premises had aged gracefully. Several motives moderate category-based and individuating processes, parsed as belonging, understanding, controlling, self-enhancing, and trusting (Stevens and Fiske, 1995). Various individual differences also predispose people toward one kind of process or the other. For example, need for structure encourages category use (Neuberg

and Newsom, 1993), whereas need for cognition encourages attribute-oriented processing (Cacioppo et al., 1996).

Two primary debates remained, both oriented to fundamental cognitive processes. One debate contrasted the CM’s apparently serial processes (first categorization, then attribute-based processes) with parallel, simultaneous processes together constraining outcomes (Kunda and Thagard, 1996). The original CM was agnostic about the underlying deep-cognitive processes. CM does claim that category processing typically will prove faster, but attribute processing could co-occur, just more slowly, so that its results appear later. Or category and attribute processing could occur serially, the latter only as needed. Most likely, because attributes and categories influence each other, parallel processes make sense. However, this deep cognitive modeling was not the CM’s main level of analysis (see Fiske et al., 1999, for more detail).

Another debate centered on the nature of the representations for category-based and attribute-based impressions. A contemporaneous dual-process impression model (Brewer, 1988), like the CM, posited two impression-formation processes, one more stereotypic, and the other more personal. Several factors differentiate the two models (Fiske, 1988), but the primary differences lie in Brewer’s model of branching decision-tree process, rather than a continuum, and in its proposal that distinct branches represent people in distinct forms (category, image, exemplar). Other differences include the CM’s explicit focus on attention and interpretation as mediators, motivation as a moderator, a common set of decision rules (information fit), and a common representational format. Brewer’s model, which posits branches rather than a continuum, proposes not only distinct representations after each decision point, but also distinct decision rules at each decision point (for more detail, see Brewer and Harasty Feinstein, 1999).

Two-mode models evidently ruled the turn of the century, as witnessed by the array of three dozen such models, too numerous to

review here but collected in an edited volume (Chaiken and Trope, 1999). The larger conceptual point is that the contrast between relatively automatic, rapid judgments ("System 1") and relatively thoughtful, slower judgments ("System 2") has permeated social cognition, attitudes, and more (see Fiske and Taylor, 2008, chapter 2, for an overview). Like other dual-process models, the CM does unify apparently separate processes, specifies conditions motivating each, and describes processes underlying both. The added value is moving from either-or to both-and.

The related theory of power-as-control generated its own controversies. For example, we had asserted (and shown) that powerholders, because of their noncontingency on others, are vulnerable to category-based processing, neglecting their subordinates (Goodwin et al., 2000). Just as our oft-presented power results got into print, other studies showed the boundaries of our predictions. For example, powerholders oriented toward social responsibility do individuate their subordinates (Overbeck and Park, 2001). Exchange-oriented powerholders stereotype, but communally oriented powerholders do not (Chen et al., 2001). The power literature then exploded with the theory of power as an approach-driven motivation. Consistent with our view but independent and much more comprehensive, a far-reaching theory showed that the powerful see the world as full of rewards, and powerholders cheerfully operate on automatic (Keltner et al., 2003). What's more, powerholders respond more constructively to what the situation affords them (Guinote, 2008), and they act on their goals (Galinsky et al., 2003). The review of this new power literature lies beyond the current chapter's scope (see Fiske, 2010), but some aspects trace back to the CM-based power-as-control theory.

Much remains to do. Future CM challenges include pursuing neural correlates of the two types of process (Lieberman et al., 2002). Early evidence indicated separate neuropsychological processes involved in generalizations (category use) and elaboration

(attribute use), differentially vulnerable to cognitive decline with aging (Mather et al., 1999). Recently we observed undergraduates' dual neural processes in the scanner, while they formed impressions of two partners, when outcome-dependent on one but not the other (Ames and Fiske, 2011). As before, they concentrated on the most informative, expectancy-inconsistent cues, but only for the partner on whom their outcomes depended. In this case only, the medial prefrontal cortex selectively activated—in an area independently identified as responsive to impression formation processes. Future research will further differentiate neural systems associated with the two kinds of impression processes.

On another promising note, in provocative early social neuroscience studies, novel other-race faces activated the amygdala, a brain region implicated in emotional vigilance (Hart et al., 2000; Phelps et al., 2000). These early returns encouraged us to pursue category-based and more individuating impressions via neuroimaging methods. After I moved to Princeton, graduate student Elizabeth Wheeler and I manipulated category-based and more individuating responses to novel cross-race yearbook photos. Based on the CM, we noted that the default conditions used in prior studies (categorizing by gender, for example) would tend to encourage category-based processing. Instead, we compared three conditions (Wheeler and Fiske, 2005): the default, categorical processing (report the pictured person's gender); a totally nonsocial condition (does the photo show a dot on the face?); and a newly-invented individuating condition (would the pictured person like a just-pictured vegetable?). Consistent with the CM, the category condition replicated the race-based amygdala activation, but neither the nonsocial nor the individuating condition did so. A conceptual replication with another outgroup (homeless people) showed that the individuating condition uniquely activated the medial prefrontal cortex (mPFC), a region implicated in social attributions (Harris and Fiske, 2006).

Describing categories: the stereotype content model

The CM privileges categories as capturing much impression formation, being the fastest, the default, and the anchoring process. If categories matter so much, which categories matter and why (Puzzle #7)? The stereotype content model (SCM) holds that, just as the processes of impression formation reflect systematic principles, so too do the contents of our categories. The SCM complements and extends the CM, in the intellectual narrative of theoretical development and empirical epic.

The SCM's fundamental insight is simple (for an overview, see Fiske et al., 2007; for detail, see Cuddy et al., 2008). The dimensions people use to categorize other people are predictable because the human problem is universal. People as perceptual objects mainly differ from nonhuman entities in having intention; that is, being autonomous agents (Fiske and Taylor, 2008: Chapter 1). So, other people need to know, first, whether another social entity (individual or group) intends the self good or ill. The sentry's query, "Friend or foe?" captures this dilemma. We call this dimension warmth (warm, friendly, trustworthy, honest). People judge this dimension of a face in less than a second (Willis and Todorov, 2006). The second question is whether the other can enact that intention: How able or unable? The competence dimension (competent, capable, skilled) also operates rapidly, only slightly more slowly than the first dimension.

Together, these two dimensions, under various names, account for 80 to 90 percent of the variance in individual impressions (Wojciszke, 2005). Solomon Asch (1946), again, was there first, but only intuitively, when his experiments held competence constant (*intelligent, skillful, industrious, determined, practical, cautious*) but contrasted warmth (*warm* versus *cold*). Fritz Heider's (1958) astute phenomenological analysis, as usual, identified the relevant factors; he described people inferring dispositions by

understanding what another person may *try* to do (goal, which includes warm or hostile intent) versus what the other *can* do (ability). On a similar but more empirical note, Seymour Rosenberg (Rosenberg and Sedlak, 1972) factor analyzed personality trait impressions, generating two dimensions: social good–bad (anchored by *warm* or *sociable* versus *unpopular*) and task good–bad (anchored by *foolish* versus *scientific*). Reflecting their breadth, similar dimensions of competence and integrity constitute impressions of presidential candidates (Abelson et al., 1982).

SCM Hypothesis 1 states that the two dimensions will differentiate societal groups into predictable quadrants. Indeed, common societal groups do spread out across the dimensions in the U.S. (Fiske et al., 2002; Cuddy et al., 2007), and all over the world (Cuddy et al., 2009). As Table 13.2 indicates, the dimensions combine to define universally recognizable stereotypes. The high-warmth/high-competence groups typically represent the ingroup and its allies, as well as societal reference groups, such as the middle class, heterosexuals, the dominant religion (e.g., in the US, Christianity), and the majority racial group. The outgroup quadrants include the lowest of the low, seen as neither warm nor competent: poor people all over the world, and in the US specifically, homeless people and drug addicts. They allegedly have no redeeming features. This much (high/high ingroups versus low/low outgroups) overlaps with standard depictions of intergroup relationships.

The SCM contributes the novel idea of mixed, ambivalent stereotypes, high on one dimension and low on the other. Some ambivalent outgroup quadrants include those high-warmth, low-competence groups (e.g., older people, disabled people), that are benign, harmless, but do not especially bear on ingroup goals. They may be likable, but not especially respected. Other ambivalent outgroups include those low-warmth, high-competence groups (e.g., rich people, outsider entrepreneurs), who earn grudging respect but also disliking.

Table 13.2 Stereotype content model

	<i>Low competence stereotypes from low status</i>	<i>High competence stereotypes from high status</i>
High warmth stereotypes from high cooperation		
Example	Old people	Middle-class people
Affect	Pity	Pride
Behavior (active ... and passive)	Active help Passive neglect	Active help Passive accommodation
Low warmth stereotypes from low cooperation		
Example	Homeless people	Rich people
Affect	Disgust	Envy
Behavior (active ... and passive)	Active harm Passive neglect	Active harm Passive accommodation

SCM Hypothesis 2 states that many or even most groups will locate in mixed quadrants. Across the world, the majority of societal groups do indeed fall into the two ambivalent clusters (Cuddy et al., 2008, 2009). However, a society's degree of ambivalence depends on its degree of inequality: the larger its income gap, the more the society apparently has to justify the divisions between its groups, so more of them are favored on one dimension (e.g., rich people may be competent) but not the other (rich people are cold) (Durante et al., submitted).

The SCM posits that stereotype contents are systematic not only by basic dimensions but also by their antecedents and consequences: *Social structure causes perceived stereotypes, which in turn generate affect, which in its turn generates behavior.* Having described the stereotype contents, what are their antecedents? The stereotypes' origins prove predictable. Social structure predicts perceived traits, based on two principles.

SCM Hypothesis 3 states that perceived competition lowers perceived warmth, and perceived status raises perceived competence. SCM predicts the competition-warmth link from a previous analysis of why people feel negative toward outgroups (Fiske and Ruscher, 1993). Negative affect toward most outgroups follows from the inherent goal conflict: the group is an outgroup precisely because their goals do not support or may even conflict with ingroup goals. Hence, they

cannot be trusted, seen as having malignant or at least not benign intent. The competition-warmth link averages a small but reliable effect across American and European samples (Cuddy et al., 2008). With improved measurement of both competition (to include value conflict as well as resource conflict) and warmth (to include not only sociability but also trustworthiness), consistent with the SCM, the correlations prove substantial (Kervyn et al., submitted).

On the other dimension, the status-competence link follows from dispositional attributions for a high-status position being due to the ability of the incumbents. This also follows from a just-world belief that groups get what they deserve, and indeed those higher on belief in a just world show even stronger status-competence correlations than people lower on this dimension (Oldmeadow and Fiske, 2007). Around the world, the status-competence correlation is reliable and averages large in size (Cuddy et al., 2008).

Turning to the consequences of stereotype content, according to *SCM hypothesis 4, each cluster provokes predictable emotional prejudices* (see Table 13.2). In the US and Europe, the supposedly both warm and competent ingroup and its allies typically elicit pride because they are responsible for positive outcomes that reflect well on the self. The low-warmth/low-competence groups are allegedly responsible for their own negative outcomes; they elicit disgust, an unadulterated negative

emotion applicable to contaminated objects as well as people.

The warm but incompetent outgroups (e.g., the elderly) elicit pity, because they have negative outcomes that are not their fault. In contrast, the competent but cold outgroups (e.g., rich people) elicit envy because they have positive outcomes that do not favor the ingroup. Indeed, the ingroup may resent them for their good fortune that relatively deprives the ingroup. Both pity and envy are mixed emotions, so these two quadrants qualify as ambivalent on two counts: stereotype content that is high on one fundamental dimension and low on the other, plus mixed emotions.

Besides social structure predicting correspondent stereotypes and both predicting unique emotional prejudices, *SCM Hypothesis 5 predicts behavioral tendencies distinct to each quadrant* (see Table 13.2; Cuddy et al., 2007). We hypothesized that warmth, being primary, would predict active responses to facilitate or interfere with those respectively allied or opposed to ingroup interests. Thus, high-warmth groups obtain active help and protection; low-warmth groups elicit active attack and fight. Both ingroup and pitied outgroups receive help and protection, for example, because both are stereotyped as warm.

The competence dimension, being secondary, should elicit passive facilitation (for competent groups) and passive interference (for incompetent groups). Thus, both ingroup and envied outgroups, both being stereotypically competent, elicit passive association, going-along-to-get-along, while disgusting outgroups and pitied outgroups, both being stereotypically incompetent, receive neglect and exclusion. A national random sample survey (Cuddy et al., 2007) supports all these predictions. For example, the ambivalently stereotyped pitied outgroups receive both active protection and passive neglect (e.g., institutionalizing the disabled or elderly). Another ambivalently stereotyped, but instead envied, cluster receives both passive association (e.g., shopping at their stores) but also active harm, at least under some

circumstances (e.g., looting and attack, under social breakdown). Across quadrants, emotions are proximate cause of behavior.

Sorting groups along these two dimensions, warmth and competence, arguably begins to look universal. If so, the historical stereotypes should show similar patterns, and they do. We replicated the SCM using recoded Princeton studies of ethnic and national stereotypes since the 1930s (Bergsieker et al., submitted). Furthermore, content coding of fascist magazine descriptions of ethnic and national groups also fits the SCM space (Durante et al., 2009).

Some national and historical variations do occur. For example, the high-high quadrant, normally reserved for the ingroup and its allies, is relatively vacant in three East Asian samples (Cuddy et al., 2009). Instead, various ingroups migrate to a more modest, middling position, in keeping with cultural norms for modesty. And among the fascists, the pity quadrant is empty, a finding that speaks for itself.

The SCM subsumes and explains particular ambivalent prejudices. For example, Peter Glick and I had previously hypothesized that sexism is not simple antipathy but instead ambivalent (Glick and Fiske, 1996). Although male status is higher in all societies, men and women universally depend on each other to procreate, as well as to form lasting bonds. This combined status difference and obligatory cooperation generates a default stereotype of traditional women, termed benevolent sexism (paternalism). The more expected form, hostile sexism, targets nontraditional women as admittedly competent but cold (e.g., lesbians, feminists, career women). Although ambivalent sexism theory (AST) predates the SCM, AST emerged from a social structural analysis of gender prejudice that anticipated the SCM's focus on cooperation-competition and societal status.

Another source of inspiration for AST were the Katz and Hass (1988) studies of racial ambivalence, which inspired us to consider the possibility that sexism too could be ambivalent, though for different reasons.

Regardless of its origins, the AST patterns occur worldwide, consistent with their origins in the fundamental relationships between men and women (Glick et al., 2000). In SCM terms, traditional women often inhabit the high-warmth, low-competence (pitied) quadrant, while nontraditional women inhabit the low-warmth, high-competence envied quadrant.

Other groups occupy combinations of quadrants, differentiated by subtypes. For example, stereotypes of black Americans differentiate by social class into black professionals, who make others proud, versus poor blacks, who disgust others (Fiske et al., 2009). Stereotypes of gay men fall into all quadrants, depending on subtype (e.g., in-the-closet, gay artist, leather-biker, cross-dresser; Clausell and Fiske, 2005). And although the default stereotype of old people is high-warmth, low-competence (Cuddy et al., 2005), subtypes of older people may try to escape, forfeiting protective pity for envious resentment directed at those who hoard their wealth, disgust at those who selfishly consume societal resources, or resistance toward those who attempt to invade younger ingroup identities (North and Fiske, submitted). Particular prejudices thus come into focus, viewed through the SCM lens. For example, anti-Asian prejudice fits the patterns of the envy quadrant (Lin et al., 2005). Dehumanization of homeless people fits the patterns of the disgust quadrant (Harris and Fiske, 2006). Envy of investment bankers and schadenfreude at their misfortunes fits the pattern of the envy quadrant (Cikara and Fiske, submitted). Likewise, the fans of mutually envied sports rivals (e.g., Yankees versus Red Sox) experience schadenfreude at their rivals losses, even to a third party (Cikara et al., 2011). Altogether, the data include self-report questionnaires, individual differences, neural patterns, and subtle facial expressions indicated by electromyography.

The original SCM research initially assessed perceived societal responses to various groups, leaving it open to criticism that

it reflected merely ideology. And all the evidence was correlational, raising questions about the proposed causal relationships. Subsequent work revealed the SCM to be both personal and causal. Experimental vignette studies showed that manipulating the social structures of interdependence and status does result in the predicted warmth and competence stereotypes (Caprariello et al., 2009). Laboratory analog experiments manipulating interpersonal interdependence and status showed that the structure of dyadic relationships also predicts personal warmth and competence expectancies (Russell and Fiske, 2008). Finally, people report personal emotional reactions to images of groups from the different quadrants, in line with SCM predictions (Harris and Fiske, 2006).

Neuroimaging data show emotional reactions consistent with self-reports so far for two quadrants, the first being disgust toward drug addicts and homeless people; insula activation typically characterizes disgust and in our data insula differentially activates to those images (Harris and Fiske, 2006). Second, the schadenfreude toward envied sports rivals appears in ventral striatum reward-area activation that correlates with self-reported harm to their fans (Cikara et al., 2011).

Future research will focus on patterns related to specific social groups, such as higher and lower classes, especially interpersonal effects of social structure, such as trust, entitlement, and imposter feelings. Where possible, we will pursue psycho-physiological as well as self-report indicators.

Application to social issues

Both the CM and the SCM instantly bolted from the ivory tower to risk real-world application. The CM particularly proved useful to discrimination lawsuits, in which I served as an expert witness. In testimony, the CM helped explain the *social framework* of bias in organizations (Fiske and Borgida, 2008), particularly how decision makers' relatively

automatic categories could lead them to discriminate in cool, cognitive, categorical ways, without any emotional animus:

- A female welder complained about workplace pornography that encouraged sexual harassment of the few women employees (*Robinson v. Jacksonville Shipyards, Inc.*¹). The prominently displayed sexualized media both primed objectified categories for women and trivialized their distress when they reported the resulting problems.
- Female applicants for jobs at a home-improvement franchise were routinely routed to dead-end cashier jobs, while comparably qualified male applicants were sent to the sales floor, on their way to bonuses and promotion (*Butler et al. and Frank et al. v. The Home Depot, Inc.*²). Interviewers categorized most men and few women as competent in construction, undervaluing the self-selected female candidates' knowledge of home renovation,
- Evaluators demanded that a female manager behave like a more stereotypic woman (*Hopkins v. Price Waterhouse*³). Although the top earner and expert in her cohort, she was advised to walk, talk, and dress in more feminine ways if she wanted to make partner.

The CM proves useful because lay people (including judges and juries) do not expect that categories, without active antipathy, could cause discrimination. For example, accounting executive Ann Hopkins was a top producer at Price Waterhouse, but she was denied partnership because of alleged interpersonal skills deficits that stemmed from being insufficiently feminine. Essentially, she made people uncomfortable because she did not fit her stereotypic category. Psychological science, including the CM, helped frame judges' understanding, all the way to the Supreme Court (Fiske et al., 1991).

The testimony in turn affected the theories. In preparing expert reports, I noticed a misfit between the available social psychological literature, which studies peers' stereotypic judgments, and real-world organizations, in which powerful decision makers' stereotypic judgments matter most. The power-as-control theory (Fiske, 1993) directly

resulted from expert testimony experiences, in interplay with the CM interdependence hypotheses. Asymmetrical control over outcomes (power) should make people vulnerable to stereotyping their subordinates, as noted earlier.

The SCM also has applied to several social issues. The potential for people to dehumanize some types of outgroups informed a *Science Policy Forum* article about the psychological dynamics of prisoner abuse at Abu Ghraib (Fiske et al., 2004). The SCM-related ambivalent sexism theory has informed expert testimony by several social psychologists (Rudman et al., 2008). And the SCM-predicted role of emotions in mediating the impact of stereotypes on behavior has informed a meta-analysis of how best to predict racial discrimination (Talaska et al., 2008).

Relevant to all the theories described here, common sense underestimates the automaticity, ambivalence, and ambiguity of prejudice. The legal implications argue against judges and juries relying on their lay psychological theories, for example, that only conscious animus causes discrimination. A *behavioral realism* approach argues for the relevance of established psychological science in the courtroom, for example, in employment law (Krieger and Fiske, 2006). Psychological theories, supported by evidence, can inform policies to avert discrimination, by suggesting organizational remedies, such as transparency about an organization's success in promoting under-represented groups (Fiske and Krieger, in press).

Social cognition theories, such as the CM and SCM, provide expertise in employment cases, particularly drawing on the high-quality science establishing (1) dual processes (automatic and controlled), (2) early bias (at attention and encoding), and (3) mental construal (interpretation) that creates category-based perception (Fiske and Borgida, 2008). The relevance of the CM is clear to all three applications, but also, respectively, are the Implicit Association Test regarding automaticity (e.g., Kang and Banaji, 2006), the

social neuroscience of prejudice at the earliest stages of encoding (e.g., Phelps et al., 2000), and the tenacity of categories despite individuating information (e.g., Heilman and Haynes, 2008). Further, a variety of psychological science usefully informs legal practice and policy, beyond common sense (Borgida and Fiske, 2008).

CONCLUSION

According to the opening personal narrative, the two theories of social category process and content – CM and SCM – have roots in the puzzles of lived experience. People are captured by categories, and consequently my work has been captivated by categories. Many theories explain the puzzles narrated earlier, but here, consider how CM and SCM explain everyday life.

Why did my by-chance dance partner in high school need the additional category of knowing my astrological sign? Because, as the CM predicts, he was motivated to go beyond gender, age, and race, and the new information evidently recategorized his impression (Puzzle #1 solved). Why did all the kids sort themselves by ethnicity in the middle-school cafeteria? Because, as the CM predicts, they were relying on easy categories, and as the SCM predicts, their ingroup elicited pride, thereby avoiding pity, envy, or disgust they might feel toward other cliques (Puzzle #2 solved).

Why are some communities, such as Hyde Park, able to integrate, while others remain segregated? At the social psychological level, part of the success is predicted by the contact hypothesis (Allport, 1954), which names interdependence as a key condition. The CM shows that cooperation makes people attend to each others' counter-stereotypic features, enabling them to overcome their categories and work together toward common goals (Puzzle #3 solved). The CM also identifies personal values, certainly relevant to that community that prizes its multicultural character.

Why did the college doorman think I was Kathleen Kennedy, based only on demographics? As the CM predicts, he was not motivated to go beyond superficial categories. And maybe some wishful thinking entered in. Puzzle #4 yields to the primacy of visual demographic categories, as predicted by the CM.

Why are likability and memory for the original evidence so separate, as in my ability to recall my first students' likability more than their identity? As the CM predicts, people form online impressions resulting in affective tags that persist beyond the original inferences. Puzzle #5 solved.

And finally, regarding the CM, how can one reconcile both category-based and attribute-based processes, as the Ohio State graduate student's question that generated Puzzle #6? Motivation and information both decide which process operates when.

As for the SCM, Puzzle #7 asked whether stereotype contents are systematic, and indeed, they are, along two fundamental dimensions of warmth and competence, with social structural predictors, associated affect, and tendencies toward behavior.

At this point, several puzzles have yielded, and some theories have developed with collaborators who inspire equal doses of curiosity, insight, persistence, and delight. Together, we faced many obstacles and dead ends, but building theory is not a solo sprint, it takes a marathon team. And that's our favorite category.

ACKNOWLEDGMENTS

The author acknowledges Russell Sage Foundation for support during the writing of this chapter and Lydia Emery for comments on a prior draft.

NOTES

¹ *Robinson v. Jacksonville Shipyards, Inc.* [M.D. Fla. 1989]; Case No. 86-927.

2 Butler et al. and Frank et al. v. The Home Depot, Inc. [1994, 1995]. US District Court, N. District of California, C 94-4335 SI and C 95-2182 SI.

3 Hopkins v. Price Waterhouse, 618 F. Supp. 1109 [D.D.C. 1985]; appeal: Price Waterhouse v. Hopkins, 825 F.2d 458 [D.C. Cir. 1987]; Supreme Court review: Price Waterhouse v. Hopkins, 109 S. Ct. 1775 [1989]; remand: Hopkins v. Price Waterhouse, No. 84-3040, slip op. [D.D.C. May 14, 1990].

REFERENCES

Abelson, R.P., Kinder, D.R., Peters, M.D. and Fiske, S.T. (1982) Affective and semantic components in political person perception. *Journal of Personality and Social Psychology*, 42, 619–630.

Allport, G.W. (1954) *The Nature of Prejudice*. Reading, MA: Addison-Wesley.

Ames, D.L. and Fiske, S.T. (2011) Outcome dependency alters the neural substrates of impression formation. Unpublished manuscript.

Andersen, S.M. and Chen, S. (2002) The relational self: An interpersonal social-cognitive theory. *Psychological Review*, 109, 619–645.

Anderson, N.H. (1981) *Foundations of Information Integration Theory*. New York: Academic Press.

Asch, S.E. (1946) Forming impressions of personality. *Journal of Abnormal and Social Psychology*, 41, 1230–1240.

Bergsieber, H.B., Leslie, L.M., Constantine, V.S. and Fiske, S.T. (submitted) Stereotyping by omission: Eliminate the negative, accentuate the positive.

Borgida, E. and Fiske, S.T. (eds) (2008) *Psychological Science in the Courtroom: Beyond Common Sense*. London: Blackwell.

Boring, E.G. (1950) *A History of Experimental Psychology*, 2nd Edition. Englewood Cliffs, NJ: Prentice-Hall.

Brewer, M.B. (1988) A dual process model of impression formation. In T.K. Srull and R.S. Wyer Jr. (eds), *Advances in Social Cognition*, 1, 1–36. Hillsdale, NJ: Erlbaum.

Brewer, M.B. and Harasty Feinstein, A.S. (1999) Dual processes in the cognitive representation of persons and social categories. In S. Chaiken and Y. Trope (eds), *Dual-process Theories in Social Psychology*, pp. 255–270. New York: Guilford Press.

Cacioppo, J.T., Petty, R.E., Feinstein, J.A. and Jarvis, W.B.G. (1996) Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin*, 119, 197–253.

Cantor, N. and Mischel, W. (1977) Traits as prototypes: Effects on recognition memory. *Journal of Personality and Social Psychology*, 35, 38–48.

Caprariello, P.A., Cuddy, J.C. and Fiske, S.T. (2009) Social structure shapes cultural stereotypes and emotions: A causal test of the stereotype content model. *Group Processes and Intergroup Behavior*, 12, 147–155.

Chaiken, S. and Trope, Y. (eds) (1999) *Dual-process Theories in Social Psychology*. New York: Guilford Press.

Chen, S., Lee-Chai, A.Y. and Bargh, J.A. (2001) Relationship orientation as a moderator of the effects of social power. *Journal of Personality and Social Psychology*, 80, 173–187.

Cikara, M., Botvinick, M.M. and Fiske, S.T. (2011) Us versus them: Neural responses to rival groups' pleasures and pains. *Psychological Science*, 22, 306–313.

Cikara, M. and Fiske, S.T. (submitted) Stereotypes and Schadenfreude: Behavioral and physiological markers of pleasure at others' misfortunes. Unpublished manuscript.

Clark, M.S. and Fiske, S.T. (eds) (1982) *Affect and Cognition: The 17th Annual Carnegie Symposium on Cognition*. Hillsdale, NJ: Erlbaum.

Clausell, E. and Fiske, S.T. (2005) When do the parts add up to the whole? Ambivalent stereotype content for gay male subgroups. *Social Cognition*, 23, 157–176.

Cuddy, A.J.C., Fiske, S.T. and Glick, P. (2007) The BIAS map: Behaviors from intergroup affect and stereotypes. *Journal of Personality and Social Psychology*, 92, 631–648.

Cuddy, A.J.C., Fiske, S.T. and Glick, P. (2008) Competence and warmth as universal trait dimensions of interpersonal and intergroup perception: The Stereotype Content Model and the BIAS Map. In M.P. Zanna (ed.), *Advances in Experimental Social Psychology*, 40, 61–149. New York: Academic Press.

Cuddy, A.J.C., Fiske, S.T., Kwan, V.S.Y., Glick, P., Demoulin, S., Leyens, J.-Ph., Bond, M.H., Croizet, J.-C., Ellemers, N., Sleebos, E., Htun, T.T., Yamamoto, M., Kim, H.-J., Maio, G., Perry, J., Petkova, K., Todorov, V., Rodríguez-Bailón, R., Morales, E., Moya, M., Palacios, M., Smith, V., Perez, R., Vala, J. and Ziegler, R. (2009) Is the stereotype content model culture-bound? A cross-cultural comparison reveals systematic similarities and differences. *British Journal of Social Psychology*, 48, 1–33.

Cuddy, A.J.C., Norton, M.I. and Fiske, S.T. (2005) This old stereotype: The pervasiveness and persistence of the elderly stereotype. *Journal of Social Issues*, 61, 265–283.

Dépret, E.F. and Fiske, S.T. (1999) Perceiving the powerful: Intriguing individuals versus threatening groups. *Journal of Experimental Social Psychology*, 35, 461–480.

Dreben, E.K., Fiske, S.T. and Hastie, R. (1979) Impression and recall order effects in behavior-based impression formation. *Journal of Personality and Social Psychology*, 37, 1758–1768.

Durante, F., Fiske, S.T., Cuddy, A.J.C. and Kervyn, N. (submitted) Inequality predicts ambivalence in stereotype content. Unpublished manuscript.

Durante, F., Volpato, C. and Fiske, S.T. (2009) Using the Stereotype Content Model to examine group depictions in fascism: An archival approach. *European Journal of Social Psychology*, 39, 1–19.

Erber, R. and Fiske, S.T. (1984) Outcome dependency and attention to inconsistent information. *Journal of Personality and Social Psychology*, 47, 709–726.

Fiske, S.T. (1980) Attention and weight in person perception: The impact of negative and extreme behavior. *Journal of Personality and Social Psychology*, 38, 889–906.

Fiske, S.T. (1982) Schema-triggered affect: Applications to social perception. In M.S. Clark and S.T. Fiske (eds), *Affect and Cognition: The 17th Annual Carnegie Symposium on Cognition*, pp. 55–78. Hillsdale, NJ: Erlbaum.

Fiske, S.T. (1988) Compare and contrast: Brewer's dual-process model and Fiske et al.'s continuum model. In T.K. Srull and R.S. Wyer (eds), *Advances in Social Cognition*, Vol. 1: A Dual Model of Impression Formation, pp. 65–76. Hillsdale, NJ: Erlbaum.

Fiske, S.T. (1993) Controlling other people: The impact of power on stereotyping. *American Psychologist*, 48, 621–628.

Fiske, S.T. (2010) Stratification. In S.T. Fiske, D.T., Gilbert and G. Lindzey (eds), *Handbook of Social Psychology*, 5th Edition. New York: Wiley.

Fiske, S.T., Bergsieker, H., Russell, A.M. and Williams, L. (2009) Images of Black Americans: Then, 'them' and now, 'Obama!' *DuBois Review: Social Science Research on Race*, 6, 83–101.

Fiske, S.T., Bersoff, D.N., Borgida, E., Deaux, K. and Heilman, M.E. (1991) Social science research on trial: The use of sex stereotyping research in *Price Waterhouse v. Hopkins*. *American Psychologist*, 46, 1049–1060.

Fiske, S.T. and Borgida, E. (2008) Providing expert knowledge in an adversarial context: Social cognitive science in employment discrimination cases. *Annual Review of Law and Social Science*, 4, 123–148.

Fiske, S.T., Cuddy, A.J.C. and Glick, P. (2007) Universal dimensions of social perception: Warmth and competence. *Trends in Cognitive Science*, 11, 77–83.

Fiske, S.T., Cuddy, A.J.C., Glick, P. and Xu, J. (2002) A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, 82, 878–902.

Fiske, S.T., Harris, L.T. and Cuddy, A.J.C. (2004) Policy Forum: Why ordinary people torture enemy prisoners. *Science*, 306, 1482–1483.

Fiske, S.T. and Krieger, L.H. (in press) Policy implications of unexamined discrimination: Gender bias in employment as a case study. In E. Shafir (ed.), *Behavioral Foundations of Policy*. Princeton: Princeton University Press and New York: Russell Sage.

Fiske, S.T., Lin, M.H. and Neuberg, S.L. (1999) The Continuum Model: Ten years later. In S. Chaiken and Y. Trope (eds), *Dual-process Theories in Social Psychology*, pp. 231–254. New York: Guilford.

Fiske, S.T. and Neuberg, S.L. (1990) A continuum model of impression formation, from category-based to individuating processes: Influence of information and motivation on attention and interpretation. In M.P. Zanna (ed.), *Advances in Experimental Social Psychology*, 23, 1–74. New York: Academic Press.

Fiske, S.T., Neuberg, S.L., Beattie, A.E. and Milberg, S.J. (1987) Category-based and attribute-based reactions to others: Some informational conditions of stereotyping and individuating processes. *Journal of Experimental Social Psychology*, 23, 399–427.

Fiske, S.T. and Pavelchak, M.A. (1986) Category-based versus piecemeal-based affective responses: Developments in schema-triggered affect. In R.M. Sorrentino and E.T. Higgins (eds), *Handbook of Motivation and Cognition: Foundations of Social Behavior*, pp. 167–203. New York: Guilford Press.

Fiske, S.T. and Ruscher, J.B. (1993) Negative interdependence and prejudice: Whence the affect? In D.M. Mackie and D.L. Hamilton (eds), *Affect, Cognition, and Stereotyping: Interactive Processes in Group Perception*, pp. 239–268. New York: Academic Press.

Fiske, S.T. and Taylor, S.E. (2008) *Social Cognition: From Brains to Culture*. New York: McGraw-Hill.

Fiske, S.T. and Von Hendy, H.M. (1992) Personality feedback and situational norms can control stereotyping processes. *Journal of Personality and Social Psychology*, 62, 577–596.

Galinsky, A.D., Gruenfeld, D.H. and Magee, J.C. (2003) From power to action. *Journal of Personality and Social Psychology*, 85, 453–466.

Glick, P. and Fiske, S.T. (1996) The Ambivalent Sexism Inventory: Differentiating hostile and benevolent sexism. *Journal of Personality and Social Psychology*, 70, 491–512.

Glick, P., Fiske, S.T., Mladinic, A., Saiz, J.L., Abrams, D., Masser, B., et al. (2000) Beyond prejudice as simple antipathy: Hostile and benevolent sexism across cultures. *Journal of Personality and Social Psychology*, 79, 763–775.

Goodwin, S.A., Fiske, S.T., Rosen, L.D. and Rosenthal, A.M. (2002) The eye of the beholder: Romantic goals and impression biases. *Journal of Experimental Social Psychology*, 38, 232–241.

Goodwin, S.A., Gubin, A., Fiske, S.T. and Yzerbyt, V. (2000) Power can bias impression formation: Stereotyping subordinates by default and by design. *Group Processes and Intergroup Relations*, 3, 227–256.

Guinote, A. (2008) Power and affordances: When the situation has more power over powerful than powerless individuals. *Journal of Personality and Social Psychology*, 95, 237–252.

Guinote, A., Brown, M. and Fiske, S.T. (2006) Minority status decreases sense of control and increases interpretive processing. *Social Cognition*, 24, 169–186.

Guinote, A. and Fiske, S.T. (2003) Being in the out-group territory increases stereotypic perceptions of outgroups: Situational sources of category activation. *Group Processes and Intergroup Relations*, 6, 323–331.

Guthrie, E.R. (1935) *The Psychology of Learning*. Oxford: Harper.

Hamilton, D.L. and Zanna, M.P. (1974) Context effects in impression formation: Changes in connotative meaning. *Journal of Personality and Social Psychology*, 29, 649–654.

Harris, L.T. and Fiske, S.T. (2006) Dehumanizing the lowest of the low: Neuro-imaging responses to extreme outgroups. *Psychological Science*, 17, 847–853.

Hart, A.J., Whalen, P.J., Shin, L.M., McInerney, S.C., Fischer, H. and Rauch, S.L. (2000) Differential response in the human amygdala to racial outgroup vs ingroup face stimuli. *Neuroreport: For Rapid Communication of Neuroscience Research*, 11, 2351–2355.

Hastie, R., Ostrom, T.M., Ebbesen, E.B., Wyer Jr., R.S., Hamilton, D.L. and Carlston, D.E. (eds) (1980) *Person Memory: The Cognitive Basis of Social Perception*. Hillsdale, NJ: Erlbaum.

Hastie, R. and Park, B. (1986) The relationship between memory and judgment depends on whether the judgment task is memory-based or on-line. *Psychological Review*, 93, 258–268.

Heider, F. (1958) *The Psychology of Interpersonal Relations*. Hoboken, NJ: Wiley.

Heilman, M.E. and Haynes, M.C. (2008) Subjectivity in the appraisal process: A facilitator of gender bias in work settings. In E. Borgida and S.T. Fiske (eds), *Psychological Science in the Courtroom: Beyond Common Sense*, pp. 127–156. London: Blackwell.

Kang, J. and Banaji, M. (2006) Fair measures: A behavioral realist revision of 'affirmative action'. *California Law Review*, 94, 1063–1118.

Katz, I. and Hass, R.G. (1988) Racial ambivalence and American value conflict: Correlational and priming studies of dual cognitive structures. *Journal of Personality and Social Psychology*, 55, 893–905.

Keltner, D., Gruenfeld, D.H. and Anderson, C. (2003) Power, approach, and inhibition. *Psychological Review*, 110, 265–284.

Kervyn, N., Fiske, S.T. and Yzerbyt, V. (submitted) Why is the primary dimension of social cognition so hard to predict? Symbolic and realistic threats together predict warmth in the stereotype content model. Unpublished manuscript.

Krieger, L.H. and Fiske, S.T. (2006) Behavioral realism in employment discrimination law: Implicit bias and disparate treatment. *California Law Review*, 94, 997–1062.

Kunda, Z. and Thagard, P. (1996) Forming impressions from stereotypes, traits, and behaviors: A parallel-constraint-satisfaction theory. *Psychological Review*, 103, 284–308.

Lieberman, M.D., Gaunt, R., Gilbert, D.T. and Trope, Y. (2002) Reflexion and reflection: A social cognitive neuroscience approach to attributional inference. In M.P. Zanna (ed.), *Advances in Experimental Social Psychology*, 34, 199–249. San Diego: Academic Press.

Lin, M.H., Kwan, V.S.Y., Cheung, A. and Fiske, S.T. (2005) Stereotype content model explains prejudice for an envied outgroup: Scale of Anti-Asian American Stereotypes. *Personality and Social Psychology Bulletin*, 31, 34–47.

Markus, H. (1977) Self-schemata and processing information about the self. *Journal of Personality and Social Psychology*, 35, 63–78.

Mather, M., Johnson, M.K. and De Leonardis, D.M. (1999) Stereotype reliance in source monitoring: Age differences and neuropsychological test correlates. *Cognitive Neuropsychology*, 16, 437–458.

Neuberg, S.L. (1989) The goal of forming accurate impressions during social interactions: Attenuating the impact of negative expectancies. *Journal of Personality and Social Psychology, 56*, 374–386.

Neuberg, S.L. and Fiske, S.T. (1987) Motivational influences on impression formation: Outcome dependency, accuracy-driven attention, and individuating processes. *Journal of Personality and Social Psychology, 53*, 431–444.

Neuberg, S.L. and Newsom, J.T. (1993) Personal need for structure: Individual differences in the desire for simpler structure. *Journal of Personality and Social Psychology, 65*, 113–131.

North, M.S. and Fiske, S.T. (submitted) A new approach to ageism: Succession, identity, and consumption.

Oldmeadow, J. and Fiske, S.T. (2007) Ideology moderates status = competence stereotypes: Roles for belief in a just world and social dominance orientation. *European Journal of Social Psychology, 37*, 1135–1148.

Operario, D. and Fiske, S.T. (2001) Effects of trait dominance on powerholders' judgments of subordinates. *Social Cognition, 19*, 161–180.

Ostrom, T.M. (1977) Between-theory and within-theory conflict in explaining context effects in impression formation. *Journal of Experimental Social Psychology, 13*, 492–503.

Overbeck, J.R. and Park, B. (2001) When power does not corrupt: Superior individuation processes among powerful perceivers. *Journal of Personality and Social Psychology, 81*, 549–565.

Pavelchak, M.A. (1989) Piecemeal and category-based evaluation: An idiographic analysis. *Journal of Personality and Social Psychology, 56*, 354–363.

Phelps, E.A., O'Connor, K.J., Cunningham, W.A., Funayama, E.S., Gatenby, J.C., Gore, J.C. and Banaji, M.R. (2000) Performance on indirect measures of race evaluation predicts amygdala activation. *Journal of Cognitive Neuroscience, 12*, 729–738.

Rosenberg, S. and Sedlak, A. (1972) Structural representations of implicit personality theory. In L. Berkowitz (ed.), *Advances in Experimental Social Psychology, 6*, 235–297. New York: Academic Press.

Rudman, L.A., Glick, P. and Phelan, J.E. (2008) From the laboratory to the bench: Gender stereotyping research in the courtroom. In E. Borgida and S.T. Fiske (eds), *Psychological Science in the Courtroom: Beyond Common Sense*, pp. 83–103. London: Blackwell.

Ruscher, J.B. and Fiske, S.T. (1990) Interpersonal competition can cause individuating processes. *Journal of Personality and Social Psychology, 58*, 832–843.

Ruscher, J.B., Fiske, S.T., Miki, H. and Van Manen, S. (1991) Individuating processes in competition: Interpersonal versus intergroup. *Personality and Social Psychology Bulletin, 17*, 595–605.

Russell, A.M. and Fiske, S.T. (2008) It's all relative: Social position and interpersonal perception. *European Journal of Social Psychology, 38*, 1193–1201.

Snyder, M. and Swann, W.B. (1978) Behavioral confirmation in social interaction: From social perception to social reality. *Journal of Experimental Social Psychology, 14*, 148–162.

Slrull, T.K. (1981) Person memory: Some tests of associative storage and retrieval models. *Journal of Experimental Psychology: Human Learning and Memory, 7*, 440–462.

Stangor, C. and McMillan, D. (1992) Memory for expectancy-congruent and expectancy-incongruent information: A review of the social and social developmental literatures. *Psychological Bulletin, 111*, 42–61.

Stevens, L.E. and Fiske, S.T. (1995) Motivation and cognition in social life: A social survival perspective. *Social Cognition, 13*, 189–214.

Stevens, L.E. and Fiske, S.T. (2000) Motivated impressions of a powerholder: Accuracy under task dependency and misperception under evaluative dependency. *Personality and Social Psychology Bulletin, 26*, 907–922.

Talaska, C.A., Fiske, S.T. and Chaiken, S. (2008) Legitimizing racial discrimination: A meta-analysis of the racial attitude-behavior literature shows that emotions, not beliefs, best predict discrimination. *Social Justice Research: Social Power in Action, 21*, 263–296.

Taylor, S.E. and Fiske, S.T. (1978) Salience, attention, and attribution: Top-of-the-head phenomena. In L. Berkowitz (ed.), *Advances in Experimental Social Psychology, 11*, 249–288. New York: Academic Press.

Taylor, S.E., Fiske, S.T., Etcoff, N.L. and Ruderman, A.J. (1978) Categorical and contextual bases of person memory and stereotyping. *Journal of Personality and Social Psychology, 36*, 778–793.

Wheeler, M.E. and Fiske, S.T. (2005) Controlling racial prejudice and stereotyping: Social cognitive goals affect amygdala and stereotype activation. *Psychological Science, 16*, 56–63.

Willis, J. and Todorov, A. (2006) First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science, 17*, 592–598.

Wojciszke, B. (2005) Morality and competence in person- and self-perception. *European Review of Social Psychology, 16*, 155–188.

Feelings-as-Information Theory

Norbert Schwarz

ABSTRACT

Feelings-as-information theory conceptualizes the role of subjective experiences – including moods, emotions, metacognitive experiences, and bodily sensations – in judgment. It assumes that people attend to their feelings as a source of information, with different feelings providing different types of information. Whereas feelings elicited by the target of judgment provide valid information, feelings that are due to an unrelated influence can lead us astray. The use of feelings as a source of information follows the same principles as the use of any other information. Most important, people do not rely on their feelings when they (correctly or incorrectly) attribute them to another source, thus undermining their informational value for the task at hand. What people conclude from a given feeling depends on the epistemic question on which they bring it to bear; hence, inferences from feelings are context-sensitive and malleable. In addition to serving as a basis of judgment, feelings inform us about the nature of our current situation and our thought processes are tuned to meet situational requirements. The chapter reviews the development of the theory, its core propositions and representative findings.

INTRODUCTION

Human thinking is accompanied by a variety of subjective experiences, including moods and emotions, metacognitive feelings (like ease of recall or fluency of perception), and bodily sensations. Feelings-as-information theory provides a general framework for conceptualizing the role of these experiences in human judgment. It was initially developed to account for the influence of happy and sad moods on evaluative judgment. However, the theoretical principles of the initial mood-as-information work (Schwarz and Clore, 1983) could be fruitfully applied to other types of feelings and developed into a more comprehensive conceptualization of the interplay of feeling and thinking. This chapter summarizes what has been learned.

The first section provides a short personal account of the theory's development and places its assumptions in their historical context (for more detailed discussions of other theoretical approaches, see Clore et al., 1994; Schwarz and Clore, 2007). The second section

presents the theory's postulates and the third section reviews representative findings.

A LOOK BACK: THE DEVELOPMENT OF FEELINGS-AS- INFORMATION THEORY

As many readers know from personal experience, our lives look better on some days than on others, even though nothing of any obvious importance has changed. In my case, the upbeat or gloomy mood induced by a sunny or rainy day is sufficient to do the trick. Trying to understand this experience as a graduate student at the University of Mannheim, Germany, in the late 1970s, I turned to what social and cognitive psychologists had learned from the experimental mood research available at the time. One account, advanced by Isen et al. (1978) and Bower (1981), held that moods increase the accessibility of mood-congruent information in memory. From this perspective, positive (negative) aspects of life are more likely to come to mind when we are in a happy (sad) mood, resulting in mood-congruent judgments. This approach was consistent with social psychology's new adoption of the information processing paradigm and its emphasis on storage and retrieval processes. However, it didn't seem "quite right" introspectively: on good days, things just "felt" better and this did not seem to involve selective recall of past events of mood-congruent valence. Phenomenological analyses in the introspective tradition of German "armchair psychology" (e.g., Bollnow, 1956), which treated moods as an integrative reflection of one's current situation, seemed closer to the mark – alas, such introspections are to be taken with a grain of salt (Nisbett and Wilson, 1977). A competing perspective, Zajonc's (1980) "affective primacy" hypothesis, had the advantage of avoiding reliance on mood-congruent retrieval processes but lacked a process model specific enough to meet the developing criteria of social cognition research.

A conversation with Bob Wyer offered a different approach. Wyer and Carlston (1979) proposed that affect can serve informational functions, "for example, one's liking for a person may be based partly on the feelings of pleasantness when the person is around" (1979: 192). In addition, they conjectured that affective states may direct our attention to information that is suitable to explain one's feelings. While their conjectures were compatible with phenomenological approaches, their conceptualization emphasized the role of cognitive representations of experience at the expense of actual current experience itself, consistent with the information processing paradigm. Research into the influence of arousal (from Schachter and Singer's emotion research [1962] to Zillman's arousal-transfer model [1978] and Zanna and Cooper's dissonance studies [1976]) suggested, however, that the online experience itself may play a crucial role. More importantly, this literature also suggested that misattribution manipulations would be suitable experimental tools to address the role of current experience in human judgment.

A post-doctoral year with Wyer and Clore at the University of Illinois provided the opportunity to pursue these issues. Clore and Byrne (1974) had proposed a reinforcement-affect model to account for affective influences on interpersonal attraction. Going beyond the learning theories of the time, their model assumed that rewards exert their influence through the positive affect they elicit. Supporting this notion, laboratory and field experiments showed that associating others with positive feelings is sufficient to increase interpersonal attraction, even when the feelings are incidental and due to an unrelated source (e.g., Griffitt and Veitch, 1971). By the late 1970s Clore began to wonder why we "don't all end up falling in love with the paymaster," as he put it, if the mere co-occurrence of people with reward is sufficient to induce attraction. Does incidental affect only influence our judgments when we are not aware of its source, as Zillman's (1978) arousal studies suggested?

Initial evidence

These converging interests resulted in a conceptually straightforward study (Schwarz and Clore, 1983: Experiment 1). We asked participants to vividly recall and describe a happy or sad event to induce a corresponding mood and crossed these mood inductions with a misattribution manipulation that took advantage of a somewhat bizarre little room, previously used for auditory research with monkeys (for the inside story see Schwarz and Clore, 2003). This allowed us to suggest to some participants that the room may induce elated feelings and to others it may induce depressed feelings. Judgments of life-satisfaction served as the dependent variable. Our procedure deliberately stacked the deck in favor of content-driven models: by inducing moods through the recall of a happy or sad event, mood-congruent recall would be facilitated both by the content of the recall task and the induced mood. The predictions were straightforward. If mood effects on judgment were a function of mood-congruent recall (Bower, 1981), participants should report higher life-satisfaction when in a happy rather than sad mood, independent of what we told them about the room. If the experience itself served informative functions, on the other hand, its impact should depend on the feeling's perceived diagnosticity. That is, mood effects on reported life-satisfaction should be attenuated when the mood is attributed to the influence of the room and hence

considered uninformative for evaluating one's life in general (a discounting effect in Kelley's [1972] terms); however, it should be enhanced when one experiences the mood despite an allegedly opposing influence of the room (an augmentation effect).

The results, shown in Table 14.1, were consistent with these predictions, although only under sad mood conditions. The latter observation was compatible with Wyer and Carlston's (1979) suggestion that sad moods may require more explanation than happy ones, which would render them more susceptible to attributional manipulations. More important, the obtained augmentation and discounting effects made it unlikely that the influence of moods was driven by mood-congruent recall. After all, we had induced moods by having participants recall happy or sad events, thus adding semantic priming to the assumed affective activation of valenced material (Bower, 1981). Nevertheless, the accessible semantic content had little impact when participants discounted the accompanying negative feelings, assigning a crucial role to the subjective experience itself. Finally, the obtained attributional effects highlighted that the path from feelings to judgment was inferential, in contrast to Zajonc's (1980) assertion that "preferences need no inferences."

A subsequent, more naturalistic study took advantage of sunny and rainy weather as a mood manipulation (Schwarz and Clore, 1983: Experiment 2). As daily experience

Table 14.1 Life-satisfaction as a function of mood and attribution

Induced mood	Expectation about room		
	Tense	None	Elated
Positive	9.6a	8.6a	9.7a
Negative	8.6a	5.7b	4.4b
Control	—	8.9a	—

Note: Shown are mean reports of life-satisfaction (11 = very satisfied). Means not sharing a common subscript differ at $p < 0.05$.

Source: Adopted from Schwarz and Clore (1983).

suggests, participants reported higher life-satisfaction (and a more positive mood) when they were called on sunny than on rainy days. However, the negative influence of bad weather was eliminated when the interviewer, who pretended to call from out of town, first inquired about the weather at respondents' place of residence. This discounting effect was not obtained under sunny weather conditions, again suggesting that sad moods are more likely to be explained than happy moods. In combination, these studies provided first evidence for several assumptions that became core themes in the development of feelings-as-information theory.

Core themes

First, our findings showed that people attend to their momentary feelings as a source of information in forming judgments, essentially asking themselves, "How do I feel about this?" Later research extended this "informative function" (Wyer and Carlston, 1979) of affective states to other feelings, including nonaffective feelings, like the metacognitive experience of ease of recall (Schwarz et al., 1991b), and bodily sensations (Stepper and Strack, 1993).

Second, the observed discounting and augmentation effects highlighted that people use their feelings like any other source of information. They do not rely on them when they become aware that their feelings may be due to an unrelated source, thus undermining their informational value for the judgment at hand. Conversely, they consider their feelings particularly informative when they experience them despite opposing forces. Later research, much of it conducted by Michel Pham and his colleagues, identified additional variables that influence how much weight we give to our feelings (see Pham, 2004).

Third, our initial studies documented more positive judgments under happy than sad moods. While this is true for the bulk of mood research (Schwarz and Clore, 2007),

Leonard Martin and colleagues (e.g., Martin et al., 1997) demonstrated that positive feelings can result in negative evaluations. For example, when we feel happy while reading a sad story, we may conclude that it is not a "good sad story" after all, or else it would make us feel sad. Such findings illustrate that the influence of feelings depends on the specific question on which the feeling is brought to bear. This theme proved particularly important in later research on metacognitive experiences (Schwarz, 2004).

Finally, the observation that misattribution effects only emerged under sad moods (Table 14.1) proved more puzzling. Because most people feel mildly positive most of the time (Matlin and Stang, 1979), we initially suggested that sad moods are deviations from one's usual state and hence more likely to require explanation. This, in turn, would direct attention to possible sources of one's mood (Wyer and Carlston, 1979), rendering sad moods more susceptible to (mis)attribution manipulations. If so, being in an unexplained sad mood should interfere with other cognitive tasks, due to the competing demands of explaining one's mood. Testing this prediction, Bless et al. (1990) exposed participants in happy or sad moods to strong or weak persuasive arguments and assumed that sad moods would reduce systematic message elaboration. To our surprise, we found the opposite: sad participants engaged in message elaboration, whereas happy participants did not, by now a familiar and frequently replicated finding (for a review see Schwarz et al., 1991a). Similarly, Sinclair (1988) reported strong evidence that being in a sad mood reduced halo effects in impression formation. Clearly, sad moods did not pose an explanation problem that interfered with other processing demands; to the contrary, sad moods increased, and happy moods decreased, systematic processing in these studies.

To account for these findings, we (Schwarz, 1987, 1990; Schwarz and Bless, 1991) suggested that the informative function of moods is more general than assumed in the initial theorizing, which had focused on evaluative judgment.

In daily life, we usually feel bad when we encounter a threat of negative or a lack of positive outcomes, and feel good when we obtain positive outcomes and are not threatened by negative ones. Hence, our moods reflect the state of our environment (Bollnow, 1956). If so, bad moods may signal a problematic situation, whereas good moods may signal a benign situation. Given the situated nature of human cognition, we may expect that our thought processes are tuned to meet the processing requirements apparently posed by the situation, resulting in systematically different processing strategies under happy and sad moods. Sad moods may foster a systematic processing style that is characterized by bottom-up processing, attention to the details at hand, and limited playfulness and creativity. Happy moods, on the other hand, may foster a top-down processing style that relies more on general knowledge structures and is accompanied by less focused attention and higher playfulness and creativity. By and large these assumptions proved compatible with the accumulating evidence (for reviews see Clore et al., 1994; Schwarz, 2002; Schwarz and Clore, 2007). Moreover, later research showed that any information that signals a benign or problematic situation – from bodily sensations (e.g., Friedman and Förster, 2000) to metacognitive experiences (Song and Schwarz, 2008) or the smiling or frowning face of a communicator (Ottati et al., 1997) – can elicit the corresponding processing style. From this perspective, the misattribution effects observed by Schwarz and Clore (1983) were limited to sad moods because sad moods facilitate the analytic reasoning needed for attributional analyses, whereas happy moods make such reasoning less likely.

In the following sections, I revisit these themes by reviewing the postulates of feelings-as-information theory and illustrative experimental evidence. If social psychologists followed the naming traditions of software engineers, this would arguably be FAIT.3. Its treatment of the use of feelings as a source of information in judgment differs

from FAIT.1 (Schwarz and Clore, 1983) by emphasizing that the specific impact of a feeling depends on the epistemic question on which it is brought to bear. Its treatment of the influence of feelings on processing style differs from FAIT.2 (Schwarz, 1990; Schwarz and Bless, 1991) by de-emphasizing the role of processing ability (consistent with Bless and Schwarz, 1999) and by extending the range of variables that influence processing style beyond the role of feelings (consistent with Schwarz, 2002). Moreover, the theory's treatment of feelings goes far beyond our initial emphasis on moods and emotions and includes nonaffective experiences, consistent with the work conducted in the quarter century since the initial studies.

FEELINGS-AS- INFORMATION THEORY: POSTULATES

The core postulates are summarized in Box 14.1; they bear on the perceived informational value of feelings, their use as a basis of judgment, and their influence on the spontaneous adoption of different processing styles.

What feelings convey

The theory postulates that people attend to their feelings as a source of information, with different types of feelings providing different types of information. This assumption has a long tradition in emotion research. As Frijda (1988: 354) put it, “*emotions* exist for the sake of signaling states of the world that have to be responded to, or that no longer need response and action.” What exactly a given emotion signals can be derived from its underlying appraisal pattern (Ellsworth and Scherer, 2003; Ortony et al., 1988). Anger, for example, is a response to a loss or lack of reward that is attributed to the causal action of another agent; when no agent attribution is

Box 14.1 Postulates

- 1 People attend to their feelings as a source of information.
 - a Different types of feelings provide different types of information.
- 2 The impact of a given feeling depends on its perceived informational value for the task at hand.
 - a People usually experience their feelings as being "about" whatever is in the focus of attention; this fosters the perception that incidental feelings are relevant.
 - b When a feeling is attributed to an incidental source, its informational value is discounted; conversely, when it is experienced despite perceived opposing forces, its informational value is augmented.
 - c Changes in one's feelings are more informative than stable states.
- 3 When feelings are used as information, their use follows the same principles as the use of any other type of information.
 - a The impact of feelings increases with their perceived relevance to the task at hand and decreases with the accessibility and consideration of alternative diagnostic inputs, which is a function of processing motivation and capacity.
 - b What people conclude from a given feeling depends on (i) the epistemic question on which they bring it to bear and (ii) the lay theory applied.
- 4 Like other information, feelings can
 - a serve as a basis of judgment
 - b influence the choice of processing strategies; feelings that signal a "problematic" situation foster an analytic, bottom-up processing style, whereas feelings that signal a "benign" situation foster a more global, top-down processing style.

made, a loss gives rise to sadness. Accordingly, anger and sadness not only inform us about a loss, but also about its cause and elicit downstream responsibility judgments that reflect this information (e.g., Keltner et al., 1993a). Because emotions arise from ongoing, implicit appraisals of situations with respect to their implications for one's goals, they have an identifiable referent (what the emotion is "about"), a sharp rise time, and limited duration. These characteristics distinguish emotions from *moods*, which lack a clear referent, may come about gradually, may last for an extended time, and are often of low intensity (Bollnow, 1956; Morris, 1989). Hence, moods are more diffuse than emotions and primarily convey generic valence information that lacks a clear referent. These differences are apparent when we say that we are angry "about" something, but "in" a bad mood.

Cognitive feelings like surprise, boredom, or feelings of familiarity provide information about the state of one's knowledge (Ortony et al., 1988). Of particular interest to social psychologists is the metacognitive experience of ease or difficulty, which can pertain to recall and thought generation (*accessibility experiences*; Schwarz, 1998) or to the processing of new, external information (*processing fluency*; Winkielman et al., 2003). Numerous variables can influence these experiences, from the amount of information a person tries to recall to the presentation format in which new information is presented (e.g., print fonts, figure-ground contrast) and the semantic context in which it is embedded. Because cognitive operations can be easy or difficult for many different reasons, the specific inferences people draw from these experiences depend on which of many lay theories of mental processes they

bring to bear on the task (Schwarz, 2004). In addition, easy processing is experienced as pleasant (as captured by psychophysiological measures; Winkielman and Cacioppo, 2001) and this affective response can itself serve as a basis of judgment (Winkielman et al., 2003).

Finally, *bodily experiences* include feelings like hunger, pain, and physiological arousal, which inform us about physical states of the organism. Other bodily experiences provide information that parallels the implications of affective and cognitive feelings. For example, furrowing one's brow (contraction of the zygomaticus) conveys a feeling of effort and affects judgment in ways that parallel the metacognitive experience of difficulty (e.g., Sanna et al., 2002; Stepper and Strack, 1993). Similarly, proprioceptive feedback from facial expressions (e.g., Strack et al., 1988) and arm flexion and extension (e.g., Friedman and Förster, 2000) influence judgment and processing style in ways that parallel affective influences.

Perceived informational value

The theory further postulates that the impact of a given feeling depends on its perceived informational value for the task at hand. When a feeling is elicited by the object of judgment ("integral" in Bodenhausen's [1993] terminology), it provides valid information about the person's own response to the target. For example, seeing Susan may elicit positive feelings in Tom and he may be well advised to consider these feelings in (some) judgments of Susan. When the feeling is due to some other source ("incidental"), however, it provides (potentially) misleading information; for example, Tom's good feelings may be due to the weather rather than Susan. Unfortunately, people are more sensitive to their feelings than to where their feelings come from. They commonly assume that any feelings they have, and any thoughts that come to mind, are "about" whatever is in the focus of their attention

(Higgins, 1998) – or why else would they have them now, in this context? Hence, they are likely to perceive incidental feelings as being "about" the target of judgment, unless their attention is drawn to a plausible incidental source.

Whenever people (correctly or incorrectly) attribute their feelings to an incidental source, the perceived informational value of their feelings for the judgment at hand is undermined. Conversely, when they perceive that they have these feelings despite opposing forces, their feelings' perceived informational value is augmented. The sad mood conditions of Table 14.1 illustrate these discounting and augmentation effects.

When the informational value of their feelings is called into question, people turn to other sources of information to arrive at a judgment. As seen above, participants in the Schwarz and Clore (1983) study who discounted their sad mood arrived at life-satisfaction judgments that did not differ from participants in the control condition, who were not exposed to a mood manipulation. Presumably, both groups could draw on extensive other information about their own lives, resulting in similar judgments. Had such alternative inputs not been available, they might have resorted to an inferential correction strategy akin to, "I feel bad about my life, but this may be due to the room – so I should adjust my judgment upward." Such theory-driven correction strategies usually result in overcorrection; that is, a bias in the opposite direction (Strack and Hannover, 1996; Wilson and Brekke, 1994). Accordingly, discounting one's feelings as a source of information can either eliminate their influence (when alternative sources of information are accessible) or elicit a bias in the opposite direction (due to overcorrection in the absence of alternative inputs).

The theory further assumes that changes in one's feelings are more informative than stable states. This assumption is consistent with numerous studies in sensory perception and the covariation principle of attribution research. However, it has rarely been explicitly

tested in feelings-as-information experiments (for exceptions see Hansen et al., 2008; Shen et al., 2010). By relying on the experimental induction of feelings, experiments always involve a change from baseline as part of the methodological routine, which contributes to the feeling's perceived informational value.

Some misunderstandings

Some common misunderstandings of these assumptions deserve attention. To disentangle the contributions of the perceiver's feelings from other information about the target, experimental tests of the feelings-as-information hypothesis rely on the induction of incidental feelings. This gave rise to the erroneous conclusion that the use of feelings as a source of information is limited to incidental feelings, which led Forgas (2001: 104) to assert that "affect can only serve as a heuristic cue due to mistaken inferences," making reliance on one's feelings "an ineffective and dysfunctional strategy." This assertion confuses the operational and theoretical level. While reliance on incidental feelings can indeed be dysfunctional, integral feelings provide valid information. Attending to this information is highly adaptive, as a large body of research on emotional intelligence and the role of feelings in decision making indicates (see Barrett and Salovey, 2002; Damasio, 1994).

Falling prey to the same confusion, Slovic and colleagues (see Slovic et al., 2002) proposed an "affect heuristic" to account for the influence of integral feelings, which they considered distinct from the influence of incidental feelings. Unfortunately, integral feelings are inherently confounded with the positive or negative target attributes that elicit them, making it impossible to determine if observed differences are driven by experiential information in the form of integral feelings (as Slovic and colleagues assume) or by declarative information in the form of different target attributes. From the

perspective of feelings-as-information theory, the use of integral and incidental feelings as a source of information reflects the same basic mechanism – and any influence of target attributes that is not mediated by the feelings they elicit is better described in terms of declarative rather than "affective" information.

Finally, some observers (e.g., Forgas, 2001) suggested that feelings-as-information effects require a conscious attribution of the feeling to the target. This is not the case. Whereas discounting and augmentation effects require some level of conscious attribution, the mere use of one's feelings as a source of information does not. As noted, people usually consider their thoughts and feelings to be "about" whatever is in the focus of their attention, rendering reliance on them the automatic default option. Accordingly, the impact of feelings increases when contextual influences, like time pressure (Siemer and Reisenzein, 1998), limit the opportunity to engage in attributional analyses, in contrast to what a conscious attribution requirement would predict.

From feelings to judgments

The theory further postulates that whenever feelings are used as a source of information, their use follows the same rules as the use of any other information. First, feelings are only used as a source of information when their informational value is not called into question (e.g., Schwarz and Clore, 1983). Second, the impact of feelings increases with their perceived relevance to the judgment at hand. For example, moods exert a stronger influence when people make decisions for themselves rather than others, whose affective response may differ from their own (Raghunathan and Pham, 1999) or when they evaluate the hedonic pleasure that can be derived from an activity rather than the activity's instrumental value for academic achievement (Pham, 1998). Third, the impact of feelings decreases the more other relevant

inputs are accessible. For example, people are less likely to rely on their feelings when they have high expertise in the domain of judgment (e.g., Ottati and Isbell, 1996; Sedikides, 1995), which presumably facilitates the assessment of the relevance of one's feelings and renders other information easily accessible. Fourth, as is the case for any other highly accessible piece of information, the impact of feelings is more pronounced under conditions of low processing capacity (e.g., Greifeneder and Bless, 2007; Siemer and Reisenzein, 1998) or motivation (e.g., Rothman and Schwarz, 1998). These conditions limit assessments of the diagnosticity and relevance of one's feelings and the search for possible alternative inputs. As these examples illustrate, the variables that govern the use and impact of experiential information as a basis of judgment parallel the variables that govern the use and impact of declarative information, consistent with the basic feelings-as-information logic.

Finally, feelings share with other information that their specific implications depend on the question asked. The observation that Bob has published a highly acclaimed book every year since his PhD can be brought to bear on many judgments of Bob, from his intelligence and ambitiousness to his professional standing and his commitment to spending time with his kids. The same holds for feelings. What people conclude from a given feeling depends on the epistemic question on which they bring it to bear. For example, Martin et al. (1993) asked happy and sad participants to list birds. When asked whether they are satisfied with what they accomplished, happy participants inferred that they are satisfied and terminated the task, whereas sad participants inferred that they are not yet satisfied and continued. This pattern reversed when participants were asked whether they enjoy what they are doing. In this case, happy participants inferred enjoyment and continued with the task, whereas sad participants inferred a lack of enjoyment and terminated the task. In both cases, their judgments were consistent with the valence information

provided by their mood, yet this valence information had diverging behavioral implications, depending on the specific question on which it was brought to bear.

Importantly, some feelings require more interpretation, and allow for a wider range of inferences, than others. As already noted, moods provide broadly applicable valence information, whereas specific emotions inform us that a specific appraisal pattern has been met, which constrains the range of plausible inferences. At the other extreme, metacognitive experiences primarily inform us that our cognitive operations are easy or difficult – and they may be so for many reasons. For example, we may find it difficult to recall information because the event happened a long time ago, because we never found it important and hence didn't pay attention, because we lack expertise in the domain, and so on. Which inferences we draw from difficulty of recall will therefore depend on which of these naïve theories of mental processes we bring to bear. Applicable theories are usually brought to mind by the judgment task (Schwarz, 2004) and the same metacognitive experience can result in differential judgments of expertise, importance or temporal distance, depending on the specific question asked.

Cognitive tuning: feelings and processing style

In addition to providing information that can serve as a basis of judgment, feelings influence *how* people process information; that is, their processing style. A number of different explanations have been offered for this observation, usually highlighting the role of one specific type of feeling (for reviews, see Schwarz and Clore, 2007, and the contributions in Martin and Clore, 2001). Feelings-as-information theory provides a unified conceptualization of these influences in the context of a situated cognition framework (Smith and Semin, 2004). It assumes that human cognition stands in the service of

action (James, 1890) and that our cognitive processes are responsive to the environment in which we pursue our goals. This responsiveness ranges from the higher accessibility of knowledge relevant to the current situation (e.g., Yeh and Barsalou, 2006) to the choice of processing strategies that meet situational requirements (e.g., Wegner and Vallacher, 1986). When things go smoothly and we face no hurdles in the pursuit of our goals, we are likely to rely on our pre-existing knowledge structures and routines, which served us well in the past. Moreover, we may be willing to take some risk in exploring novel solutions. Once things go wrong, we abandon reliance on our usual routines and focus on the specifics at hand to determine what went wrong and what can be done about it.

Feelings play a crucial role in this tuning process by providing a fast and parsimonious indicator of whether our current situation is “benign” or “problematic.” The influence of feelings on processing style is eliminated when the informational value of the feeling is called into question (e.g., Sinclair et al., 1994) and can be overridden by the individual’s goals or explicit task demands (e.g., Bless et al., 1990).

in between. This influence is not observed when the informational value of the mood is called into question through (mis)attribution manipulations (Schwarz and Clore, 1983; for conceptual replications see Gorn et al., 1993; Savitsky et al., 1998; Siemer and Reisenzein, 1998, among others). Even when one’s mood is considered informative, its impact depends on its perceived relevance to the judgment at hand (e.g., Pham, 1998) and the accessibility of competing inputs (e.g., Sedikides, 1995), as discussed above. Importantly, mood effects are not limited to inconsequential judgments. Instead, moods have been found to influence highly consequential decisions, from medical school admissions (Redlmeier and Baxter, 2009) to stock market investments. For example, Hirshleifer and Shumway (2003) observed a reliable influence of the weather on stock market returns in 26 countries: the market is more likely to go up when the sun shines in the city that hosts the country’s major stock exchange. Presumably, the upbeat mood associated with sunny weather makes investors more optimistic about the future of the economy, paralleling observations in experiments.

Whereas the bulk of the research shows more positive (negative) judgments under happy (sad) mood, moods can also result in mood-incongruent judgments under specific conditions. First, mood incongruent judgments can result from the logic of discounting effects themselves (e.g., Ottati and Isbell, 1996). Suppose, for example, that you are evaluating a job candidate and are aware that you are in a miserable mood due to an earlier event. To which extent are your bad feelings an integral part of your reaction to the candidate and to which extent are they due to the earlier event? If you fully discount your bad feelings, you may arrive at an unduly positive evaluation of the candidate. Second, mood-inducing events can elicit contrast effects in the evaluation of closely related targets by serving as extreme standards of comparison. For example, Schwarz et al. (1987) conducted an experiment in a very pleasant or unpleasant room. Replicating earlier findings, their

REPRESENTATIVE FINDINGS

Next, I review representative findings pertaining to the influence of moods, emotions, and metacognitive experiences on judgment and processing style and highlight some real-world implications (for more extensive reviews of findings see Clore et al., 1994; Schwarz and Clore, 2007).

Feelings as a basis of judgment

Moods

As discussed, moods convey valence information that usually results in more positive judgments when people are in a happy rather than sad mood, with neutral moods falling

student participants reported higher life-satisfaction when the pleasant room induced a positive mood than when the unpleasant room induced a negative mood. When asked about their housing-satisfaction, however, this pattern reversed, presumably because even modest dorm rooms seemed luxurious compared to the salient standard introduced by the unpleasant room. Finally, the target of judgment can carry affective expectations to which one's current feelings are compared; for example, Martin and colleagues (1997) observed that happy participants rated a sad story less favorably than sad participants. Presumably, their happy feelings implied that the sad story failed to achieve its goal of making them feel sad, leading them to conclude that it was a poor sad story.

Emotions

As observed for moods, the impact of specific emotions is eliminated when they are attributed to an incidental source. For example, Schwarz et al. (1985) found that a fear-arousing communication did not affect participants' attitudes when they attributed their feelings to allegedly arousing side-effects of a pill; conversely, expecting the pill to have tranquilizing effects enhanced the message's impact. However, the informational value of specific emotions differs from the informational value of global moods in ways that can be traced to the role of appraisals.

Recall that emotions reflect the person's appraisal of a specific event (Ellsworth and Scherer, 2003; Ortony et al., 1988), which is in the focus of the person's attention. This makes emotions less likely to be misread as bearing on unrelated targets than is the case for diffuse moods. Indeed, merely labeling one's current feelings with specific emotion terms is sufficient to elicit an event attribution and has been found to be as efficient in eliminating effects on unrelated judgments as a standard misattribution manipulation (Keltner et al., 1993b). Note that this observation has important methodological implications: using detailed emotion terms as manipulation checks invites causal

attributions to determine the specific emotion, which can eliminate the expected effect.

Moreover, experiencing an emotion implies that a specific set of appraisal criteria has been met. Anger, for example, informs us that somebody did us wrong and hence provides more specific information than a diffuse negative mood. Accordingly, the influence of emotions can be predicted on the basis of the underlying appraisals (e.g., Lerner and Keltner, 2000). For example, Lerner and colleagues (2003) observed in a national survey during the immediate aftermath of the terrorist attacks of September 11, 2001, that inducing participants to focus on the experienced fear increased risk estimates and plans for precautionary behavior, whereas focusing on the experienced anger did the reverse.

Metacognitive experiences

Compared with the appraisal information conveyed by emotions, the information conveyed by metacognitive experiences of ease or difficulty is relatively diffuse. All the experience, by itself, conveys is that "this" is easy or difficult – and most cognitive operations can be so for many different reasons. Hence, the same experience can give rise to different inferences, depending on which of many lay theories of mental processes comes to mind (Schwarz, 2004). As observed for moods and emotions, the influence of metacognitive experiences is eliminated when they are attributed to an incidental source (e.g., Schwarz et al., 1991b).

Accessibility experiences As an example, consider the ease or difficulty with which information can be brought to mind. According to most models of judgment, an object should be evaluated more favorably when we recall many rather than few positive attributes; similarly, an event should seem more likely when we generate many rather than few reasons for its occurrence. Empirically, the opposite is often the case. For example, people consider themselves less assertive after recalling many rather than

few examples of their own assertive behavior (Schwarz et al., 1991b); hold an attitude with less confidence after listing many rather than few supporting arguments (Haddock et al., 1999); and consider an event less likely after listing many rather than few reasons for its occurrence (Sanna and Schwarz, 2004). Throughout, their inferences are consistent with the implications of accessible thought content when thought generation is easy (few), but opposite to these implications when thought generation is difficult (many). This pattern reflects a lay theory of mental processes that is at the heart of Tversky and Kahneman's (1973) availability heuristic: "The more exemplars exist, the easier it is to bring some to mind." Hence, the difficulty of generating many reasons or examples suggests that there aren't many, giving rise to the above conclusions. When participants attribute the experienced difficulty to an incidental influence, like music played in the background, its informational value is undermined and they turn to accessible thought content as an alternative input. In this case, the otherwise observed pattern reverses and they infer, for example, that they are more assertive, the more examples of assertive behaviors they recall (Schwarz et al., 1991). Similarly, yoked participants, who merely read the thoughts generated by someone else and are hence deprived of the generation experience, are more influenced when their partner lists many rather than few arguments, in contrast to the person who lists them (Wänke et al., 1996). These observations highlight that the thought content by itself is compelling once it is not qualified by a subjective difficulty experience.

Other lay theories hold, for example, that details of recent events are easier to recall than details of distant events, and details of important events easier than details of unimportant ones. Which of these (or many other) theories comes to mind depends on the question posed. Schwarz and Xu (2010) had participants recall details of the Oklahoma City bombing. When first asked to date the event, participants inferred that it was more

recent after recalling two rather than ten details; when first asked how important they found the event at the time, they inferred higher importance after recalling two rather than ten details. Thus, the same accessibility experience informed judgments of temporal distance or of importance, depending on the question posed. More important, application of a given theory entails an attribution of the experience to a specific cause (here, recency or importance), which changes the implications of the experience for other judgments (Schwarz, 2004). Accordingly, participants who initially attributed the difficulty of recalling many details to the event's temporal distance subsequently reported that the event was very important to them – after all, they could still recall details even though the event had apparently happened long ago, so it must have been quite important. Conversely, participants who initially attributed difficulty of recall to low personal importance subsequently dated the event as closer in time – after all, they could still recall details despite the event's low personal importance, so it must have been quite recent. Such findings (for a review see Schwarz, 2010) show that inferences from metacognitive experiences are highly malleable, presumably because people are aware that cognitive operations can be easy or difficult for many different reasons, each of which provides a different inference rule.

Paralleling the findings for other feelings, people are more likely to rely on their accessibility experiences under conditions that commonly foster heuristic processing, but turn to accessible content under conditions that commonly foster systematic processing. The latter conditions include high personal relevance (e.g., Rothman and Schwarz, 1998), high need for cognition (e.g., Greifeneder et al., 2010) and being in a sad rather than happy mood (e.g., Ruder and Bless, 2003).

Processing fluency Just like information can be easy or difficult to bring to mind, new information we encounter can be easy or difficult to process. Numerous variables can

influence ease of processing, ranging from presentation characteristics (like figure-ground contrast, print font, or rhyming versus nonrhyming form) to the semantic relatedness of the material and the frequency and recency of previous exposure. Because these variables result in similar phenomenological experiences, the meaning of the experience is open to interpretation. Which interpretation people choose, and which inferences they draw, again depends on the naïve theory they bring to bear (Schwarz, 2004, 2010).

One naïve theory that is of particular importance to social psychological phenomena is the (usually correct) assumption that familiar material is easier to process than unfamiliar material. Hence, fluently processed material seems more familiar than fluently processed material, even when the fluency experience is solely due to incidental variables, like the print font or color contrast in which the material is presented. As observed for other feelings, drawing people's attention to these incidental sources of fluency undermines the informational value of the experience and eliminates the otherwise observed effects (e.g., Novemsky et al., 2007). In the absence of such attribution manipulations, however, the fluency-familiarity association affects numerous judgments of everyday importance, including judgments of social consensus, truth, and risk.

As Festinger (1954) noted, we often rely on social consensus information to determine whether an assertion is true or false – if many people believe it, there's probably something to it. One heuristic to estimate social consensus is to assess whether the assertion seems familiar. Accordingly, fluency of processing gives rise to increased estimates of social consensus (Weaver et al., 2007) and facilitates the acceptance of a statement as true (for a review see Schwarz et al., 2007). For example, statements like "Osorno is a city in Chile" are more likely to be judged "true" when they are presented in colors that make them easy rather than difficult to read against the background (Reber and Schwarz, 1999).

Familiarity also figures prominently in intuitive assessments of risk – if a stimulus is familiar and elicits no negative memories, it presumably hasn't hurt us in the past. Accordingly, incidental variables that affect processing fluency also influence peoples' risk assessments. For example, ostensible food additives are perceived as more hazardous when their names are difficult (e.g., Fluthoractnip) rather than easy (e.g., Magnalroxate) to pronounce (Song and Schwarz, 2009) and stocks with easy to pronounce ticker symbols attract more investors at their initial public offering (Alter and Oppenheimer, 2006). In addition to the mediating role of perceived familiarity observed by Song and Schwarz (2009), intuitive assessments of risk may be further affected by perceivers' positive affective response to fluently processed stimuli (addressed below), consistent with the observation of mood effects on judgment of risk (Johnson and Tversky, 1983) and the beneficial influence of sunny weather on the stock market (Hirshleifer and Schumway, 2003).

Fluency and affect As known since Zajonc's (1968) pioneering mere exposure studies, repeated exposure to an initially neutral stimulus, without any reinforcement, leads to gradual increase in liking. However, repeated exposure is just one of many variables that can increase processing fluency and any other variable that facilitates fluent processing has the same effect. For example, people like the same stimulus more when it is preceded by a visual (Reber et al., 1998) or semantic (Winkielman et al., 2003) prime that facilitates fluent processing, and less when it is preceded by primes that impede fluent processing. In fact, the influence of many variables long known to affect liking and aesthetic preference – from figure-ground contrast to symmetry and prototypicality – can be traced to increased processing fluency (Reber et al., 2004).

This fluency-liking link reflects that fluent processing itself is experienced as pleasant and elicits a positive affective response that

can be captured with psychophysiological measures (Winkielman and Cacioppo, 2001). If this affective response mediates the influence of fluency on liking, it should be eliminated when the positive affect is attributed to an incidental source. Empirically, this is the case as Winkielman and Fazendeiro (reported in Winkielman et al., 2003) demonstrated with misattribution procedures.

Summary and applications

As the reviewed examples illustrate, people attend to a wide range of feelings as a source of information. However, they are more sensitive to their feelings than to where these feelings come from and routinely consider incidental feelings relevant to the task at hand. What exactly they conclude from a given feeling depends on the epistemic task they face. Different epistemic tasks bring different lay theories to mind, which link the feeling to the task at hand and serve as inference rules. When feelings are used as a source of information, their use follows the same rules as the use of any other information; hence, the impact of feelings increases with their perceived relevance and applicability and decreases with the consideration of alternative inputs. Whenever people become aware that their feelings may be due to an incidental source, the informational value of the feeling is discredited and people turn to alternative inputs to arrive at a judgment. These regularities hold for moods, emotions, metacognitive experiences, and bodily sensations (Schwarz and Clore, 2007).

The observed use of feelings as a source of information pervades daily life and is not limited to any particular “applied” domain. From the quality of their own lives (Schwarz and Clore, 1983) to consumer products (Isen et al., 1978) and the daily risks they face (Johnson and Tversky, 1983), people arrive at more upbeat evaluations when they are in a positive mood (Schwarz and Clore, 2007) – even when the stakes are high, as illustrated by pervasive mood effects on the stock market (Hirshleifer and Shumway, 2003). Similarly, emotions like anger or fear can

shift risk perception and policy preferences (Lerner et al., 2003), as can differences in processing fluency (Song and Schwarz, 2009). Moreover, assessments of truth (Reber and Schwarz, 1999) and the spread of rumors (Schwarz et al., 2007) are profoundly affected by metacognitive experiences, as are consumer (Novemsky et al., 2007) and investment (Alter and Oppenheimer, 2006) decisions. Our feelings are part and parcel of how we think about life and their influence can be observed in any area of investigation.

Feelings and processing style

The theory further predicts that feelings or environmental cues that signal a “problematic” situation foster an analytic, bottom-up processing style with considerable attention to detail, whereas feelings or environmental cues that signal a “benign” situation allow for a less effortful, top-down processing style and the exploration of novel (and potentially risky) solutions (Schwarz, 1990, 2002). This does not imply that people in a happy mood, for example, are unable or unwilling to engage in analytic processing (in contrast to what an earlier version of the theory suggested; Schwarz and Bless, 1991). Instead, it merely implies that happy feelings (and other “benign” signals) do not convey a need to do so; when task demands or current goals require bottom-up processing, happy individuals are able and willing to engage in it. A study that addressed the influence of moods on people’s reliance on scripts (Schank and Abelson, 1977) illustrates this point.

Employing a dual-task paradigm, Bless et al. (1996a) had participants listen to a tape-recorded restaurant story that contained script consistent and script inconsistent information. While listening to the story, participants also worked on a concentration test that required detail-oriented processing; in contrast, the restaurant story could be understood by engaging either in script-driven top-down processing or in data-driven bottom-up processing. Happy participants relied on the

script, as indicated by the classic pattern of schema-guided memory: they were likely to recognize previously heard script-inconsistent information, but also showed high rates of intrusion errors in form of erroneous recognition of script-consistent information. Neither of these effects was obtained for sad participants, indicating that they were less likely to draw on the script to begin with. Given that top-down processing is less taxing than bottom-up processing, we may further expect that happy participants' reliance on the script allows them to do better on a secondary task. Confirming this prediction, happy participants outperformed sad participants on the concentration test. In combination, these findings indicate that moods influence the spontaneously adopted processing style under conditions where different processing styles are compatible with the individual's goals and task demands, as was the case for comprehending the restaurant story. Under these conditions, sad individuals are likely to spontaneously adopt a systematic, bottom-up strategy, whereas happy individuals rely on a less effortful top-down strategy. But when task demands (like a concentration test) or explicit instructions (e.g., Bless et al., 1990) require detail-oriented processing, happy individuals are able and willing to live up to the task.

Numerous findings pertaining to a broad range of feelings (from moods and emotions to bodily experiences and processing fluency) and cognitive tasks (from creative and analytic problem solving to persuasion and stereotyping) are consistent with the predictions of feelings-as-information theory (for reviews see Schwarz, 2002; Schwarz and Clore, 2007). Here I focus on two domains of particular applied interest, namely persuasion and stereotyping.

Applied implications

Persuasion In general, strong arguments are more persuasive than weak arguments when recipients engage in systematic message elaboration, whereas argument strength exerts little influence when they do not (Petty and

Cacioppo, 1986). Accordingly, the impact of argument strength can serve as a diagnostic tool for assessing processing strategy. Studies using this strategy consistently found that happy recipients engage in less, and sad recipients in more, elaboration of counter-attitudinal messages than recipients in a non-manipulated mood (see Schwarz et al., 1991a, for a review). Hence, happy recipients are moderately and equally persuaded by strong as well as weak arguments, whereas sad recipients are strongly persuaded by strong arguments, and not persuaded by weak arguments. Consistent with the feelings-as-information logic, these effects are eliminated when recipients are aware that their mood is due to an unrelated source (Sinclair, et al., 1994). Moreover, the spontaneously adopted processing strategy can be overridden by other variables, such as explicit instructions to pay attention to the arguments (e.g., Bless et al., 1990) or the promise that carefully thinking about the message would make one feel good (e.g., Wegener et al., 1995).

Paralleling the influence of moods, the experience of low processing fluency (which is associated with negative affect and a sense that the processed material is unfamiliar) also fosters the adoption of detail-oriented bottom-up processing, whereas high processing fluency fosters top-down processing (e.g., Song and Schwarz, 2008). Not surprisingly, numerous environmental cues can serve the same informational functions. For example, the same message is less likely to be scrutinized when presented by a communicator with a smiling, happy face than when presented by a communicator with a neutral, somber face (Ottati et al., 1997). Further illustrating the power of contextual cues, Soldat and Sinclair (2001) printed persuasive messages on colored paper. Their recipients were persuaded by strong arguments, but not by weak arguments, when the paper had a depressing blue hue, whereas both types of arguments were similarly persuasive when the paper had an upbeat red hue.

From an applied perspective, these findings suggest that communicators with strong

and compelling arguments have little to gain from putting their audience into a good mood; to the contrary, a somber audience is more likely to elaborate on the substantive implications of the message, facilitating its long-term impact. On the other hand, happy feelings make spontaneous message scrutiny less likely, making smiles, jokes, and upbeat colors promising tools when we have nothing compelling to say.

Stereotyping and impression formation We can form impressions of others by attending to their specific behaviors (bottom-up processing) or by drawing on stereotypic knowledge about social categories (top-down processing). Reiterating the observations from persuasion research, perceivers in a sad mood are more likely to elaborate individuating information about the target person, whereas perceivers in a happy mood are more likely to draw on the person's category membership. This results in more stereotypical judgments under happy than under sad moods (e.g., Bodenhausen et al., 1994; for a review see Bless et al., 1996b). Related research into the influence of brands on product evaluation similarly shows higher reliance on brand information under happy than sad moods (e.g., Adaval, 2001). Note that increased reliance on general knowledge about a brand or group can have positive as well as negative consequences for individual exemplars (be they products or persons), helping exemplars of liked and hurting exemplars of disliked categories. Paralleling the persuasion findings, happy individuals' reliance on category membership information can be overridden by manipulations that increase their processing motivation, such as personal accountability for one's judgment (Bodenhausen, et al., 1994).

CONCLUSION

In sum, internal and external cues that signal a benign or problematic situation have cognitive and motivational consequences

(Schwarz, 2002). Human cognition is tuned to meet situational requirements and problem signals foster vigilance and the adoption of a detail-oriented bottom-up processing style, which is usually adaptive. Signals that characterize the situation as benign, on the other hand, are not by themselves associated with particular processing requirements. They foster reliance on pre-existing knowledge structures and top-down processing, *unless* goals or task demands require otherwise. Which processing strategy facilitates or impedes performance depends on the specific task. In general, internal or external "problem" signals improve performance on tasks that require analytic reasoning and attention to detail, but impair performance on insight and creativity tasks that require divergent thinking and the exploration of novel solutions (Schwarz, 2002).

ACKNOWLEDGMENTS

Completion of this chapter was supported by a fellowship at the Center for Advanced Study in the Behavioral Sciences.

REFERENCES

Adaval, R. (2001) Sometimes it just feels right: The differential weighting of affect-consistent and affect-inconsistent product information. *Journal of Consumer Research*, 28, 1–17.

Alter, A.L. and Oppenheimer, D.M. (2006) Predicting short-term stock fluctuations by using processing fluency. *Proceedings of the National Academy of Science*, 103, 9369–9372.

Barrett, L.F. and Salovey, P. (eds) (2002) *The Wisdom in Feelings*. New York: Guilford.

Bless, H., Bohner, G., Schwarz, N. and Strack, F. (1990) Mood and persuasion: A cognitive response analysis. *Personality and Social Psychology Bulletin*, 16, 331–345.

Bless, H., Clore, G.L., Schwarz, N., Golisano, V., Rabe, C. and Wölk, M. (1996a) Mood and the use of scripts: Does being in a happy mood really lead to

mindlessness? *Journal of Personality and Social Psychology*, 71, 665–679.

Bless, H. and Schwarz, N. (1999) Sufficient and necessary conditions in dual process models: The case of mood and processing style. In S. Chaiken and Y. Trope (eds), *Dual-process Models in Social Psychology*, pp. 423–440. New York: Guilford.

Bless, H., Schwarz, N. and Kemmelmeier, M. (1996b) Mood and stereotyping: The impact of moods on the use of general knowledge structures. In M. Hewstone and W. Stroebe (eds), *European Review of Social Psychology*, 7, 63–93.

Bodenhausen, G.V. (1993) Emotions, arousal, and stereotype-based discrimination: A heuristic model of affect and stereotyping. In D.M. Mackie and D.L. Hamilton (eds), *Affect, Cognition, and Stereotyping: Interactive Processes in Group Perception*, pp. 13–35. San Diego: Academic Press.

Bodenhausen, G.V., Kramer, G.P. and Süsser, K. (1994) Happiness and stereotypic thinking in social judgment. *Journal of Personality and Social Psychology*, 66, 621–632.

Bollnow, O.F. (1956) *Das Wesen der Stimmungen*. Frankfurt: Klostermann.

Bower, G.H. (1981) Mood and memory. *American Psychologist*, 36, 129–148.

Clore, G.L. and Byrne, D. (1974) A reinforcement-affect model of attraction. In T.L. Huston (ed.), *Foundations of Interpersonal Attraction*, pp. 143–170. New York: Academic Press.

Clore, G.L., Schwarz, N. and Conway, M. (1994) Affective causes and consequences of social information processing. In R.S. Wyer, and T.K. Srull (eds), *Handbook of Social Cognition*, Vol. 1, 2nd Edition, pp. 323–418. Hillsdale, NJ: Erlbaum.

Damasio, A.R. (1994) *Descartes' Error: Emotion, Reason and the Human Brain*. New York: Grosset/Putnam.

Ellsworth, P.C. and Scherer, K.R. (2003) Appraisal processes in emotion. In R.J. Davidson, K.R. Scherer, and H.H. Goldsmith (eds), *Handbook of Affective Sciences*, pp. 572–595. New York: Oxford University Press.

Festinger, L. (1954) A theory of social comparison processes. *Human Relations*, 7, 123–146.

Forgas, J.P. (2001) The affect infusion model (AIM): An integrative theory of mood effects on cognition and judgment. In L.L. Martin and G.L. Clore (eds), *Theories of Mood and Cognition: A User's Guidebook*, pp. 99–134. Mahwah, NJ: Erlbaum.

Friedman, R.S. and Förster, J. (2000) The effects of approach and avoidance motor actions on the elements of creative insight. *Journal of Personality and Social Psychology*, 79, 477–492.

Frijda, N.H. (1988) The laws of emotion. *American Psychologist*, 43, 349–358.

Gorn, G.J., Goldberg, M.E. and Basu, K. (1993) Mood, awareness, and product evaluation. *Journal of Consumer Psychology*, 2, 237–256.

Greifeneder, R. and Bless, H. (2007) Relying on accessible content versus accessibility experiences: The case of processing capacity. *Social Cognition*, 25, 853–881.

Greifeneder, R., Müller, P., Stahlberg, D., Van den Bos, K. and Bless, H. (2010) Guiding trustful behavior: The role of accessible content and accessibility experiences. *Journal of Behavioral Decision Making*. DOI:10.1002/bdm.705.

Griffitt, W. and Veitch, R. (1971) Hot and crowded: Influences of population density and temperature on interpersonal behavior. *Journal of Personality and Social Psychology*, 17, 92–98.

Haddock, G., Rothman, A.J., Reber, R. and Schwarz, N. (1999) Forming judgments of attitude certainty, importance, and intensity: The role of subjective experiences. *Personality and Social Psychology Bulletin*, 25, 771–782.

Hansen, J., Dechêne, A. and Wänke, M. (2008) Discrepant fluency increases subjective truth. *Journal of Experimental Social Psychology*, 44, 687–691.

Higgins, E.T. (1998) The aboutness principle: A pervasive influence on human inference. *Social Cognition*, 16, 173–198.

Hirshleifer, D. and Shumway, T. (2003) Good day sunshine: Stock returns and the weather. *Journal of Finance*, 58, 1009–1032.

Isen, A.M., Shalker, T.E., Clark, M.S. and Karp, L. (1978) Affect, accessibility of material in memory, and behavior: A cognitive loop? *Journal of Personality and Social Psychology*, 36, 1–12.

James, W. (1890) *Principles of Psychology* (Vols. 1 and 2). New York: Dover.

Johnson, E. and Tversky, A. (1983) Affect, generalization, and the perception of risk. *Journal of Personality and Social Psychology*, 45, 20–31.

Kelley, H.H. (1972) *Causal Schemata and the Attribution Process*. Morristown, NJ: General Learning Press.

Keltner, D., Ellsworth, P. and Edwards, K. (1993a) Beyond simple pessimism: Effects of sadness and anger on social perception. *Journal of Personality and Social Psychology*, 64, 740–752.

Keltner, D., Locke, K.D. and Audrain, P.C. (1993b) The influence of attributions on the relevance of negative feelings to satisfaction. *Personality and Social Psychology Bulletin*, 19, 21–30.

Lerner, J.S. and Keltner, D. (2000) Beyond valence: A model of emotion-specific influences on

judgment and choice. *Cognition and Emotion*, 14, 473–493.

Lerner, S.J., Gonzalez, R.M., Small, D.A. and Fischhoff, B. (2003) Effects of fear and anger on perceived risks of terrorism: A national field experiment. *Psychological Science*, 14, 144–150.

Martin, L.L., Abend, T., Sedikides, C. and Green, J.D. (1997) How would it feel if...? Mood as input to a role fulfillment evaluation process. *Journal of Personality and Social Psychology*, 73, 242–253.

Martin, L.L. and Clore, G.L. (eds) (2001) *Theories of Mood and Cognition: A User's Guidebook*. Mahwah, NJ: Erlbaum.

Martin, L.L., Ward, D.W., Achée, J.W. and Wyer, R.S. (1993) Mood as input: People have to interpret the motivational implications of their moods. *Journal of Personality and Social Psychology*, 64, 317–326.

Matlin, M.W. and Stang, D. (1979) *The Pollyanna Principle: Selectivity in Language, Memory, and Thought*. Cambridge, MA: Shenkman.

Morris, W.N. (1989) *Mood: The Frame of Mind*. New York: Springer Verlag.

Nisbett, R.E. and Wilson, T.D. (1977) Telling more than we know: Verbal reports on mental processes. *Psychological Review*, 84, 231–59.

Novemsky, N., Dhar, R., Schwarz, N. and Simonson, I. (2007) Preference fluency in choice. *Journal of Marketing Research*, 44, 347–356.

Ottney, A., Clore, G.L. and Collins, A. (1988) *The Cognitive Structure of Emotions*. New York: Cambridge University Press.

Ottati, V., Terkildsen, N. and Hubbard, C. (1997) Happy faces elicit heuristic processing in a televised impression formation task: A cognitive tuning account. *Personality and Social Psychology Bulletin*, 23, 1144–1156.

Ottati, V.C. and Isbell, L.M. (1996) Effects of mood during exposure to target information on subsequently reported judgments: An on-line model of misattribution and correction. *Journal of Personality and Social Psychology*, 71, 39–53.

Petty, R. and Cacioppo, J. (1986) *Communication and Persuasion: Central and Peripheral Routes to Attitude Change*. New York: Springer-Verlag.

Pham, M.T. (1998) Representativeness, relevance, and the use of feelings in decision making. *Journal of Consumer Research*, 25, 144.

Pham, M.T. (2004) The logic of feeling. *Journal of Consumer Psychology*, 14, 360–369.

Raghunathan, R. and Pham, M.T. (1999) All negative moods are not equal: Motivational influences of anxiety and sadness on decision making. *Organizational Behavior and Human Decision Processes*, 79, 56–77.

Reber, R. and Schwarz, N. (1999) Effects of perceptual fluency on judgments of truth. *Consciousness and Cognition*, 8, 338–342.

Reber, R., Schwarz, N. and Winkielman, P. (2004) Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality and Social Psychology Review*, 8, 364–382.

Reber, R., Winkielman P. and Schwarz N. (1998) Effects of perceptual fluency on affective judgments. *Psychological Science*, 9, 45–48.

Redlmeier, D.A. and Baxter, S.D. (2009) Rainy weather and medical school admission interviews. *Canadian Medical Association Journal*, 181, 933.

Rothman, A.J. and Schwarz, N. (1998) Constructing perceptions of vulnerability: Personal relevance and the use of experiential information in health judgments. *Personality and Social Psychology Bulletin*, 24, 1053–1064.

Ruder, M. and Bless, H. (2003) Mood and the reliance on the ease of retrieval heuristic. *Journal of Personality and Social Psychology*, 85, 20–32.

Sanna, L. and Schwarz, N. (2004) Integrating temporal biases: The interplay of focal thoughts and accessibility experiences. *Psychological Science*, 17, 474–481.

Sanna, L., Schwarz, N. and Small, E. (2002) Accessibility experiences and the hindsight bias: I-knew-it-all-along versus It-could-never-have-happened. *Memory and Cognition*, 30, 1288–1296.

Savitsky, K., Medvec, V.H., Charlton, A.E. and Gilovich, T. (1998) What, me worry? Arousal, misattribution, and the effect of temporal distance on confidence. *Personality and Social Psychology Bulletin*, 24, 529–536.

Schachter, S. and Singer, J.E. (1962) Cognitive, social, and physiological determinants of emotional state. *Psychological Review*, 69, 379–399.

Schank, R. and Abelson, R. (1977) *Scripts, Plans, Goals and Understanding: An Inquiry into Human Knowledge Structures*. Hillsdale, NJ: Lawrence Erlbaum.

Schwarz, N. (1987) *Stimmung als Information* [Mood as information.] Heidelberg: Springer Verlag.

Schwarz, N. (1990) Feelings as information: Informational and motivational functions of affective states. In E.T. Higgins and R. Sorrentino (eds), *Handbook of Motivation and Cognition: Foundations of Social Behavior*, 2, 527–561. New York: Guilford Press.

Schwarz, N. (1998) Accessible content and accessibility experiences: The interplay of declarative and experiential information in judgment. *Personality and Social Psychology Review*, 2, 87–99.

Schwarz, N. (2002) Situated cognition and the wisdom of feelings: Cognitive tuning. In L. Feldman Barrett

and P. Salovey (eds), *The Wisdom in Feelings*, pp. 144–166. New York: Guilford.

Schwarz, N. (2004) Meta-cognitive experiences in consumer judgment and decision making. *Journal of Consumer Psychology*, 14, 332–348.

Schwarz, N. (2010) Meaning in context: Metacognitive experiences. In L.F. Barrett, B. Mesquita and E. Smith (eds), *The Mind in Context*, pp. 105–125. New York: Guilford.

Schwarz, N. and Bless, B. (1991) Happy and mindless, but sad and smart? The impact of affective states on analytic reasoning. In J. Forgas (ed.), *Emotion and Social Judgment*, pp. 55–71. Oxford: Pergamon.

Schwarz, N., Bless, H. and Bohner, G. (1991a) Mood and persuasion: Affective states influence the processing of persuasive communications. *Advances in Experimental Social Psychology*, 24, 161–199.

Schwarz, N., Bless, H., Strack, F., Klumpp, G., Ritterauer-Schatka, H. and Simons, A. (1991b) Ease of retrieval as information: Another look at the availability heuristic. *Journal of Personality and Social Psychology*, 61, 195–202.

Schwarz, N. and Clore, G.L. (1983) Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology*, 45, 513–523.

Schwarz, N. and Clore, G.L. (2003) Mood as information: 20 years later. *Psychological Inquiry*, 14, 296–303.

Schwarz, N. and Clore, G.L. (2007) Feelings and phenomenal experiences. In A. Kruglanski and E.T. Higgins (eds), *Social Psychology. Handbook of Basic Principles*, 2nd Edition, pp. 385–407. New York: Guilford.

Schwarz, N., Sanna, L., Skurnik, I. and Yoon, C. (2007) Metacognitive experiences and the intricacies of setting people straight: Implications for debiasing and public information campaigns. *Advances in Experimental Social Psychology*, 39, 127–161.

Schwarz, N., Servay, W. and Kumpf, M. (1985) Attribution of arousal as a mediator of the effectiveness of fear-arousing communications. *Journal of Applied Social Psychology*, 15, 74–78.

Schwarz, N., Strack, F., Kommer, D. and Wagner, D. (1987) Soccer, rooms and the quality of your life: Mood effects on judgments of satisfaction with life in general and with specific life-domains. *European Journal of Social Psychology*, 17, 69–79.

Schwarz, N. and Xu, J. (2010) Constructing heuristics on the spot: Divergent inferences from ease of recall. Unpublished manuscript, University of Michigan.

Sedikides, C. (1995) Central and peripheral self-conceptions are differentially influenced by mood. *Journal of Personality and Social Psychology*, 69, 759–777.

Shen, H., Jiang, Y. and Advaval, R. (2010) Contrast and assimilation effects of fluency. *Journal of Consumer Research*, 36, 876–889.

Siemer, M. and Reisenzein, R. (1998) Effects of mood on evaluative judgements: Influence of reduced processing capacity and mood salience. *Cognition and Emotion*, 12, 783–805.

Sinclair, R.C. (1988) Mood, categorization breadth, and performance appraisal: The effects of order of information acquisition and affective state on halo, accuracy, information retrieval, and evaluations. *Organizational Behavior and Human Decision Processes*, 42, 22–46.

Sinclair, R.C., Mark, M.M. and Clore, G.L. (1994) Mood-related persuasion depends on misattributions. *Social Cognition*, 12, 309–326.

Slovic, P., Finucane, M., Peters, E. and MacGregor, D.G. (2002) The affect heuristic. In T. Gilovich, D. Griffin and D. Kahneman (eds), *Heuristics and Biases: The Psychology of Intuitive Judgment*, pp. 397–420. New York, NY: Cambridge University Press.

Smith, E.R. and Semin, G.R. (2004) Socially situated cognition: Cognition in its social context. *Advances in Experimental Social Psychology*, 36, 53–117.

Soldat, A.S. and Sinclair, R.C. (2001) Colors, smiles, and frowns: External affective cues can directly affect responses to persuasive communications in a mood-like manner without affecting mood. *Social Cognition*, 15, 55–71.

Song, H. and Schwarz, N. (2008) Fluency and the detection of distortions: Low processing fluency attenuates the Moses illusion. *Social Cognition*, 26, 791–799.

Song, H. and Schwarz, N. (2009) If it's difficult-to-announce, it must be risky: Fluency, familiarity, and risk perception. *Psychological Science*, 20, 135–138.

Stepper, S. and Strack, F. (1993) Proprioceptive determinants of emotional and nonemotional feelings. *Journal of Personality and Social Psychology*, 64, 211–220.

Strack, F. and Hannover, B. (1996) Awareness of influence as a precondition for implementing correctional goals. In P.M. Gollwitzer and J.A. Bargh (eds), *The Psychology of Action*, pp. 579–595. New York: Guilford.

Strack, F., Martin, L.L. and Stepper, S. (1988) Inhibiting and facilitating conditions of the human smile: A non-obtrusive test of the facial feedback hypothesis. *Journal of Personality and Social Psychology*, 53, 768–777.

Tversky, A. and Kahneman, D. (1973) Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5, 207–232.

Wänke, M., Bless, H. and Biller, B. (1996) Subjective experience versus content of information in the construction of attitude judgments. *Personality and Social Psychology Bulletin*, 22, 1105–1113.

Weaver, K., Garcia, S.M., Schwarz, N. and Miller, D.T. (2007) Inferring the popularity of an opinion from its familiarity: A repetitive voice can sound like a chorus. *Journal of Personality and Social Psychology*, 92, 821–833.

Wegener, D.T., Petty, R.E. and Smith, D.M. (1995) Positive mood can increase or decrease message scrutiny: The hedonic contingency view of mood and message processing. *Journal of Personality and Social Psychology*, 69, 5–15.

Wegner, D.M. and Vallacher, R.R. (1986) Action identification. In R.M. Sorrentino and E.T. Higgins (eds), *Handbook of Motivation and Cognition: Foundations of Social Behavior*. New York: Guilford.

Wilson, T.D. and Brekke, N. (1994) Mental contamination and mental correction: Unwanted influences on judgments and evaluations. *Psychological Bulletin*, 116, 117–142.

Winkielman, P. and Cacioppo, J.T. (2001) Mind at ease puts a smile on the face: Psychophysiological evidence that processing facilitation leads to positive affect. *Journal of Personality and Social Psychology*, 81, 989–1000.

Winkielman, P., Schwarz, N., Fazendeiro, T. and Reber, R. (2003) The hedonic marking of processing fluency: Implications for evaluative judgment. In J. Musch and K.C. Klauer (eds), *The Psychology of Evaluation: Affective Processes in Cognition and Emotion*, pp. 189–217. Mahwah, NJ: Erlbaum.

Wyer, R.S. and Carlston, D.E. (1979) *Social Cognition, Inference, and Attribution*. Hillsdale, NJ: Erlbaum.

Yeh, W. and Barsalou, L.W. (2006) The situated nature of concepts. *American Journal of Psychology*, 119, 349–384.

Zajonc, R.B. (1968) Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology: Monograph Supplement*, 9, 1–27.

Zajonc, R.B. (1980) Feeling and thinking. Preferences need no inferences. *American Psychologist*, 35, 151–175.

Zanna, M.P. and Cooper, J. (1976) Dissonance and the attribution process. In J.H. Harvey, W.J. Ickes and R.F. Kidd (eds), *New Directions in Attribution Research*, 1, 199–217. Hillsdale, NJ: Erlbaum.

Zillman, D. (1978) Attribution and misattribution of excitatory reactions. In J.H. Harvey, W.I. Ickes and R.F. Kidd (eds), *New Directions in Attribution Research*, 2, 335–368. Hillsdale, NJ: Erlbaum.

The Linguistic Category Model

Gün R. Semin

ABSTRACT

How did a model of the distinctive terms (predicates) that we use to describe interpersonal events emerge? What were the circumstances that coalesced different inquiries into a model of interpersonal language and how did its applications prove that this was to become an important model of the cognitive properties of language. These are the types of questions that are answered in this chapter, along with how the significance of this model was underlined by its application to how people use language strategically when describing positive and negative behaviors of their ingroups and outgroups. The linguistic category model, which was developed jointly with Klaus Fiedler, has found many applications in answering both fundamental and applied questions. One of the central fundamental issues that this model has revealed is to reformulate the language–cognition interface into a functional: “what is language *for*?” instead of the classic formulation: “what is language?” The answer to the functional question is: “to direct attention.” This formulation opens new vistas for research, such as how people, when pursuing inquiries in question–answer contexts, formulate their choices of predicates that reveal their biases and shape their interviewees’ answers. New directions that can be explored by the linguistic category model are discussed in conclusion.

INTRODUCTION

Language is the most pervasive feature of our social environment. It is a tool that we use to implement our goals, intentions, and desires (Semin, 2000a, 2000b). Language extends our capabilities for action irrespective of whether this involves ordering a drink at McDonald’s, engaging in a detailed exchange with a sommelier or reviewing a paper, to navigating our ways in a new town with the aid of street names to negotiating our identities in a new relationship. It is undoubtedly the case that there are other incredibly rich nonverbal resources that we recruit consciously and unconsciously in situated interaction (cf. Semin, 2007) but language places us in a world that allows us to travel in time both to the future and the past, to create imaginary worlds that we can share with others, lives that can give us joys and horrors. Language has the potency of ripping us out of our immediate reality and throwing us into different worlds, which can become powerfully possessive realities in their own right.

It is therefore not surprising that at some point of their life histories people take a reflective stance upon what language is, means, and does to us. Some of us experience the significance of language only when they visit a country and do not understand a single word. Others enter with an intense interest as either avid readers or writers. Some start as chemical engineers, and end up as linguists proposing intriguing hypotheses such as: do our habitual ways of using language shape the way we think, or vice versa (Whorf, 1956).

It is not surprising in my case that language has always constituted a reality that I had to reflect upon because much like the analogy of the fish who realizes the significance of water after landing on the sand, I have been thrown for extended periods from one linguistic community to another and learned that language was important. Indeed, playing with language and writing used to be one of the activities I enjoyed considerably during my high school period. At a very early age, I spent a year in the States and learned English and kept it up fluently after returning to my hometown, Istanbul, because I continuously had British or American friends. I developed a very good working knowledge of French, because my parents spoke fluent French and did so whenever they had a secret to hide from me. They had lots of secrets and I learned fast. I was then sent to the “Deutsche Oberrealschule” in Istanbul and acquired German, and ended up studying psychology in Germany. After diverse extended stays in different European research centers while firmly anchored on British soil at Sussex University, I arrived in the Netherlands where I acquired Dutch. So, I had to learn to swim in diverse linguistic communities and one would think that that is a good reason to reflect on language. However, I was only interested in language as a tool that had to be mastered and masterly applied in communication until I was well into my PhD research.

It was the research problems I encountered during my thesis research that began to direct me towards language. My PhD was on

research that was remotely related to language, namely group decision-making and risk-taking. Trying to understand the dynamics of this question; namely, why do groups make riskier decisions than individuals, I discovered that the question was not quite right. I stumbled upon the semantics of the stimulus items that constituted the subject for the decisions that groups had to make. Indeed, after content-analyzing group discussions and systematically varying different words in the formulation of the dilemmas, it became evident that specific terms in the stimulus items would anchor the focus for discussions. In a dilemma where the group had to advise on whether, for example, an electrical engineer should take a risky job or not, a simple adjective such as “young” was sufficient to focus the discussion and lead to a riskier decision since the group generated a set of beliefs and arguments that were associated with being “young” and were related to risk-taking. If that single word in the dilemma was modified to “old” then the group was likely to generate arguments against taking a risk. These words “primed” the discussion and the direction of the group discussion and decision. This realization, namely that a single word had such a powerful effect in directing the focus of a discussion became important in channeling my attention to the significance of language and raising questions about what role, if any, was played by group processes, since I could also get individuals to generate arguments and produce the same “shifts” (or polarization as it was later labeled) as groups did.

I wanted to learn more about semantics and some of the research I started conducting at the time necessitated this. So, I consulted Gerald Gazdar who was also recently appointed at Sussex University. We had diverse discussions about numerous things, but if my memory serves me well, he recommended not wasting time reading anything on semantics and that I should get on with it on my own, because there were as many semantic theories as there were people writing on it.

At the time, I took his advice to heart and proceeded with some research which was designed to uncover the degree to which different psychological theories were actually anchored in semantic regularities (e.g., Semin and Chassein, 1985; Semin and Krahé, 1987, 1988). To the extent that the semantic regularities that one discovered mirrored the regularities proposed by theories in personality or otherwise, then the question became: what is the psychological component of the theory? I think that a study I published in 1989 illustrates the nature of my research interest during this period perfectly and reveals that the properties of language can be mistaken for properties of psychological processes. The research I am referring to was a replication of Asch's classic impression formation studies (1946). The method I used to uncover the misattribution of regularities (Semin, 1989) relied on an analysis of a synonym and antonym dictionary and involved computing the semantic overlap between the different stimulus sets (including warm, cold, polite, blunt) and the response scales used by participants in the studies. By developing an index of the strength of semantic associations between the composite stimulus sets and the response items, I was able to show that with the distinctive patterns revealed by actual Asch data could be replicated simply by using a dictionary and no subjects. There was a 90 percent overlap between the dictionary-derived index and the Asch data. This finding suggests that our studies are often examining regularities of the types of tools that we use – particularly language – rather than regularities due to psychological processes.

It was around the time that this research was conducted that my then PhD student, Liam Greenslade, and I (Semin and Greenslade, 1985) became interested in the so-called "systematic distortion hypothesis" by Rick Shweder (e.g., Shweder and D'Andrade, 1980). His contention was that inferences about personality are based on a systematic memory bias whereby co-occurrence frequencies between properties and behaviors of a person are substituted by *similarity*

propositions. Liam Greenslade and my take on this supposed general cognitive bias contention for memory-based ratings of persons was that it did not account for functional differences between different types of linguistic forms and that while the argument may hold true for certain linguistic forms it was unlikely to hold true for others. At the time, we made a distinction between *immediate terms* (e.g., verbs, behaviors) and *mediate terms* (e.g., adjectives, traits). This in fact was the turning point in my thinking in that the research we did reflected our realization that different linguistic forms fulfilled different functions in the description of behaviors and persons. As we showed, while the application of adjectives (mediate terms) was driven by the conceptual relationships (semantics) between the adjectives and were unaffected by situated features that were also relevant, this was not the case for verbs (immediate terms). The use and application of verbs displayed considerable systematicity as a function of the situated features in which people found themselves. These were sensitive to the contingencies of the environment unlike the proposition advanced by the systematic distortion hypothesis.

This research (Semin and Greenslade, 1985) constituted a turning point for me. I moved from my focus on investigating whether assumed psychological processes were merely reproductions of semantic relationships found in everyday language to developing an increased interest in the metasemantic properties of language. The person who was responsible in accelerating this perspective was my friend and colleague Klaus Fiedler with whom we had extended discussions on the subject. The outcome of different ideas turned out to be the paper, which is the metasemantic model of interpersonal language, namely the linguistic category model (LCM). Our discussions focused on specifying the crude taxonomy that Liam Greenslade and I had advanced by, first, limiting the domain to interpersonal language and, second, refining what we had termed "immediate terms." Around this time, Brown and

Fish's (1983) fascinating paper on the causality implicit in language had also appeared and a visiting American colleague with whom we had discussed our ideas drew our attention to it. Reading this paper influenced our thinking and shaped the way we developed our taxonomy of interpersonal language.

We developed our ideas to investigate the linguistic devices that are used in the description of social events and persons – the actors in such situations. The question was: What are the distinctive properties that such devices have? This resulted in what is now known as the LCM (Semin and Fiedler, 1988, 1991) which is a classificatory approach to the domain of interpersonal language and consists of interpersonal (transitive) verbs that are tools used to describe *actions* (help, punch, cheat, surprise) or psychological *states* (love, hate, abhor) and *adjectives* and *nouns* that are employed to characterize persons (e.g., extroverted, helpful, religious). The question addressed by the LCM is: what are the distinctive interpersonal lexical units that are available and are there systematic cognitive properties that they share in differing degrees?

The idea of examining language as possessing cognitive properties in the late 1980s was an unusual one and indeed the publication of the first paper on the LCM took an extensive reviewing and revision process of over two years until it was accepted for publication by the then editor-in-chief of the *Journal of Personality and Social Psychology*, Jim Sherman. The fortunate aspect of this extended scholarly discussion was that the reviewers and editor were constructively involved in drawing attention to ways of improving the taxonomy, its criteria, and the clarity of the presentation. We were very fortunate in Jim Sherman's informed selection of reviewers, of whom one – we think – was Roger Brown, and the other was Bill McGuire. There were other constructive reviewers but our evidence was not strong enough to infer their identity. And, yet at the time when the information processing based social cognition tradition was at its height, presenting a paper about the cognitive properties of language was

unusual. So, let me now turn to the LCM to detail the nature of this taxonomy, which was driven by an idea that language is a tool.

THE LINGUISTIC CATEGORY MODEL

This model of interpersonal language furnishes the means by which it is possible to identify the nuances of how people use interpersonal terms when they are representing social events in communication, and thus is informative about how verbal behavior is driven strategically by psychological processes and communication constraints. This is made possible by providing a systematic model of the linguistic terms (verbs, adjectives, and nouns) that we use in communicating about social events and their actors.

In this model, a distinction is made between five different categories of interpersonal terms, namely Descriptive Action Verbs (DAV), Interpretative Action Verbs (IAV), State Action Verbs (SAV), State Verbs (SV), and Adjectives (ADJ) (cf. Semin and Fiedler, 1991). The distinction between the categories is obtained on the basis of a number of converging grammatical tests and semantic contrasts (cf. Bendix, 1966; Brown and Fish, 1983; Miller and Johnson-Laird, 1976).

DAVs are the most concrete terms and are used to convey a description of a single, observable event and preserve the perceptual features of the event (e.g., "A punches B" whereby punching is always achieved by means of a fist). The second category (IAVs) describes specific observable events. However, these verbs are more abstract in that they refer to a general class of behaviors and do not preserve the perceptual features of an action (e.g., "A hurts B"). SAVs and SVs refer to psychological states while DAVs and IAVs do not. SAVs refer to the affective consequences of actions that are not specified any further (to amaze, surprise, bore, thrill, etc.) but can be supplied when asked (e.g., "Why was she surprised?"). SVs typically describe an unobservable affective or mental

state and not a specific event (e.g., “A hates B”). Finally, adjectives (e.g., “A is aggressive”) constitute the most abstract category. These generalize across specific events and objects and describe only the subject. They show a low contextual dependence and a high conceptual interdependence in their use. In other words, the use of adjectives is governed by abstract, semantic relations rather than by the contingencies of contextual factors. The opposite is true for action verbs (e.g., Semin and Fiedler, 1988; Semin and Greenslade, 1985). The most concrete terms retain a reference to the contextual and situated features of an event. This dimension of abstractness-concreteness of interpersonal predicates has been operationalized (e.g., Semin and Fiedler, 1988) in terms of a number of different inferential features or properties with DAVs (hit, kiss) being the most concrete category. IAVs (help, cheat) are more abstract. SVs (like, abhor) follow next, and ADJs (friendly, helpful) are the most abstract predicates.

It is important to note that the properties by which abstractness-concreteness had been operationalized are *generic to the entire predicate classes* represented in the LCM. The types of meanings or implications as defined by the distinctive inferential properties of the LCM are different from the more conventional study of meaning, namely semantics. The more conventional approaches in linguistics are the study of meaning in terms of *semantic fields*, *semantic relations*, or the analysis of lexical items in terms of *semantic features* to investigate the semantic component of a grammar’s organization. One may refer to the meaning domain identified by the LCM as meta-semantic since the inferential properties apply across semantic fields and are also distinctive in that they escape conscious access (Franco and Maass, 1996, 1999; Semin, 2006; Von Hippel et al., 1997).

LCM: The early applications

The development of the LCM was followed by a rich series of experimental studies

which were all concerned with examining systematic use of the different predicate categories and had an eye to either explaining existing phenomena as a consequence of strategic language use (e.g., actor-observer discrepancy and fundamental attribution error, Semin and Fiedler, 1989), or introducing new phenomena that resulted from systematic differences in language use. A superb example for the latter is the research field initiated by Anne Maass on strategic language use by members of in- and outgroups (Maass, 1999; Maass et al., 1989).

Design processes for stereotypes

What are the “design processes” involved in communicating about stereotypes? Which lexical categories of the LCM do we choose when we are conveying events involving members of our in- or outgroups? This was the issue that was addressed by Anne Maass (1999 for a review) and her colleagues. Stereotypes are the emergent results of socially situated interactions between individuals, rather than a product that resides within the head of an individual, an idea whose roots can be dated in Lippman’s (1922) famous “pictures in the head” metaphor. The research on strategic language use and stereotypes, initiated by Anne Maass and her colleagues, on how language is used as a tool reveals that the way we represent events involving in- and outgroups is a dynamic design process driven by cognitive and motivational factors.

The linguistic intergroup bias (LIB, Maass and Arcuri, 1992; Maass et al., 1989, 1995) involves a linguistic strategy for individuals to describe positive ingroup and negative outgroup behaviors in relatively abstract terms (e.g., adjectives), implying that the behavior in question is due to enduring dispositions or the actor’s stable characteristics. Conversely, negative ingroup and positive outgroup behaviors are typically described in relatively concrete terms, implying the incidental or situational specificity of the

behavior, and hence an external attribution of the behavior.

Two mechanisms have been postulated to account for these systematic differences that people display when they are talking about positive and negative out- and ingroup behaviors. One possible mechanism of the LIB could be motivational (Maass, 1999; Maass et al., 1989) having to do with the fact that abstract descriptions of positive ingroup behaviors and of negative outgroup behaviors portray the ingroup in favorable and the outgroup in unfavorable terms implying that these behaviors are due to enduring characteristics. Similarly, concrete depictions of negative ingroup behaviors minimize their significance as evidence for corresponding group characteristics, as do concrete depictions of positive outgroup behaviors. In other words, those linguistic strategies serve to protect the perception that the ingroup is superior to the outgroup. Another mechanism to account for these patterns of language use is provided by a cognitive or expectancy account which states that expected behaviors are described with abstract language and unexpected behaviors by the use of concrete predicates (e.g., Rubini and Semin, 1994). Both processes appear to be operative depending on the motivational circumstances under which the strategic language is produced (Maass, 1999; Maass et al., 1995). The expectancy driven mechanism has been termed the linguistic expectancy bias (LEB; Wigboldus et al., 2000).

The extensive research on the LIB/LEB extends the language social cognition interface, by showing precisely how both cognitive and motivational processes systematically influence language as a tool that renders these processes as action. Indeed, subsequent research investigating the impact of these messages upon third parties has shown that third parties draw the implications for which the messages are designed (e.g., Werkman et al., 1999; Wigboldus et al., 2000). Extending this line of thinking, and building on their earlier work (2003), Douglas and Sutton have shown that linguistic choices

implicitly convey to listeners of messages the type of relationship that holds between the producers of such messages and the targets of their message (Douglas and Sutton, 2006).

The communication cycle between psychological processes, message production, and message impact is closed in a series of studies reported by Reitsma-van Rooijen, et al. (2007). These authors have shown that when participants receive messages that are either concretely or abstractly formulated about their *own* positive and negative behaviors (e.g., success or failure) from another person, then this affects their perceived social distance to the sender of the message. Similar to the findings reported by Rubini and Kruglanski (1997) in an entirely different context, abstract messages about positive behaviors and concrete messages about negative behaviors produced perceived social proximity. In contrast, receiving concrete messages about positive behaviors and abstract messages about negative behaviors led to judgments of social distance to the sender of the message.

There are different ways of looking at these studies that reveal strategic language use in intergroup and interpersonal contexts. These studies take the language-cognition interface outside of the boundaries of the individual "mind" and investigate the effects of psychological processes upon strategic language use in the first instance. How do more broadly conceived cognitive and motivational processes influence the "design process" people use when producing a message or utterances? The next step is to investigate the impact of such messages. This can be done in a number of ways as mentioned earlier. How do neutral participants decipher what is in a message? How do involved parties whose doings are talked about react to such senders? More importantly, how does language use contribute to the regulation of social behavior?

The language-cognition (or rather psychological processes) question acquires an entirely different complexion as can be seen

in the studies I summarized above. The question becomes an investigation of how psychological processes contribute to the design of language use, and consequently how different utterances as the result of design processes direct attention to different facets of the same event, thereby influencing the way a receiver perceives the same event and how such a process contributes to the regulation of social interaction. This constitutes a major shift in the way the language question has been posed in psychology. Instead of posing the question “What is language?” and examining the language psychological process interface as interconnected interpsychological inquiry, the question becomes, “What is language for?,” and transforms our approach to language as a tool that is used to extend cognition into communication. This is the transition that the research on the linguistic intergroup bias and other related research with the LCM introduced in my thinking.

Thus, the central mission that emerged from this research led to a reformulation of the classic language–cognition interface question that has been with us for a very long time. The contribution of a functional take on language, namely, “What is language for?,” leads to a translation of the very same classic puzzles into contextually situated problems that open increasingly inclusive research challenges. The central mission that emerges is interfacing psychological processes within a communication context with a view towards developing a better understanding of the regulation of social behavior. Thus, before proceeding with how the work on the LCM transformed my thinking and the future output from this model, I shall detail the distinction between the classic and functional approach towards understanding the interface between language and cognition.

What is language (for)?

There are many questions that one can pose to address the different puzzles that language presents, one of the most distinctive

human artifacts. Not surprisingly, the question one asks shapes the nature of the answer, and nearly without exception, this *a priori* determination of an answer by the question escapes our awareness. Every question is the result of a set of presuppositions, and it is no different when questions about language are formulated. Very subtle differences may have substantial consequences. We can ask what language *is*, or alternatively, what language is *for*. The answer to the former would treat language as an entity with an existence independent of its “natural habitat” (communication), namely an object that is timeless, subjectless, and with an internal logic. The latter formulation, which introduces an apparently minor modification, namely what is language *for*, all of a sudden changes the nature of our inquiry dramatically, and our interest shifts substantially from an inquiry about a decontextualized abstract entity to a live and active instrument serving diverse purposes in a dynamic interpersonal context. Indeed, these two questions about language represent the classic and the more recent approaches towards understanding language and their psychological implications. In the following, I shall present these two perspectives in turn and conclude with some issues about the implications arising from the second question.

What is language?

“What is language?” typifies the traditional approach to language – a view that still informs much research in social cognition. In this view, language is a disembodied structure and dissociated from real time. Language is regarded as a set of symbols and “rules” about how to combine them and both are “virtual and outside of time” (Ricoeur, 1955). Language is considered in the abstract, and “without a subject.” This lens presents language as an extra-individual and systematic set of abstract properties and a life of its own. Consequently, it is “subjectless,” “timeless,” and disembodied. This assumption enables the examination of language in terms of the relationships between its specific properties

in detached abstraction (e.g., lexical semantics, grammatical categories, word order).

Such an approach to language does not lend itself readily for a *social* examination of how language interfaces with psychological processes, but it does pair with a view of cognition, which treats cognition in a manner comparable to the way language is treated in the “What is language?” perspective. In the classic (and largely current) approach, cognitive processes are also thought of in a disembodied, timeless, and subjectless manner. In this view, both language and cognition are separated from real time.

Detaching both language and cognition from situated action has led to a perspective where both are located in and examined from an *individual perspective*. The central question then becomes the mutual influence between language and social cognition. This mutual influence is understood in terms of two sets of “inner representational” systems, rather than interaction in a social world (e.g., dialogue, etc.). Given these assumptions, the interface between language and social cognition becomes an investigation of the interaction between two different *representational* modules in an intrapersonal “forum,” with social cognition referring to individual processes – encoding, representing, thinking, retrieving, and so on. The classic and ongoing debate on linguistic relativity (Whorf, 1956) in psychology is located precisely at the center of this classic viewpoint. The research agenda was conceived by Lenneberg and his colleagues (Brown and Lenneberg, 1954; Lenneberg, 1953; Lenneberg and Roberts, 1956). The central question: *Does the structure of a given language affect the thoughts (or thought potential), the memory, the perception, the learning ability of those who speak that language?* (Lenneberg, 1953: 463, italics here). This early crystallization of how to tackle the topic defined and determined the research focus on the interface between language and social cognition, and it is current in both psychology and social psychology: *Do linguistic structures influence nonlinguistic categorization, memory,*

perception, attention, thinking, and so on? The research questions are seated in the “mind” and the language and social cognition relationship becomes an analysis of the construction and manipulation of inner representations. The more recent versions of the view also share the same individual centered representational perspective (e.g., Hardin and Banaji, 1993; Hoffman et al., 1986; Hunt and Agnoli, 1991). The research question can be illustrated with the following example. If two cultures code the color spectrum differently in their respective linguistic communities, do they then perceive and represent colors incommensurably or not (cf. Özgen, 2004)? Not surprisingly, thinking of language and cognition in this way leads to the fascinating and classic issues that have occupied many scholars about the relationship between language and cognition, and their mutual influence (cf. Whorf, 1956).

What is language? Language locked in the mind

An interesting instance for the case of “What is language?” was provided by a “phenomenon” that has attracted considerable research interest in social cognition and termed by Brown and Fish (1983) “the *psychological causality implicit in language*.” As I shall note shortly, this is a much broader field than the focus that Brown and Fish have introduced. The question that the research in this field attempted to answer was: What types of systematic “inferences” do transitive verbs that describe interpersonal events (e.g., to help, to cheat, to respect, to surprise) convey?

The research emerging from different strands has focused on the inferences that people make as a function of the verb type in minimal sentences (e.g., “subject–verb–object” sentences such as “John helps David” or “John likes David”). The difference between the two example sentences is the verbs the first referring to actions (e.g., hit, help, cheat) and the second to states (e.g., like, respect, hate). This research has revealed a multitude of different types of inferences

over a period of more than 40 years (see, for detail, Semin, 2007). Indeed, the origin of this question can be found in a research program initiated by Abelson and his collaborators (e.g., Gilson and Abelson, 1965, Abelson and Kanouse, 1966) much before Brown and Fish's work (1983). The phenomenon was investigated independently in linguistics with a "pronoun disambiguation" paradigm (e.g., Caramazza et al., 1977; Ehrlich, 1980; Garvey and Caramazza, 1974; Garvey et al., 1976). Another perspective focused on the role that interpersonal verbs play in *text comprehension* (e.g., Garnham and Oakhill, 1985; Garnham et al., 1992, 1996; Greene and McKoon, 1994; McKoon et al., 1993).

The interest in interpersonal verbs has focused on the systematic and different inferences mediated by *action* and *state verbs* has revealed a range of inferences varying from the assignment of cause or agency to the event (e.g., Brown and Fish, 1983), different inductive-deductive inferences (e.g., Gilson and Abelson, 1965), implicit quantifiers (Kanouse, 1972), differential salience of sentence subject and object (Kasoff and Lee, 1993), differential disambiguations of personal pronouns (e.g. Garvey et al., 1976; Garvey and Caramazza, 1974), different foci of what preceded and antecedent the event in the stimulus sentence (Au, 1986; Fiedler and Semin, 1989), *inter alia*. The list is long (cf. Semin, 2007 for an overview).

The nature of these findings point to a set of paradoxes that emerge if one approaches the language-cognition interface in this domain with a "What is language?" perspective. The sheer fact of the great number of demonstrated systematic inferences mediated by interpersonal verbs poses a number of problems. One of these has to do with the assumptions about the supposed psychological processes (inferences) that are "activated." If all of the recorded systematic inferences were automatic and autonomous as they are assumed to be (e.g., Brown and Fish), and were to be coactivated, then the sheer cognitive load that this would introduce would in all likelihood bring the

"cognitive apparatus" to a grinding halt and lead to a rather pathological state in the subject. This problem arises due to the assumption that both language and cognition are intrapersonal processes whereby a function perspective, and consequently the adaptive action implications of both language and cognition, is absent. This is not to deny the empirical reality of the findings, but to question their conceptual grounding. The more interesting question that arises from a situated and functional perspective is what the multitude of systematic inferences mean.

The tacit treatment of language and cognition as inner representational systems rather than being "for" adaptive interaction with the world renders the puzzle of inferential processes an individual centered one. Language and cognition are assumed to "happen" *within an individual*. Language and cognitive processes remain disembodied, timeless, and subjectless. These findings do not make much sense if seen from a traditional perspective on language and cognition, which is not informed about a communicative or interpersonal context, and the chief function that language serves.

The paradoxical problems that this type of conceptual framing faces arises as a consequence of the traditional or geocentric view that dismisses the inherently variable processes that characterize the adaptive flexibility of cognition as well as language as a property that emerges in the interaction between the organism a social environment (Smith and Semin, 2004), and the function of language as a tool in extending cognition in real or imagined communicative contexts (cf. Semin, 2000b, 2001, 2007). As I have elaborated elsewhere (Semin, 1998, 2007) the varied experimental scenarios in which the different inferences are revealed constitute different situations where the dependent variables (about causal agency, inductive or deductive inferences, etc.) constitute different situated contexts that demonstrate the malleability of language rather than some underlying psychological process driving different types of inferences. To understand what this means

more specifically, we have to adopt a different more heliocentric view of language. This invites the question: What is language for?

What is language for? Language as a tool for action

Posing the question, “What is language for?,” introduces an alternative way of looking at this relationship, namely a functional view by considering the language–cognition interface in a *language use context*. In this functional view, language is treated as a *tool* and a means to extend cognition in the implementation of action (e.g., Semin, 2000a, 2001). From this perspective, language is for use, and in general terms language use is a *design process* that extends cognitive and motivation processes of a speaker with a view to direct the attention of a listener to some aspect of social, physical, or psychological reality. In other words, language is used in a communicative context with a view to structure the cognition of an addressee by driving his or her attention. Obviously, this is an interactive process and not unidirectional. Seen this way, cognition can refer to both (1) processes which contribute to how a speaker shapes a communicative act (production processes), (2) to those processes that contribute to how a communicative act (a message) is received by an addressee (comprehension processes), and (3) the entirety of communication itself, namely independent of the individual productions, as a regulator of joint action (see Hutchins, 1996).

This way of conceptualizing language introduces a way of thinking about language that is somewhat alien to how it is conceptualized traditionally in psychology. It effectively means that *language* – as a tool – *has cognitive properties*. Treating language as a tool (Semin, 1998) means literally that: different linguistic units and their conventionalized combinatorial possibilities constitute devices to do things. The distinctive characteristic that language as a tool shares with devices in general is that tools have evolved to meet the dual demands involved in solving a problem by adapting to the features or

properties of a particular task (e.g., cutting a piece of paper) as well as the constitutional features of human capabilities and body (e.g., the shape of the hand). This *dual adaptation* gives us capabilities or powers to do things that we do not have by nature. For instance, a pair of scissors consists of two blades with a ring shaped handle, which is adapted to the features of a handgrip. These blades are so pivoted that their sharp edges move one against the other, which is the perfect adaptation to perform a neat cutting operation on paper or materials. The distinctive feature of any tool, as in the case of a pair of scissors, is that *knowledge is downloaded onto it*. The way it has been engineered, the sharp blades, the way the blades are pivoted, the rings on the blades constitute a remarkable achievement in terms of facilitating the interface or the coupling between the human constitution and the human goal. The distinctive feature of any device is that it is embodied (Semin and Smith, 2008). Tools are shaped by the constraints of human biological constitution on one hand and the types of adaptation that the tasks at hand require on the other. This way of looking at language takes it from the exclusively representational perspective and grounds it into an embodied framework. Language as a *tool* contains functional knowledge and the type of “task” faced in a communication context introduces a “design process,” which results in unique communicative acts to resolve the task. One particular implication of this view is to examine the types of tools and their characteristics in terms of the knowledge downloaded onto them.

What is language for? Driving attention!

Any social event can be represented in a number of different ways with a number of different linguistic devices that give different nuances to the aspect of the reality to which we would like to draw the listener’s attention. Concretely, this means to draw attention to specific aspects of the physical, psychological, and social environment in communication (e.g., an event whereby David’s fist

connects with John's jaw). To do so, the speaker has a variety of options in terms of the devices that she can use to draw attention to the very same reality (e.g., He punched John; He hurt John; He hates John; He is aggressive). Particular psychological processes and the social context drive the choice of specific lexical categories and contribute to the design of communicative acts (utterances) directing attention to different aspects of reality. Accordingly, different linguistic devices (e.g., punch, hurt, hate, aggressive) serve different perceptual functions.

This particular take on the relationship between language and cognition assigns different roles upon language and cognition. Language becomes the tool by means of which cognition is implemented in action whereby language functions as an attention-driving device in an interpersonal context.

I shall now turn to another application of the LCM – to illustrate the “What is language for?” perspective. This application is the examination of how “interpersonal verbs” can be used in the formulation of questions and how they shape answers by drawing attention to different thematic foci and contribute to the shape of inferences.

LCM: Applications in interview situations – the question–answer paradigm

What are the implications of a “What is language for?” perspective for the paradoxical literature on interpersonal verbs referred to above? What is the consequence of taking interpersonal verbs out of representational contexts as mentioned earlier, and inserting them into an experimentally induced dialogical context? Dialogical contexts are situated and thus driven by specific goals or tasks. Take a specific event, such as witnessing a “sailor at La Scala to see a soprano.” One can *direct a listener's attention* to different facets of this event by choosing one's verbs as well as who occupies the sentence subject or object. For instance, we may say “Why does

the sailor *visit* the soprano?,” or “Why does the sailor *like* the soprano?,” thereby changing the focus of attention to the sentence subject or object depending on the respective sentences used. Saliency, but also respectively agency, for the event is thus modified by the choice of verb in the question and indeed it has been shown that answers to such questions focus respectively on the sentence subject (question with action verb) or object (question with state verb) (Semin, 2000b; Semin et al., 1995). While the sailor is in the semantic role of agent in the former, he occupies the semantic role of experiencer in the second sentence. Alternatively, we can formulate questions that have been used in prior research. Posing the question: “How likely is it that sailors like sopranos?” is inductive generalization. The question “How likely is it that the sailor will like the Maria Callas?” is deductive generalization, and so on. Thus, given the very same event (stimulus sentence), the formulation of the sentence (in these particular instances the question) is used as an attention-driving device. This formulation “compels” the respondent with the implicitly operating rules of conversation (Grice, 1975) and obliges her to follow the social contract that is issued – thus having to produce the “appropriate” answer. The appropriate answer is driven by the conventional cognitive properties of interpersonal verbs (e.g., Semin and Fiedler, 1988) and these are not determinate because language is not a determinate tool. As von Humboldt observed, language “makes infinite use of finite media” (1836: 70) whose “synthesis creates something that is not present *per se* in any of the associated constituents” (1836: 67). The point of this exercise is to simply demonstrate that the type of question focuses attention on a distinctive feature or aspect of the social event observed and drives the answer of the recipient of the question.

While the above considerations remain thought experiments, there is a collection of research that has investigated the attention-driving or thematic-focusing function of interpersonal verbs (action and state verbs) in

experimentally situated contexts. The two types of interpersonal verbs function as structural devices that differentially focus attention (Stapel and Semin, in press) upon either the sentence subject or object in the question. This function of interpersonal verbs is what has been termed by some (e.g., Kay, 1996) “perspectivization” or “topicalization” (Fillmore, 1968). This attention-driving function of interpersonal verbs and their psychological implications has been examined in an experimentally induced communication paradigm termed the “question–answer paradigm” (QAP; Semin, 2000b). This paradigm was introduced by Semin et al. (1995) who systematically manipulated the verb in question formulation and examined whether these two verb types (action versus state verbs) systematically bias a target’s answer as well as the implications of this answer (cf. De Poot and Semin, 1995; Rubini and Kruglanski, 1997; Semin and De Poot, 1997a, 1997b). The research by Semin et al. (1995) indicated that questions formulated with action verbs (e.g., “to help,” “to write”) cue the logical *subject* of a question as the causal origin of answers. Questions formulated with state verbs (e.g. “to love” or “to like”) cue the logical *object* of a question as the causal origin for answers. Thus, if asked such simple and mundane question as, “Why do you *own* a dog?” (using an action verb) participants are prompted to respond by referring to themselves (the subject of the question) as the causal agent in the answer; for example, by stating, “Because *I* enjoy the companionship that dogs provide.” If one is asked “Why do you *like* dogs?” however, one is prompted to respond by referring to the object itself, for example, “Because *dogs* are good companions” (e.g., Semin and de Poot, 2007a). Furthermore, the abstractness level of questions was also shown to influence the abstractness level of the answers (cf. Semin, 2000b, for a review). Thus, more abstractly formulated questions tend to elicit more abstract answers. Interestingly, Rubini and Kruglanski (1997) set out to investigate the further implications of such differences in

abstractness not only for how such verbs steer thematic focus but also the implications of dialogue sessions controlled over experiments for perceived interpersonal proximity and distance. They did so by examining if individuals had the impression that they disclosed more about themselves when asked questions formulated with action-verbs (concrete) as opposed to state-verbs (abstract) as well as examining the moderating role of need for closure. In their first experiment designed to investigate these issues, Rubini and Kruglanski (1997) had participants under high (versus low) need for closure (operationalized via ambient noise) rank order questions out of a list in terms of their likelihood of using them in a real interview. The list included 32 questions, eight questions on each of four different topics. It was found that participants under noise (versus no noise) assigned higher ranks to questions characterized by higher (versus lower) level of abstractness. In a follow-up study, questions selected by participants under high (versus low) need for closure were found to elicit more abstract answers from respondents, and ones focused more on the logical *object* (versus *subject*) of the question. In addition, respondents reported that they felt less friendly toward the interviewer whose questions were more (versus less) abstract. Finally, in a third study the results of the previous two experiments were replicated in a free-interaction context. Interviewers with high (versus low) need for closure asked more abstract questions, which in turn elicited more abstract answers and ones focused more on the logical *object* (versus *subject*) of the question, and elicited lesser friendliness from the interviewee. These results suggest that the permanence tendency induced by the need for nonspecific closure may affect the level of linguistic abstractness, and in so doing may imbue the nascent social relations among conversation partners.

In this section, I reviewed some of the research that reveals the attention driving properties of interpersonal verbs in terms of both the types of answers they induce when

used in question formulation as well as their psychological consequences for interpersonal relationships. This stands in contrast to the traditional or geocentric approach in which the interpersonal context of language is absent. Admittedly, the traditional approach has yielded a wealth of findings and yet the very nature of the findings gives rise to internal inconsistencies and paradoxes as I noted earlier.

CONCLUSION

The chief point that emerged in the pursuit of the different ways in which the LCM can serve research was to shape a different way of formulating the relationship between language and psychological processes in general and cognition in particular. The driving idea that emerged in the course of the research that I and others engaged in was to understand the issues that connect language and human function in terms of: “What is language for?,” namely its function. This perspective was developed from the following arguments.

- Language is a tool.
- As any tool it is characterized by dual adaptation, namely to (1) the specific communicative tasks at hand, and (2) the human constitution.
- As a consequence of dual adaptation, language provides an ecological niche that contains information.
- The main function of language is to implement psychological processes (cognitive and motivational) in action.
- The composition of utterances (Semin, 2006) is the result of a “design process,” which is driven by psychological processes, situated demands, and conversational rules.
- The main function of language use is driving attention.

In closing this contribution, I shall elaborate briefly upon the last argument above, which is probably the most important argument deriving from the perspective adopted on language and psychological processes.

In a sense, the last argument may be self-evident, particularly since everyday examples of strategic use of language, in other words *political spin*, have become an integral part of our folklore. Political spin is in principle the strategic use of language in order to attract attention or detract attention from certain facts of life. Let me briefly illustrate. One can label specific people as *freedom fighters* and thereby direct attention to specific aspects of these people including sowing the associated seeds of images that have to do with virtue, self-sacrifice, patriotism, which help go beyond the “label” given – to paraphrase Jerome Bruner. By labeling the very same people *terrorists* we conjure entirely different images and associations (cruel fanatics, possessed psychopaths, etc.). This is probably the obvious attention driving function of politically driven design processes and one that operates at a semantic level.

At this semantic level of analysis, the attention of the receiver is driven by specific features and properties associated with a category label. This is perhaps an obvious point from the traditional approach to language (What is language?). Language directs people’s attention and perceptual focus, and different linguistic devices direct attention to different aspects of reality. This is an idea that is also at the core of Whorf’s (1956: 221) linguistic relativity hypothesis, which suggests that the use different “grammars” direct people to different types of observations and evaluation of events. But what is distinctive about this assumption is that the focus is upon the content of attention driving which is more about associations and specific topics, themes, or beliefs. Moreover, the entire process is supposed to be played out in a person’s head. This is a *surface-level process* driven by lexical meaning or meaning as moderated by the “rules” governing the permissible word compositions in utterances. This is in fact the way the classic work on the language–cognition interface has proceeded and produced considerable insights by comparative studies across linguistic communities. Examples are abundant, such as the types of lexical categories

that are available to describe persons in English and Chinese (e.g., Hoffman et al., 1986); the domain of language responsible for spatial locations (Majid et al., 2004), variations in the availability of basic color terms across a diversity of linguistic communities (e.g., Berlin and Kay, 1969); differences between languages in gender marking (e.g., English versus Turkish) or pronoun drop (e.g., Kashima and Kashima, 2003), or lexical categorizations emotional states (Semin et al., 2002). Notably, all these cultural differences are domain specific and can be seen as culturally marked manners by which attention is driven to specific aspects of the social, physical, and psychological reality.

The pursuit of the implications furnished by the LCM drove me to a different perspective on this research. If language is an attention-driving device, then it may be the case that specific linguistic devices, such as the metasemantic properties of the linguistic categories identified by the LCM (Semin and Fiedler, 1988) are themselves functionally organized to drive attention in a generic fashion. This leap, although implicitly if not explicitly always inherent in the conceptualization of the LCM, was never the direct object of empirical investigation.

Diederik Stapel and I, after lengthy discussions over dinner in Amsterdam, used the LCM as a conceptual framework to investigate the hypothesis that different linguistic devices within the same language may have generic, metasemantic effects on cognition (Stapel and Semin, 2007). In our studies, we focused on the abstractness-concreteness dimension of the LCM, and suggested that if concrete terms such as action verbs (to kiss, to hit, to push) are used predominantly in situated contexts and refer to the specific details of a social event, then their obvious function – aside from providing a semantic representation of the event – is to *draw attention to the situated, local features of the event* (Stapel and Semin, 2007). At the other end, abstract terms such as adjectives detract from any transient situated features of an event and direct global focus (John is aggressive).

From this it follows that different predicate classes are likely to direct attention to different features of an object. Now, the research question is not about specific semantic categories (e.g., freedom fighter) and their semantic associates. It is about metasemantic features of interpersonal language (Semin, 2000). We (Stapel and Semin, 2007) translated this by postulating that concrete terms (e.g., action verbs) generically direct attention to details of an event or object. Contrastively, abstract terms (e.g., adjectives) were postulated to draw attention to the global properties of an object. In four experiments with an unrelated task paradigm we exposed participants to either abstract predicates (adjectives) or concrete predicates (action verbs). This was achieved by asking participants to use either verbs or adjectives in a spontaneous narrative or by subtly priming them with these categories supraluminally (Experiments 2 and 3) or subliminally (Experiment 4). This was followed by an ostensibly unrelated second task. In this second task, the participants were instructed to complete one or more dependent measures. These were designed to examine the attention-driving implications of the concrete versus abstract linguistic categories and consisted of either self-report (Experiments 1 and 4), Kimchi and Palmer's (1982) perceptual global-specific focus task (Experiments 1 and 4), Isen and Daubman's (1984) categorical inclusiveness task (Experiment 2), or Kitayama et al.'s (2003) framed line test (Experiment 3). Consistently across all four experiments and all dependent variables, we were able to show that abstract predicates induced a global perceptual focus, while concrete predicates induced a local perceptual focus.

This final research example diverges from the traditional approach comparing different linguistic communities to the linguistic relativity debate. Traditionally, in this debate, the focus has been specific semantic domains, such as color (e.g., Özgen, 2004), space (e.g., Majid et al., 2004), gender (e.g., Stahlberg et al., 2001) to name a few. In contrast, our

research (Stapel and Semin, 2007) opens an entirely different way of looking at the language–cognition interface that is made possible by taking a functional approach to language. The implications of this “new look” are wide ranging. One such implication is for the scaffold this functional view provides for the *culture–language–perception* interface. The above-summarized findings show that different linguistic categories in general (i.e., action verbs and adjectives) drive attention to different aspects of a stimulus. However, if two different linguistic communities (say Italian and Japanese) were to have differential accessibility in their daily use of adjectives and action verbs, then what would the implications of the findings reported by Stapel and Semin (2007) be? Let me illustrate with only one example.

There is recent evidence that Japanese use preferentially concrete language (i.e., verbs) rather than abstract language (i.e., adjectives) as has been shown by Maass et al. (2006). Based on these findings and those obtained by Stapel and Semin (2007), one would predict that participants from interdependent cultures with a preference for concrete language use (e.g., Japan) should make more errors in a perceptual task, such as the framed line task (Kitayama et al., 2003) compared with participants from a culture that is independent and use more abstract language habitually (e.g., Americans). This is precisely what Kitayama et al. (2003) have demonstrated. The metasemantic attention-driving function of the language approach would therefore appear to open ways of looking and linking research findings not solely within the same linguistic community, but also furnishing a way of actually grounding cross-linguistic research as well as anchoring these findings in a more general level of explanation. In concluding, I should note that these are the early steps emerging from modifying a classic question from “What is language?” to “What is language for?” and this minor modification – in my view – opens a major reorientation full of research promise.

ACKNOWLEDGMENTS

Supported by the Royal Netherlands Academy of Arts and Sciences ISK/4583/PAH. I would like to express my thanks to Tory Higgins whose detailed and helpful comments improved the chapter considerably.

REFERENCES

Abelson, R.P. and Kanouse, D.E. (1966) Subjective acceptance of verbal generalizations. In S. Feldman (ed.), *Cognitive Consistency: Motivational Antecedents and Behavioral Consequents*, pp. 171–197. New York: Academic Press.

Asch, S.E. (1946) Forming impressions of personality. *Journal of Abnormal and Social Psychology*, 41, 258–290.

Au, T.K. (1986) A verb is worth a thousand words: The causes and consequences of interpersonal events implicit in language. *Journal of Memory and Language*, 25, 104–122.

Bendix, E.H. (1966) *Componential Analysis of General Vocabulary: The Semantic Structure of a Set of Verbs in English, Hindu and Japanese*. The Hague: Mouton.

Berlin, B. and Kay, P. (1969) *Basic Color Terms: Their Universality and Evolution*. Berkeley: University of California Press.

Brown, R. and Fish, D. (1983) The psychological causality implicit in language. *Cognition*, 14, 233–274.

Brown, R. and Lenneberg, E.H. (1954) A study in language and cognition. *Journal of Abnormal and Social Psychology*, 49, 454–462.

Caramazza, A., Grober, E., Garvey, C. and Yates, J. (1977) Comprehension of anaphoric pronouns. *Journal of Verbal Learning and Verbal Behavior*, 16, 601–609.

de Poot, C. and Semin, G.R. (1995) Pick your verbs with care when you formulate a question!. *Journal of Language and Social Psychology*, 14, 351–369.

Douglas, K.M. and Sutton, R.M. (2006) When what you say about others says something about you: Language abstraction and inferences about describer's attitudes and goals. *Journal of Experimental Social Psychology*, 42, 500–508.

Ehrlich, K. (1980) Comprehension of pronouns. *Quarterly Journal of Experimental Psychology*, 32, 247–255.

Fiedler, K. and Semin, G.R. (1989) On the causal information conveyed by different interpersonal verbs: The role of implicit sentence context. *Social Cognition*, 6, 21–39.

Fillmore, C.J. (1968) The case for case. In E. Bach and R.G. Harms (eds), *Universals in Linguistic Theory*. pp. 1–87. New York: Holt, Rinehard and Winston.

Franco, F.M. and Maass, A. (1996) Implicit versus explicit strategies of out-group discrimination: The role of intentional control in biased language use and reward allocation. *Journal of Language and Social Psychology*, 15, 335–359.

Franco, F.M. and Maass, A. (1999) Intentional control over prejudice: When the choice of the measure matters. *European Journal of Social Psychology*, 29, 469–477.

Garnham, A. and Oakhill, J. (1985) On-line resolution of anaphoric pronouns: Effects of inference making and verb semantics. *British Journal of Psychology*, 76, 385–393.

Garnham, A., Oakhill, J. and Cruttenden, H. (1992) The role of implicit causality and gender cue in the interpretation of pronouns. *Language and Cognitive Processes*, 7, 231–255.

Garnham, A., Traxler, M., Oakhill, J. and Gernsbacher, M.A. (1996) The locus of implicit causality effects in comprehension. *Journal of Memory and Language*, 35, 517–543.

Garvey, C. and Caramazza, A. (1974) Implicit causality in verbs. *Linguistic Inquiry*, 5, 459–464.

Garvey, C., Caramazza, A. and Yates, J. (1976) Factors influencing assignment of pronoun antecedents. *Cognition*, 3, 227–243.

Greene, S. and McKoon, G. (1994) Initiating/reacting verbs and the accessibility of discourse referents. Unpublished manuscript, Princeton University.

Gilson, C. and Abelson, R.P. (1965) The subjective use of inductive evidence. *Journal of Personality and Social Psychology*, 2, 301–310.

Grice, H.P. (1975) Logic and conversation. In P. Cole and J. Morgan (eds), *Syntax and Semantics*, pp. 41–58. New York: Academic Press.

Hardin, C. and Banaji, M.R. (1993) The influence of language on thought. *Social Cognition*, 11, 277–308.

Hoffman, C. and Tschir, M.A. (1990) Interpersonal verbs and dispositional adjectives: The psychology of causality embodied in language. *Journal of Personality and Social Psychology*, 58, 765–778.

Hunt, E. and Agnoli, F. (1991) The Whorfian hypothesis: A cognitive psychology perspective. *Psychological Review*, 98, 377–389.

Hutchins, E. (1996) *Cognition in the Wild*. Cambridge, MA: MIT Press.

Isen, A.M. and Daubman, K.A. (1984) The influence of affect on categorization. *Journal of Personality and Social Psychology*, 47, 1206–1217.

Kanouse, D.E. (1972) Verbs as implicit quantifiers. *Journal of Verbal Learning and Verbal Behavior*, 11, 141–147.

Kashima, Y. and Kashima, E.S. (2003) Individualism, GNP, climate, and pronoun drop: Is individualism determined by affluence and climate, or does language use play a role? *Journal of Cross Cultural Psychology*, 34, 125–134.

Kasoff, J. and Lee, J.Y. (1993) Implicit causality as implicit salience. *Journal of Personality and Social Psychology*, 65, 877–891.

Kay, P. (1996) Inter-speaker relativity. In J.J. Gumperz and S.C. Levinson (eds), *Rethinking Linguistic Relativity*, pp. 97–114. Cambridge: Cambridge University Press.

Kimchi, R. and Palmer, S.E. (1982) Form and texture in hierarchically constructed patterns. *Journal of Experimental Psychology: Human Perception and Performance*, 8, 521–535.

Kitayama, S., Duffy, S., Kawamura, T. and Larsen, J.T. (2003) Perceiving an object and its context in different cultures. *Psychological Science*, 14, 201–207.

Lenneberg, E.H. (1953) Cognition in ethnolinguistics. *Language*, 29, 463–471.

Lenneberg, E.H. and Roberts, J.M. (1956) The language of experience: A study in methodology. *International Journal of American Linguistics*, 22 (2, part2, Memoir 13).

Lippman, W. (1922) *Public Opinion*. New York: Hartcourt, Brace and Co.

Maass, A. and Arcuri, L. (1992) The role of language in the persistence of stereotypes. In G. Semin and K. Fiedler (eds), *Language, Interaction and Social Cognition*, pp. 129–43. Newbury Park, CA: Sage.

Maass, A., Karasawa, M., Politi, F. and Suga, S. (2006) Do verbs and adjectives play different roles in different cultures? A cross-linguistic analysis of person representation. *Journal of Personality and Social Psychology*, 90, 734–750.

Maass, A., Milesi, A., Zabbiini, S. and Stahlberg, D. (1995) The linguistic intergroup bias: Differential expectancies or in-group protection? *Journal of Personality and Social Psychology*, 68, 116–126.

Maass, A., Salvi, D., Arcuri, L. and Semin, G. (1989) Language use in intergroup contexts: The linguistic intergroup bias. *Journal of Personality and Social Psychology*, 57, 981–993.

Majid, A., Bowerman, M., Kita, S., Haun, D.B.M. and Levinson, S.C. (2004) Can language restructure

cognition? The case for space. *Trends in Cognitive Sciences*, 8, 108–114.

McKoon, G., Geene, S.B. and Ratcliff, R. (1993) Discourse models, pronoun resolution, and the implicit causality of verbs. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 1040–1052.

Miller, G.A. and Johnson-Laird, P. (1976) *Language and Perception*. Cambridge: Cambridge University Press.

Özgen, E. (2004) Language, learning, and color perception. *Current Directions in Psychological Science*, 13, 95–98.

De Poot, C. and Semin, G.R. (1995) Pick your verbs with care when you formulate a question! *Journal of Language and Social Psychology*, 14, 351–368.

Reitsma-van Rooijen, M., Semin, G.R. and van Leeuwen, E. (2007) The effects of linguistic abstraction on interpersonal distance. *European Journal of Social Psychology*, 37, 817–823.

Rubini, M. and Kruglanski, A.W. (1997) Brief encounters ending in estrangement: Motivated language use and interpersonal rapport in the question-answer paradigm. *Journal of Personality and Social Psychology*, 72, 1047–1060.

Rudolph, U. and Försterling, F. (1997) The psychological causality implicit in verbs: A review. *Psychological Bulletin*, 121, 192–218.

Sapir, E. (1951) In D.G. Mandelbaum (ed.), *Selected Writings of Edward Sapir in Language, Culture, and Personality*. Berkeley: California University Press.

Semin, G.R. (1989) The contribution of linguistic factors to attribute inferences and semantic similarity judgments. *European Journal of Social Psychology*, 19, 85–100.

Semin, G.R. (1995) Interfacing language, and social cognition. *Journal of Language and Social Psychology*, 14, 182–194.

Semin, G.R. (1998) Cognition, language, and communication. In S.R. Fussell and R.J. Kreuz (eds), *Social and Cognitive Psychological Approaches to Interpersonal Communication*, pp. 229–257. Hillsdale, NJ: Laurence Erlbaum.

Semin, G.R. (2000a) Agenda 2000: Communication: language as an implementational device for cognition. *European Journal of Social Psychology*, 30, 595–612.

Semin, G.R. (2000b) Language as a cognitive and behavioral structuring resource: Question-answer exchanges. In W. Stroebe and M. Hewstone (eds), *European Review of Social Psychology*, pp. 75–104. Chichester: Wiley.

Semin, G.R. (2001) Language and social cognition. In A. Tesser and N., Schwarz (eds), *Handbook of Social Psychology, Intraindividual Processes*, 1, 159–180. Oxford: Blackwell.

Semin, G.R. (2006) Modeling the architecture of linguistic behavior: Linguistic compositionality, automaticity, and control. *Psychological Inquiry*, 17, 246–255.

Semin, G.R. (2007) Grounding communication: Synchrony. In A. Kruglanski and E.T. Higgins (eds), *Social Psychology: Handbook of Basic Principles*, 2nd Edition, pp. 630–649. New York: Guilford Publications.

Semin, G.R. and Chassein, J. (1985) The relationship between higher order models and everyday conceptions of personality. *European Journal of Social Psychology*, 15, 1–15.

Semin, G.R. and Fiedler, K. (1988) The cognitive functions of linguistic categories in describing persons: Social cognition and language. *Journal of Personality and Social Psychology*, 54, 558–568.

Semin, G.R. and Fiedler, K. (1989) Relocating attributional phenomena within a language-cognition interface: The case of actors' and observers' perspectives. *European Journal of Social Psychology*, 19, 491–508.

Semin, G.R. and Fiedler, K. (1991) The linguistic category model, its bases, applications and range. In W. Stroebe and M. Hewstone (eds), *European Review of Social Psychology*, 2, 1–30. Chichester: Wiley.

Semin, G.R. and Fiedler, K. (1992) The configuration of social interaction in interpersonal terms. In G.R. Semin and K. Fiedler (eds), *Language, Interaction and Social Cognition*. London: Sage Publications.

Semin, G.R., Göts, C., Nandram, S. and Semin-Goossens, A. (2002) Cultural perspectives on the linguistic representation of emotion and emotion events. *Cognition and Emotion*, 16, 11–28.

Semin, G.R. and Greenslade, L. (1985) Differential contributions of linguistic factors to memory based ratings: Systematizing the systematic distortion hypothesis. *Journal of Personality and Social Psychology*, 49, 1713–1723.

Semin, G.R. and Krahé, B. (1987) Lay conceptions of personality: Eliciting tiers of a scientific conception of personality. *European Journal of Social Psychology*, 17, 199–209.

Semin, G.R. and Krahé, B. (1988) Explaining perceived cross-situational consistency: Intuitive psychometrics or semantic mediation? *European Journal of Personality*, 2, 239–252.

Semin, G.R. and De Poot, C.J. (1997a) The question-answer paradigm: You might regret not noticing how a question is worded. *Journal of Personality and Social Psychology*, 73, 472–480.

Semin, G.R. and De Poot, C.J. (1997b) Bringing partiality to light: Question wording and choice as indicators of bias. *Social Cognition*, 15, 91–106.

Semin, G.R. and Smith, E.R. (2008) *Embodied Grounding: Social, Cognitive, Affective, and Neuroscientific Approaches*. New York: Cambridge University Press.

Semin, G.R., Rubin, M. and Fiedler, K. (1995) The answer is in the question: The effect of verb causality on locus of explanation. *Personality and Social Psychology Bulletin*, 21, 834–841.

Smith, E.R. and Semin, G.R. (2004) Socially situated cognition: Cognition in its social context. *Advances in Experimental Social Psychology*, 36, 53–117.

Stahlberg, D., Sczesny, S. and Braun, F. (2001) Name your favorite musician: Effects of masculine generics and of their alternatives in German. *Journal of Language and Social Psychology*, 20, 464–469.

Stapel, D. and Semin, G.R. (2007) The magic spell of language. *Journal of Personality and Social Psychology* (in press).

Von Hippel, W., Sekaquaptewa, D. and Vargas, P. (1997) The linguistic intergroup bias as an implicit indicator of prejudice. *Journal of Experimental Social Psychology*, 33, 490–509.

Von Hippel, W., Sekaquaptewa, D. and Vargas, P. (2007) Armed only with paper and pencil: 'low-tech' measures of implicit attitudes. In B. Wittenbrink and N. Schwarz (eds), *Implicit Measures of Attitudes: Procedures and Controversies*, pp. 103–124. New York: Guilford Press.

von Humboldt, W. (1836) *Über die Verschiedenheit des menschlichen Sprachaufbaues und ihren Einfluß auf die geistige Entwicklung des Menschengeschlechts*. Berlin: Royal Academy of Sciences: Republished (trans. G.C. Buck and F. Raven) (1999) *Linguistic Variability and Intellectual Development*. Baltimore, MD: University of Miami Press.

Werkman, W.M., Wigboldus, D.H. and Semin, G.R. (1999) Children's communication of the Linguistic Intergroup Bias and its impact upon cognitive inferences. *European Journal of Social Psychology*, 29, 95–104.

Whorf, B.L. (1956) *Language, Thought, and Reality*. Cambridge, MA: MIT Press.

Wigboldus, D., Semin, G.R. and Spears, R. (2000) How do we communicate stereotypes? Linguistic bases and inferential consequences. *Journal of Personality and Social Psychology*, 78, 5–18.

Action Identification Theory

Robin R. Vallacher and Daniel M. Wegner

ABSTRACT

The theory of action identification is a system of three principles explaining how people's thoughts of what they are doing relate to what they do. In a sense, the principles suggest an "operating system" for a human being – a program linking thought to action. The principles suggest simply that people do what they think they are doing, that when they can think of doing something more, they do that – but that when they can't do what they were thinking of doing, they think of doing less. Stated this way, the principles may seem perhaps too simple, but their interaction and ramifications are remarkably complex and predictive of a rich array of phenomena. The theory has things to say about how people can conceptualize their actions optimally, it offers insights into how people regulate actions through thinking, it provides a template for understanding how action connects to emotion and to self-concept, and it suggests how social influence occurs by changing how actions are understood. These consequences of the theory were discovered early on, but have recently been supplemented by extensions to encompass the dynamics of action and the role of action identification in the understanding of own and others' minds.

INTRODUCTION

Life is full of big and little things. One moment you might be planning your career; the next moment you might be turning on your computer or searching for a pen. Sometimes the little things are fairly automatic, requiring little conscious attention, and are irrelevant to the bigger things we have in mind. And sometimes the little things are simply subordinate components to the big things, the details by which we accomplish goals, implement plans, or demonstrate traits and values. So although everyday life can be characterized as the ebb and flow of actions, big and little, psychologists typically ignore the little ones and concentrate instead on people's broader action tendencies. Theories in social psychology are replete with references to goals, values, personal standards, schemas, chronic concerns, and personality traits. Searching for a pen or turning on a computer may be unavoidable details of everyday life but not all that significant in their own right.

The theory of action identification (Vallacher and Wegner, 1985, 1987; Wegner and Vallacher, 1986) grants a far bigger role to the little things. Or more precisely, the theory posits a deep and reciprocal connection between the details of what we do and the larger meanings we have in mind. This dynamic interplay between mind and action is important in its own right, but it is also highly germane to a number of issues of concern to social psychologists. Not all of these implications were on the radar screen for us when we developed the theory in the 1980s. Our aim at the time was far more circumscribed: to generate a parsimonious model of the mental control of action. Since that time, the theory has evolved in two ways. First, it has been extended to topics that go beyond mental control *per se*, from personality and self-concept to social influence and conflict resolution. Second, the theory has morphed into forms that resonate with contemporary perspectives in psychology, including cognitive science and dynamical systems. The emergent nature of the theory is apparent in some of the work the two of us have pursued independently, often without explicit reference to the original theory.

Our big aim in this chapter is to describe action identification theory and trace its trajectory from the 1980s to the present. We begin by outlining the basic theory and its initial range of application. We then indicate how the theory's scope has broadened over the years to account for important domains of social experience. In the next part, we discuss how the theory has been reframed in recent years in light of new developments in psychology and science generally. In a concluding section, we throw caution to the wind and suggest that the dynamic scenario at the heart of the theory provides a template for human experience. When liberated a bit from its specific principles and couched in more general terms, action identification theory can be viewed as a metatheory that expresses the essence of personal and interpersonal dynamics.

THE BASIC THEORY

Back in the early and mid 1980s, we were blissfully unaware of where our actions would take us. We were certainly trying to think big, but we were also caught up in lots of little things. In big terms, we looked upon our collaborative effort as constructing a theory about mental control. We also saw our mission as integrating insights from several perspectives that were coming into their own at the time. Attribution theory and cognitive dissonance phenomena were never far from our minds, but of more immediate interest were cybernetic models of action (Carver and Scheier, 1981), unitization in behavior perception (Newtson, 1976), and the endogenous-exogenous distinction in lay epistemology (Kruglanski, 1975). Our reading and thinking even strayed from psychology to encompass philosophical approaches to the understanding of action – like those of Goldman (1970) and Hampshire (1960). Each of these perspectives conveyed something fundamental about the relationship between mind and action, and we aspired to synthesize these insights in a parsimonious framework.

But we also did lots of little things, many of which seemed at odds with constructing a new theory. We stayed up late, occasionally drank scotch, made goofy jokes, offered wild-eyed speculations on the human condition, and even constructed a few faux psychological instruments (e.g., the Hidden Brain Damage Scale) to make each other (and others with suspicious standards) laugh. Out of these seemingly irrelevant asides, we experienced a number of insights and breakthroughs that transformed our thinking about mind and action – although these probably paled in number with the all-too-frequent “What were we thinking?!?” false starts. In a reflexive approach to theory construction that borders on narcissism, our reflections on the role of these little acts in promoting a bigger sense of what we were doing became central to the theory itself. We realized that an action, even a highly important one like

theory construction, tends to become transformed mentally as well as mechanically over time and we suspected that there might be a reliable and meaningful pattern to such transformations. Perhaps, we thought, the movement between big and little things is central to the mind-action relationship, satisfying the dual and sometimes conflicting criteria of effectiveness and global understanding. At that point, of course, we promptly opened another bottle of scotch.

Central to our thinking at the time was the inherently ambiguous nature of human action and the perils this posed for theory construction. But we came to see this issue as a promise, more than a dilemma, and proceeded to develop the theory around it. Below we describe this progression from the uncertainty of action to the certainty of action understanding conveyed in the theory's three principles.

The uncertain act

In science, the phenomenon to be explained is usually well defined. When chemists theorize about chemical reactions, they don't have to guess about the nature of molecules. The same certainty lies at the heart of physics. Sure, quarks and dark energy are inferred, but at least the inferences, however controversial, are clear and precise. Even in the social sciences, the fundamental units are well defined. Economists develop models concerning the flow of money and tangible resources, anthropologists construct theories about religious beliefs and cultural traditions, and political scientists theorize about power structures and voting practices.

Social psychology is certainly a science, but the building blocks for theory construction in our field are not as well defined as are the building blocks in other scientific realms. Social psychological theories focus on a wide variety of constructs – norms, roles, beliefs, values, and so forth – but underlying all these constructs is the concept of action. Social psychology is ultimately concerned with what people *do*, whether privately or in

explicit social contexts such as relationships, groups, and crowded streets. But therein lies the rub: what exactly is the unit of action that corresponds to the physicist's protons, the biologist's genes, or the economist's dollar? What is an action and how can it be defined unequivocally so that it can provide a solid foundation for theory construction?

This question, we soon learned, was a bit of an obsession in philosophy. Coming from distinct traditions and armed with diverse assumptions, many philosophers have noted that an action does not admit to a single, unequivocal definition, but rather can be identified in many different ways (e.g., Danto, 1963; Goldman, 1970; Ryle, 1949). Something as straightforward as "driving to work" could be just as accurately identified as "operating a motor vehicle," "burning gas," "obeying traffic laws," "following a daily routine," or simply as "stepping on pedals" and "turning a wheel." When the focus turns to actions that are of greater interest to social psychologists – harmful versus helpful behavior, fairness versus injustice, and so on – the availability of multiple identities for the same action can prove problematic for theory construction. The act of "criticizing someone," for instance, could be viewed by observers, or even by the actor him or herself, as "acting rudely," "asserting dominance," "offering a different point of view," "providing constructive feedback," or simply as "uttering words." Which identity is deemed the "real" one makes a big difference in the explanation that is generated.

Even if there is consensus in the meaning attached to someone's action (e.g., the critical comment is widely perceived to be mean-spirited), the action still does not have an unequivocal definition. The same action in a different context might promote a very different consensus regarding its meaning (e.g., the critical comment in a policy-making group might be seen as constructive). It is also the case that actions are commonly characterized by equifinality (functional equivalence), further complicating the task of establishing stable building blocks for

theory construction. Mean-spirited, for example, provides functional equivalence for both verbal criticism and the silent treatment. Criticism, in turn, can be instantiated in very different ways (e.g., pointed comments or sarcasm). This problem exists even if one reduces action to button pushes or other simple operational definitions. A button can be pushed by a wide variety of physical motions involving different limbs and digits. Beyond that, when an action is reduced in this fashion, the number of potential identities for it increases in a proportional manner. A button push can mean virtually anything depending on the program to which it is attached or the context in which it occurs. Protons and dollars don't pose this problem.

To a critic, it might seem that social psychology is built on shifting sands, without a clear foundation for theory construction. Just such a criticism was in ascendance during the period in which we were busy developing our theory. Gergen (1985) was especially vocal in this regard, arguing that a science of social psychology was impossible in principle because of the inherent ambiguity of action. He, and others who shared this sentiment, seemed to suggest that we should give up the pretense and concentrate instead on how to make people happy (whatever that is!). But we were mindful of Kurt Lewin's remark that "one man's artifact is another man's theory." In this spirit, it occurred to us that far from being a problem for theory construction in psychology, the inherent ambiguity of action could be viewed as the key to understanding the relationship between mind and action. After all, despite the multiple identities available for any action, people seem to know in unequivocal terms what they are doing, have done, or intend to do. Somehow people sidestep the uncertainty of action and navigate the stream of potential act identities one at a time. How do they do it?

Levels of action identification

Ironically, we can make sense of action certainty by looking into the human mind that is

ultimately responsible for the ambiguity of action. The mind is designed to identify (or create) patterns in the real world. Lacking this pervasive tendency, people would be overwhelmed by the complex and nuanced information that continually bombards their senses on a moment-to-moment basis. Meaningful actions exist because we find or impose patterns on the specific behaviors we observe or otherwise learn about. The patterns are constructions, but once generated, they are maintained because they disambiguate reality and thereby provide coherent understanding and a stable platform for subsequent thought and behavior. Because they are constructions, however, they can admit to tremendous variability across people and contexts. Hence, the certainty of action that exists for each individual embedded in a particular context coexists with the uncertainty of action across individuals and contexts. Philosophers and psychologists live in the latter world, at least when wearing their professional hats and when talking about other people's realities.

That said, there is one metric for disambiguating action that seems solid and reflects a shared reality. The multiple act identities for an action tend to be organized in a hierarchical manner. Lower-level identities in the hierarchy convey the details or specifics of the action and thus indicate *how* the action is performed. Higher-level identities provide a more general understanding of the action; they indicate *why* the action is performed or what its effects and implications are. Higher-level identities are less movement-defined than lower-level identities, and provide a more abstract and comprehensive understanding of the action. Identification level is relative, so whether a particular act identity is a high or low level depends on the identity with which it is compared. What looks like a goal (e.g., "getting married") can be identified as a means with respect to a higher-level identity (e.g., "starting a family"). An act identity that seems molecular (e.g., "turning on a computer"), meanwhile, could represent a goal with respect to a yet lower-level identity (e.g., "pushing a button").

A simple criterion is useful for sorting an action's multiple identities into a hierarchy. One act identity is higher-level than another identity if it makes sense to say that one does the former *by* doing the latter (Goldman, 1970). Thus, one "goes to work" *by* "driving a car," one "drives a car" *by* "stepping on pedals and turning the steering wheel," and one "steps on pedals and turns the wheel" *by* "moving one's arms and leg." People appreciate the notion of an asymmetric *by* relation and demonstrate a very high degree of consensus in sorting act identities hierarchically in this fashion (Vallacher and Wegner, 1985).

Some act identities are not asymmetric in this sense, but rather have a *level-indequate* relation with one another. "Being rude" and "being constructive" are both plausible higher-level identities for "criticizing someone," for example, but they do not have an obvious hierarchical relation. To distinguish among such act identities in a meaningful manner, we have employed various psychometric methods, including multidimensional scaling and, more commonly, factor analysis. In the factor analytic approach, participants rate a wide variety of identities for an action (generated in pilot research) according to how well they personally describe the action. Factor analyses of these ratings typically reveal a single low-level factor that captures the most rudimentary act identities and two or more higher-level factors that reflect differences among identities with respect to valence and content-relevant dimensions. "Being rude," for example, loads on the same factor as other negative implications of criticism (e.g., "showing disrespect," "disregarding someone's point of view"), whereas "being constructive" loads on the same factor as other positive higher-level identities of criticism (e.g., "offering a different point of view," "providing useful feedback"). The dimensionality of higher-level act identities has theoretical relevance for a host of personal and social phenomena. These implications follow from the principles of action identification.

Theoretical principles

It is not surprising that people can distinguish among an action's available identities by virtue of an asymmetric *by* relation. But how do people pick one level over another to identify what they are doing, have done, or intend to do? Why does someone look upon his or her action as, say, "going to work" rather than "driving a car," or as "playing a piano" rather than "expressing feelings"? And after an action is identified in a particular way, what determines whether it is maintained under that identity or instead is reidentified later in different terms? Why do higher-level identities (e.g., "being cooperative") sometimes resist change in the face of social pressure, but at other times change dramatically in content and valence (e.g., "being competitive")? What is the relationship between how an action is identified and how effectively the action is performed? The answers to these questions are conveyed in three principles.

Principle 1: Action is maintained with respect to its prepotent identity

Action identification would hardly be worthwhile for people, let alone worth theorizing about by social psychologists, in the absence of this principle. It is central to models of self-regulation, of course, and in a broader sense to any theory that posits a link between mind and action (e.g., Carver and Scheier, 2002; Higgins, 1998; James, 1890; Miller et al., 1960). A person's prepotent act identity functions as an intention to initiate an action, a frame of reference for performing the action, and a criterion to assess how well the action has been performed. Because an action can be identified at different levels, the principle directly implies that people can maintain action at different levels. A person may intend to "send an e-mail," for instance, and monitor his or her subsequent action to see whether this intention was fulfilled. Alternatively, the person may intend to "contact a colleague" (a higher-level act identity) or "propose a collaborative writing venture" (a yet higher-level act identity), and monitor

the attainment of whichever act identity is prepotent.

Principle 2: When both a lower- and a higher-level act identities are available, there is a tendency for the higher-level identity to become prepotent

The idea here is that people are sensitive to the meanings, consequences, and implications of what they do. Within social psychology, this principle is reflected in the emphasis on goals, plans, values, and other global constructs that are said to motivate personal and interpersonal behavior. The preference for higher-level understanding is not unique to social psychology, but rather is central to many theoretical traditions, including learning under reinforcement contingencies, the mastery of skilled action, pattern formation and recognition in Gestalt psychology, inductive reasoning in cognitive psychology, and the “search for meaning” in existential psychology. These different schools of psychology share the assumption that act representations expand to encompass broader effects and meanings. In learning, a basic act expands to incorporate the reinforcing effects of the action. In the development of mastery, discrete acts become automated and integrated into a larger action unit. In Gestalt psychology, parts are unified to produce a perceptual whole. In induction, separate observations or pieces of information are organized into an explanatory account. And in existentialism, patterns discerned in distinct actions are recognized as manifestations of a larger action tendency.

Principle 2 enables people to choose one of many plausible act identities and actually do something. Without this principle, people would be “buried in thought,” entertaining a multitude of possible intentions in a given context. So when two or more plausible identities are available, people are inclined to choose the identity that provides the most comprehensive understanding of what they are doing, plan to do, or have done. A person could think about his or her behavior as “using eating utensils,” for example, but is more

inclined to gloss over such details and identify the behavior as “eating dinner.” If cues to yet higher-level identities are available, the person is likely to embrace them over the now-lower-level “eating identity.” Thus, he or she may look upon the meal as “getting nutrition,” “satisfying my appetite,” or “putting on weight.”

Principle 3: When an action cannot be performed in terms of its prepotent identity, there is a tendency for a lower-level identity to become prepotent

If Principle 2 was the only basis for action identification, people’s minds would be populated by abstractions, fantasies, hopes, and fears, as increasingly higher-level identities emerge as ways of thinking about one’s action. Even the most rudimentary act could be charged with high-level significance in this manner. Such progressive integration would be possible if people lived in a world in which every thought was feasible and easily enacted.

Reality is not so accommodating. There are obstacles to enacting goals and plans, and even in the absence of such disruptions, the personal difficulty of achieving one’s ends can derail an action undertaken with respect to a higher-level identity. A person might set out to “demonstrate tennis prowess,” for example, only to disappoint him or herself, as well as observers, as he or she loses point after point during the tennis match. To regain control of the action, the person is inclined to adopt a more manageable lower-level act identity, such as “hitting cross-court shots.” If this act identity is not effectively performed, the person might drop to a yet lower-level act identity, thinking about “preparing the racket” or “getting in proper position before swinging the racket.” Whereas Principle 2 can lead to progressively higher levels of action identification, Principle 3 pulls people in the opposite direction, leading to progressively finer gradations of detail, with attention devoted to increasingly molecular features of what they are doing. The potential for flights of fancy that is inherent in the second principle, then, is unlikely to

pose a serious problem for most people because of the reality orientation inherent in the third principle.

Several factors dictate how detailed a person's act identity is likely to be. Actions that are complex, unfamiliar, or time-consuming tend to be identified in low-level terms. An easy or familiar action may also be identified in lower-level terms if it is disrupted. And any action – even one that is fairly easy and free from disrupting influences – tends to be identified in lower-level terms as the time for enactment approaches, a point that has been further documented in temporal construal theory (Trope and Lieberman, 2003). Engaged couples, for example, identify the act of getting married in high-level terms several weeks in advance of the wedding, but think about progressively finer details (e.g., saying vows, walking down the aisle) as the day approaches (Vallacher and Wegner, 1985: Chapter 4). Finally, simply instructing people to think about their action in molecular terms can prove sufficient to induce a low-level mindset.

The emergence process

Taken together, the three principles impart a dynamic interplay to the connection between mind and action. Low-level identification is a relatively unstable state that is adopted out of necessity rather than preference. Movement to a lower-level state (Principle 3) thus provides the precondition for adopting a higher-level identity (Principle 2) that restores coherence in the mental system and provides a stable platform for action (Principle 1).

Sometimes this process is akin to getting around a roadblock. After a brief detour to lower-level details, the person is back on track implementing the original higher-level identity. If this were always the case, though, people would never develop new ways of acting. But people do develop new insights into their actions and often chart new courses of action. When a higher-level meaning has been abandoned in order to regain control of

an action at a lower level, for example, the person becomes sensitive to cues to higher-level meaning in the action context, and these may provide an avenue of emergence to a new way of understanding the action. Lacking the lower-level state, the change from one high-level identity to a different one would not occur.

Our initial experiments on action identification were attempts to validate the emergence process. Wegner et al. (1984, Experiment 1), for example, investigated whether experienced coffee drinkers could be led to think about coffee drinking in a new way if they were induced to focus on the details of drinking. Some participants read an essay arguing that coffee drinking makes people *seek out stimulation*; others read an essay arguing instead that coffee drinking makes people *avoid stimulation*. After reading the essay, participants listened to musical passages through headphones, with the expectation that they would subsequently rate the passages. They were free to adjust the volume of the music if it was too loud or not loud enough. Half the participants drank from normal coffee cups, but the others drank from heavy cups covered with duct tape (ostensibly as a safeguard against electric shock because of the electronic equipment in the music appreciation task).

Participants who drank coffee from the normal cups did not adjust the volume reliably one way or the other in response to the essay they read. Presumably, they already had a high-level identity for coffee drinking (which we verified in a pilot study), so the identity for coffee drinking provided in the essay did not provide a new way to think about the act. But participants who drank coffee from the unwieldy cups were influenced by the essay: those who read the “seeking stimulation” essay turned the volume up (increasing stimulation), whereas those who read the “avoiding stimulation” essay turned the volume down (decreasing stimulation). The unwieldy cups induced lower-level act identities for coffee drinking (verified in pilot research), such as “lifting a

cup" and "swallowing liquid." The essay then provided an avenue of emergence to a higher-level identity (seeking versus avoiding stimulation) of what they were doing.

Emergence also takes place on a longer timescale. As people become increasingly competent at an action, for example, they tend to identify the action in terms of its consequences, self-evaluative implications, and other forms of meaning, rather than in terms of its lower-level details. This "sealing off" of lower-level act identities, which is consistent with research on skill acquisition, has been demonstrated for a variety of actions, including piano playing, essay writing, tennis, karate, and videogames (Vallacher and Wegner, 1985). In all cases, people initiate the act with a relatively high-level identity in mind, move to lower-level identities as they learn the action, and then move to a higher-level act identity as the action becomes more-or-less automated and mastered. We discovered, however, that the emergent identity was rarely the same high-level identity that motivated the people to begin with. Playing the piano, for example, was identified initially by many people as "impressing my friends" but after a sustained period of low-level maintenance, the now-proficient piano players came to identify piano playing as "relaxing myself." The tendency for an emergent act identity to differ from the action's antecedent identity suggests a scenario by which people develop new motives, interests, concerns, and insights into their mental make-up.

Negative high-level identities can also be embraced in this fashion. People can deflect an undesirable characterization of their behavior through a variety of interchangeable cognitive mechanisms (cf. Tesser et al., 1996) as long as they have a more flattering depiction available at the same identification level. Someone informed that he or she has acted rudely or demonstrated insensitivity, for example, may be uninfluenced by this feedback if he or she looks upon the action in question as offering constructive feedback. But if the person is induced to think about the

lower-level aspects of the action, he or she is primed for emergence and thus is more likely to accept the unflattering higher-level characterization. In this way, people are capable of accepting responsibility for actions with negative consequences and implications, and are open to new insights into their motives and personality dispositions.

The optimality hypothesis

The tension between preference (Principle 2) and necessity (Principle 3) is manifest as an oscillation over time between higher and lower levels of identification. This oscillation eventually dissipates, with the prepotent act identity converging on a restricted range of identities at a level that provides a balance between the two tendencies. This level represents the *optimal level of identification*. A person may prefer to think about his or her behavior at a party as "demonstrating wit and charm," for example, but reality constraints may promote a lower-level orientation involving such identities as "think of funny comments," "smile and maintain eye contact," and "look for gullible people." The optimal level represents a compromise between comprehensive understanding and effective action. As such, it signifies the base-rate difficulty of the action for people generally, or the level of action mastery for an individual.

Despite the tendency toward optimality in action identification, people's mental dynamics are not always in perfect resonance with their overt behavior. Indeed, the lay (and scientific) fascination with psychology is attributable in large part to problems in the feedback between mind and action conveyed in the optimality hypothesis. People routinely fail to do what they intend, and often fail to profit from their mistakes in subsequent planning and behavior. To a certain extent, lapses in mental control reflect a lack of skill or experience with respect to the action in question. If failures in action signified only a lack of preparation or ability, however, the interest in psychology would center

primarily on issues of learning and skill development, with the study of personality and social psychology occupying a subsidiary role. Fortunately for our field, people make mistakes despite having the requisite skill, experience, and motivation to perform the action effectively (cf. Baumeister and Heatherton, 1996).

Dysfunctional action occurs when people monitor and control what they are doing with respect to a level of identification that is non-optimal for the action's performance. This potential exists because an action's available identities are constrained by the context in which the action occurs. Most of the contexts defining everyday life are stacked in favor of relatively high-level identities, since there are usually salient cues to an action's causal effects, socially labeled meanings, and potential for self-evaluation. When a person is offered a reward or threatened with punishment, for instance, it may prove impossible for him or her not to define what he or she is doing in these terms. In similar fashion, situations involving competition, audience evaluation, or other pressures to do well may keep the person mindful of high-level identities of a self-evaluative nature (e.g., "demonstrating my skill," "trying to win," "impressing others") at the expense of the action's more molecular representations. If the action is personally difficult, the context-induced high-level identities will lack sufficient detail and coordination of components necessary for optimal performance.

This reasoning is consistent with research demonstrating that unfamiliar or complex actions are adversely affected by factors that charge the action with significance. Thus, a complex task suffers when it is linked to salient rewards, when it is performed in the presence of an evaluative audience, and when it is performed in a competitive context (cf. Baumeister and Heatherton, 1996). These factors each impart a fairly high-level representation to the task (earning a reward, impressing an audience, demonstrating one's skill), so it is not surprising that they impair performance when the task is difficult or

unfamiliar and hence best enacted with respect to more molecular representations. In effect, the basic action elements are discharged without the degree of conscious control necessary to assure their moment-to-moment coordination.

On the other hand, some contexts can induce a level of action control that is too *low* for effective performance. This is likely to occur when there are distractions, obstacles, or other sources of disruption that render the action's details prepotent. A jammed keyboard, for instance, can reduce "preparing a chapter" to a series of discrete key-strokes. Low-level identities also become prepotent in novel settings lacking familiar cues to an action's higher-level meaning. In yet other contexts, a person's attention may be drawn to the details of his or her behavior, inducing him or her to experience a lower level of identification than would normally be the case. If the action is personally easy for the person and thus best performed with relatively high-level identities in mind, the prepotence of lower-level identities resulting from disruption, novelty, or instruction can undermine the quality of the person's performance.

Consistent with this possibility, research has shown that performance can be disrupted when attention is drawn to the over-learned details of an action (e.g., Langer and Imber, 1979). The more over-learned the action is, the greater the performance impairment engendered by a conscious concern with how to perform the act (Kimble and Perlmuter, 1970). A proficient piano player, for instance, can become derailed in the middle of a familiar piece if he or she looks at the music and starts to think about which sequence of keys to hit with which fingers. This form of non-optimality represents consciousness trying to micromanage a problem that is best left to lower-level echelons to work out among themselves.

Support for both manifestations of nonoptimality is provided in a study of speech fluency (Vallacher et al., 1989). Participants were asked to deliver a prepared speech to

either an easy-to-persuade audience or a difficult-to-persuade audience. Half the participants were induced to think about the action in high-level terms (“try to persuade the audience”), and half were induced to focus on the lower-level details of delivering the speech (monitoring their voice quality). In line with the optimality hypothesis, the participants made fewer speech errors when their level of identification matched the personal difficulty of the act. Specifically, speech fluency was maximized when the task was personally easy (i.e., the audience was believed to be easy to persuade) and identified at high level and when the task was personally difficult (the audience was hard to persuade) and identified at low level. When the match between the action’s difficulty and level of identification was nonoptimal (the easy act identified in low-level terms and the difficult act identified in high-level terms), participants tended to stumble over their words and experience frequent speech pauses. The nonoptimal participants also rated their speech performance as relatively poor compared with participants in the optimality conditions.

THE EXTENDED THEORY

The theory in its original form was intended to address the ambiguous nature of action and to develop principles concerning the mental control of action. It soon became apparent, however, that the reciprocal feedback between mind and action had implications for a variety of other topics of interest to social psychologists, including emotion, personality, self-concept, and social influence.

Emotion

Perhaps the most widely accepted division of psychological processes is the tripartite distinction among action, cognition, and emotion. The theory had a lot to say about the

connection between thought and action, but barely mentioned emotion. After developing the theory and devising empirical tests of emergence and optimality, the role that emotion might play in action identification dynamics came into sharp relief. Far from being irrelevant to action identification, emotion plays a key role in calibrating the mind–action connection. It does so by providing the hedonic basis for achieving and maintaining an optimal level of identification.

Emotion serves this function by signaling when mind and action are not well calibrated. In general terms, this contention is consistent with theories suggesting that negative emotion commonly results from the interruption of goal-directed action (cf. Carver and Scheier, 2002; Simon, 1967; Vallacher and Nowak, 1999). Interruption can take many forms, from environmental obstacles to the interference of other people. Our perspective suggests that nonoptimal identification also qualifies as an interrupt – an identity that is either too low- or too high-level disrupts the successful coordination of elements necessary for action implementation. Interruption promotes arousal, which in turn engages epistemic concerns regarding the source of the arousal, presumably in service of choosing an appropriate course of action. In like manner, a breakdown in mind–action coordination can promote heightened arousal which alerts the performer to possible reasons for the breakdown. In essence, the ineffective performer examines his or her mental content with respect to the action, with an eye toward finding another way to think about what he or she is doing. Arousal stemming from faulty calibration, then, becomes manifest phenomenologically as a special form of self-scrutiny.

Research has established a link between heightened arousal and self-focused attention (Wegner and Giuliano, 1980) and various models concur that attention to the self tends to impair task performance. There is little consensus, however, regarding the means by which self-focused attention produces this effect. In some models, to be self-conscious

is to be aware of the potential self-evaluative implications of what one is doing (e.g., Wicklund and Frey, 1980). Defined in this way, self-focused attention is said to impair performance by reducing attention to task-relevant features. Test anxiety, for instance, is said to occur when the test taker is preoccupied with self-evaluative thoughts (e.g., "I'm going to fail," "I'll be embarrassed") to the relative exclusion of attention to the subtleties of the task at hand (Sarason, 1972; Wine, 1971). But in other formulations (e.g., Kimble and Pelmutter, 1970), self-consciousness refers to heightened awareness of the processes or mechanics underlying the execution of behavior (e.g., the physical movements involved or the coordination of such movements). Models that embrace this definition of self-focused attention argue that a conscious concern with the process of performance essentially disintegrates the action, robbing it of its fluidity and rhythm (cf. Baumeister and Heatherton, 1996).

These contrasting views of self-consciousness map directly onto the optimality hypothesis. The self-evaluative implications of behavior constitute a special class of relatively high-level action identification, whereas the mechanics of behavior reflect considerably lower-level act identities. From the optimality perspective, neither orientation is inherently linked to self-consciousness; rather, each can give rise to such a state depending on the personal difficulty of the action. Specifically, the experience of self-consciousness arises whenever one's conscious level of action control is either too high or too low, given the action's personal difficulty. When performing a simple act or a complex one that has become fairly automated, people will feel self-conscious to the extent that they are conscious of the molecular features of the action. For a difficult act that is best performed with such features in mind, however, self-consciousness is associated instead with sensitivity to the act's larger meanings and effects.

Evidence for this perspective on self-consciousness and its role in performance is

provided in the study of speech fluency (Vallacher et al., 1989) discussed above. In addition to tracking participants' speech errors and self-rated performance quality, we asked them to indicate their degree of self-consciousness (as well as anxiety, tension, etc.) while delivering the speech. Results showed that these ratings paralleled the pattern obtained for speech fluency: self-consciousness and related feelings of aversive arousal were greatest when the easy task was identified in low-level terms and the difficult task was identified in high-level terms.

The arousal and negative emotion associated with nonoptimal identification diminishes as people bring their conscious representation of what they are doing into line with the action's difficulty. This does not mean, however, that people's affective state goes flat when they control an action with respect to the optimal level of identification. Rather, people experience positive affect when there is a perfect match between the demands of a task and their mental and behavioral readiness to perform the task. Csikszentmihalyi (1990), for example, has observed that when people experience congruence between their capabilities and the demands of the task – a state he refers to as "flow" – they report diminished self-awareness. The link between optimality and positive emotion is also consistent with research on perceptual fluency (e.g., Winkielman and Cacioppo, 2001), which shows that stimuli are viewed positively to the extent that their features are processed easily. In analogous manner, one can speak of *action fluency*, in which the person has a well-orchestrated and adaptive understanding of an action that facilitates smooth execution of the action. An optimal level of identification enhances the degree to which a person can experience action fluency.

Personality

The optimality hypothesis holds that people will identify their actions at the highest possible

level that still affords effective action control. People clearly differ in their respective expertise in specific domains and thus are also likely to show corresponding differences in the optimal level of identification. A professional tennis player will identify “playing tennis” at a considerably higher level (“winning a match,” “improving my ranking”) than will (or should) a week-end warrior (“get the ball over the net,” “bend my knees but not too much”). But it also occurred to us that people may differ in their competence more generally, across a variety of action domains. Some people undertake more hobbies during their youth, for example, and become adept at a wider ranger of activities than do others. There may also be variation in the degree to which individuals encounter information pertaining to the distal consequences of action. Positions of responsibility, for example, may sensitize people to the higher-level implications of action. It could be, too, that learning to appreciate the higher-level identities of action in a few domains creates a readiness to see actions generally in these terms.

We referred to cross-domain differences among people in level of action identification as individual variation in *level of personal agency* (Vallacher and Wegner, 1989). At one end of this dimension is the low-level agent, a person who functions in different domains in a relatively molecular or detailed manner. This person’s base-rate tendency is to focus on the mechanistic details of actions. At the other extreme is the high-level agent, someone whose base-rate tendency is to view what he or she does in terms of the action’s causal effects, social meanings, and self-evaluative implications. Everyone is likely to have a stable, domain-specific identification level for certain actions, but level of personal agency influences how an individual identifies action across a wide range of domains.

This dimension of individual variation does not represent a trait in the usual sense of the term. A personality trait typically refers to a tendency to emit behaviors from within a content-defined class, such as “sociability”

or “conscientiousness.” Because any action can be identified in many ways, however, this approach to personality is problematic. “Criticizing an acquaintance,” for example, could be viewed as an instance of a trait such as unfriendly, but this classification may miss what the action really meant to the actor. Perhaps he or she was “offering constructive feedback” or “expressing an opinion.” Behavioral dispositions reify one particular identity at a relatively high level and thus may fail to capture what people really do. Level of personal agency goes beyond specific behavioral dispositions to address whether the person has trait-like dispositions at all. High-level agents can be expected to enact many of their behaviors under the guidance of higher-level meanings such as self-conceived traits, goals, and values. In contrast, low-level agents tend to engage in actions that are not personally connected to such larger meanings. Level of personal agency, in other words, represents the degree to which an individual has organized his or her actions into abstract, meaningful categories that can channel behavior into dispositional tendencies.

To explore whether there are cross-domain individual differences in level of action identification, we developed the Behavior Identification Form (BIF) (Vallacher and Wegner, 1989). It consists of 25 act identities, each followed by two alternative identities, one lower and one higher in level. “Making a list,” for example, is followed by “Getting organized” (higher-level) and “Writing things down” (lower-level). “Resisting temptation” is followed by “Saying ‘no’” (lower-level) and “Showing moral courage” (higher-level). Participants are asked to choose the alternative identity that best describes the action for them. Their level of personal agency is simply the number of high-level identities chosen across the 25 items.

We found that level of personal agency was reliably correlated with several aspects of people’s psychology, including action effectiveness, action planning, impulsivity,

action stability, self-monitoring, internal versus external locus of control, and key aspects of self-concept. Compared to low-level agents, the high-level agents were more skilled at a variety of actions (e.g., planning a party, teaching tricks to a pet), had more hobbies, were less impulsive and more intentional in their everyday behavior, were lower in self-monitoring, and had an internal locus of control. High-level agents were also more likely to describe themselves in terms of traits, ascribed greater importance to traits in their self-definition, had higher self-concept certainty, and were less susceptible to social feedback regarding their dispositional qualities. High- and low-level agents did not differ in their level of self-esteem, however, so the relevance of personal agency to certainty and malleability of self-concept is not mediated by the valence of people's self-concept.

Level of personal agency sheds light on the issue of personal versus situation causation, which periodically surfaces as an obsession in social psychology. Low-level agents' behavior tends to be under situational control, in that they enter action contexts with little sense of the action's implications in mind and thus are primed to accept cues to higher-level meaning found in social feedback or situational pressures. High-level agents' behavior, in contrast, tends to be under the control of personal goals and self-conceived tendencies. As a result, they are able to maintain their actions with respect to meaningful representations they carry with them across action contexts. Most people exist between the extremes of this dimension, so it is not surprising that behavior for people in general reflects a combination of personal and social influences.

Self-concept

A prevailing wisdom in social psychology is that people will go to great lengths to maintain their self-concept (Tesser et al., 1996), even if the self-concept is unflattering (Baumeister, 1993; Swann, 1990). Thus, people resist new

information that might promote a change in their self-perceived qualities, clinging even more tightly to their prevailing self-view. On the other hand, there is reason to believe that self-assessments are strongly impacted by feedback from significant others or even casual acquaintances (cf. Felson, 1989; Mead, 1934). The emergence scenario suggests that both generalizations are valid, but under different circumstances.

A person is likely to deflect or discount social feedback if he or she has a coherent high-level perspective on his or her behavior. If the person knows that he or she is cooperative, for example, he or she is unlikely to embrace feedback suggesting that he or she is really competitive. The self is clearly a familiar object of thought, so the base-rate tendency is to think about one's self-relevant behavior in high-level terms, thereby making self-concept change difficult. But under conditions that promote lower-level identities for an action, a person should show susceptibility to feedback from others that provides an avenue of emergence to higher-level understanding. The emergent identity may be quite different from prior identifications, and thus may provide the person with new "insight" into his or personal make-up.

We investigated the relevance of emergence for self-concept change and stability (Wegner et al., 1986). We arranged for participants to have a "computer analysis" of their personalities. The input for this analysis was participants' description of five things they had done in a recent interaction with someone of the same sex. In the high-level condition, they were asked to describe five things they had done that reflected their opinions, values, and personality traits. In the low-level condition, they were asked to indicate five specific actions reflecting concrete movements and utterances. The computer responded to these inputs with one of two personality profiles – one indicating that the participant was cooperative, the other that he or she was competitive. Participants were then asked to judge the validity of the computer feedback, describe themselves on various

traits dimensions including cooperative and competitive, and rank order several activities (including both a cooperative activity and a competitive activity) in order of preference for future participation.

By all three measures, results confirmed the emergence scenario. Participants in the high-level condition were skeptical of the computer program and their self-descriptions showed a slight reactance effect – those described as cooperative rated themselves as somewhat competitive and those described as competitive tended to rate themselves as cooperative. This is what one would expect from models that stress self-concept defense (e.g., Swann, 1990). The high-level participants' activity rankings did not show a preference for the activity that reflected the feedback (cooperative or competitive) they had received. High-level identification, then, provided a shield against social feedback. The results were far different for participants in the low-level condition. They judged the computer program to be valid and they rated themselves in accordance with the feedback the program provided – as highly cooperative in response to cooperative feedback but as highly competitive in response to competitive feedback. And their ranking of the future activities reflected their emergent self-understanding. Those provided with cooperative feedback gave a higher ranking to the activity that called for cooperative behavior, but those provided with competitive feedback ranked the competitive activity more favorably.

This scenario of self-concept change has straightforward implications for the effectiveness of psychotherapy intended to change a person's dysfunctional attitudes concerning him or herself. People resist changing their ideas of what they are like, even when these ideas paint a rather gloomy or depressing picture (e.g., Swann et al., 1992). Faced with a person's dysfunctional self-view, a therapist (or a well-intended friend) might be tempted to challenge the person's view directly, encouraging him or her to adopt a more positive attitude. Such a frontal assault

may have a temporary impact (e.g., Swann et al., 1990) but is destined to fail or even backfire in relatively short order because it runs counter to the person's coherent and comprehensive self-assessment.

The challenge is to disassemble this high-level identity by getting the person to focus on specific aspects of his or her behavior, and then provide cues to alternative higher-level identities that paint a more flattering self-portrait. Instructing the person to think about the behavioral evidence for his or her self-assessment, as is done in certain brands of cognitive therapy (e.g., Beck and Weishaar, 2000), trades on this idea. Cognitive behavioral therapy goes a step further by encouraging the client to *do* things rather than simply think (e.g., Meichenbaum, 1977). Because concrete actions require at least some attention to lower-level details, this approach is especially likely to create the low-level mindset that is the precondition for emergent understanding.

Social influence

Social influence is widely considered to be the pivotal process in social psychology. Any theory worth its salt, then, must have something worthwhile to say about the factors that determine whether or not an individual will change the way he or she thinks or acts with respect to a particular topic or domain. Of course, people can be induced to change their actions and expressed opinions through the application of strong forces, whether rewarding or punitive. But change is likely to be more enduring when it goes beyond enforcing overt behavior to changing the internal dynamics of the target. To promote such changes, it is necessary to disassemble or otherwise destabilize the target's way of thinking, priming him or her for reconfiguration in line with cues to the new message provided by the influence agent.

The disassembly-reconfiguration scenario clearly follows the contours of the emergence process. When people have a coherent

high-level identity for someone's behavior, they are relatively immune to alternative interpretations and evaluations. Metaphorically, people "freeze" when they have a high-level understanding that provides a sense of cognitive closure (cf. Kruglanski and Webster, 1996; Lewin, 1936). But when people identify someone's behavior in relatively low-level terms, they become receptive to coherent perspectives on the behavior provided by social sources and other external factors. In effect, people are motivated to "seize" a higher-level interpretation that provides personal closure concerning the action's meaning.

The extrapolation to social influence is straightforward. The influence agent first induces the target to consider the relevant action in concrete, low-level terms. Simply describing the action in terms of its details can induce low-level identification, as can presenting the target with a surplus of concrete information regarding the action. From this low-level state, the target experiences a heightened press for coherence. On his or her own, the target might emerge with a higher-level identity that reflects past perspectives or perhaps one that reflects a new integration. But if the influence agent offers a message that provides the missing integration before the target has demonstrated emergence on his or her own, the target may embrace this message as an avenue of emergent understanding, even if it conflicts with his or her prior conception.

The emergence scenario in social influence was tested by asking participants to allocate blame for an alleged rape incident (Vallacher and Selz, 1991). The nature of the incident was such that the motives and intentions of the alleged rapist and victim were open to different interpretations. The incident was presented in the form of a police interview with either the alleged rapist or the victim. Participants read the interview under either a low-level set (reading for detail) or a high-level set (reading for meaning). They then read a police summary concluding either that the perpetrator should be charged with rape or that there were insufficient grounds to

press charges. The participants were then asked to allocate responsibility for the event between the perpetrator and the victim. Participants (both males and females) in the low-level condition assigned blame in line with the police summary they read, whereas those who read the interview under a high-level set were not influenced by the police summary. Focusing on "just the facts" in a case of alleged wrongdoing may reduce the influence of one's personal biases, but this attention to detail can make one all the more vulnerable to influence by other people with biases of their own.

The disassembly-reconfiguration perspective on social influence has been embraced by others in recent years, albeit with considerable refinement and extension to different domains of influence (cf. Knowles and Linn, 2004). Social influence comes in diverse forms (compliance, persuasion, guilt, seduction, etc.), but perhaps these forms are built to a certain extent on a shared platform that reflects people's press for coherent higher-level understanding (Vallacher et al., 2003).

THE EMERGENT THEORY

The extensions we have described were largely unanticipated in the 1980s when we were preoccupied with getting the principles straight. The theory may be poised for yet further growth, but change this time is likely to take the form of transformation in light of new ideas and methods in psychological science. In particular, two lines of theory and research – *mind perception* and *dynamical social psychology* – represent new ways of framing the dynamics of action identification. The mind perception perspective generalizes action identification principles to the understanding of the minds and actions of other people. It is consistent with several emphases in contemporary psychology, including theory of mind, neural bases of empathy, and the perception (and illusion) of agency. Dynamical

social psychology adapts principles and methods of dynamical systems and complexity, which are at the forefront of contemporary theory and research in the natural sciences, to the investigation of personal and social processes. Within this perspective, the principles of action identification capture a basic dynamic scenario in the mind-action relationship.

Mind perception

The experience of high-level identification says a lot about the actor's mind. In particular, when the action is complex and extended in time, a high-level mindset suggests both intentionality (acting on purpose, having a plan, working toward a goal) and executive cognitive processes (conscious control of the action, working memory, mindfulness). Actions performed under lower-level identities also require an active mind, of course, but they don't seem to demand the same caliber of mental states. One can move a finger without a great deal of intention and thought, after all, but appreciating the consequences of the finger movement (e.g., sending an e-mail, firing a gun) implicates mental processes rather directly. A recent neuroimaging (fMRI) study has in fact demonstrated greater activation of brain regions associated with higher-order cognitive processes (e.g., temporo-parietal junction) when people are identifying action in high- as opposed to low-level terms (Marsh et al., 2009).

The connection between a person's level of action identification and his or her mental control of the action provided the original focus of the theory. Recently, though, the connection between identification level and mental states has been extended to the perception of other people's actions and minds (Kozak et al., 2006). People are quite willing and able to infer how other people's minds work and such inferences are central to a host of issues in social judgment, including liking, the attribution of personal versus situational causation, personality judgment, and

allocation of responsibility (e.g., Carruthers and Smith, 1996; Epley and Waytz, 2009; Frith and Frith, 2003; Idson and Mischel, 2001; McPherson-Franz and Janoff-Bulman, 2000; Wegner, 2002). Do people show the same variability in identifying the actions of others as they do in thinking about their own behavior? Is this variability related to the inferences people make about others' mental states? What factors shape each inference – action identification and attribution of mind – and the relationship between such inferences?

To answer these questions, Kozak et al. (2006) modified the BIF (Vallacher and Wegner, 1989) to allow participants to identify a target person's actions rather than their own. They also developed a Mind Attribution Scale (MAS), consisting of ten items assessing participants' inferences about another person's capacity to act with intention, engage in complex cognition, and experience emotion. The BIF and MAS were then employed in several experiments involving vignettes about various hypothetical target persons. Each experiment focused on a particular aspect of the link between action identification, mind attribution, and person perception. Kozak et al. found, first of all, that high-level identification and attribution of mind (intention and complex cognitive processes) were both associated with liking for a target person who was described in fairly neutral terms. Indeed, the level of action identification for liked targets was often higher than the identification level participants' indicated for their own actions. People are constrained by reality (e.g., personal difficulty, unfamiliarity) in identifying their own actions, but the sky is the limit when thinking about the actions of other people.

By itself, the link between liking and inferences about mental states is open to interpretation; liking may cause high-level identification and mind attribution, but the reverse causal path is also plausible. Kozak et al. (2006) untangled this issue in other experiments by having participants make

judgments about a target person who was clearly likable versus unlikable. The results made clear that evaluation was primary: The liked person was credited with higher-level identities and a more complex mind than the disliked person. This is not cause for concern when likability centers on obvious qualities (e.g., honesty versus dishonesty, sociability versus aloofness). But they found the same relationship when liking was manipulated by the target's fortune versus misfortune. There is evidence that victims of misfortune tend to be derogated by others, even when the misfortune is not of their own doing (Lerner, 1980). When participants read a vignette about a male student in financial straits who could only afford one meal a day, they reported unfavorable evaluations of him, identified his behavior in lower-level terms, and credited him with less complex cognitions. Victims, it seems, do not have minds like the rest of us.

There is an exception to the connection between liking and high-level identification. Sometimes the people we like do bad things and sometimes the people we don't like do good things. Kozak et al. found that high-level identities were inferred when the valence of the actor matched the valence of the action. So for liked target persons, positive actions were identified at higher levels than were negative actions, but for disliked target persons, negative actions were identified at higher levels. Because high-level identification is linked to personal responsibility, these results are consistent with the tendency to credit liked others (e.g., friends, heroes) for good behavior and to blame disliked others (enemies) for bad behavior. The results are also consistent with research on people's self-presentation of their own actions (Vallacher et al., 1987). People tend to describe their mistakes in terms of lower-level details but to emphasize personal attributes and goals when describing their successes and noteworthy deeds.¹

If higher-level action identification promotes mind perception, it might also undermine processes that lead to the devaluation of

minds and, ultimately, to dehumanization. The tendency to see others as animals (Epley and Waytz, 2009), robots (Haslam, 2006), objects (Fredrickson and Roberts, 1997), or as otherwise less than fully human (Harris and Fiske, 2006) may thus depend in part on the deconstruction of mind perception through low-level action identification. It is not yet clear which particular aspects of mind perception are reduced in lower-level action identification – whether, for example, minds are seen as less capable of having experiences or being agentic (Gray et al., 2007), or as less capable of serving as moral agents or moral patients (Gray and Wegner, 2009). What is clear is that identification level influences our perception of minds as worth preserving, and so may ultimately be instrumental in leading people to treat each other as less than human. Action identification may be an initial step toward both the moral regard we give to our most respected conspecifics and as well as to the abject disregard we visit on those we fail to recognize as fellow members of the human race.

Dynamical social psychology

The dynamic interplay between higher and lower levels of action identification has a natural resonance with the way dynamical systems in other areas of science evolve, function, and change (cf. Guastello et al., 2009). In its most basic sense, a dynamical system is a set of interconnected elements that influence each other to achieve a common or coordinated state. The resultant higher-order state typically has emergent properties, which simply means that the qualities of the state cannot be reduced to the properties of the constituent elements. Once such a state emerges, it constrains the behavior of the elements that gave rise to it. Because it 'attracts' the system's dynamics, the coherent state is referred to as an *attractor*. A system's attractor stabilizes the system and actively resists change due to outside influences. If change occurs, it is because feedback loops among

the elements are weakened, increasing the degrees of freedom in the system and thereby undermining the coherence of the higher-order state. From this disassembled state of affairs, the system is primed for emergence to a new higher-order state that provides a different configuration of the lower-level elements.

Over the past decade, these hallmarks of a dynamical system have been identified for a variety of social processes, including social judgment, self-concept, social influence, societal transition, and intractable conflict (cf. Vallacher and Nowak, 2007). In self-concept, for example, elements of self-relevant information become integrated to form a coherent perspective on the self, which then constrains the processing of subsequent input, enabling the system to resist change when exposed to inconsistent information or social feedback (Nowak et al., 2000). Self-concept change occurs when the lower-level elements are singled out or decoupled from one another, setting the stage for their reconfiguration with respect to a new and possibly quite different self-view. This basic scenario, which has been detailed as well for the other phenomena indicated above, reflects the essence of the emergence scenario of action identification theory. Not until these manifestations of dynamical social psychology were developed, however, was their genesis in action identification principles appreciated.

Reframing action identification in dynamical terms suggests two refinements of the emergence process. In the basic theory, emergence occurs when people in a low-level state are provided cues (e.g., social feedback, vivid consequences) that signal an action's higher-level meaning. In a dynamical system, though, emergence can occur without external influence due to self-organization among system elements (Vallacher and Nowak, 1997). This means that the intrinsic dynamics of the mental system can promote emergent meaning – an attractor – for an action's lower-level identities. As a person thinks about and performs a sequence of basic acts,

a higher-level identity may spontaneously emerge. Because intrinsic dynamics of mind can take place outside of conscious attention (Port and van Gelder, 1995), new insights into one's action can pop into awareness without warning, in a manner reminiscent of an "Aha!" experience.

The second refinement concerns the potential for multistability in a psychological system (Vallacher and Nowak, 2007). As a system's elements become organized with respect to one coherent state (attractor), the elements that are excluded may form an alternative attractor that can compete for prepotence with the original attractor. When a person develops a highly favorable judgment of someone, for example, inconsistent (i.e., unflattering) information about him or her tends to be discounted. If enough elements of information undergo this fate, they may become organized into an alternative perspective on the target person. If conditions change (e.g., the other person finally goes too far), the judgment system could show a catastrophic shift to the previously latent attractor. The idealized assessment, in other words, could transform into a negative view without going through intermediate steps of disinterest or mild disapproval.

With respect to action, the potential for multistability suggests that although a high-level identity resists change, at some threshold of inconsistent information, the person may suddenly embrace a wholly different high-level identity that has formed by virtue of self-organization among elements that had been discounted in service of maintaining the original identity. A person who stubbornly sees his or her critical comments as constructive despite being told otherwise, for example, might suddenly recognize this behavior as mean-spirited.

The potential for self-organization and multistability have been invoked to understand the nature of social conflicts that have become protracted to the point of seeming intractability, and to suggest new means of resolving such conflicts (e.g., Coleman et al., 2007; Vallacher et al., 2010). When a

conflict develops, the opposing parties each experience a press for integrative understanding that can provide a coherent and stable platform for action. So although the parties may have a wealth of specific knowledge regarding one another, their respective judgments lose complexity – the separate elements of information become linked by positive feedback loops and take on the same (e.g., negative) higher-level meaning. The resultant coherent state functions as an attractor that incorporates new information and resists external forces that threaten to undermine it.

Some conflict-relevant information, though, cannot be interpreted in line with the attractor. The other party may act in an unambiguously positive manner, for example, or previous positive acts by the party may be made salient in memory. Instances of inconsistent information may be discounted or suppressed when first encountered, but over time they may provide the seeds for an alternative attractor associated with positive thoughts and action possibilities. If conditions should change and promote ‘ripeness’ for peace, there may be a sudden and dramatic change to this latent attractor. The potential for sudden transitions in the relations between groups mired in conflict has counter-intuitive implications for conflict resolution. Rather than addressing the issues that launched the conflict, a more effective strategy is to create the basis for an alternative way of thinking and behaving that is likely to be dismissed in the short run but which creates an alternative coherent state that can become manifest in the long run.

THE METATHEORY

At the most basic level, action identification theory is a set of principles concerning the representation and control of action. The principles capture the conflicting forces that interact to promote a particular form of understanding in the face of a multitude of

equally plausible identities for one’s actions. This basic identity for the theory gives rise to a number of implications that expand the possible ways in which the theory can be seen. The interplay of the theory’s principles is manifest in several notable phenomena, including emotion, stability versus change in self-concept, social judgment, social influence, and individual variation in the mind–action relationship. Which of these aspects of the theory are prepotent depends on the social context surrounding the lay person and the research agenda of the psychologist.

At a yet higher level of understanding, action identification theory can be looked upon as a basic dynamic for mind and action that defines human experience. The theory captures the interplay between the often-competing concerns with comprehensive understanding and effective action that underlies personal functioning across social contexts. Because life itself is dynamic, this interplay is iterated continuously on different (embedded) timescales, ensuring complexity and growth as people go about their daily lives. Action identification, in this light, is a specific, lower-level manifestation of a pervasive dynamic that coordinates the interplay between mind and reality in people’s lives. Our goal in this chapter was to illustrate this dynamic. Of course, we were also focused on keystrokes and coming up with a concluding thought. We just did both.

NOTES

1 The results are also consistent with research on intergroup biases in the language used to describe action (Maass et al., 1995). People describe the positive behaviors of ingroup members at more abstract levels than they do the positive behaviors of out-group members, but describe the negative behaviors of outgroup members at more abstract levels than they do the negative behaviors of ingroup behaviors. This resemblance should be viewed cautiously, though, because identification levels center on means–ends relations, not on levels of abstraction *per se*. For example, “pushing a button” is higher-level than “moving a finger” by virtue of the *by*

relation, but the former does not seem to be all that more abstract than the latter.

REFERENCES

Baumeister, R.F. (ed.) (1993) *Self-esteem: The Puzzle of Low Self-regard*. New York: Plenum.

Baumeister, R.F. and Heatherton, T.F. (1996) Self-regulation failure: An overview. *Psychological Inquiry*, 7, 1–15.

Beck, A.T. and Weishaar, M. (2000) Cognitive therapy. In R.J. Corsini and D. Wedding (eds), *Current Psychotherapies*, 6th Edition, pp. 241–272. Itasca, IL: F.E. Peacock.

Carver, C.S. and Scheier, M.F. (1981) *Attention and Self-regulation: A Control-theory Approach to Human Behavior*. New York: Springer-Verlag.

Carver, C.S. and Scheier, M.F. (2002) Control processes and self-organization as complementary principles underlying behavior. *Personality and Social Psychology Review*, 6, 304–315.

Carruthers, P. and Smith, P.K. (eds) (1996) *Theories of Theories of Mind*. Cambridge: Cambridge University Press.

Coleman, P.T., Vallacher, R.R., Nowak, A. and Bui-Wrzosinska, L. (2007) Intractable conflict as an attractor: A dynamical systems approach to conflict escalation and intractability. *American Behavioral Scientist*, 50, 1454–1475.

Csikszentmihalyi, M. (1990) *Flow: The Psychology of Optimal Experience*. New York: Harper & Row.

Danto, A. (1963) What we can do. *Journal of Philosophy*, 40, 435–445.

Epley, N. and Waytz, A. (2009) Mind perception. In D.T. Gilbert, and S.T. Fiske (eds), *The Handbook of Social Psychology*, 5th Edition. New York: Oxford University Press.

Felson, R.B. (1989) Parents and the reflected appraisal process: A longitudinal analysis. *Journal of Personality and Social Psychology*, 56, 965–971.

Fredrickson, B.L. and Roberts, T.A. (1997) Objectification theory. *Psychology of Women Quarterly*, 21, 173–206.

Frith, U. and Frith, C.D. (2003) Development and neurophysiology of mentalizing. *Philosophical Transactions: Biological Sciences*, 358, 459–473.

Gergen, K.J. (1985) The social constructionist movement in modern psychology. *American Psychologist*, 40, 266–275.

Goldman, A.I. (1970) *A Theory of Human Action*. Princeton, NJ: Princeton University Press.

Gray, H.M., Gray, K. and Wegner, D.M. (2007) Dimensions of mind perception. *Science*, 315, 619.

Gray, K. and Wegner, D.M. (2009) Moral typecasting: Divergent perceptions of moral agents and moral patients. *Journal of Personality and Social Psychology*, 96, 505–520.

Guastello, S., Koopmans, M. and Pincus, D. (eds) (2009) *Chaos and Complexity in Psychology: The Theory of Nonlinear Dynamical Systems*. New York: Cambridge University Press.

Hampshire, S. (1960) *Thought and Action*. London: Chatto and Windus.

Harris, L.T. and Fiske, S.T. (2006) Dehumanizing the lowest of the low: Neuro-imaging responses to extreme outgroups. *Psychological Science*, 17, 847–853.

Haslam, N. (2006) Dehumanization: An integrated review. *Personality and Social Psychology Review*, 10, 252–264.

Higgins, E.T. (1998) Promotion and prevention: Regulatory focus as a motivational principle. In M.P. Zanna (ed.), *Advances in Experimental Social Psychology*, Vol. 30, 1–46. New York: Academic Press.

Idson, L.C. and Mischel, W. (2001) The personality of familiar and significant people: The lay perceiver as a social-cognitive theorist. *Journal of Personality and Social Psychology*, 80, 585–596.

James, W. (1890) *Principles of Psychology*. New York: Holt.

Kimble, G.A. and Perlmuter, L.C. (1970) The problem of volition. *Psychological Review*, 77, 361–384.

Knowles, E.S. and Linn, J.A. (eds) (2004) *Resistance and Persuasion*. Mahwah, NJ: Erlbaum.

Kozak, M.N., Marsh, A.A. and Wegner, D.M. (2006) What do I think you're doing? Action identification and mind attribution. *Journal of Personality and Social Psychology*, 90, 543–555.

Kruglanski, A.W. (1975) The endogenous–exogenous partition in attribution theory. *Psychological Review*, 82, 387–406.

Kruglanski, A.W. and Webster, D.M. (1996) Motivated closing of the mind: 'Seizing' and 'freezing'. *Psychological Review*, 103, 263–283.

Langer, E.J. and Imber, L.G. (1979) When practice makes imperfect: Debilitating effects of overlearning. *Journal of Personality and Social Psychology*, 37, 2014–2024.

Lerner, M.J. (1980) *The Belief in a Just World: A Fundamental Delusion*. New York: Plenum.

Lewin, K. (1936) *Principles of Topological Psychology*. New York: McGraw-Hill.

Maass, A., Milesi, A., Zabbiini, S. and Stahlberg, D. (1995) Linguistic intergroup bias: Differential expectancies or in-group protection? *Journal of Personality and Social Psychology*, 68, 116–126.

Marsh, A.A., Kozak, M.N., Reid, M., Yu, H., Wegner, D.M. and Blair, R.J.R. (2009) The neural substrates of action identification. Unpublished manuscript, Georgetown University.

McPherson-Franz, C. and Janoff-Bulman, R. (2000) Considering both sides: The limits of perspective-taking. *Basic and Applied Social Psychology*, 22, 31–42.

Mead, G.H. (1934) *Mind, Self, and Society*. Chicago: University of Chicago Press.

Meichenbaum, D. (1977) *Cognitive Behavior Modification*. New York: Plenum.

Miller, G.A., Galanter, E. and Pribram, K.H. (1960) *Plans and the Structure of Behavior*. New York: Holt.

Newtonson, D. (1976) Foundations of attribution: The perception of ongoing behavior. In J.H. Harvey, W.J. Ickes, and R.F. Kidd (eds), *New Directions in Attribution Research*, 1, 223–247. Hillsdale, NJ: Erlbaum.

Nowak, A., Vallacher, R.R., Tesser, A. and Borkowski, W. (2000) Society of self: The emergence of collective properties in self-structure. *Psychological Review*, 107, 39–61.

Port, R.F. and van Gelder, T. (eds) (1995) *Mind as Motion: Explorations in the Dynamics of Cognition*. Cambridge, MA: MIT Press.

Ryle, G. (1949) *The Concept of Mind*. London: Hutchinson.

Sarason, I.G. (1972) Experimental approaches to test anxiety: Attention and the uses of information. In C.D. Spielberger (ed.), *Anxiety: Current Trends in Theory and Research*, 2, 383–403. New York: Academic Press.

Simon, H.A. (1967) Motivational and emotional controls of cognition. *Psychological Review*, 74, 29–39.

Swann, W.B., Jr. (1990) To be adored or to be known? The interplay of self-enhancement and self-verification. In E.T. Higgins and R.M. Sorrentino (eds), *Handbook of Motivation and Cognition*, 2, 408–448. New York: Guilford.

Swann, W.B., Jr., Hixon, J.G., Stein-Seroussi, A. and Gilbert, D. (1990) The fleeting gleam of praise: Cognitive processes underlying behavioral reactions to self-relevant feedback. *Journal of Personality and Social Psychology*, 59, 17–26.

Swann, W.B., Jr., Wenzlaff, R.M., Krull, D.S. and Pelham, B.W. (1992) Allure of negative feedback: Self-verification strivings among depressed persons. *Journal of Abnormal Psychology*, 101, 293–306.

Tesser, A., Martin, L. and Cornell, D. (1996) On the substitutability of self-protective mechanisms. In P.M. Gollwitzer and J.A. Bargh (eds), *The Psychology of Action*, pp. 48–68. New York: Guilford.

Trope, Y. and Lieberman, N. (2003) Temporal construal. *Psychological Review*, 110, 403–421.

Vallacher, R.R. and Nowak, A. (1997) The emergence of dynamical social psychology. *Psychological Inquiry*, 4, 73–99.

Vallacher, R.R. and Nowak, A. (1999) The dynamics of self-regulation. In R.S. Wyer, Jr. (ed.), *Advances in Self-regulation*, 12, 241–259. Mahwah, NJ: Erlbaum.

Vallacher, R.R. and Nowak, A. (2007) Dynamical social psychology: Finding order in the flow of human experience. In A.W. Kruglanski and E.T. Higgins (eds), *Social Psychology: Handbook of Basic Principles*, pp. 734–758. New York: Guilford.

Vallacher, R.R., Nowak, A. and Miller, M.E. (2003) Social influence and group dynamics. In T. Millon and M.J. Lerner (eds), *Handbook of Psychology, Vol. 5: Personality and Social Psychology*, 383–417. New York: Wiley.

Vallacher, R.R. and Selz, K. (1991) Who's to blame? Action identification in allocating responsibility for alleged rape. *Social Cognition*, 9, 194–219.

Vallacher, R.R. and Wegner, D.M. (1985) *A Theory of Action Identification*. Hillsdale, NJ: Erlbaum.

Vallacher, R.R. and Wegner, D.M. (1987) What do people think they're doing? Action identification and human behavior. *Psychological Review*, 94, 3–15.

Vallacher, R.R. and Wegner, D.M. (1989) Levels of personal agency: Individual variation in action identification. *Journal of Personality and Social Psychology*, 57, 660–671.

Vallacher, R.R., Wegner, D.M. and Frederick, J. (1987) The presentation of self through action identification. *Social Cognition*, 5, 301–322.

Vallacher, R.R., Wegner, D.M. and Somoza, M.P. (1989) That's easy for you to say: Action identification and speech fluency. *Journal of Personality and Social Psychology*, 56, 199–208.

Wegner, D.M. (2002) *The Illusion of Conscious Will*. Cambridge, MA: MIT Press.

Wegner, D.M. and Giuliano, T. (1980) Arousal-induced attention to self. *Journal of Personality and Social Psychology*, 38, 719–726.

Wegner, D.M. and Vallacher, R.R. (1986) Action identification. In R.M. Sorrentino and E.T. Higgins (eds), *Handbook of Motivation and Cognition*, 1, 550–582. New York: Guilford.

Wegner, D.M., Vallacher, R.R., Kiersted, G. and Dizadji, D. (1986) Action identification in the emergence of social behavior. *Social Cognition*, 4, 18–38.

Wegner, D.M., Vallacher, R.R., Macomber, G., Wood, R. and Arps, K. (1984) The emergence of action. *Journal of Personality and Social Psychology*, 46, 269–279.

Wicklund, R.A. and Frey, D. (1980) Self-awareness theory: When the self makes a difference. In D.M. Wegner and R.R. Vallacher (eds), *The Self in Social Psychology*, pp. 31–54. New York: Oxford University Press.

Wine, J.D. (1971) Test anxiety and the direction of attention. *Psychological Bulletin*, 76, 92–104.

Winkielman, P. and Cacioppo, J.T. (2001) Mind at ease puts a smile on the face: Psychophysiological evidence that processing facilitation leads to positive affect. *Journal of Personality and Social Psychology*, 81, 989–1000.

Social Cognitive Theory

Albert Bandura

ABSTRACT

The present chapter traces the evaluation of social cognitive theory toward a psychology of human agency. The nature of human agency and the mechanisms through which it operates is analyzed in the context of major changes in the conception of human behavior. These theoretical orientations include behaviorism, the tight hold of the psychoanalytic movement on psychology and the pop culture, conceptions of mind as symbol manipulator in the likeness of the linear computer, eliminative physicalism contending that human behavior is shaped and regulated by neurophysiological mechanisms that operate outside one's awareness and control, and the growing ascendancy of human agency in the coevolution process. The utility of a psychological theory is judged by three criteria: its explanatory power, its predictive power and, in the final analysis, its operative power to effect personal and social change. Social cognitive theory lends itself readily to social applications. Three illustrative applications document the transformative changes in the field of psychotherapy, development of large-scale health promotion systems founded on a shift in the health field from a disease model to a health model, and global applications that address some of the most urgent worldwide problems.

INTRODUCTION

The present chapter traces the evolution of social cognitive theory toward a psychology of human agency. To be an agent is to influence the course of events by one's actions. In this view, people are contributors to their life circumstances, not just products of them (Bandura, 2006a, 2008a). The agentic theoretical perspective serves as the integrative principle in human self-development, adaptations, and change. Human functioning is rooted in social systems. Therefore, personal agency operates within a broad network of sociostructural influences. In these agentic transactions, people create social systems and the practices of social systems, in turn, influence personal development and functioning. Given this dynamic bidirectional influence, social cognitive theory rejects a dualism between personal agency and a social structure disembodied from human activity.

The theoretical framework guiding my work was originally labeled *social learning theory*.

I later relabeled the theory as *social cognitive theory* for several reasons (Bandura, 1986). A variety of theories founded on divergent tenets – Miller and Dollard's drive theory, Rotter's outcome expectancy theory, Gewirtz's operant conditioning theory, and Patterson's functionalist theory – were all christened social learning theory. This created untold confusion in the literature concerning the theory being tested. Moreover, the theory under discussion had always been much broader than the initial descriptive label as a theory of learning. It not only addressed how people acquire knowledge and competencies but also how they motivate and regulate their behavior and create social systems that organize and structure their lives. In the more fitting appellation as *social cognitive theory*, the *social* portion of the title acknowledges the social origins of much human thought and action; the *cognitive* portion recognizes the influential contribution of cognitive processes to human motivation, affect, and action.

When I began my career, behaviorism had a stranglehold on the field of psychology. It focused primarily on learning by direct experience through paired stimulation and rewarding and punishing response consequences. Methodological reductionism prescribed the research strategy, typically with rats and pigeons, on the assumption that understanding rudimentary learning processes would explain complex human behavior. Cognitive processes were dismissed as redundant inner links in the chain of causation or as explanatory fictions.

The behavioristic theorizing was discordant with the evident social reality that much of what we learn is through the power of social modeling. I found it difficult to conceive of a culture in which its intricate competencies, language, mores, customs, and familial, educational, occupational, religious, and political practices were laboriously shaped by rewarding and punishing consequences of trial-and-error performances. This tedious and potentially hazardous process is short cut by social modeling.

CENTRALITY AND Pervasiveness of Social Modeling

Despite the centrality and pervasiveness of social modeling there was little theorizing and research on the nature, scope, and mechanisms governing this basic mode of social influence. Some of the early accounts of social modeling characterized it as imitation and marginalized it as simply mimicry of specific acts. This narrow conception discouraged interest in the phenomenon. For example, Miller and Dollard (1941), viewing social modeling from the behavioristic perspective, treated it as a special case of discrimination learning. A model provides a social cue, the observer performs a matching response, and its reinforcement strengthens the tendency to behave imitatively. Personality and developmental theorists conceptualized it as identification involving wholesale incorporation of modeled attributes. The defining properties of identification were too diffuse, arbitrary, and empirically questionable either to clarify modeling processes or to guide scientific inquiry (Bandura, 1969). I found these early conceptions seriously wanting on the determinants, mechanisms, and scope of social modeling. We launched a program of research on social modeling as it typically occurs observationally in the absence of reinforced performance.

In a chapter entitled "Vicarious Processes: A Case of No-trial Learning" (Bandura, 1965), I presented the findings of our studies showing that observational learning through exposure to models requires neither response enactment nor reinforcement. Social modeling operated through four cognitive sub-functions encompassing attentional, representational, enactive translational, and motivational processes (Bandura, 1971a). I came under heavy fire from operant conditioners for whom nonreinforced modeling posed a major problem for their explanatory system (Baer et al., 1967). They contended that reinforcement of some matching responses established imitation as a conditioned reinforcer. We demonstrated that generalized imitation is governed by beliefs

about how to influence the models' behavior and outcome expectations for modeled activities rather than by conditioned reinforcers.

Social cognitive theory broadened the scope of modeling influences and the functions it serves (Bandura, 1986). In addition to cultivating cognitive and behavioral competencies, modeling influences were shown to alter motivation, create and modify emotional proclivities, serve as social prompts that activate, channel, and support given styles of behavior, and shape images of reality.

POWER AND REACH OF SYMBOLIC MODELING

A growing influential source of social learning is the varied and pervasive symbolic modeling through the electronic media. A major advantage of symbolic modeling is that it can transmit information of virtually limitless variety to vast populations simultaneously in widely dispersed locales. Indeed, the extraordinary advances in the technology of communication are transforming the nature, reach, speed, and loci of human influence (Bandura, 2002a). Modeled new ideas, values, and styles of conduct are now being rapidly spread worldwide in ways that foster a globally distributed consciousness.

Social cognitive theory addressed the personal and social structural factors that determine the adoption of innovation and the social networks through which the influence diffuses (Bandura, 2006b). The evolving information technologies increasingly serve as a vehicle for building social networks that transcend the barriers of time and space. The Internet provides instant communicative access worldwide. Global broadcasts now show sociopolitical conflicts, the strategies and countermeasures used, and their effects as they are happening. This makes electronic modeling a powerful vehicle for transcultural and sociopolitical change (Bandura, 2002a; Braithwaite, 1994).

Our traditional theories of human behavior were formulated long before these revolutionary

advances in communication technologies. The theoretical explanations of human behavior were heavily rooted in influences from the local social environment and the effects of enactive experiences. The social reality of contemporary society is markedly different with growing ascendancy of the symbolic environment and increased opportunities for the exercise of personal and collective agency in self-development, adaptation, and change.

CORRECTING MISCONCEPTIONS ABOUT THE NATURE AND SCOPE OF MODELING

There were a number of entrenched misconceptions about the nature and scope of modeling that put a damper on research and social applications of this powerful mode of learning and social influence. Progress in this area, therefore, required research designed not only to elucidate the determinants and mechanisms of social modeling but also to put the misconceptions to rest. One such misconception was that modeling, construed as "imitation," could produce only response mimicry. This is largely the legacy of the early narrow conceptions of modeling. Exemplars usually differ in content and other details but embody the same underlying principle. To cite a simple example, the passive linguistic form may be embodied in any variety of utterances. Research on abstract modeling showed that modeling involves abstracting the information conveyed by specific exemplars about the structure and the underlying principles governing the behavior, rather than simply mimicking the specific exemplars (Bandura, 1986; Rosenthal and Zimmerman, 1978). Once individuals learn the guiding principle, they can use it to generate new versions of the behavior that go beyond what they have seen or heard. They can tailor the behavior to suit changing circumstances.

There was another oft-repeated misconception regarding the scope of modeling. Many activities

involve cognitive skills on how to acquire and use information for predicting and solving problems. Critics argued that modeling cannot build cognitive skills because thought processes are covert and are not adequately reflected in modeled actions, which are the end-products of the cognitive operations. This was a limitation of conceptual vision rather than an inherent limitation of modeling. Cognitive skills can be readily exemplified and cultivated by cognitive modeling in which models verbalize aloud their reasoning strategies as they engage in problem-solving activities. The thoughts guiding their decisions and actions are thus made observable. Cognitive modeling was shown to be more powerful in enhancing perceived self-efficacy and building complex cognitive skills than the commonly used tutorial methods.

Another misconception requiring retirement held that modeling is antithetical to creativity. We were able to show how innovation can emerge through modeling. When exposed to models who differ in their styles of thinking and behavior, observers rarely pattern their behavior exclusively after a single source. Nor do they adopt all the attributes even of preferred models. Rather, observers combine various features of different models into new amalgams that differ from the individual modeled sources. Thus, two observers can construct new forms of behavior solely through modeling that differ from each other by selectively blending different features from the variant models. In many social and technological innovations, individuals adopt modeled aspects found to be effective, improve upon them, synthesize them into new forms, and tailor them to their particular circumstances. In short, selective modeling is often the mother of innovation.

BAPTISM IN POWER POLITICS

At the time that I was conducting the modeling experiments in the late 1950s, television had diffused rapidly throughout society. The advent of television vastly expanded the

range of models available to the general public. The broadcast industry traded heavily on gratuitous violence in the belief that violence sells. Television provided viewers with unlimited opportunities day in and day out to learn the whole gamut of homicidal conduct within the comfort of their homes. There was growing public concern about the possible effects of televised violence on children.

Among the different experimental methods I used to study observational learning was the oft-cited Bobo doll studies on the acquisition of novel forms of aggression through modeling (Bandura et al., 1963). The theory in vogue at the time contended that exposure to modeled aggression is cathartic. It reduces aggression by draining aggressive impulses. We found otherwise. Children who had observed an adult aggress in unique verbal and physical ways toward an inflated Bobo doll modeled the aggressive styles of conduct. They also were less restrained in expressing, in their play activity, aggression they had learned elsewhere, such as attraction to guns.

I was invited to testify before the Senate Communications Committee, the Federal Trade Commission, and the National Commission on the Causes and Prevention of Violence prompted by the assassination of Robert Kennedy. The Federal Trade Commission was troubled by increasing reports of serious injuries suffered by children who modeled hazardous activities in televised advertisements. The commission used our research findings on modeling to get advertisers to alter ads depicting injurious feats by children on bicycles and dune buggies, ads for headache remedies in which the characters induce splitting headaches by pounding each other on the head with mallets, and other types of ads showing children performing activities that risk serious injury.

This excursion into the public policy arena provided a sobering glimpse into the power of the broadcast industry, some of which was directed at me personally. I got my first inkling into the exercise of this power at a

meeting convened by the National Institute of Mental Health (NIMH) to draft a research agenda on television's effects. Surprisingly, we met at the plush Waldorf Towers in New York rather than in Washington for what turned out to be essentially a production staged by the broadcast industry under the auspices of NIMH. After we identified the different lines of research that could advance the understanding of television's effects, the research community was invited to submit grant proposals. A review panel, meeting in a luxurious Caribbean setting, rejected my proposal.

Look magazine invited me to write a piece on the social influence of television for a special issue they were publishing on youth. When it appeared, the Television Information Office, a subsidiary of the National Association of Broadcasters, sent a large packet to its sponsor stations explaining why my research on social modeling should be disregarded. This was just the beginning of a multipronged offense. Psychologist Ruth Hartley prepared a document commissioned by CBS in which she took me to task and criticized the relevance of other experimental studies demonstrating a positive relation between exposure to violent fare and aggressive behavior. In an editorial in *TV Guide* titled, "A Child is Not Rat," Edith Efron (1969a), senior editor of *TV Guide*, misrepresented the modeling studies as "conditioning studies." She cited Hartley, whom she dubbed the "best-known attacker," as the authoritative critic of experimental research on the effects of televised violence. In an expansive indictment, Efron not only took issue with my study, but included "virtually all of his colleagues" as well in the critique of experimental studies of aggressive modeling.

With financial sponsorship and coproduction by CBS, Milgram and Shotland (1973) conducted studies showing that exposure to modeled thievery does not lead viewers to steal money from a charity box labeled *Good Ship Hope* for a medical charity that treats poor children worldwide. The charity box

was mounted on a poster showing a physician treating a little girl and a picture of the hospital ship with the words, "Where there is Hope there is life." This experimental setup is analogous to demonstrating television null effects by showing that viewers will not rip off charitable contributions to Mother Theresa. As the saying goes, there is honor even among thieves. The studies were published as a book and distributed free of charge by the network. They did not survive conceptual and empirical scrutiny. In an author editorial in *TV Guide* under the title "The Man in the Eye of the Hurricane," Edith Efron (1969b) dismissed the modeling studies, complained that the research by members of the "Bandura school ... won them center stage in Washington," and criticized the Surgeon General's office for acting "as if Rome were burning and Dr. Bandura were a fire extinguisher" (1969: 37).

One evening I received a call from one of my graduate students telling me to turn on my television set to see the character playing my role undergoing a blistering cross-examination concerning the modeling studies. I wasn't doing too well! In the plot-line of this televised movie, a beleaguered wife of a screenwriter defends him as he is being unmercifully victimized by a haranguing press and a vindictive mother who claims her son's crime was prompted by a similar act in one of the screenwriter's televised plots. The cross-examiner was disputing evidence that televised violence affects aggressive behavior. As I was being pummeled by media-commissioned critiques, sponsored studies, paid consultants, and fictionalized dramas, I began to feel a kinship with the battered Bobo doll!

Failure to distinguish between the diverse effects of televised violence and the appropriate methodologies for elucidating them provided a fertile ground for disputes. Different lines of research identified four major effects of exposure to televised violence: it can teach novel aggressive styles of conduct; weaken restraints over interpersonal aggression by legitimizing, glamorizing, and

trivializing violent conduct; desensitize and habituate viewers to human cruelty; and shape public images of reality. Each of these separable effects requires a different methodology (Bandura, 2009).

I had to address misunderstandings and misrepresentations of the research using novel modeled aggression to study observational learning. The mistaken critique, which continues to be repeated in our textbooks, is that the study used a nonhuman target and Bobo dolls are for punching. The Bobo doll laboratory experiments were designed to clarify the processes governing observational learning. The methodology for measuring learning effects requires conditions in which viewers feel free to reveal all they have learned. This requires simulated targets rather than retaliative ones. To use human targets to assess the instructive function of televised influence would be as nonsensical as to require bombardiers to bomb San Francisco, New York, or some other inhabited locations to test the extent to which they had acquired bombing skills.

We were not interested in whether children punched the Bobo doll. Rather, we measured whether children assaulted it in the novel modeled ways, such as pummeling it with a mallet and voiced the novel aggressive neologisms as they assaulted the doll. Children in the control condition never exhibited the highly novel form of aggression. Although modeled aggression was only one among a variety of experimental paradigms we used to clarify the mechanisms governing diverse modeling effects, it is the one that is featured in portrayals of social cognitive theory.

There are more chapters to the exercise of political leverage regarding research on media effects. The National Commission on the Causes and Prevention of Violence (1969) was about to release its report concluding, in the mass media section, that the empirical evidence taken as a whole was supportive of a positive relation between televised violence and aggressive behavior. In a surprise move, Senator Pastore, a supporter of the broadcast industry (Paisley, 1972) who chaired the Communications Subcommittee, instructed the Surgeon General, with President Nixon's

endorsement, to assemble a committee of experts to evaluate the effects of televised violence and to allocate a million dollars for new research on this topic. The first meeting of the evaluation committee took place at the Center for Advanced Study at Stanford. Ed Parker, who coauthored a book on *Television in the Lives of Children*, and I were invited to sit in on the meeting. We were surprised to find that 40 percent of the committee membership were tied to the broadcast industry – two network researchers, two network consultants, and a former research executive at CBS.

We enlisted Senator Metcalf, a Stanford graduate, to obtain information on the selection procedure. Health, Education, and Welfare Secretary Finch explained that each network was allowed to veto, without explanation, any of the nominees on the list submitted by professional associations and the broadcast networks. I was one of eight researchers, including Len Berkowitz, Percy Tannenbaum, child psychiatrist Leon Eisenberg, and sociologists Leo Bogart and Otto Larsen, who were vetoed. Finch provided two justifications for the veto procedure – precedent and objectivity. He explained that the tobacco industry was given veto power in the formation of the committee to evaluate the health effects of smoking. The media report would have greater impact, he claimed, if the committee members were entirely objective. Senator Metcalf was astonished to learn that the tobacco industry was also given sole veto power. He questioned the selective privilege of veto power given to the broadcast industry and how stacking the committee with folks tied to the television industry accomplished impartiality.

Writing the report created headaches for the broadcast-linked members because the empirical data were not friendly to the conclusion of cathartic or null effects. The report was written in opaque technobabble that was better suited to confuse than to inform the public. Rose Goldsen (1972), a Cornell sociologist, dubbed the report “science in wonderland.” Before the report was released, a copy was leaked to Jack Gould (1972) of the *New York Times*, which published a column

on the report under the misleading headline, “TV Violence Held Unharmful to Youth.”

Researchers who conducted the studies for the Committee were incensed at the misrepresentation of their findings. They protested to Senator Pastore, who then scheduled an open Senate hearing on the committee’s report. After years of obfuscation, negation, and disparagement of research programs by the broadcast industry, their own chief researcher, Joseph Klapper, acknowledged at the hearings, “There were indications of a causal relationship … The catharsis theory had no empirical support.” No US network reported on the Senate hearing. Because of concern over the spillover of US televised violence into Canada, the Film Board of Canada (1972) filmed the entire Senate hearing. Several social scientists reported on the perversion of the scientific review process. Mathilda Paisley (1972) wrote a piece on violence done to TV violence research. In a book devoted to this controversial episode, *TV Violence and the Child*, Cater and Strickland (1975) traced the evolution and fate of the report. *Science* published a lead article documenting and condemning the misuse of the scientific advisory system for policy initiatives (Boffey and Walsh, 1970).

The late President Johnson once remarked that politics is like sausage making. You don’t want to examine what goes into it. Social scientists seek to advance knowledge that can inform public policy. As revealed in the stealthy workings of the sociopolitical forces swirling around the issue of television violence, we also need to study how politics and power, which shape public policy, determine how our knowledge is used. Policy research is difficult to conduct, and we do little of it.

TRANSFORMATION OF THE FIELD OF PSYCHOTHERAPY

While behaviorism ruled over general psychology when I entered the field, psychodynamic theory, especially the psychoanalytic

form, reigned over the fields of personality, developmental, psychotherapy, and the pop culture. The mid-1950s witnessed growing disillusionment with this line of theorizing and its mode of treatment. The theory lacked predictive power and did not fare well in therapeutic effectiveness. Dick Walters and I provided an alternative view of human behavior in the book, *Social Learning and Personality Development* (Bandura and Walters, 1963).

During this period, I was teaching courses on personal and social change at Stanford. I was intrigued by cases in which direct modification of problem behavior not only produced lasting improvements in people’s lives but fostered generalized benefits in nontreated areas of functioning. Drawing on an emerging literature on psychosocial change, I published the article “Psychotherapy as a Learning Process” in the *Psychological Bulletin* (Bandura, 1961). It was organized around six basic principles of personal change.

The time was apparently ripe for a new direction in the conceptualization and treatment of behavior. I was flooded with reprint requests from home and abroad across specialties and disciplinary domains. Eysenck invited me to contribute a chapter to a volume he was editing. The chapter kept enlarging until it outgrew the assignment. Instead, it turned into the volume *Principles of Behavior Modification* (Bandura, 1969). It addressed the influential role of cognitive, vicarious, and self-regulatory mechanisms in human functioning.

We were devising new modes of treatment for phobic conditions using guided mastery experiences as the principal vehicle of change. With appropriate mastery aids seemingly unachievable changes become doable. (Bandura et al., 1969). This proved to be a consistently powerful treatment that instilled a robust sense of coping efficacy; transformed attitudes toward the phobic objects from abhorrence to liking and wiped out anxiety, biological stress reactions, and phobic behavior. These people had been

plagued by recurrent nightmares for 20 or 30 years. One of the most striking changes was the power of mastery experiences to transform dream activity and wipe out chronic nightmares. For example, as one woman gained mastery over her snake phobia, she dreamt that the boa constrictor befriended her and was helping her to wash the dishes! Reptiles soon faded from her dreams. The changes endured.

The 1960s ushered in remarkable transformative changes in the explanation and modification of human functioning and change (Bandura, 2004b). Causal analysis shifted from unconscious psychic dynamics to transactional psychosocial dynamics. Human functioning was construed as the product of the dynamic interplay between personal, behavioral, and environmental influences. Action-oriented treatments replaced interpretive talk therapies. The modes of treatment were altered in the content, locus, and agents of change. Within a decade, the field was transformed by a major paradigm shift. New conceptual models and analytic methodologies were created. New sets of periodicals were launched for the rising stream of interest. New organizations were formed for the advancement of behaviorally oriented approaches. New professional conventions provided a forum for the exchange of ideas.

Not all the critics of the psychodynamic model worshipped at the same theoretical alter, however. Some thought the operant conditioning route provided the best glimpse of the promised land. Others adopted Hullian theory. I took the social cognitive route, emphasizing the influential role of agentic capabilities in self-development, adaptation, and change. Vigorous battles were fought over cognitive determinants and their scientific legitimacy (Bandura, 1995b, 1996; Catania, 1975; Skinner, 1977).

The popular media were deluging the public with repugnant imagery of brainwashing and frightful scenarios of *1984* and *Brave New World* dominated by social engineers wielding powerful methods of behavioral control.

The hit movie, *A Clockwork Orange*, graphically portrayed the fiendish nature of behavior modifiers physically shocking people into submission. In his movie *Sleeper*, Woody Allen amusingly outwits the ironclad control by despotic social engineers who reduce humans to mindless zombies. Skinner's (1971) publication, *Beyond Freedom and Dignity*, alarmed the public that the application of these new psychological methods would strip people of their dignity and deprive them of their freedom. The Unabomber targeted Jim McConnell at the University of Michigan as his first victim with a tirade about the evils of behavior modification. Lyndon La Rouche, who became a perennial candidate for the US presidency, branded the practitioners of behavioral approaches as "Rockefeller Nazis," formally tried some of the leading figures in his tribunal for crimes against humanity, stormed classes at the University of New York at Stony Brook, and issued threats requiring police surveillance of the Associates for the Advancement and Behavior Therapy (AABT) convention in Chicago. As in any professional practice, there were some reprehensible applications of behavioral principles, especially in coercive institutional systems, that affirmed and fueled the public's fears.

At the height of this media frenzy, I began my term as president of the American Psychological Association (APA). A responsible social science must concern itself not only with the advancement of knowledge but also with the effects of its social applications. In keeping with this dual commitment, we formed an APA interdisciplinary task force to examine the way in which knowledge on behavioral modification was being used both at the individual and institutional level. Its wide-ranging analysis, which was published in the volume, *Ethical Issues in Behavior Modification* (Stolz, 1978), provided a thoughtful evaluation of existing applications and a set of standards for ethical practice that helped to dispel the frightful misconceptions propagated by the mass media. Growing applications of

cognitive behavioral treatments not only won public acceptance, but are now used as the evidence-based method of choice for diverse maladies of the human condition. This fascinating odyssey involved dual transformative changes – a paradigm shift in theory and practice as well as a sweeping change in public acceptance.

SELF-EFFICACY COMPONENT IN SOCIAL COGNITIVE THEORY

My entry into self-efficacy was serendipitous. In the development and evaluation of the guided mastery treatment, we focused on three fundamental processes: the power of the treatment to promote psychosocial changes; the generality or scope of the effected changes; and their durability or maintenance. Having demonstrated the power of this mode of treatment on each of these evaluative dimensions, I explored the possibility of a further function, the power of a treatment to build resilience to adverse experiences. We tested the proposition that a lot of neutral or positive experiences after functioning is fully restored can neutralize the negative impact of an aversive event and curtail the spread of negative effects.

In a follow-up assessment, the participants expressed deep gratitude to be rid of their phobia, but then explained that the treatment had a more profound impact. Their lives had been debilitated intrapsychically, socially, recreationally, and occupationally for 20 to 30 years. They were plagued by recurrent nightmares and perturbing ruminations. To overcome, within a few hours, a phobic dread that had constricted and tormented their lives was a transformational experience that radically altered their beliefs in their efficacy to exercise control over their lives. They were acting on their new self-efficacy belief and enjoying their successes, much to their surprise. These preliminary findings pointed to a common mechanism through which personal agency is exercised.

I mounted a multifaceted program of research to gain a deeper understanding of the nature and function of this belief system. To guide this new mission, the theory addressed the key aspects of perceived self-efficacy (Bandura, 1997). These include the origins of efficacy beliefs, their structure and function, their diverse effects, the processes through which they work, and the modes of influence by which a resilient sense of efficacy can be created and strengthened for personal and social change. Diverse lines of research, conducted by a variety of investigators, provided new insights into the role of perceived self-efficacy in the fields of education, health promotion and disease prevention, clinical dysfunctions such as anxiety disorders, depression, eating disorders, substance abuse, personal and team athletic attainments, organizational functioning, and the efficacy of our social and political systems to make a difference in people's lives.

A major question in any theory of cognitive regulation of motivation, affect, and action concerns the issue of causality. A variety of experimental strategies were used to verify that beliefs of personal efficacy are contributors to performance, not merely reflectors of it (Bandura, 1997; Bandura and Locke, 2003). The issues raised by devout proponents of competing theories were also addressed both conceptually and empirically (Bandura, 2008a).

I receive a steady flow of e-mails requesting my all-purpose measure of self-efficacy or a couple of trait-like items that could be inserted into an omnibus questionnaire. There is no all-purpose measure of self-efficacy. The requesters are referred to a detailed instructional manual on how to construct psychometrically sound self-efficacy scales (Bandura, 2006d). Self-efficacy assessments are tailored to spheres of functioning and the realities people have to manage. Another entry in the research agenda was to differentiate an agentic model of personality from a trait model. It also required purging misconceptions of constructs. *Self-efficacy* is a judgment of personal capability; *self-esteem* is a

judgment of self-worth. These are entirely different constructs. Nor is self-efficacy the same as *locus of control*, which is a belief about whether outcomes are contingent on behavior or on extraneous factors. Belief that outcomes are determined by one's performance (internal locus) can be motivating under high self-efficacy but demoralizing under low self-efficacy to produce the required performance.

The field of personality is deeply entrenched in trait thinking. The currently-in-vogue "Big-Five" model shrinks personality to five clusters of behaviors disconnected from the vast body of knowledge on the development, organization, regulation, and modification of behavior (McCrae and Costa, 1996). These traits are measured by decontextualized behavioral descriptors in a one-size-fits-all questionnaire. Human functioning is too multifaceted, contextualized, and conditionally manifested to be reduced to a small number of behavioral descriptors reified as personality determinants.

A five-fold behavioral taxonomy is hardly a theory of personality. The traits comprising this approach should be renamed behavioral traits rather than personality traits because the items are mainly clusters of behaviors. One needs a theory of personality to explain how intrapersonal factors contribute to the development and adoption of conscientious, agreeable, receptive, and socially outgoing behavior. In a chapter entitled "Social Cognitive Theory of Personality," I argued that personality determinants reside in intrapersonal factors and psychosocial processes not in behavioral clusters (Bandura, 1999b). The convenience of all-purpose global tests of personal attributes is gained at the cost of explanatory and predictive power. All too often, personality psychology is marginalized as simply a supplier of handy off-the-shelf trait measures. Such measures are being appended, often with little conceptual rationale, to whatever one is studying under the illusion that it represents the contribution of "personality" to human functioning.

EVOLUTION OF AN AGENTIC THEORY OF HUMAN BEHAVIOR

There are four core properties of human agency (Bandura, 2006a, 2008a). One such property is *intentionality*. People form intentions that include action plans and strategies for realizing them. The second property involves the temporal extension of agency through *forethought*. This includes more than future-directed plans. People set themselves goals and anticipate likely outcomes of prospective actions to guide and motivate their efforts anticipatorily. When projected over a long time course, a forethoughtful perspective provides direction, coherence, and meaning to one's life. The third feature is *self-reactiveness*. Agents are not only planners and forethinkers. They are also self-regulators. They adopt personal standards and monitor and regulate their actions by evaluative self-reactions. They do things that give them satisfaction and a sense of self-worth, and refrain from actions that bring self-censure. The fourth feature is *self-reflexiveness*. People are not only agents of action. They are self-examiners of their own functioning. Through functional self-awareness they reflect on their personal efficacy, the soundness of their thoughts and actions, the meaning of their pursuits, and try to make corrective adjustments if necessary.

LOCUS OF CAUSATION

Over the years, theorists engaged in vigorous debates on whether the causes of human behavior reside in the environment as the situationists claim, or in the individual as the dispositionalists claim. Although most theorists now adopt an interactionist model of causation, we are still arguing over whether the person or the situation is the locus of causation. Sociologists and many social psychologists argue for the power of the situation, while personologists argue for the power of the person. In the interactionist view,

human functioning is a product of the interaction between personal and environmental influences. However, there are three types of interactionism, two of which subscribe to one-way causation in the link to behavior.

In the *unidirectional* model, persons and situations are treated as independent influences that combine in unspecified ways to produce behavior. The major weakness with this causal model is that personal and environmental influences do not function as independent determinants. They affect each other. People create, alter, and destroy environments. The changes they produce in environmental conditions, in turn, affect them personally. The causally unidirectional relation to behavior is another serious deficiency of this model of interactionism.

The *partially bidirectional* model of interaction, which is widely adopted in personality theory, acknowledges that persons and situations affect each other, but still treats influences relating to behavior as flowing in only one direction. The person-situation interchange is said to produce behavior unidirectionally, but the behavior itself does not affect the ongoing transaction between the person and the situation. A major limitation of this causal model is that behavior is not procreated by an intimate interchange between a behaviorless person and the environment. Such a feat would be analogous to immaculate conception. Except through their social stimulus value, people cannot affect their environment, other than by their actions. Behavior is an interacting determinant rather than a detached byproduct of a behaviorless person-situation interchange.

The partially bidirectional model of interactionism is typically evaluated by partitioning the average amount of variance in behavior attributed to persons, environments, and their interaction. In this conception, the interactional form is statistical not an interdependently dynamic one in the causal structure. The statistical partitioning reminds one of the nonswimming statistician who drowned while crossing a river that averaged two feet in depth.

Social cognitive theory conceptualizes the interactional causal structure as *triadic reciprocal causation*. In this conception, human functioning is a product of a reciprocal interplay of intrapersonal, behavioral, and environmental determinants (Bandura, 1978). Phillips and Orton (1983) claimed that I was proposing a theory of holistic causation in which the three major classes of determinants act simultaneously as a fused whole. They likened the triadic reciprocity to Haldane's holistic doctrine that William James (1884) criticized. This characterization came as a surprise to me because at no time have I advocated a doctrine of simultaneous holistic reciprocity in which the triadic determinants operate "at the same moment in time."

In a rejoinder, I explained that a causal process in which the interactants are influencers and influenced at the same time is not only illogical but functionally impossible (Bandura, 1983). A causal factor takes some time to exert influence and experience a reciprocal effect. In a verbal interchange, for example, questions and answers cannot occur simultaneously. Mutual influences and their reciprocal effects do not spring forth all at once. They work their mutual effects over

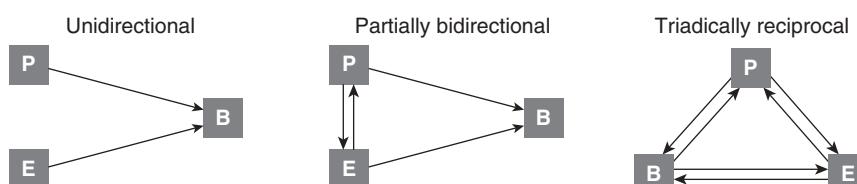


Figure 17.1 Schematization of different causal models of interaction between key classes of determinants. Social cognitive theory is founded on the causal model of triadic reciprocal determination

variable time courses. Time lags between events permit clarification of how different segments of reciprocity operate. Knowledge of how the various subsystems work advances understanding of how the superordinate system operates. It is the subsystems and their various interactions rather than the entirety that are analyzed. In the analytic decomposition of triadic causation, different subspecialties of psychology center their inquiry on selected segments of the reciprocal interplay (Bandura, 1986). A herculean effort to examine every possible interactant at the same time would beget investigatory paralysis.

Staddon (1984), a proponent of the operant conditioning view, got in on the act as well. He argued that internal determinants are unmeasurable so causal structures should be confined to stimulus inputs. Scientific advances are promoted by two kinds of theories (Nagel, 1961). One form seeks relations between directly observable events but shies away from the mechanisms subserving the observable events. The second form focuses on the mechanisms that explain the functional relations between observable events. The problem with the orthodox version of environmental determinism is that behavior is neither always cued by the stimuli that precede it nor always controlled by the stimuli that follow it. Hence, Staddon expanded this model to one in which behavior is under dual stimulus control – by current external stimuli and the internal residuum of past stimulus inputs. The implanted history of reinforcement carried the major explanatory burden. In a commentary, I explained that an implanted history is an inferred inner determinant, not a directly observable one (Bandura, 1984). The dispute is no longer about inner causes but the form they take. I documented the diverse ways in which cognition breaks the chain of stimulus control. Forethought can enhance, neutralize, or override the impact of environmental inputs.

There is much advice on how to conduct informative research and get it published, but

little on how to manage challenges to, and misunderstandings of, posited theories and the findings of verification tests. This is part and parcel of a scientific career that can command a lot of one's attention. If the theory addresses diverse aspects of human functioning and is widely cited, the disputational aspect of professional life can crowd a busy academic schedule. I discuss elsewhere issues regarding the process of theory building (Bandura, 2005a).

HUMAN CAPACITY FOR SELF-REGULATION

As previously noted, people's capacity to regulate their own functioning and shape the course their lives take is a core feature of social cognitive theory. The road I have traveled is very much in keeping with the agentic perspective toward human development, adaptation, and change. It underpins social cognitive theory (Bandura, 2006c). I grew up in a tiny rural hamlet in northern Alberta. The only school in town, which housed first grade through high school, was woefully short of teachers and educational resources. Two teachers had to teach the entire high school curriculum. They tried their best but were not always fully informed in key subject areas. We had to take charge of our own learning. Self-directed learning was an essential means of academic self-development, not a theoretical abstraction debated in arcane language in learned circles. However, the paucity of educational resources turned out to be enabling rather than handicapping. The content of courses is perishable, but self-regulatory skills have lasting functional value whatever the pursuit might be. These formative experiences spoke to an agentic view of human behavior that accented positive enabling factors rather than debilitating risk factors (Bandura, 2006a, 2008b).

In our excursion into the nature of self-directedness, our laboratory studies explored the mechanisms of self-regulation (Bandura,

1986, 1991c). Some of the studies clarified how personal standards are constructed from the myriad social influences. Others documented the regulatory power of self-reactive influences. Rational models of human behavior embraced the centrality of agency but they too provided a truncated view of self-regulation rooted in the market metaphor. Behavior was said to be regulated by self-interest construed almost entirely in terms of material costs and benefits. We demonstrated that human motivation and performance attainments are governed not only by material incentives, but also by social incentives, and self-evaluative incentives linked to personal standards. People often settle for alternatives of marginal utility or even sacrifice material gain to preserve their positive self-regard. Some of our studies examined self-regulation under conflictual conditions where individuals are rewarded for behavior they devalue, or are punished for activities they personally value. Principled dissenters often find themselves in the latter predicament. Their sense of self-worth is so strongly invested in certain convictions that they will submit to maltreatment rather than accede to what they regard as unjust or immoral.

Operant conditioners defined self-regulation out of existence by rechristening it as “stimulus control” and locating it in the external environment. They took me to task for introducing self-referent factors into the determination of behavior (Catania, 1975). In rejoinders I relocated self-management in the sentient person and reviewed the growing body of evidence on the means by which individuals exercise self-directedness (Bandura, 1976).

This was not a hospitable time to present an agentic theory of human behavior. Psychodynamicists depicted behavior as driven unconsciously by impulses and complexes. Behaviorists portrayed behavior as shaped and shepherded by environmental forces. The cognitive revolution was ushered in on a computer metaphor. This conception stripped humans of agentic capabilities, a functional consciousness, and a self-identity.

The mind as a symbol manipulator in the likeness of a linear computer became the conceptual model for the times. It was not individuals but their subpersonal parts that were orchestrating activities nonconsciously. Control theories of motivation and self-regulation (Carver and Scheier, 1981; Lord and Levy, 1994) drew heavily on Powers’ (1973) perceptual control theory, which is an outgrowth of the mechanical cybernetic model. In addition, as will be shown later, eliminative materialists likened cognitive factors, such as beliefs, goals, and expectations, to the phlogiston of yesteryear.

A good share of the conceptual sparring over the nature of self-regulation involved differences between social cognitive theory and Powers’ control theory. The core feature of Powers’ theory is the negative feedback loop. Discrepancies between a programmed reference standard and the perceived input from the output automatically trigger action to match the standard. Error correction is the driving force. I regarded self-regulation by negative discrepancy as only half the story and, in many respects, the less challenging half to explain.

Viewed from the agentic perspective, self-regulation operates through dual control systems – proactive discrepancy production working in concert with reactive discrepancy reduction (Bandura, 1991c). People are aspiring and proactive organisms, not just reactive ones. Their capacity to exercise forethought enables them to wield adoptive control anticipatorily rather than being simply reactive to the effects of their efforts. We demonstrated that people motivate and guide their actions through proactive control by setting themselves challenging goals and performance standards that create negative discrepancies to be mastered. They then mobilize their effort and resources to fulfill those challenges. After people attain the goals they have been pursuing, those of high perceived self-efficacy set a higher standard for themselves. The adoption of further challenges creates new motivating discrepancies to be mastered. In short, people are motivated and

guided by foresight of goals, not just hindsight of shortfalls.

A theory of self-regulation governed by forethought and affective self-reactions did not sit well with Powers (1991), the foremost advocate of control theory. In his view, the human organism is “nothing more than a connection between one set of physical quantities in the environment (input quantities) and another set of physical quantities in the environment (output quantities)” (Powers, 1978: 421). Cognitive and affective processes were dismissed as irrelevant because “we are not modeling the interior of the subject” (1978: 432). We questioned the conceptual and empirical adequacy of this austere mechanistic model, as well as derivative control theories that grafted on the negative feedback loop a variety of intrapersonal properties (Bandura, 1991b; Bandura and Locke, 2003; Locke, 1994). These adjuncts violated the parent control theory but befit a sentient being.

Social cognitive theory lends itself readily to social applications. Collaboration with Robert De Busk and Kate Lorig in the Stanford Medical School provided an opportunity to extend the agentic model of self-regulation to health promotion and disease prevention. These applications were rooted in a shift of emphasis from a disease model to a health model. It promotes effective self-management of health habits by cultivating self-regulatory skills that enable people to live healthier and retard the process of aging. This work led to the development of a self-management health system that promotes health and reduces risk of disease on a large scale at relatively low cost (Bandura, 2004c, 2005b).

Vast populations have no access to services that promote health and timely help in changing habits that impair it. By linking the interactive aspects of the self-management model to the Internet, one can vastly expand its availability to people wherever they may live. The goal of this implementational extension is to develop interactive online systems that enable people worldwide to improve the quality of their health.

SELF-REGULATION IN THE EXERCISE OF MORAL AGENCY

In areas of functioning involving achievement strivings and cultivation of competencies, the personal standards that serve as the mark of adequacy are progressively altered as knowledge and skills are acquired and challenges are met. In many areas of social and moral conduct, the internal standards are relatively stable. That is, people do not change from week to week in what they regard as right or wrong or good or bad. The agentic theory of self-regulation encompasses not only aspirational self-management but moral self-regulation as well in the exercise of moral agency (Bandura, 1991, 2004a). The self-regulation operates through the same basic set of subfunctions across aspirational and moral spheres of functioning. The verified explanatory commonality of the self-regulatory theory across markedly diverse spheres of activities is in keeping with Occam’s maxim advocating theoretical parsimony.

Psychological theories of morality focus heavily on the acquisition of moral standards and the structure of abstract moral reasoning, often to the neglect of moral conduct. A complete theory of moral agency must link moral knowledge and reasoning to moral action. This requires an agentic theory of morality rather than one confined mainly to cognitions about morality. Adoption of moral standards does not create an immutable internal moral control system. Indeed, large-scale humanities are typically perpetrated by people who are considerate and compassionate in other aspects of their lives. A full theory of moral agency must, therefore, explain how otherwise considerate people can behave inhumanely.

In the social cognitive theory of moral agency, there are eight psychosocial mechanisms operating at four loci in at which moral self-sanctions can be selectively disengaged from harmful practices (Bandura, 1999a). At the *behavior locus*, worthy ends are used to sanctify harmful means by social and moral justification, exonerative comparison that

renders the practices righteous, and sanitizing and convoluted language that disguise what is being done. At the *agency locus*, people obscure personal responsibility by displacement and diffusion of responsibility. This absolves them of accountability for the harm they cause. At the *outcomes locus*, the detrimental social effects of one's actions are ignored, minimized, or disputed. At the *victim locus*, perpetrators dehumanize and blame recipients for bringing the maltreatment on themselves. These mechanisms operate, often in concert, at both individual and social system levels.

The philosopher Seneca once portrayed seeming serendipity as: "Luck is what happens when preparation meets opportunity." This account characterizes well my varied partnerships in research designed to advance understanding of the determinants, mechanisms, and effects of moral disengagement in diverse spheres of life. One such collaborative effort is a multifaceted longitudinal project with Gian Caprara and his associates, Claudio Barbaranelli and Tina Pastorelli at the University of Rome "La Sapienza," in which we are studying child development from an agentic theoretical perspective. Part of this research centered on the development and exercise of moral agency. We created scales for measuring facility in moral disengagement and demonstrated that moral disengagement increases proclivity for aggressive and antisocial activities.

A San Francisco newspaper ran a story on the program of research by Lisa Bero, at the University of California Medical School, showing that corporate funding of research biases the findings. I checked with Lisa on whether she would be interested in a large-scale analysis of moral disengagement in the corporate world. A joint project we mounted verified the widespread moral disengagement at the social systems level in the tobacco, lead, vinyl chloride, and silicosis-producing industries. Through Caprara's acquaintance with Lazlo Zsolni, at the Business Ethics Center at Budapest University, we extended the analysis of corporate moral disengagement to the massive Bhopal chemical disaster,

Nestlé's aggressive marketing of infant formula products to third-world countries despite serious health effects, the defective Ford Pinto that took a heavy toll of lives, and the linguistic sanitizing of the mishap at the nuclear power plant on Three Mile Island.

Michael Osofsky, one of my undergraduate advisees at Stanford, often accompanied his father who chaired the Psychiatry Departments at Louisiana State University, on consulting visits to a Southern maximum security penitentiary. Because of this relationship, the warden not only granted us permission to study members of the execution team but also gained us access to two other Southern penitentiaries. This research gave us a better understanding of how executioners disengage moral self-sanctions from the taking of human life. Haney (1997) documented the morally disengaging ways in which capital trials are structured and conducted to enable jurors to sentence a person to death. A collaborative public survey study with Alfred McAlister at the University of Texas testified to how moral disengagement eases the public's qualms about the use of state executions. This set of studies furthered understanding of how moral disengagement is enlisted at each of the three levels in the application of the death penalty – at the societal, judicial, and execution levels.

In another line of study in our research partnership, McAlister and I were conducting a national survey on moral disengagement in support of military force. Midway through this study the terrorists struck the World Trade Center and the Pentagon. The terrorist attacks raised the level of moral disengagement. In path analysis, moral disengagement completely mediated the impact of the terrorist attack and sociodemographic characteristics on public support for the use of military force. These diverse collaborative projects illustrate the multiplicative advances in knowledge that can be achieved by courting opportunities when they arise. I also analyzed independently moral disengagement in the commercialization of gratuitous violence by the broadcast

industry (Bandura, 1973), in terrorism and in the enlistment of public support for, and conduct of, the war in Iraq (Bandura, 2004d).

My laboratory research on the moral disinhibitory function of dehumanization and diffusion of responsibility predated the studies cited above of moral disengagement at the social systems level (Bandura et al., 1975). While demonstrating how easy it is to bring out the worst in others, these experiments also revealed the power of humanization to curb inhumane conduct. However, I discontinued the laboratory studies on the disengagement of moral agency because it required the experimenters to disengage their own moral qualms to be able to them.

AFFIRMATION OF AGENCY IN A REDUCTIONISTIC ERA

There are two theoretical developments that minimize or dismiss the exercise of human agency. The first detractor is physical eliminationism. In this view, human behavior is governed by intricate neural networks that operate outside one's awareness and control. The second is the growing geneticization of human behavior. Social roles and human practices are increasingly being proclaimed as driven by prehistoric biological programming. In both of these theoretical orientations, human behavior is shaped and regulated at the subpersonal level. I addressed elsewhere, at some length, these two lines of theorizing and critically analyzed some of the widely cited evidence for them (Bandura, 2008a). A brief summary of this analysis is presented in the sections that follow.

Physicalistic theory of human agency

Agentic contributions to human functioning are dismissed in some quarters on the grounds that human behavior is regulated by neuronal mechanisms operating at a subpersonal level.

Thoughts are construed as epiphenomenal events that create an illusion of control but actually have no effect on how one behaves. In this view, humans are essentially conscious hosts of automata that dictate their behavior subpersonally. In support of an agentic theory of human functioning, I argued that physical eliminationists frame the issue of personal regulation in the wrong terms at the wrong level of control.

In acting as agents, individuals obviously are neither aware of, nor directly control, their neuronal mechanisms. Rather, they exercise *second-order control*. They do so by intentionally engaging in activities at the macrobehavioral level known to be functionally related to given outcomes. In pursuing these activities, over which they can exercise control, they shape the functional circuitry and enlist the neurophysiological events subserving their pursuits. Cognitions are higher-level cerebral events involving deliberative, reflective, self-referential, and evaluative processes operating in a top-down fashion through highly interconnected brain systems within the same material entity rather than in a physicalistic dualism.

Consider the following analogy. In driving an automobile to a desired place, the driver engages in coordinated acts of shifting gears, steering, manipulating the gas pedal, and applying brakes. These deliberate acts, which the driver controls directly, regulate the mechanical machinery to get safely to where the driver wants to go. But the driver has neither awareness nor understanding of the correlative microcombustion, transmission, steering, and braking processes subserving the driver's purposes. The deliberate planning of where to go on a trip, what route to take, what to do when one gets there, and securing reservations for these diverse activities far in advance requires considerable proactive top-down cognitive regulation. The temporal structuring of behavior by goals and purposes sets the course for one's activities. Proximal self-regulation provides the guides, strategies, and motivators in the here and now to get to where one is going

(Bandura, 1991c). Having constructed a trip, travelers cannot sit back and wait for lower-level sensory-motor activity to consummate the arrangements unconsciously.

Similarly, in second-order control over their cardiovascular functioning, individuals obviously do not intentionally direct their atrial and ventricular cardiac muscle fibers to fire and their aortic and pulmonary valves to open and close. However, by intentionally engaging in an exercise routine and controlling their activity level, they can enhance their cardiac functioning and raise and lower their heart rate without having the foggiest idea of how they indirectly recruited the subserving neurophysiological mechanisms. In short, enactments of functional activities at the controllable macrobehavioral level serve as the means for agentic recruitment of the subserving events at the microneural level.

Almost everyone adopts the ontological view that cognitive events are brain activities not immaterial entities. It is the epistemological form of reductability that is most in contention. The major argument against it is that each level of complexity – atomic, molecular, biological, psychological, and social structural – involves emergent new properties that are distinct to that level and, therefore, must be explained in their own right; for example, knowing the locality and brain circuitry subserving learning can say little about how best to devise conditions of learning in terms of level of abstractness and challenge; how to provide incentives to get people to attend to, process, and organize relevant information; and whether learning is better achieved independently, cooperatively, or competitively. The optimal conditions must be specified by psychological principles. There is little at the subatomic or neuronal level that can tell us how to develop efficacious parents, teachers, and social reformers or how to build and run social systems.

People are contributors to their activities, not just onlooking hosts of subpersonal networks autonomously shaping and regulating their performances. An aspiring violinist, for example, has to practice tenaciously to train the brain, build muscular strength and dexterity,

and hone sensory acuity to realize a virtuoso performance. Tell an aspiring violinist, who has spent countless hours training the brain and manual fingering and bowing dexterity to execute the pyrotechnical wizardry of a Paganini violin concerto, that the neural network is really the violinist and the indefatigable musician is just a self-aggrandizing illusionist.

The sensory, motor, and cerebral systems are tools people use to accomplish the tasks and goals that give meaning, direction, and satisfaction to their lives (Bandura, 2008a; Harré, 1983). To make their way successfully through a complex world people have to make sound judgments about their capabilities, anticipate the probable effects of different events and courses of action, size up sociostructural opportunities and constraints, and regulate their behavior accordingly. These belief systems are a working model of the world that enables people to achieve desired futures and avoid untoward ones.

There is growing unease about progressive divestiture of different aspects of psychology to biology and subpersonal neuroscience. It is feared that as we give away more and more psychology to disciplines lower down the conceptual food chain, there will be no core psychological discipline left. Psychology will become merely a branch of biology. Contrary to the proclamations of the divestitive oracles, psychology is the one discipline that uniquely encompasses the complex interplay among intrapersonal, biological, interpersonal, and sociostructural determinants of human functioning. Psychology is best suited to advance understanding of the integrated biopsychosocial nature of humans, and how they manage and shape the everyday world around them. The field of psychology should be articulating a broad vision of human beings, not a reductive fragmentary one.

Growing primacy of human agency in the co-evolution process

The conceptions of human nature regarding the capacity to exercise some measure of

control have changed markedly over time. In the early theological conceptions, human nature was ordained by original divine design. Evolutionism transformed the conception to one in which human nature is shaped by environmental pressures acting on random gene mutations and reproductive recombinations. This process of natural selection is devoid of deliberate plans or purposes. The evolutionary emergence of language and abstract and regulative cognitive capacities transformed the nature of the coevolution process. It provided the neuronal structure for supplanting aimless environmental selection with cognitive agency. Human forebears evolved into a sentient agentic species. Their advanced symbolizing capacity enabled humans to transcend the dictates of their immediate environment and made them unique in their power to shape their circumstances and life. Although not limitless, learnability and agentic capability became the hallmark of human nature.

Biology provides the information-processing architectures and potentialities and sets constraints. But in most spheres of functioning, biology permits a broad range of cultural possibilities. As Jay Gould (1987) notes, the major explanatory battle is not between nature and nurture as commonly framed. But whether nature operates as a determinist, that has culture on a “tight leash,” as Wilson (1998) contends, or as a potentialist that has culture on a “loose leash,” as Gould (1987) maintains. Biological determinists support a conservative view of society. It emphasizes the rule of nature, inherent constraints, and limitations. They argue that people should not try to remake themselves and their societies against the rule of nature, however they construe it. Biological potentialists give greater weight to enabling social conditions for personal development and societal change.

Evidence supports the potentialist view. Humans have created societies of diverse natures: aggressive and pacific ones, egalitarian and despotic ones, altruistic and selfish ones, individualistic and collectivistic ones, enlightened and backward ones. The human

species exhibits transformative changeability as well as interculture and intraculture diversity. People have changed little genetically over the past millennium but they have changed markedly even over the recent decades in their beliefs, mores, social and occupational roles, cohabiting arrangements, family practices, and styles of behavior. They have done so through rapid cultural and technological evolution. Cultures evolve over generations and shape the ways people need to live in the particular cultural milieu in which they are immersed.

Other species are heavily innately programmed for stereotypic survival in a particular habitat. In contrast, as an agentic species, humans devise ways of adapting flexibly to remarkably diverse geographic, climatic, and social environments. Consider the many ways in which the psychosocial side of coevolution is gaining ascendancy through agentic ingenuity. People create technologies to transcend their biological limitations. For example, humans have not evolved morphologically to fly but they are soaring through the air and even in the rarified atmosphere of outer space at breakneck speeds, despite the biological unachievability. Agentic inventiveness overrode biological design in getting them airborne. People also use their ingenuity to circumvent and insulate themselves from selection pressures. They create devices that compensate immensely for their sensory and physical limitations. They transcend time, place, and distance as they interact globally with the virtual environment of the cyberworld.

Through genetic engineering, humans are creating biological natures for better or for worse, rather than waiting for the slow process of natural evolution. They are now changing the genetic make-up of plants and animals that evolved over eons. Not only are humans cutting and splicing nature’s genetic material, but, through synthetic biology, they are also creating new types of genomes. In short, humans are an agentic species that is altering evolutionary heritages and shaping their future.

The notion in vogue is that biological evolution provides the potential and culture can do only so much within those constraints. This view flies in the face of the growing primacy of human agency in the coevolution process. It is not that social cognitive theory minimizes biological endowment. Quite the contrary. People have evolved the complex biological systems required for the very agentic activities by which personal and social changes are realized. These include deliberative and generative thought, forethoughtful self-regulation, and reflective self-evaluation. Neither the agentic human ascendancy in the coevolution process nor the rapid transformational societal changes would be possible without the biological endowment of abstract cognitive capabilities.

Social cognitive theory highlights the forward-looking impact of our biological endowment, rather than backward-looking speculation about adaptation to primitive conditions of prehistoric times. The study of how humans are changing endowed heritages, circumventing biological constraints, and shaping their future through social and technological evolution is more fruitful than spinning fanciful stories about prehistoric mating patterns in drafty caves.

Were Darwin writing today, he would be documenting the overwhelming human domination of the environment. As the unrivaled ruling species atop the food chain, we are degrading the earth's life support systems and drafting the requiem for biodiversity. By wielding powerful technologies that amplify control over the environment, driven by a foreshortened perspective, humans may be well on the road to outsmarting themselves into irreversible ecological crises that they can no longer control.

EXERCISE OF AGENCY IN CULTURAL CONTEXT

Contentious dualisms pervade our field, pitting autonomy against interdependence,

individualism against collectivism, and social structure against agency. Cultures are dynamic and internally diverse systems, not static monoliths (Bandura, 2002b, Kim et al., 1994). Analyses across activity domains and classes of social relationships reveal that people behave communally in some aspects of their lives and individualistically in many other aspects. They express their cultural orientations conditionally depending on incentive conditions, rather than invariantly. The categorical approach masks extensive diversity.

Not only are cultures not monolithic entities but they are no longer insular. Transnational interdependencies and global market forces are restructuring national economies and shaping the political and social life of societies. Advanced telecommunications technologies are disseminating ideas, values, and styles of behavior globally at an unprecedented rate. The symbolic environment, feeding off communication satellites, is altering national cultures and homogenizing collective consciousness. People are spending much of their time in the ever-expanding cyberworld. This is furthering the globalization of culture. In addition, mass migrations of people and high global mobility of entertainers, athletes, journalists, academics, and employees of multinational corporations are changing cultural landscapes. This intermixing creates new hybrid cultural forms, blending elements from different ethnicities. These social forces are homogenizing some aspects of life, polarizing other aspects, and fostering a lot of cultural hybridization.

It is widely claimed that Western theories lack generalizability to non-Western cultures (Sampson, 1977). One must distinguish between basic human capacities and how culture shapes these potentialities into diverse forms. For example, social modeling through observational learning figures prominently in social cognitive theory. Humans have evolved an advanced capacity for learning through observation of modeled attitudes, values, and styles of behavior. It is essential for their

self-development and functioning regardless of the culture in which they reside. Modeling is a universalized human capacity. But what is modeled, how modeling influences are structured, and the purposes they serve, varies in different cultural milieus.

The same distinction in levels of analysis applies to perceived efficacy. A common duality inappropriately equates self-efficacy with self-centered individualism and pits it against collectivism (Schooler, 1990). Because efficacy beliefs involve self-referent processes, self-efficacy is often misconstrued as aggrandizement of an individualistic autonomous self and contrasted with an interdependent collectivistic self. Self-efficacy does not come in only an individualistic form nor with a built-in value system. People's belief in their efficacy is exercised in individual, proxy, and collective forms. Social cognitive theory is, therefore, just as relevant to human attainments realized through interdependent collective effort as to those achieved individually. Moreover, self-efficacy can serve communal purposes as well as individualistic ones.

Being immobilized by self-doubt about one's capabilities and belief in the futility of effort has little adaptive advantage. A growing body of research shows that, indeed, a resilient sense of efficacy has generalized functional value regardless of whether one resides in an individualistically oriented culture or a collectivistically oriented one (Bandura, 1995a, 2002b). But how efficacy beliefs are developed, the form they take, the ways in which they are exercised, and the purposes to which they are put vary cross-culturally. The cross-cultural findings debunk the misconception that belief in one's efficacy is an egocentric orientation wedded to Western individualism. Personal efficacy is valued, not because of reverence for individualism, but because a resilient sense of efficacy has generalized functional value regardless of whether activities are pursued individually or by people working together for common cause.

In short, there is a cultural commonality in basic agentic capacities and mechanisms of

operation, but diversity in the culturing of these inherent capacities. In this dual-level analysis, universality is not incompatible with manifest cultural plurality. Cultural variations emerge from universalized capacities through the influence of social practices reflecting shared values and norms, incentive systems, role prescriptions, and pervasive modeling of distinctive styles of thinking and behaving. Kluckholn and Murray summarized eloquently the blend of universality, commonality, and uniqueness of human qualities (Muñoz and Mendelson, 2005): "Every person is in certain aspects like all other people. Like some other people. Like no other person."

GOING GLOBAL WITH SOCIAL COGNITIVE THEORY

Global applications of social cognitive theory to promote society-wide change testify to the effectiveness of social modeling and enhanced individual and collective efficacy to improve the quality of life in diverse cultural milieus (Bandura, 2006b, 2006c). These applications, which reach millions of people in Africa, Asia, and Latin America, address some of the most urgent global problems. Soaring population growth tops the list because of its heavy strain on natural and social resources and on ecological supports of life. Developed nations are stabilizing their population, but developing ones, where most of the growth is occurring, are rapidly doubling their populations. We are currently at 6.7 billion, adding about one billion every fifteen years, and heading toward a population of nine to ten billion by mid-century. We are witnessing spiraling destruction of ecological systems that have evolved over eons that keep our planet cool and habitable. Burgeoning population growth also fuels civil strife. Most violent conflicts are in countries with large youth populations that are uneducated, unemployed, and living in poverty under autocratic rulers plagued by corruption. Youth living

under such dismal life conditions are easily recruited for violent activities.

Another widespread problem is the pernicious gender inequality. In these societies, women are marginalized, devalued, disallowed aspiration, and denied their liberty and dignity. The spreading AIDS epidemic is another mounting global problem with devastating societal consequences. Some societies present unique problems that require special social themes tailored to the detrimental cultural practices.

Long-running serialized dramas serve as the vehicle for promoting personal and social changes. To change deeply held beliefs and social practices requires strong emotional bonding to enabling models that exemplify a vision of a better future and realistic paths to it. There are no quick fixes. The dramatic productions are not just fleeting fanciful stories. They dramatize people's everyday struggles, the impediments they face, and the effects of different societal practices. The storylines speak to people's fears and hopes, help them see a better life, and inform and enable them to take the steps to realize it.

The storylines model family planning, women's equality, degrading dowry systems, spouse abuse, environmental conservation, AIDS prevention, and varied life skills. Hundreds of episodes get people emotionally engaged in the evolving lives of the models and are inspired and enabled by them. This is a highly flexible format, which contributes to its generalizability, versatility, and power. By including multiple intersecting plotlines, one can address diverse aspects of life at both the individual and social structural level.

Major advances in science and its applications require the combination of the expertise of diverse sources. This is best achieved, not through bureaucratic interdisciplinary prescription, but through collaboration fueled by common interest and purpose. These ambitious global applications testify to the value of informally created partnerships. Promoting large-scale changes requires three component models – theoretical, translational, and social

diffusion models. The first component is a *theoretical model*. It specifies the determinants of psychosocial change and the mechanisms through which they produce their effects. This knowledge provides the guiding principles for effecting change. The second component is a *translational and implementation model*. It converts theoretical principles into an innovative operational model and specifies the content, strategies of change, and their mode of implementation. We often do not profit from our theoretical successes because we lack effective translational and social diffusion models to disseminate proven psychosocial approaches. The knowledge languishes in our periodicals.

One morning I received a call from Miguel Sabido, a creative dramatist and producer at Televisa in Mexico City. He explained that he is creating engrossing serial dramas, founded on our modeling principles from the Bobo doll studies, to promote literacy and family planning practices nationally. He had devised a remarkably innovative translational model that was producing impressive society-wide changes. We had an informative theory and a creative translational model. But we lacked the expertise and resources to disseminate it widely. The third component, a *social diffusion model* on how to promote adoption of psychosocial programs in diverse cultural milieus, was the missing link. David Poindexter, the director of Population Communications International, heard about the success of this model and became the global diffuser to diverse cultural milieus. Everett Rogers, the foremost theorist in the field of diffusion of innovation, was also an influential player in this global effort. Because social cognitive theory encompasses social network factors in the spread of social influence, this was a synergistic collaboration (Bandura, 1986, 2006b).

These are not social programs foisted on nations by outsiders in pursuit of their self-interest. The dramatic serials are created only on invitation by countries seeking help with intractable problems. Social cognitive

principles are generalizable but their application has to be tailored to the cultural practices and the types of desired changes. This requires functional adaptations. The diffusion center works in partnership with personnel in the host countries to create serial dramas appropriate to their culture and address their particular needs.

In the formative phase, extensive cultural and value analyses are conducted to identify the problems of major concern, the impediments people face, and the improvements they seek in their lives. The dramatizations are grounded in the internationally endorsed values codified in United Nations covenants and resolutions. These values embody respect for human dignity, equitable opportunities, and social practices that support common human aspirations.

Many worldwide applications of this model in Africa, Asia, and Latin America are promoting personal and society-wide changes that are bettering the lives of millions of people (Bandura, 2006c; Rogers, et al., 1999; Vaughan et al., 2000). They are raising national literacy, enhancing the status of women, reducing unplanned childbearing that contribute to soaring population growth and perpetuate the cycle of poverty, impeding child trafficking for inhumane labor, curtailing the spread of the AIDS epidemic, preserving biodiversity, promoting environmental conservation practices, and in other ways bettering of people's lives.

These global applications illustrate how the power of psychosocial knowledge can be amplified by blending different types of expertise that no one discipline can provide. We often cite examples in the physical and biological sciences where knowledge pursued for its own sake has unforeseen human benefits. The knowledge gained from the Bobo doll experiments conducted years ago spawned unimagined global applications designed to alleviate some of the most pressing global problems. That's a pretty good shelf life.

As I reflect on my journey at this octogenarian milepost, I am reminded of the saying that it is not the miles traveled but the amount

of tread remaining that is important. When I last checked, I still have too much tread left to gear down or to bring this engaging psychological odyssey to a close. In a decision that take goodly control of one's life have begun work on a book o the role of moral disengagement in the exercise of moral agency.

REFERENCES

Baer, D.M., Peterson, R.F. and Sherman, J.A. (1967) The development of imitation by reinforcing behavioural similarity to a model. *Journal of the Experimental Analysis of Behaviour*, 10, 405–416.

Bandura, A. (1961) Psychotherapy as a learning process. *Psychological Bulletin*, 58, 143–159.

Bandura, A. (1965) Vicarious processes: A case of no-trial learning. In L. Berkowitz (ed.), *Advances in Experimental Social Psychology*, 2, 1–55. New York: Academic Press.

Bandura, A. (1969) *Principles of Behavior Modification*. New York: Holt, Rinehart and Winston.

Bandura, A. (ed.) (1971a) *Psychological Modeling: Conflicting Theories*. New York: Aldine-Atherton.

Bandura, A. (1973) *Aggression: A Social Learning Analysis*. Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1976) Self-reinforcement: Theoretical and methodological considerations. *Behaviorism*, 4, 135–155.

Bandura, A. (1978) The self system in reciprocal determinism. *American Psychologist*, 33, 344–358.

Bandura, A. (1983) Temporal dynamics and decomposition of reciprocal determinism. *Psychological Review*, 90, 166–170.

Bandura, A. (1984) Representing personal determinants in causal structures. *Psychological Review*, 91, 508–511.

Bandura, A. (1986) *Social Foundations of Thought and Action: A Social Cognitive Theory*. Englewood Cliffs: Prentice-Hall.

Bandura, A. (1991a) Social cognitive theory of moral thought and action. In W.M. Kurtines and J.L. Gewirtz (eds), *Handbook of Moral Behavior and Development*, Vol. 1. Hillsdale, NJ: Erlbaum.

Bandura, A. (1991b) Human agency: The rhetoric and the reality. *American Psychologist*, 46, 157–162.

Bandura, A. (1991c) Self-regulation of motivation through anticipatory and self-reactive mechanisms. In R.A. Dienstbier (ed.), *Perspectives on Motivation: Nebraska Symposium on Motivation*, 38, 69–164. Lincoln: University of Nebraska Press.

Bandura, A. (ed.) (1995a) *Self-efficacy in Changing Societies*. Cambridge: Cambridge University Press.

Bandura, A. (1995b) Comments on the crusade against the causal efficacy of human thought. *Journal of Behavior Therapy and Experimental Psychiatry*, 26, 179–190.

Bandura, A. (1996) Ontological and epistemological terrains revisited. *Journal of Behavior Therapy and Experimental Psychiatry*, 27, 323–345.

Bandura, A. (1997) *Self-efficacy: The Exercise of Control*. New York: Freeman.

Bandura, A. (1999a) Moral disengagement in the perpetration of inhumanities. *Personality and Social Psychology Review*, 3, 193–209.

Bandura, A. (1999b) A social cognitive theory of personality. In L. Pervin and O. John (eds), *Handbook of Personality*, 2nd Edition, pp. 154–196. New York: Guilford Publications.

Bandura, A. (2002a) Growing primacy of human agency in adaptation and change in the electronic era. *European Psychologist*, 7, 2–16.

Bandura, A. (2002b) Social cognitive theory in cultural context. *Journal of Applied Psychology: An International Review*, 51, 269–290.

Bandura, A. (2004a) Selective exercise of moral agency. In T.A. Thorkildsen and H.J. Walberg (eds), *Nurturing Morality*, pp. 37–57. New York: Kluwer Academic.

Bandura, A. (2004b) Swimming against the mainstream: The early years from chilly tributary to transformative mainstream. *Behavioral Research and Therapy*, 42, 613–630.

Bandura, A. (2004c) Health promotion by social cognitive means. *Health Education and Behavior*, 31, 143–164.

Bandura, A. (2004d) The role of selective moral disengagement in terrorism and counterterrorism. In F.M. Mogahaddam and A.J. Marsella (eds), *Understanding Terrorism: Psychological Roots, Consequences and Interventions*, pp. 121–150. Washington, DC: American Psychological Association Press.

Bandura A. (2005a) The evolution of social cognitive theory. In K.G. Smith and M.A. Hitt (eds) *Great Minds in Management*. pp. 9–35, Oxford: Oxford University Press.

Bandura, A. (2005b) The primacy of self-regulation in health promotion. *Applied Psychology: an International Review*, 54, 245–254.

Bandura, A. (2006a) Toward a psychology of human agency. *Perspectives on Psychological Science*, 1, 164–180.

Bandura, A. (2006b) On integrating social cognitive and social diffusion theories. In A. Singhal and J. Dearing (eds), *Communication of Innovations: A Journey With Ev Rogers*, pp. 111–135. Beverley Hills: Sage Publications.

Bandura, A. (2006c) Going global with social cognitive theory: From prospect to paydirt. In S.I. Donaldson, D.E. Berger, and K. Pezdek (eds), *Applied Psychology: New Frontiers and Rewarding Careers*, pp. 53–79. Mahwah, NJ: Lawrence Erlbaum.

Bandura, A. (2006d) Guide to the construction of self-efficacy scales. In F. Pajares and T. Urdan (eds), *Self-efficacy Beliefs of Adolescents*, 5, 307–337. Greenwich, CT: Information Age Publishing.

Bandura, A. (2008a) The reconstrual of 'free will' from the agentic perspective of social cognitive theory. In J. Baer, J.C. Kaufman and R.F. Baumeister (eds), *Are We Free? Psychology and Free Will*, pp. 86–127. Oxford: Oxford University Press.

Bandura, A. (2008b) An agentic perspective on positive psychology. In S.J. Lopez (ed.), *Positive Psychology: Exploring the Best in People*, 1, 167–196. Westport, CT: Greenwood Publishing Company.

Bandura, A. (2009) Social cognitive theory of mass communications. In J. Bryant and M.B. Oliver (eds), *Media Effects: Advances in Theory and Research*, 2nd Edition, pp. 94–124. Mahwah, NJ: Lawrence Erlbaum.

Bandura, A., Blanchard, E.B. and Ritter, B. (1969) Relative efficacy of desensitization and modeling approaches for inducing behavioral, affective, and attitudinal changes. *Journal of Personality and Social Psychology*, 13, 173–199.

Bandura, A. and Locke, E. (2003) Negative self-efficacy and goal effects revisited. *Journal of Applied Psychology*, 88, 87–99.

Bandura, A., Ross, D. and Ross, S.A. (1963) Imitation of film-mediated aggressive models. *Journal of Abnormal and Social Psychology*, 66, 3–11.

Bandura, A., Underwood, B. and Fromson, M.E. (1975) Disinhibition of aggression through diffusion of responsibility and dehumanization of victims. *Journal of Research in Personality*, 9, 253–269.

Bandura, A. and Walters, R.H. (1963) *Social Learning and Personality Development*. New York: Holt, Rinehart and Winston.

Boffey, P. and Walsh, J. (1970) Study of TV violence: Seven top researchers blackballed from panel. *Science*, 168, 949–952.

Braithwaite, J. (1994) A sociology of modeling and the politics of empowerment. *British Journal of Sociology*, 45, 445–479.

Carver, C.S. and Scheier, M.F. (1981) *Attention and Self-regulation: A Control-theory Approach to Human Behavior*. New York: Springer-Verlag.

Catania, C.A. (1975) The myth of self-reinforcement. *Behaviorism*, 3, 192–199.

Cater, D. and Strickland, S. (1975) *TV Violence and the Child*. New York: Russell Sage Foundation.

Ephron, E. (1969a) A Child Is Not A Rat. *TV Guide*, November 26.

Efron, E. (1969b) The man in the eye of the hurricane. *TV Guide*, November 15.

Film Board of Canada (1972) *The Question of Violence: The US Senate Hearings*. [Motion Picture]. United States: Phoenix Films. (Available from Phoenix Films, Inc., 470 Park Avenue South, New York, NY 10016).

Gewirtz, J.L. (1971) Conditional responding as a paradigm for observational, imitative learning and vicarious-reinforcement. In H.W. Reese (ed.), *Advances in Child Development and Behavior*, 6, 273–304. New York: Academic Press.

Goldsen, R. (1972) Science in wonderland. Unpublished manuscript, Cornell University, Ithaca, NY.

Gould, J. (1972) Violence Held Unharmful to Youth. *New York Times*, p. 27.

Gould, S.J. (1987) *An Urchin in the Storm*. New York: Norton.

Haney, C. (1997) Violence and the capital jury: Mechanisms of moral disengagement and the impulse to condemn to death. *Stanford Law Review*, 49, 1447–1486.

Harré, R. (1983) *Personal Being: A Theory for Individual Psychology*. Oxford: Blackwell.

James, W. (1884) Absolutism and empiricism. *Mind*, 9, 281–286.

Kim, U., Triandis, H.D., Kâğıtçibasi, C., Choi, S. and Yoon, G. (1994) *Individualism and Collectivism: Theory, Method, and Applications*. Thousand Oaks, CA: Sage.

Locke, E.A. (1994) The emperor is naked. *International Review of Applied Psychology*, 43, 367–370.

Lord, R.G. and Levy, P.E. (1994) Moving from cognition to action: A control theory perspective. *Applied Psychology: An International Review*, 43, 335–398.

McCrae, R.R. and Costa, P.T. (1996) Toward a new generation of personality theories: Theoretical contexts for the five-factor model. In J. S. Wiggins (ed.), *Five-factor model of personality: Theoretical perspectives*, 51–87. New York: Guilford Press.

Milgram, S. and Shotland, R.L. (1973) *Television and Antisocial Behavior: Field Experiments*. New York: Academic Press.

Miller, N.E. and Dollard, J. (1941) *Social Learning and Imitation*. New Haven: Yale University Press.

Muñoz, R. and Mendelson, T. (2005) Toward evidence-based interventions for diverse populations: The San Francisco General Hospital prevention and treatment manuals. *Journal of Consulting and Clinical Psychology*, 73, 790–799.

Nagel, E. (1961) *The Structure of Science*. New York: Harcourt, Brace and World.

National Commission on the Causes and Prevention of Violence (1969) *Commission statement on violence in television entertainment programs*. Washington, DC: US Government Printing Office.

Paisley, M. (1972) *Social Policy Research and the Realities of the System: Violence Done to TV Research*. Stanford, CA: Stanford University, Institute for Communication Research.

Patterson, G.R. (1982) *A Social Learning Approach* (Vol 3). Coercive family process. Eugene, OR: Castalia.

Phillips, D.C. and Orton, R. (1983) The new causal principle of cognitive learning theory: Perspectives on Bandura's 'reciprocal determinism'. *Psychological Review*, 90, 158–165.

Powers, W.T. (1973) *Behavior: The Control of Perception*. Chicago: Aldine.

Powers, W.T. (1978) Quantitative analysis of purposive systems: Some spadework at the foundations of scientific psychology. *Psychological Review Monograph Supplements*, 85, 417–435.

Powers, W.T. (1991) Comment on Bandura's 'human agency'. *American Psychologist*, 46, 151–153.

Rogers, E.M., Vaughan, P.W., Swalehe, R.M.A., Rao, N., Svenkerud, P. and Sood, S. (1999) Effects of an entertainment-education radio soap opera on family planning behavior in Tanzania. *Studies in Family Planning*, 30, 1193–1211.

Rosenthal, T.L. and Zimmerman, B.J. (1978) *Social Learning and Cognition*. New York: Academic Press.

Rotter, J.B. (1954) *Social Learning And Clinical Psychology*. Englewood Cliff, NJ: Prentice-Hall.

Sampson, E.E. (1977) Psychology and the American ideal. *Journal of Personality and Social Psychology*, 35, 767–782.

Schooler, C. (1990) Individualism and the historical and social-structural determinants of people's concerns over self-directedness and efficacy. In J. Rodin, C. Schooler and K.W. Schaeie (eds), *Self-directedness: Cause and Effects Throughout the Life Course*, pp. 19–58. Hillsdale, NJ: Erlbaum.

Skinner, B.F. (1971) *Beyond Freedom and Dignity*. New York: Knopf.

Skinner, B.F. (1977) Why I am not a cognitive psychologist. *Behaviorism*, 5, 1–10.

Staddon, J.E.R. (1984) Social learning theory and the dynamics of interaction. *Psychological Review*, 91, 502–507.

Stoltz, S.B. (1978) *Ethical Issues in Behavior Modification*. San Francisco: Jossey-Bass.

Vaughan, P.W., Roger, E.M., Singhal, A. and Swalehe, R.M. (2000) Entertainment-education and HIV/AIDS

prevention: A field experiment in Tanzania. *Journal of Health Communications*, 5, 81–100.

Wilson, E.O. (1998) *Consilience: The Unity of Knowledge*. New York: Knopf.

AUTHOR'S NOTE

Some sections of this chapter include revised, updated, and elaborated material from the following publications:

Bandura, A. (2005) The evolution of social cognitive theory. In K.G. Smith and M.A. Hitt (eds) *Great*

Minds in Management. pp. 9–35. Oxford: Oxford University Press.

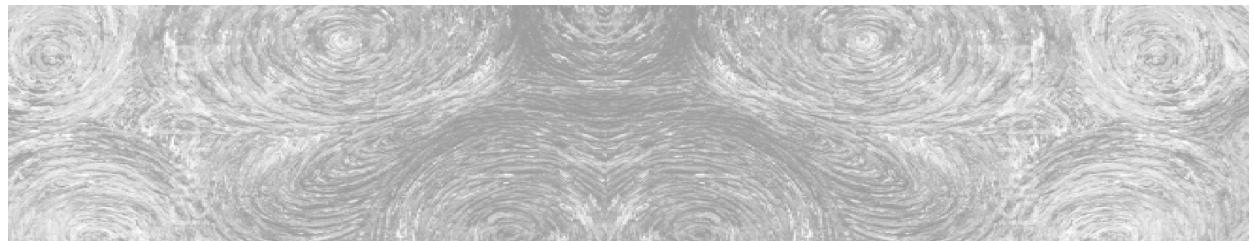
Bandura, A. (2006) Toward a psychology of human agency. *Perspectives on Psychological Science*, 1, 164–180.

Bandura, A. (2006) Autobiography. M.G. Lindzey and W.M. Runyan (eds) *A History of Psychology in Autobiography*, IX, 42–75. Washington, DC: American Psychological Association.

Chapters and research articles that address issues discussed in this publication are available in electronic form in the following website: <http://des.emory.edu/mfp/Bandura>.

PART III

Motivational/Affective Level of Analysis



Cognitive Dissonance Theory

Joel Cooper

ABSTRACT

Cognitive dissonance theory has been a major pillar of social psychology for decades. In this chapter, I discuss some of the reasons that Festinger's straightforward proposition about the relationship among cognitions created the pointed controversy that propelled it into the forefront of the discipline. Theories with the broad approach of cognitive dissonance often need modification and this chapter traces that journey. I then present the New Look and Self-Standards models that attempted to integrate extant data and changed our understanding of the motivational foundation that underlies dissonance. Current perspectives on cognitive dissonance focused on vicarious dissonance arousal by which individuals experience dissonance based on the behavior of fellow group members. Finally, the chapter examines the potential use of personal cognitive dissonance for optimizing the effectiveness of psychotherapy and the use of vicarious dissonance for increasing positive behaviors to protect health and wellbeing.

IN THE BEGINNING

The theory of cognitive dissonance has been a staple of social psychology for more than half a century. In this chapter, I will present

my own perspective on the birth of dissonance and then take the responsibility for two alternative approaches that modified our view of the meaning of cognitive dissonance: the New Look model (Cooper and Fazio, 1984) and the subsequent Self-Standards model of dissonance (Stone and Cooper, 2001). The story begins with Festinger's observations about cognitive inconsistency and will move to our newer views of the motivation underlying cognitive dissonance.

The late Leon Festinger molded the original theory of cognitive dissonance from his interest in people's susceptibility to field forces (Lewin, 1951), including pressures from groups. He had published a number of major statements about the pressures that groups place on individuals to achieve attitudinal consensus (Festinger, 1950). In 1954, he shifted his focus to the study of the individual. Instead of viewing pressure from the vantage point of the group's needs and goals, he took the perspective of the individual who was driven to use others as a benchmark to measure his or her own standing in a group. He proposed that people were driven to

compare their opinions and abilities with similar others and that there was pressure either to conform to the attitudes of similar others or to convince others to hold attitudes similar to oneself (Festinger, 1954).

In the theory of cognitive dissonance, Festinger (1957) completed the task of viewing the world from the perspective of the individual. In dissonance theory, cognitive consistency was represented inside the head of the person. To think of mental life as a set of cognitive representations was a radical departure from the mainstream view of the 1950s. For the first time, people's views of their social world, their appraisals of their fellow group members, their own opinions about the world, and their observations of their own and others' behaviors could all be projected onto a common screen. All were cognitive representations inside the head. Moreover, some of those cognitive representations bore a relationship to each other. The birth of cognitive dissonance theory occurred at that instant. The new theory – cognitive dissonance – became the most productive of all of his creative insights, and dissonance theory charted a research agenda that would last for half a century.

COGNITIVE DISSONANCE IN A LEARNING THEORY WORLD

The major tenets of the original version of dissonance theory are well known and straightforward. The state of cognitive dissonance occurs when people perceive that a pair of cognitions is inconsistent. Formally, Festinger defined a pair of cognitions as dissonant if the actor believed that one cognition followed from the obverse of the other. He postulated that dissonance is experienced as an unpleasant drive and, like other unpleasant drive states, needs to be reduced. The reduction occurs by changing the cognition least resistant to change or by adding cognitions that minimize the perceived magnitude of the discrepancy. In keeping with Festinger's

philosophical assumption that the dissonance battle was played out inside the head of the perceiver, he reasoned that inconsistency itself is a psychological state – that is, two cognitions are dissonant if the perceiver believes they are dissonant. The psychology of the perceiver, not the philosophical rules of logic, determines the existence of dissonance.

The idea that people prefer consistency to inconsistency was not new. Fritz Heider (1946) and Theodore Newcomb (1956) had written about such ideas previously and they were also consistent with field theoretical notions of Festinger's advisor, Kurt Lewin. To my knowledge, Festinger's 1957 book outlining his ideas did not raise controversy until the publication two years later of Festinger and Carlsmith's (1959) now classic study showing that people experienced dissonance after an attitude-inconsistent statement. As we know, Festinger and Carlsmith had people participate in a task that was specifically designed to be tedious and dull. Those participants then agreed to make a statement to a person, who they believed was a fellow student, extolling the excitement of the task. Few would have had difficulty with the finding that making the statement about the excitement of the task induced people to change their attitudes in the direction of their statement. If we accept the premise that people do not like inconsistency and that they are motivated to reduce the incongruence between their behavior and attitude, then the finding makes sense and would have been predicted by any consistency theory.

The provocative element in Festinger and Carlsmith's (1959) research was the role played by the magnitude of the incentive that participants were offered to make a public statement that was contrary to their beliefs. Some were offered a large incentive while others were offered a meager incentive. Festinger and Carlsmith reasoned that the large incentive would be a sufficiently important cognition consistent with the behavior to keep dissonance low, whereas people offered a small incentive would still be in the throes

of dissonance. The finding that behavior associated with small incentives could create more change than behavior associated with large incentives had the effect of prodding a large sleeping animal with a small stick. It woke the animal and, to continue the metaphor, allowed the provocateur not only to be noticed but, in the end, to assume a leadership position in the metaphorical jungle.

The year 1957 was the realm of learning theories. The number of people considering themselves social psychologists was scant. The “science” of psychology was focused on the rules of sensation, perception, and learning. The last of these topics was particularly vibrant, with lively debates among followers of Hull, Spence, Tolman, and Skinner filling the pages of the professional literature. They disagreed about many issues such as the importance of habits and the proper role of drive states. What they all agreed upon, however, was the role of rewards and reinforcements. Although they conceptualized them differently, larger rewards led to more behavior change; smaller rewards led to less change. This was the gospel according to learning theory.

Dissonance theory, and Festinger and Carlsmith’s findings in particular, put the brakes on that assumption. The playing field changed. Suddenly, attitudinal and behavioral change were at the service of smaller rewards rather than larger ones. Large incentives merely reduced the dissonance drive state and led to less change than smaller rewards, or perhaps no rewards at all. We should also not underestimate Festinger’s use of the drive state as the motivation for change. By positing that people were motivated by what was essentially a drive, he positioned dissonance theory alongside the major learning theories in which drive-reduction played the critical role. It is unclear whether Festinger ever believed that we would ever find evidence for the drive concept, but by using it as the motivational metaphor, his findings were instantly recognized as a challenge to all who wanted to see

social behavior as merely a carryover of the behavioral rules that applied to rats and pigeons.

Other findings continued to goad the conventional learning theory wisdom. The more people suffered, the more they liked what they suffered to attain (Aronson and Mills, 1959). The less punishment with which children were threatened to refrain from playing with an attractive toy, the more they devalued the toy (Aronson and Carlsmith, 1963) and the longer they refrained from playing with it (Freedman, 1965). Learning theorists rose to the challenge of these findings, criticizing the methodology, the conclusions, and the theory. Journals, including the new *Journal of Personality and Social Psychology*, were filled with debates about the intriguing new ideas predicted from cognitive dissonance theory. The early controversial years of provocative dissonance results were followed by challenges from critics, and followed again by studies from a growing number of dissonance researchers. By the early 1970s, dissonance had become a “movement” and very few questioned the existence of dissonance as a powerful principle of human social behavior.

TROUBLE IN DISSONANCE HEAVEN

My own activity as a dissonance researcher came from an unlikely source. I attended graduate school at Duke University, partly to study with Jack W. Brehm, one of Festinger’s original students and the researcher responsible for the very first published experiment based on dissonance theory (Brehm, 1956). However, Jack was busy with his new theory of psychological reactance and I was assigned to work with Edward E. Jones, who was soon to become a towering figure in the world of person perception and attribution. Nonetheless, Jones was a fan of dissonance research and was perplexed by a study that had been published by Rosenberg (1965) that seemed to spell trouble for the fledgling theory.

Rosenberg had argued that the reason for Festinger and Carlsmith's startling result was that participants had been worried that they were being evaluated on their level of psychological consistency. He coined the term *evaluation apprehension* to capture participants' concern that the person who had asked them to say that the task was fun was also privy to their attitudinal confessions when they responded to the dependent measure. Rosenberg provided evidence that attitude change was a direct function of level of reward (i.e., more attitude change for higher compensation) when the experimenter who collected the dependent measure was not in any way connected to the person who made the request for the counterattitudinal behavior. Jones suspected that the key element missing in Rosenberg's critical study was the element of choice. Although Festinger had not paid much attention to the need for a person to *choose* to engage in counterattitudinal behavior, he had built it into his original experiment (Festinger and Carlsmith, 1959). So, my advisor sent me out to devise a study to show that how people perceived their freedom to engage in attitude discrepant behavior made all of the difference in the arousal of dissonance.

The results of our study were published as Linder, Cooper, and Jones (1967) – my very first published dissonance article. In two experiments, we showed that attitude change can be a direct function of incentive magnitude *or* an inverse function, depending on whether dissonance is ever aroused by the procedure. If people perceive that they were free to accept or decline the invitation to make a counterattitudinal statement, then dissonance is aroused. We found that when people had been given the freedom to decline, then just like Festinger and Carlsmith had found, attitudes were an inverse function of incentive magnitude: the lower the reward, the greater the attitude change. It is only in the absence of dissonance – brought about by the absence of decision freedom – that attitude change became a direct function of the magnitude of the reward. In other words, reinforcement works to elicit attitude change

but is trumped by dissonance arousal. When dissonance exists – in this case, facilitated by the perception of decision freedom – it provides the underpinning for attitude change following counterattitudinal advocacy. (Later, Jones told me that I got his idea completely wrong, ran a different experiment from the one he thought I was running, but had no complaints about the outcome!)

THE THEORY EVOLVES: THE SEARCH FOR MODIFIERS BECOMES THE NEW LOOK MODEL OF DISSONANCE

A broad-stroke theory almost always needs modifiers. Many of the studies in the first decade of dissonance theory did exactly that. Linder et al.'s (1967) results were of that genre. Dissonance is aroused and leads to attitude change when decision freedom is high but not when it is low. We thus learned that choice was a modifier of dissonance. Carlsmith et al. (1966) as well as Davis and Jones (1960) had shown that people need to feel committed to their counterattitudinal behavior. If participants thought they would have a chance to "take back" their counterattitudinal statements, then dissonance would not occur. We thus learned that commitment was a modifier of cognitive dissonance.

A few years later, Steve Worchel and I wondered aloud whether *any* counterattitudinal statement expressed in *any* circumstance produced dissonance or whether something had to happen as a result of the behavior. Put another way, would an attitude discrepant statement uttered in the dark of one's own residence with no one to hear the remark cause the arousal of dissonance? We suspected that the solitary utterance would not lead to dissonance. We designed a study that replicated the procedure of Festinger and Carlsmith's (1959) with the added factor of whether or not the student in the waiting room seemed convinced by the participant's statement that the task was fun and interesting. We predicted and found that the inverse

relationship between incentive magnitude and attitude change following counterattitudinal behavior occurred only when there was a consequence pursuant to the behavior – in this case, misleading a fellow student to have a false expectation (Cooper and Worchel, 1970). We had found another modifier: inconsistent cognitions lead to dissonance in the presence, but not the absence, of an unwanted consequence.

In 1980, a visiting professor at Princeton University, Paul Secord, expressed his consternation about the state of play of dissonance theory. He told me that he loved the theory when it was broad, simple, and comprehensible – that is, inconsistent cognitions lead to the state of dissonance. However, with two decades of research under the dissonance tent, he felt he needed a scorecard to know when it is true that inconsistent cognitions arouse dissonance and when it is not. He expressed the thought that the modifiers of dissonance had become the face of the theory and that someone needed to make sense of the modifiers. That was the challenge that my former graduate student, Russell Fazio, and I took up in the paper that became our signature modification of dissonance theory: the “new look” (Cooper and Fazio, 1984).

The New Look model of dissonance

Fazio and I examined the published data. Secord had a point. Inconsistent cognitions aroused dissonance

- *but only* if the actions were freely chosen;
- *but only* if the actor was committed to the discrepant cognition;
- *but only* if there was an aversive event following the dissonance;
- *but only* if the aversive consequences were foreseeable.

And the list went on. The theory was sorely in need of a cure for the *but onlys*.

It occurred to us that the important quest to find the modifiers of dissonance had contributed to a more meaningful understanding

of dissonance. Looked at from a slightly different perspective, the research of the prior two decades had transformed dissonance into a different kind of theory. Rather than maintaining that dissonance is a function of inconsistency, the data had actually called out for a new statement about the meaning of dissonance. The data were there. The *New Look* just had to tell the new story in different words.

As Festinger had surmised, dissonance is an arousing, uncomfortable tension state that motivates change. However, it is not brought about by cognitive inconsistency per se, but rather by the perception that one is responsible for bringing about an unwanted event. That perception, rather than the perception of inconsistency, is what results in the experience of cognitive dissonance.

Dissonance arousal and dissonance motivation

According to the New Look, dissonance begins with a behavior. In order for that behavior to lead to a cognitive or attitude change, a set of processes must unfold that can best be divided into two stages: *dissonance arousal* and *dissonance motivation*. Dissonance arousal occurs when people take responsibility for bringing about an aversive event. This conclusion may come about quickly, but not easily. There are several decision points that need to be crossed in order for a behavior to bring about dissonance arousal.

First, a behavior has to be perceived to have an unwanted consequence. Almost every behavior has a consequence. The question at issue for the actor is whether the consequence is unwanted and, if so, how unwanted is it? For example a person may favor a particular solution to the healthcare dilemma in the United States but advocate for a bill that is a bit different. If the person successfully convinces her friends, colleagues, or senators to support the bill, is this consequence sufficiently unwanted to lead to dissonance? Here, we argued that the consequence has to

fall outside a person's "latitude of acceptance" of possible positions that she can accept in order for dissonance to be aroused, otherwise dissonance will not follow from the behavior. Empirical evidence supports this first link in the dissonance process (Fazio et al. 1977).

The second decision point is the acceptance of *personal responsibility* for the consequences of the behavior. We defined responsibility as a combination of two factors: freely choosing the behavior in question and the ability to foresee the consequences of that behavior. Accepting responsibility leads to dissonance; denial of responsibility allows people to avoid the unpleasant state of dissonance. Decision freedom is crucial because it is necessary for the acceptance of responsibility.

Although decision freedom is necessary, it is not sufficient to lead to acceptance of responsibility. Imagine that the healthcare advocate in our example purchased a book that examined the potential uses and abuses of privatized healthcare. After the purchase, she found out that the proceeds from the book were donated to an advocacy group that condemned any government-supported healthcare plan. Would our advocate feel personally responsible for supporting a group that she finds unacceptable and contrary to her values? The New Look model argues that the answer is no. Despite the fact that she really gave money in support of the disliked group, she had no way to foresee that her actions would lead to the unwanted consequence. Foreseeability, then, is the second element that combines with decision freedom to determine whether one is responsible for an aversive event (Goethals et al., 1979). It is the acceptance of personal responsibility that provides the necessary and sufficient link to dissonance arousal.

On the radical nature of the New Look

We did not intend the New Look to be a radical departure from Festinger's original

notion, but we soon saw that it was. The arousal of dissonance no longer depended on what had been its major tenet – that is, the presence of inconsistency. It is true that inconsistency is typically a reasonable proxy for the active ingredients that lead to dissonance arousal. When people behave in ways that are inconsistent with cherished beliefs, they are typically simultaneously taking personal responsibility for bringing about at least the potential for an unwanted event. They may never find out if someone will be convinced by an attitude-discrepant statement, but the potential for bringing about an unwanted event is apparent. Thus, according to the New Look, acting inconsistently often brings with it the features that are actually the active ingredients in the dissonance process. Similarly, when people choose one attractive item over another in the free-choice research paradigm (Brehm, 1956) of dissonance, they are responsible for the aversive consequence: that is, rejecting all of the attractive elements of the unchosen alternative and accepting all of the unattractive features of the chosen alternative. And when people in an effort justification study (Aronson and Mills, 1959) choose to suffer in order to attain a goal, they are responsible for engaging in the unpleasant work, embarrassment or effort that is, in itself, an unwanted consequence. The problem with relying on inconsistent cognition as the underpinning of cognitive dissonance is that it requires a long list of caveats and exceptions (the *but onlys*) that moderate the effect. The underpinnings of the New Look model permit a more comprehensive view of the process.

Testing the New Look

Most of the work on the New Look model of dissonance preceded the publication of the New Look. As explained, the New Look was a way of understanding the theoretical underpinnings of dissonance, given what we already knew from published work on dissonance.

I would postulate that there were four major categories of findings on which the New Look relied. First, research must show that people experience dissonance arousal when they choose to engage in attitude-discrepant behaviors and not when they were forced to do so. This finding had already been well established prior to the New Look and continues to be a reliable research result. Second, attitude discrepant behavior must be shown to lead to dissonance when and only when the potential for aversive consequences exists. We had already shown this phenomenon on many occasions prior to the New Look model. Cooper and Worchel (1970) were the first to demonstrate it, but replications of the effect abound (Cooper et al., 1974; Goethals and Cooper, 1972; Goethals et al., 1979; Norton et al., 2003). (In fairness, this feature has emerged as a more controversial issue with some research questioning not its importance but its ubiquity [Harmon-Jones et al., 1996]). A third critical feature of the model was that dissonance arousal relies on the consequences of a behavior being foreseeable at the time they commit to their behavior. This, too, was well established (Cooper and Goethals, 1974; Goethals and Cooper, 1972; Goethals et al. 1979).

The fourth crucial feature had not been demonstrated when the model was first introduced. The responsibility-for-consequences model predicts that any behavior – not just inconsistent behavior – will arouse cognitive dissonance if it foreseeably leads to an aversive event. The radical nature of the New Look lay in its position that inconsistency was not necessary for dissonance to occur. In a study designed to collect evidence for this hypothesis, Scher and Cooper (1989) had people commit to attitude-consistent or attitude-inconsistent behavior. A cover story led participants to believe that a university committee was considering a policy in which student health records would become open to parents to peruse. Some were asked to write counterattitudinal essays that supported this unpopular and unwanted policy while others were asked to write proattitudinal essays.

The students were led to believe that the essays might either convince the committee or might boomerang and convince the committee of the opposite of what was written. In this way, the behavior (counterattitudinal versus proattitudinal) was manipulated orthogonally with the side of the issue that the committee was likely to believe (a desired versus an unwanted consequence of the essay).

When attitudes were measured following the essay, the results showed an effect for the consequence of the essay and not for the discrepancy of the arguments. Counterattitudinal advocacy led to attitude change only if it had the potential to produce an aversive event, thus replicating the finding of Cooper and Worchel (1970). But so did proattitudinal advocacy. Writing in the direction of their own beliefs, students who discovered that their essays were likely to produce a boomerang and thus help to bring about an unwanted event changed their attitudes. The consequence and not the inconsistency of behavior and attitude produced dissonance arousal that ultimately motivated attitude change.

Is dissonance arousing?

One of the hallmarks of Festinger's brilliance was his adoption of the drive model to introduce the theory of cognitive dissonance. Clark Hull, Kenneth Spence and their colleagues were debating the proper mix of drives and habits to understand how organisms learn. M.L. Skinner (1953) had made the unique proposal that drives were not necessary at all to understand learning. The drive concept was highly accessible and controversial as Festinger introduced cognitive dissonance to the community.

By postulating a drive state underlying the dissonance process, he made his work relevant to the learning community both inside and outside of social psychology. By postulating that dissonance reduction in social behavior is often an inverse function of the

magnitude of incentive and reward, he upset the foundation that the learning community had laid for social psychologists. I believe that the drive and reinforcement notions were both key to the influence that dissonance had in the 1950s. I was never certain that Festinger ever expected the drive concept to receive a direct test. It was a virtual metaphor, a way to think about the process that had the added benefit of making his fledgling theory both noticed and controversial. I believe he was pleasantly surprised when subsequent research found that he had guessed correctly.

Converging evidence has now demonstrated that dissonance has, as Festinger (1957) called it, "drive-like properties." Waterman and Katkin (1967) reasoned that if dissonance was a drive, then it should have the effect that drives typically have on learning: it should facilitate simple learning and interfere with complex learning. Waterman and Katkin (1967) found evidence for the former but not for the latter. That evidence was supplied a few years later by Pallak and Pittman (1972) who found that dissonance following counterattitudinal advocacy interfered with people's ability to learn a complex task.

Using a different logic, Zanna and Cooper (1974) showed that if people believed their arousal following counterattitudinal advocacy was due to something other than the advocacy (such as a pill), they did not show attitude change following the advocacy. Apparently, attitude change is directed toward lowering the uncomfortable drive state. If the arousal was thought to be due to some other agent, no attitude change ensued. Moreover, reducing the degree of bodily arousal by means of a sedative was shown to decrease attitude change following counterattitudinal advocacy while ingesting an arousing agent was shown to increase attitude change (Cooper et al. 1978).

A third line of evidence was the measurement of arousal following counterattitudinal advocacy. Croyle and Cooper (1983) found skin conductance differences between

participants in high dissonance and low dissonance conditions. Losch and Cacioppo (1990) replicated that finding and also showed that dissonance reduction is directed at reducing the uncomfortable affect that dissonance produces. Elliot and Devine (1994) added to the literature on the uncomfortable motivational state by asking participants how they felt following a dissonance-arousing act. Participants reported significantly more negative discomfort than participants in low dissonance conditions.

The New Look complements the original version of dissonance theory by accepting the notion that cognitive changes are motivated by psychological discomfort and physiological arousal. What may have begun as a metaphor to predict change has received substantial support from a variety of research perspectives. Dissonance, we believe, is motivated by being responsible for bringing about an unwanted consequence of behavior. It is experienced as discomfort and motivates cognitive change.

SELF-STANDARDS: MOVING THE NEW LOOK FORWARD

The New Look model of dissonance was not without its critics (see Harmon-Jones and Mills, 1999). Elliot Aronson, one of dissonance theory's most innovative pioneers, advocated a perspective on dissonance that focused on an individual's violating his or her concept of self. Good people expect that they will do good things; bad people expect they will do bad things and when the twain mixes, dissonance is aroused. Early work (e.g., Aronson and Carlsmith, 1962) had shown that people who expect to fail will choose to fail as a way of staying consistent with their self-concept. To do otherwise would result in the unpleasant inconsistency known as dissonance. Aronson (e.g., 1992; Thibodeau and Aronson, 1992) argued that inconsistency with self was sufficient to produce dissonance and the New Look's

insistence on aversive consequences was unnecessary.

In order to demonstrate that dissonance can occur in the absence of aversive events, Thibodeau and Aronson (1992) introduced what they called the 'hypocrisy paradigm'. In this research, people were asked to make statements that were consistent with their private beliefs but were reminded of times in the past that they had acted inconsistently. The general finding from this research is that dissonance occurs as a function of the proattitudinal statement when the statements are made freely and participants are made mindful of their prior inconsistent behavior. Typically, participants engage in subsequent behavior that is more in line with their proattitudinal advocacy. For example, college students in a study by Stone et al. (1994) were asked to make public statements to a group of high school students advocating the use of condoms when engaging in sex. This speech was consistent with the participants' attitudes about the use of condoms. In a key hypocrisy condition, the participants were made mindful of occasions in which they had not practiced what they had just preached – that is, they were asked to recall times that they had not used condoms. When the study was allegedly finished, the participants were allowed to purchase as many condoms as they wished. The results showed that when making the statement under conditions of free choice and with the reminder of their own discrepant behavior, participants purchased more condoms than in any other condition in the experiment. In order to reduce their dissonance, participants exaggerated their behavior to bring it in line with their attitude and their proattitudinal statement.

This intriguing line of research raised the question of the necessity of an aversive event for the dissonance process (Aronson, 1992). On the other hand, it seemed to me (Cooper, 1992) that the aversive consequence was intrinsically enmeshed in the mindfulness manipulation. Being reminded of not having worn condoms or, in other studies, being

reminded of times that you wasted water (Dickerson et al., 1992) or failed to recycle (Fried and Aronson, 1995) are all aversive consequences. They happen to be in the past, but they are consequences nonetheless. The results seemed consistent with the New Look model.

The set of hypocrisy studies accomplished a great deal, however. It added to the number of ways that researchers can investigate dissonance arousal; the dissonance resulting from hypocrisy is frequently channeled to behavior change rather than the more typical attitude change and the behavior change is usually in a direction that promotes constructive social and personal values – an issue that we will return to later in this chapter. From a theoretical standpoint, the studies laid bare an issue about which we had said little in the New Look model: What do we mean by an aversive consequence?

In the New Look, Fazio and I defined an aversive consequence as an occurrence that one would rather not have occur. That is, if you can think of something you would not like to bring about, such as convincing a person to believe in an unwanted position or suffering embarrassment or being stuck with unattractive features of a choice alternative, then that is what we meant by "aversive consequence." We believed that there was no adequate *a priori* definition of an aversive consequence. Whatever a person thought was unwanted and yet acted in a way that caused that event to occur was grist for the dissonance mill. We had no disagreement with Aronson's (1969) notion that violations of self-expectancies can cause dissonance – as long as people find it aversive to violate their self-expectancies. The New Look differed from self-expectancy view because we did not believe that self-expectancy violations were the *only* route to dissonance. Whatever a person finds aversive or unwanted, whether it is a violation of self-expectancy or a behavior that brings about any other unwanted consequence fits the New Look's understanding of an aversive event and serve to arouse cognitive dissonance.

Jeff Stone, who had been a graduate student with Elliot Aronson and then a postdoctoral fellow working with me, helped to integrate the New Look's emphasis on *any* aversive consequence and self-expectancy's emphasis on violations of the self as necessary for dissonance. In what I believe to be the most recent full model of dissonance processes, Stone and I (Stone and Cooper, 2001) advanced the Self-Standards model of dissonance. What was missing in the New Look was an explicit way to judge the meaning of a behavior. In Stone and Cooper (2001) we argued that what arouses dissonance is an initial assessment of a behavior against a particular behavioral standard. All behaviors have consequences. To judge the desirability of those consequences requires a comparison to a standard of judgment. In the Self-Standards model we spelled out those judgment standards.

Normative and personal standards

We reasoned that there are two major categories of standards that a person can use to assess the meaning of the consequences of his or her behavior – normative and personal. There are some outcomes that we can create in the world that most people would agree are of a particular valence. Most people would agree that contributing to a charity or helping a roommate study for an exam are positive events. We know there may be occasions in which helping a roommate and/or contributing to charity may have complicated mixed motives, but by and large, such actions are considered positive. Similarly, there are consequences that most people would agree are negative or undesired. Running into someone on the street and knocking him down would be generally aversive. So, too, is lying to someone, especially when the person believes you and is influenced by your lie (e.g., Festinger and Carlsmith, 1959). Granted, there may be some odd times when those outcomes are positive, but, typically, most people would agree they are negative.

When a standard of judgment is based on a perception of what *most people* perceive to be foolish, immoral, or otherwise negative, we can say people are using a *normative standard of judgment*. The main thrust of this definition is that the standards are based on a shared understanding of good and bad, wanted or unwanted, foolish or clever (Higgins, 1989). The other broad category of standards of judgment are those that are based on the unique characteristic of the individual. These are *personal standards of judgment*. They refer solely to the judgments people make when they consider only their own values or desires. Consider a casual runner who runs a mile in 4.5 minutes. By the standards of most casual runners, this is an extraordinary experience. However, this particular runner expected to cross the mile marker in closer to 4 minutes. Regardless of whether it is rational or not, regardless of whether others would agree with this runner's judgment, the outcome fails to meet the runner's personal standard of judgment. The achievement, when compared with the runner's personal standard of judgment, is not an achievement at all but an unwanted, aversive event.

The self-standards dissonance model asserts that people can use either a normative standard of judgment or a personal standard of judgment in order to assess their behavior. Which standard they use is a function of the standard that is accessible at the time of their behavior. If the situation makes normative standards accessible, then people will use their concept of what most people would find desirable as the way to assess the consequences of their behavior. Conversely, if people are induced to have their personal standards accessible, then they will use their self-expectations as the standard of judgment to determine whether or not an outcome is aversive.

Personal and normative self-standards can also be chronically accessible for particular individuals. If a person thinks of herself as being a deceitful person, she will not be upset at all by convincing a waiting fellow

student that a boring task was actually interesting. She carries with her a chronically available self-standard and compares her behavior to that standard. Another person may think of herself as honest and carry that self-standard as a chronic measuring stick against which to judge her behavior. She would be in the throes of dissonance after agreeing to dupe a fellow student. For these two hypothetical students, their judgment is measured against a personal standard that overrides the effect of the social circumstance.

The predictions of the self-standard model have been supported in a number of studies (Weaver and Cooper, 2002; Stone, 1999; Stone and Cooper, 2003). When people compare their behavior to normative standards of judgment, then they assess consequences to be aversive in a manner similar to most people in the culture. We would not expect dissonance to be moderated by their sense of self – for example, their level of self-esteem. By contrast, when ideographic dissonance is aroused by comparison to personal standards of judgment, then what is considered aversive varies by self-esteem. People with a high sense of self-esteem expect to make good and rational choices. They are upset when their choices lead to a consequence that is negative. When people with chronically low self-esteem make choices, they expect those choices to have negative results and are not upset by what other people would consider negative outcomes.

In a study reported by Stone (1999), participants were divided by median split into those with high and low self-esteem. They were asked to make choices between two attractive music albums. Half of the students were primed to make their personal standards accessible while the other half were primed to have their normative standards accessible. Following the decision of which album to keep, the participants rerated the albums. The prediction was that normative-primed participants would experience dissonance whenever they made a difficult

choice and would show the classic dissonance finding of raising their evaluation of the chosen album and reducing the attractiveness of the rejected album. Stone predicted that self-esteem would not enter into the findings because these students were measuring the consequences of their behavior against a normative standard of judgment. By contrast, self-esteem was expected to play a role when participants had been primed to use their personal standards. People with high self-esteem are far more likely than people with low self-esteem to believe that elements discrepant with their choice are aversive – but only because they are assessing those consequences against a personal standard of judgment. Stone found that when normative standards were primed, self-esteem made no difference and participants changed their attitudes toward the albums as predicted by dissonance. However, when personal standards were primed, participants with high self-esteem changed their attitudes far more than participants with low self-esteem.

Progress report on a classic theory

Festinger thought that dissonance was a function of cognitive inconsistency. As I look back at more than a half century of theory and research on this now-classic theory, two facets of Festinger's genius are palpable from his writing. One facet was about form and the other about substance. Festinger formed his theory around the great issues of the time. Learning was king of the psychology literature in the 1950s and Festinger adapted the Hullian drive concept for use in social psychology. In social psychology, learning and reinforcement concepts were the assumed girders of much of the research in persuasion and attitude change (Hovland et al., 1949, 1953). Festinger aligned his theory to make precisely the opposite predictions about persuasion than would have been predicted by learning, therefore creating instant controversy and research.

From a substantive perspective, Festinger thought that inconsistency among cognitions led to the uncomfortable arousal state he called dissonance. From what we now know, he was partially correct. "All theories are wrong," Festinger (1987) once wrote, "One asks, 'How much of the empirical realm can it handle and how must it be modified and changed as it matures?'" In the New Look and Self-Standards models, my colleagues and I tried to right the ship when it veered away from its path and to find new solutions to the *but only* dilemmas that taught us that the theory did not capture the totality of the dissonance phenomenon. To be clear, we were not the only investigators to notice that the theory needed additional concepts and perspectives to capture the richest range of phenomena and data. For example, Beauvois and Joule (1999), Harmon-Jones (1999), and Steele (1988) are among the creative scholars who have used alternate lenses to analyze the progress of dissonance through the decades. There is consensus that Festinger set us out on the path to understand how people view the 'fit' of their cognitions. There is consensus that one of his brilliant and ever-lasting insights was to allow us to consider all cognitions – whether mental representation of the world or mental representations of internal states – on the same grid and therefore subject to the rules of the dissonance process. There is consensus that his straightforward set of principles stimulated research in a way that was unprecedented in the field of social psychology.

I believe in the dissonance processes that we outlined in Cooper and Fazio (1984) and enhanced in Stone and Cooper (2001). As Festinger once taught us, however, our own version of the work will one day be proven wrong (only in part, I hope). Festinger wrote, "The only kind of theory that can be proposed and ever will be proposed that absolutely will remain inviolate for decades ... is a theory that is not testable. If a theory is at all testable, it will not remain unchanged. It has to change."

NEW AVENUES OF DISSONANCE RESEARCH

From personal dissonance to vicarious dissonance

People's selves are integral to the dissonance process. Recent theorizing has made clear that the self is both personal and social (Leary and Tangney, 2003). It is about one's own personal characteristics and simultaneously about one's interconnectedness with others and with social groups (e.g., Brewer and Gardner, 1996), yet prior research connecting the experience of cognitive dissonance to membership in social groups was scant during the formative period of dissonance research. Ironically, the first study of cognitive dissonance ever reported was the study of disconfirmed expectancies by members of a doomsday cult who believed that the world would end in a cataclysmic flood. Their reactions to the disconfirmed expectancy formed the basis of *When Prophecy Fails* (Festinger et al., 1956). However, it would take researchers several decades to systematically vary group membership and assess the impact of participants' acting as individuals compared with their acting as members of a small group (Cooper and Mackie, 1983; Zanna and Sande, 1987).

In the theory of vicarious dissonance (Cooper and Hogg, 2007), we went more to the heart of the meaning of group membership and considered its impact on dissonance. We considered the effect of one group member's counterattitudinal advocacy on the attitudes and behaviors of other members of one's group. Social identity theory was the vehicle that helped us to link the dissonant behavior of one group member with the attitudes of other members of the group. Because of the impact of social identity on members of social groups, we reasoned that dissonance aroused in one group member could cause other group members to experience dissonance vicariously and result in attitude change by the other members of the social group.

As we know from the important work in social identity theory (Tajfel, 1970) and social categorization theory (Turner and Hogg, 1987), people in groups forge a common identity. When thinking of themselves as group members, there is a tendency toward depersonalization and intersubjectivity such that members assimilate toward the prototypical member of their group. The more strongly they feel about their group, the more they share in intersubjectivity and the more they take on the characteristics and emotions of fellow group members. Put simply, happiness or fear or sadness experienced by one member of a group spreads via intersubjectivity to other members (Mackie et al., 2007).

Michael Hogg and I wondered whether one group member's dissonance could spread to other group members in the same way. Suppose you are a member of a conservative antitax group and you observe a fellow group member make a public speech advocating an increase in the progressive income tax in order to support social programs. You know that the person voluntarily made the speech and that it was played before a potentially convincible audience. The situation has all of the ingredients to create cognitive dissonance in the speaker. But what about you, the witness? We reasoned that you will experience cognitive dissonance vicariously. Because of your common group membership, your view of your self is partly determined by your membership in the group to which you and the speaker belong. Your identity is wrapped up with your fellow group members and intersubjectivity makes you and the speaker fuse toward a common identity. The speaker's experience of discomfort will become your experience of discomfort. The speaker's attitude change will become your attitude change. The speaker's reduction of dissonance will become yours as well.

Norton et al. (2003) provided evidence for vicarious arousal of dissonance. A fictitious cover story created a rationale for a student to witness a fellow student agree to write a counterattitudinal message and to learn whether or not the student was a member of

the participant's social group. At Princeton University, all entering undergraduate students are assigned at random to one of five residential colleges. Each student lives and eats in one of the colleges and each college has its own social and academic activities. The student's residential college served as the crux of the ingroup versus outgroup manipulation because each participant believed that he or she was witnessing an interaction with a student who happened to be a fellow resident of his or her residential college (ingroup member) or happened to be a resident of a different residential college (outgroup member).

The students arrived for a study of "linguistic subcultures" in groups of two, although each reported to a separate room, separated by two-way mirrors. We told the students that we were interested in how people in different residential colleges come to speak in slightly different ways, learning to use slightly different inflections or terms in their spoken behavior. For example, we know that people who live in the Midwest develop a slightly different pattern of speech than people who live in South Carolina or Massachusetts. The experimenter explained that the purpose of the students' participation in the current study was to see if these speech patterns occur in microcosms – that is, small groups within a larger context. We told the students that, in this study, we wanted to see if the speech patterns of students in the different residential colleges at Princeton University were different from one another and whether we could measure them.

We explained that one of the two students, selected at random, was going to deliver a speech on a given topic and the other student was going to listen carefully, and then respond to several questions about the speaker's speech patterns. Each participant was told that he or she was the one who had been randomly picked to rate the speech, while the student in the other room was assigned to give the speech. The procedure allowed us to make the student's residential college group salient and manipulate systematically whether

the speaker's residential college group was the same (ingroup) or different (outgroup) from the participant's. The experimenter found a pretext to turn the lights on briefly, which allowed the participants to see that there truly was another student in the other room. The illumination was kept low so that the students' identities could not be accurately discerned. What students did not realize was that each of them had been assigned the role of listener. All information about what the alleged other student said or did was manipulated by instruction or audiotape.

The experimenter left the room, ostensibly to instruct the other participant about the speech he or she was to make. During the intervening period, participants filled out various measures, including measures of how much they liked and felt identified with their residential college on a scale developed by Hogg and his associates (Hogg et al., 1998). In a few minutes, the experimenter returned with a tape recording that included the completed speech and the experimenter's alleged conversation with the other student. On the tape, the experimenter explained that he was fortunate to be able to combine two studies into one. The dean's office had asked for a study trying to assess student opinion about the possibility of raising tuition fees by a more than typical amount. The experimenter then asked the student to write a strong and forceful speech advocating a spike in tuition fees. He explained that this would be the speech that the other subject (i.e., the real participant) would rate for its linguistic features and that it would then be sent on to the dean's office. The experimenter also asked the alleged other student how he or she felt about raising tuition and the student responded, "Well ... I'd be against it."

The participants thus had a credible, albeit fabricated, story that allowed them to overhear an ingroup or an outgroup member make a counterattitudinal speech on a controversial topic. The tape recorder was stopped while the writer supposedly organized his or her thoughts, and then restarted for the participant to hear the alleged speech. The speech was a relatively brief exposition on how

higher tuition rates could allow the university to hire more faculty staff, purchase more books for the library, and so forth. Before rating the speech for its linguistic properties, participants were asked about their own attitudes toward tuition increases at the university. This served as the dependent measure of our study.

The results of the study showed that observing a fellow group member behave in a counterattitudinal fashion caused the participant to change his or her attitude in the direction of the group member's counterattitudinal advocacy. As predicted by vicarious dissonance, this effect *only* occurred when the participant strongly identified with his or her group. In the absence of a strong affinity with one's group, observing an ingroup or an outgroup member did not affect participants' attitudes (see also Monin et al., 2004).

Dissonance, vicarious dissonance, and culture

The concept of vicarious can help us unravel some of the cultural differences that have been identified in the expression of cognitive dissonance across cultures. Joan Miller (1984) was among the first to suggest that cultural differences may lead to different expressions of social psychological processes. She analyzed the difference between holistic and agentic cultures that broadly corresponded to Western European and North American cultures on the one hand and Asian and Indian on the other. In agentic cultures, people see themselves as responsible for their own actions and decisions. They make personal attributions for events, viewing behavior as a window on their own traits and characteristics. People from holistic cultures view the self in relationship to others. They see situations and social roles as determining their behavior and view behavior as a means toward harmonious social relationships.

In a subsequent seminal paper, Markus and Kitayama (1991) expanded the analysis of cultural differences by drawing a distinction between collectivist and individualist

cultures. Collectivist cultures are concerned with relationships among people. Social harmony is a key goal; attitudes and behavior primarily serve the goal of collective harmony. Individualist cultures are concerned with self-actualization. People's attitudes and behaviors are their own and expressing them truly and honestly is an important component of the self-actualization process.

By raising the question of whether different cultures have values that differentially affected the expression of attitudes, Markus and Kitayama opened an entirely new research direction for cognitive dissonance. They suggested that that dissonance is a uniquely Western or individualistic phenomenon. In individualistic cultures, people express opinions that are supposed to accurately reflect their judgment. They are supposed to say what they believe and believe what they say. Inconsistent cognitions do not fit an individualistic culture's notion of attitude expression. On the other hand, the expression of attitudes in collectivist cultures is only partly self-description but also constitutes expressions that affect the degree of harmony among people or groups. A member of a collectivist culture may not find it aversive to express attitudes that differ from their behaviors, but they would find it aversive if expressing opinions that disrupted interpersonal or intergroup harmony.

Although this chapter will not review the considerable research that has been conducted on dissonance in collectivist and individualist cultures during the last two decades, the conclusion of that research has revealed interesting and important aspects of the dissonance process itself. Heine and Lehman (1997) collected data in Canada comparing dissonance processes among Canadians of European descent and people of Japanese descent. Using a typical free-choice paradigm, Heine and Lehman found that, unlike the European Canadians, Japanese participants did not show the spreading of alternatives effect that had been found numerous times in the social psychology literature.

Did this mean that cognitive dissonance is not experienced by people in collectivist cultures? Hoshino-Browne and colleagues (2005)

reported a series of creative experiments in which they elucidated the impact of culture on cognitive dissonance. They showed that people from collectivist cultures showed dissonance reduction following a choice if they made that choice for a friend rather than for themselves. That is, when attitudes and behaviors were inconsistent within a social network of relationships, it produced dissonance. When the inconsistency had no social referent, it did not.

Research on culture and dissonance provides a window into the important social values that, when disrupted, create the aversive event that leads to cognitive dissonance. For collectivist cultures, the value is interpersonal harmony. When people behave in ways that upset the social order, it produces the aversive consequence that leads to cognitive dissonance. In Western cultures, when people act in ways that produce unwanted consequence to the individual actor, it leads to dissonance. Research by Kitayama and his colleagues has also shown that the same action that leads to dissonance arousal in individualists can also lead to dissonance among collectivists, when the presence of social others is subtly primed (Imada and Kitayama, 2010; Kitayama et al., 2004).

Vicarious dissonance research provides another perspective to consider the differences between individualist and collectivist cultures. Chong and Cooper (2007) reported a study using the induced compliance paradigm of dissonance. Students in South Korea wrote essays that could bring about an unwanted policy change at their university. Chong and Cooper found that Korean students did not change their attitudes following the counterattitudinal speech, even though they acted with free choice and their action had the potential of bringing about an unwanted policy. However, Korean students did change their attitudes when the situation was changed into a vicarious dissonance study. When they witnessed students from their group writing a counterattitudinal essay, they changed their own opinion just as the participants in Norton et al.'s (2003) study had done.

Vicarious dissonance is intrinsically a social phenomenon. It is an arousal that is

experienced on behalf of someone else in an important social network. Taken together with the research of Hoshino-Browne et al. and Kitayama et al., we now understand that dissonance is aroused in both collectivist and individualistic cultures. For collectivists, more so than for individualists, a social consequence seems necessary for an act to be considered aversive and lead to the tension state of dissonance.

DISSONANCE IN A SOCIAL WORLD

Why does it matter?

One of the characteristics of cognitive dissonance is its ubiquity. When we make choices or suffer embarrassment or expend effort, we are in a dissonant state. It is difficult to go through a day without arousing dissonance. Attitudes ranging from consumer preferences (Menasco and Hawkins, 1978) to military service (Staw, 1974) have been viewed from a dissonance perspective. We have gained insight into why people feel passionately about social groups such as sororities and fraternities that they suffered to get into from the vantage point of cognitive dissonance. And what professor has not considered whether some students seem to love very difficult courses because of the effort justification intrinsic to cognitive dissonance?

To conclude this chapter, I will address two particular areas of social impact that have been addressed systematically with the principles of cognitive dissonance. The first is the potential of dissonance to effect change through psychotherapy and the second is the use of vicarious cognitive dissonance to induce positive changes in health behavior.

Dissonance as psychotherapy

Can psychotherapy be considered an instance of cognitive dissonance? In the 1980s, Danny Axsom and I conducted a series of studies in which we showed how dissonance could

induce people to change their attitudes and behaviors psychotherapeutically (Axsom, 1989; Axsom and Cooper, 1985; Cooper, 1980). We noted the parallel between most psychotherapies and the principles of effort justification that Aronson and Mills had introduced in 1959. All therapies require effort. People engage in the effortful work freely, although the goal for which they are working is something for which they have some trepidation. It may be that they came to therapy to reduce their fear toward an object or their anxiety toward performing a particular behavior. Whatever the goal may be, prospective patients have ambivalence toward it, then engage in an effortful set of therapy sessions designed to overcome it.

We reasoned that the choice to engage in an effortful procedure was akin to the high choice conditions of an effort justification study. We decided to run studies in which people would attempt to reach a goal using a high degree of effort totally unrelated to any bona fide theory of psychotherapy. In one study, we invited people who wanted to reduce their fear of snakes to come to the laboratory where we measured how close they could come to our six-foot boa constrictor. They then participated in an effortful therapy that they believed was related to overcoming the fear. In truth, it contained a set of physical exercises designed to be difficult, embarrassing, and exhausting. We found that the participants were able to overcome their fear following this physical exercise therapy. Moreover, we also varied participants' choice to engage in the effortful therapy. Consistent with dissonance theory predictions, phobic participants who freely chose to engage in the effortful therapy overcame their fear significantly more than participants who were not given a choice (Cooper, 1980).

In a similar study, Axsom and I (Axsom and Cooper, 1985) put dissonance to use against the problem of obesity. People who were overweight and who had tried numerous weightloss programs volunteered for our experimental research. In a series of

five sessions, we asked our experimental group to engage in tasks that required a great deal of cognitive effort. They made perceptual judgments, read tongue-twisters and recited nursery rhymes, backward, for an hour. A second group engaged only in a low degree of effort, spending their time making simple judgments and relaxing. Six months after the end of the sessions, the participants were weighed. As we had predicted from dissonance theory, the high effort group lost more weight than the low effort group (8.6 lbs versus 0.8 lbs respectively) and kept the weight off for a year.

We do not claim that all psychotherapy is cognitive dissonance but we do believe that cognitive dissonance is one of the active ingredients of most psychotherapies. With knowledge of the conditions that give rise to maximal dissonance, we should be able to design psychotherapies in a way that allows dissonance to be helpful for psychotherapeutic change. Whichever approach a particular therapist adopts, maximizing the impact of dissonance in the therapeutic program can only enhance the accomplishments of the therapy. Therapists are advised to focus patients' attention on the effortful nature of the patient's work. Moreover, therapists should emphasize the patient's personal responsibility for engaging in the therapeutic effort. If these elements are included in psychotherapy, then the arousal and motivation that accrue from the dissonance process will be put to productive use for the patient.

Vicarious dissonance can lead to a healthier society

I believe that vicarious dissonance opens up a vast array of possibilities to put dissonance to use for a better world. We know that exposure to group members who are engaging in a dissonant act creates dissonance in other members (Monin et al. 2004; Norton et al, 2003). Consider the following scenario: A person observes a fellow group member advocate the use of risk-protective health

behaviors; for example, eating a healthy diet, using condoms during sexual activity, refraining from smoking or applying sunscreen to protect against cancer. If the person observes the fellow group member admit to prior instances in which she did not practice what she preached, the elements for the vicarious experience of dissonance will be met. The vicarious dissonance should be resolved by the person's committing to more healthy future behaviors.

We (Fernandez et al., 2007) conducted such a study. Students at the University of Arizona were asked to listen to a speech made by a student that encouraged people to use sunblock as a preventative measure for skin cancer. They were led to believe that the speech had been made by a student at their university (ingroup condition) or by a student at a rival university (outgroup condition.) The speech was consistent with the participants' and the speaker's attitude toward sunscreen. It concluded, "No matter how busy you think you are with work or school you can and should always wear sunscreen to reduce your risk of cancer."

Through the use of an appropriate cover story, participants heard the ingroup or outgroup speaker reminded that she did not use sunscreen herself *every time* she went outdoors. We predicted that the experience of vicarious dissonance would be aroused for students who heard an ingroup member admit to hypocrisy but not an outgroup member and the more identified a student was with her own group (University of Arizona), the more she would experience vicarious dissonance.

We predicted that the vicarious dissonance would lead to prohealth behaviors and pro-health attitude change on the part of the participants, and that is what we found. Women in the study changed their own attitudes to become more ardent in their opinion that sunscreen should be used on all occasions. Moreover, when given the opportunity to return a coupon for a free bottle of sunscreen, 74 percent of the women in the vicarious hypocrisy condition requested their bottle

whereas only 54 percent of the women in the low vicarious hypocrisy (outgroup speaker) condition requested their sample.

There is good reason to believe that vicarious hypocrisy can be recruited by institutions from schools to the workplace to help its members live a healthier lifestyle and make healthier, less risky life decisions. Vicarious dissonance is a multiplier. A person who is induced to express vicarious hypocrisy creates his or her own change as an individual but may multiply that change throughout the groups to which he or she belongs. Schools, for example, can provide an opportunity for students to observe a fellow student speaking forcefully of her commitment to a healthy behavior regimen; for example, a commitment to a frequent exercise program. If the student acknowledges some occasions in which she failed to get to the gym, the conditions for vicarious dissonance will have been met. The dissonant cognitions of the speaker will spread to group members who will reduce their vicarious dissonance by adopting the exercise regimen that the single group member espoused. Similarly, a work place might bring group members together to witness a fellow worker advocate for healthy eating choices. By acknowledging his own dietary transgressions, fellow group members will experience vicarious dissonance, which they can reduce by making healthier food choices. The dissonance expressed by the single group member can spread through the group, affecting all of the members of the group. We can further speculate that this will be especially true if the members feel strongly identified with their group.

CONCLUSION

I came upon the theory of cognitive dissonance when it was in its childhood. It was a precocious child, already having generated fans and enemies, proponents and critics. A half-century later, the theory continues to inspire. Although only a few of the most

ardent skeptics still doubt the existence of dissonance, the precise mechanism continues to be elusive. My own joy from being part of the dissonance cohort comes from having helped move the theory to a new level of understanding. In my own thinking, the reliance on inconsistency gave way to an understanding of the role of responsibility-for-consequences that is the core of the New Look model and for the importance of the self-concept that is the foundation of the Self-Standard model.

As Festinger tried to teach us, anyone's perspective on dissonance process will ultimately prove inadequate as data continue to be collected. All theories are at least partly wrong and all theories must change. Nonetheless, the search for change is part of the science and part of the fun. In the half century since Festinger first brought dissonance to our attention, we have not only moved toward a deeper understanding of this ubiquitous process, but we have also seen the theory spawn new ideas and relationships. Such major theoretical approaches as Kunda's motivated reasoning (Kunda, 1990), Higgins' (1989) self-discrepancy and Tesser's (1988) self-evaluation maintenance are but a few examples of that search. There will be more. The theoretical stability of dissonance and the change it continues to inspire are the twin legacies of cognitive dissonance theory.

REFERENCES

Aronson, E. (1969) The theory of cognitive dissonance: A current perspective. In L. Berkowitz (ed.), *Advances in Experimental Social Psychology*, 7, 1-34. New York: Academic Press.

Aronson, E. (1992) The return of the repressed: dissonance theory makes a comeback. *Psychological Inquiry*, 3, 303-311.

Aronson, E. and Carlsmith, J.M. (1962) Performance expectancy as a determinant of actual performance. *Journal of Abnormal and Social Psychology*, 65, 178-182.

Aronson, E. and Carlsmith, J.M. (1963) The effect of the severity of threat on devaluation of forbidden

behavior. *Journal of Abnormal and Social Psychology*, 66, 584–588.

Aronson, E. and Mills, J. (1959) The effect of severity of initiation on liking for a group. *Journal of Abnormal and Social Psychology*, 59, 177–181.

Axsom, D. (1989) Cognitive dissonance and behavior change in psychotherapy. *Journal of Experimental Social Psychology*, 25, 234–252.

Axsom, D. and Cooper, J. (1985) Cognitive dissonance and psychotherapy: The role of effort justification in inducing weight loss. *Journal of Experimental Social Psychology*, 21, 149–160.

Beauvois, J. and Joule, R.V. (1999) A radical point of view on dissonance theory. In E. Harmon-Jones and J. Mills (eds), *Cognitive Dissonance: Progress on a Pivotal Theory in Social Psychology*, pp. 43–70. Washington, DC: American Psychology Association.

Brehm, J. (1956) Postdecision changes in the desirability of alternatives. *Journal of Abnormal and Social Psychology*, 52, 384–389.

Brewer, M.B. and Gardner, W. (1996) Who is the 'We'? Levels of collective identity and self representation. *Journal of Personality and Social Psychology*, 71, 83–93.

Carlsmith, J.M., Collins, B.E. and Helmreich, R.L. (1966) Studies in forces compliance: I. The effect of pressure for compliance on attitudes change produced by face to face role playing and anonymous essay writing. *Journal of Personality and Social Psychology*, 4, 1–13.

Chong, J. and Cooper, J. (2007) Cognitive dissonance and vicarious dissonance in East Asia: Can I feel your discomfort but not my own? Poster presentation at a Meeting of Society for Personality and Social Psychology, Memphis, TN.

Cooper, J. (1980) Reducing fears and increasing assertiveness: The role of dissonance reduction. *Journal of Experimental Social Psychology*, 16, 199–213.

Cooper, J. (1992) Dissonance and the return of the self-concept. *Psychological Inquiry*, 3, 320–323.

Cooper, J. and Fazio, R.H. (1984) A new look at dissonance theory. In L. Berkowitz (ed.), *Advances in Experimental Social Psychology*, 17, 229–262. Hillsdale, NJ: Erlbaum.

Cooper, J. and Goethals, G.R. (1974) Unforeseen events and the elimination of cognitive dissonance. *Journal of Personality and Social Psychology*, 29, 441–445.

Cooper, J. and Hogg, M.A. (2007) Feeling the anguish of others: A theory of vicarious dissonance. In M.P. Zanna (ed.), *Advances in Experimental Social Psychology*, 39, 359–403. San Diego, CA: Academic Press.

Cooper, J. and Mackie, D.M. (1983) Cognitive dissonance in an intergroup context. *Journal of Personality and Social Psychology*, 44, 536–544.

Cooper, J. and Worchel, S. (1970) The role of undesired consequences in the arousal and cognitive dissonance. *Journal of Personality and Social Psychology*, 19, 312–320.

Cooper, J., Zanna, M.P. and Goethals, G.R. (1974) Mistreatment of an esteemed other as a consequence affecting dissonance reduction. *Journal of Experimental Social Psychology*, 10, 224–233.

Cooper, J., Zanna, M.P. and Taves, P. (1978) Arousal as a necessary condition for attitude change following induced compliance. *Journal of Personality and Social Psychology*, 36, 1101–1106.

Croyle, R. and Cooper, J. (1983) Dissonance arousal: Physiological evidence. *Journal of Personality and Social Psychology*, 45, 782–791.

Davis, K.E. and Jones, E.E. (1960) Change in interpersonal perception as a means of reducing cognitive dissonance. *Journal of Abnormal and Social Psychology*, 61, 402–410.

Dickerson, C.A., Thibodeau, R., Aronson, E. and Miller, D. (1992) Using cognitive dissonance to encourage water conservation. *Journal of Applied Social Psychology*, 22, 841–854.

Elliot, A.J. and Devine, P.G. (1994) On the motivational nature of cognitive dissonance: Dissonance as psychological discomfort. *Journal of Personality and Social Psychology*, 67, 382–394.

Fazio, R.H., Zanna, M.P. and Cooper, J. (1977) Dissonance and self-perception: An integrative view of each theory's proper domain of application. *Journal of Experimental Social Psychology*, 13, 464–479.

Fernandez, N., Stone, J., Cascio, E., Cooper, J. and Hogg, M.A. (2007) Vicarious hypocrisy: The use of attitude bolstering to reduce dissonance after exposure to a hypocritical ingroup member. Paper presented at the Meeting of the Society for Personality and Social Psychology, Memphis, TN.

Festinger, L. (1950) Informal social communication. *Psychological Review*, 57, 271–282.

Festinger, L. (1954) A theory of social comparison processes. *Human Relations*, 7, 117–140.

Festinger, L. (1957) *A Theory of Cognitive Dissonance*. Evanston, IL: Row, Peterson.

Festinger, L. (1987) Reflections on cognitive dissonance: Thirty years later. Paper presented at the 95th Annual Convention of the American Psychological Association, August, 1987, New York.

Festinger, L. and Carlsmith, J.M. (1959) Cognitive consequences of forced compliance. *Journal of Abnormal and Social Psychology*, 58, 203–210.

Festinger, L., Riecken, H.W. and Schachter, S. (1956) *When Prophecy Fails*. Minneapolis, MN: University of Minnesota Press.

Freedman, J.L. (1965) Long-term behavioral effects of cognitive dissonance. *Journal of Experimental Social Psychology*, 1, 145–155.

Fried, C.B. and Aronson, E. (1995). Hypocrisy, misattribution, and dissonance reduction. *Personality and Social Psychology Bulletin*, 21, 925–933.

Goethals, G.R. and Cooper, J. (1972) Role of intention and postbehavioral consequences in the arousal of cognitive dissonance. *Journal of Personality and Social Psychology*, 23, 292–301.

Goethals, G.R., Cooper, J. and Naficy, A. (1979) Role of foreseen, foreseeable, and unforeseeable behavioral consequences in the arousal of cognitive dissonance. *Journal of Personality and Social Psychology*, 37, 1179–1185.

Harmon-Jones, E. (1999) Toward an understanding of the motivation underlying dissonance effects: Is the production of aversive consequences necessary? In E. Harmon-Jones and J. Mills (eds), *Cognitive Dissonance: Progress on a Pivotal Theory in Social Psychology*, pp. 71–99. Washington, DC: American Psychological Association.

Harmon-Jones, E. and Mills, J. (eds) (1999) *Cognitive Dissonance: Progress on a Pivotal Theory in Social Psychology*, Washington, DC: American Psychological Association.

Harmon-Jones, E., Brehm, J.W., Greenberg, J., Simon, L. and Nelson, D.E. (1996) Evidence that the production of aversive consequences is not necessary to create cognitive dissonance. *Journal of Personality and Social Psychology*, 70, 5–16.

Heider, F. (1946) Attitudes and cognitive organization. *The Journal of Psychology*, 21, 107–112.

Heine, S.J. and Lehman, D.R. (1997) Culture, dissonance, and self-affirmation. *Personality and Social Psychology Bulletin*, 23, 389–400.

Higgins, E.T. (1989) Self-discrepancy theory: what patterns of self-beliefs cause people to suffer? In L. Berkowitz (ed.), *Advances in Experimental Social Psychology*, 22, 93–1360. San Diego, CA: Academic Press.

Hogg, M.A., Haines, S.C. and Mason I. (1998) Identification and leadership in small groups: Salience, frame of reference, and leader stereotypicality effects on leader evaluations. *Journal of Personality and Social Psychology*, 75, 1248–1263.

Hoshino-Browne, E., Zanna, A.S., Spencer, S.J., Zanna, M.P., Kitayama, S. and Lackenbauer, S. (2005) On the cultural guises of cognitive dissonance: The case of Easterners and Westerners. *Journal of Personality and Social Psychology*, 89, 294–310.

Hovland, C.I., Lumsdaine, A.A. and Sheffield, F.D. (1949) *Experiments on Mass Communication (Studies in Social Psychology in World War II*, Vol. 3. Princeton, NJ: Princeton University Press.

Hovland, C.I., Janis, I.L. and Kelley, H.H. (1953) *Communication and Persuasion*. New Haven: Yale University Press.

Kunda, Z. (1990) The case for motivated reasoning. *Psychological Bulletin*, 8, 480–498.

Imada, T. and Kitayama, S. (2010) Social eyes and choice justification: Culture and dissonance revisited. *Social Cognition*, 28, 589–608.

Kitayama, S., Snibbe, A.C., Markus, H.R. and Suzuki, T. (2004) Is there any "free" choice? Self and dissonance in two cultures. *Psychological Science*, 15, 527–533.

Leary, M.R. and Tangney, J.P. (2003) *Handbook of Self and Identity*. New York: Guilford Press.

Lewin, K. (1951) *Field Theory in Social Psychology*. New York: Harper.

Linder, D.E., Cooper, J. and Jones, E.E. (1967) Decision freedom as a determinant of the role of incentive magnitude in attitude change. *Journal of Personality and Social Psychology*, 6, 245–254.

Losch, M.E. and Cacioppo, J.T. (1990) Cognitive dissonance may enhance sympathetic tonus, but attitudes are changed to reduce negative affect rather than arousal. *Journal of Experimental Social Psychology*, 26, 289–304.

Mackie, D.M., Maitner, A.T. and Smith, E.R. (2007) Intergroup emotions theory. In T.D. Nelson (ed.), *Handbook of Prejudice, Stereotyping, and Discrimination*, pp. 285–307. Mahwah, NJ: Erlbaum.

Markus, H. and Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, 98, 224–253.

Menasco, M.B. and Hawkins, D.L. (1978) A field test of the relationship between cognitive dissonance and state anxiety. *Journal of Marketing Research*, 15, 650–655.

Miller, J.G. (1984) Culture and the development of everyday social explanation. *Journal of Personality and Social Psychology*, 46, 961–978.

Monin, B., Norton, M.I., Cooper, J. and Hogg, M.A. (2004) Reacting to an assumed situation vs. conforming to an assumed reaction: The role of perceived speaker attitude in vicarious dissonance. *Group Processes and Intergroup Relations*, 7, 207–220.

Newcomb, T.M. (1956) The prediction of interpersonal attraction. *American Psychologist*, 11, 575–586.

Norton, M.I., Monin, B., Cooper, J. and Hogg, M.A. (2003) Vicarious dissonance: Attitude change from the inconsistency of others, *Journal of Personality and Social Psychology*, 85, 47–62.

Pallak, M.S. and Pittman, T.S. (1972) General motivational effects of dissonance arousal. *Journal of Personality and Social Psychology*, 21, 349–358.

Rosenberg, M. (1965) *Society and the Adolescent Self-image*. Princeton, NJ: Princeton University Press.

Scher, S. and Cooper, J. (1989) The motivational basis of dissonance: The singular role of behavioral consequences. *Journal of Personality and Social Psychology*, 56, 899–906.

Skinner, M.L. (1953) *Science and Human Behavior*. Oxford: Macmillan.

Staw, B.M. (1974) Attitudinal and behavioral consequences of changing a major organizational reward: A natural field experiment. *Journal of Personality and Social Psychology*, 29, 742–751.

Steele, C.M. (1988) The psychology of self-affirmation: sustaining the integrity of the self. In L. Berkowitz (ed.), *Advances in Experimental Social Psychology*, 21, 261–302. San Diego, CA: Academic Press.

Stone, J. (1999) What exactly have I done? The role of self-attribute accessibility in dissonance. In E. Harmon-Jones and J. Mills (eds), *Cognitive Dissonance: Progress on a Pivotal Theory in Social Psychology*, pp. 175–200. Washington, DC: American Psychological Association.

Stone, J. and Cooper, J. (2001) A self-standards model of cognitive dissonance. *Journal of Experimental Social Psychology*, 37, 228–243.

Stone, J. and Cooper, J. (2003) The effect of self-attribute relevance on how self-esteem moderates attitude change in dissonance processes, *Journal of Experimental Social Psychology*, 39, 508–515.

Stone, J., Aronson, E., Crain, A.L., Winslow, M.P. and Fried, C.B. (1994) Inducing hypocrisy as a means of encouraging young adults to use condoms, *Personality and Social Psychology Bulletin*, 20, 116–128.

Tajfel, H. (1970) Experiments in intergroup discrimination. *Scientific American*, 223, 96–102.

Tesser, A. (1988) Toward a self-evaluation maintenance model of social behavior. In L. Berskowitz (ed.), *Advances in Experimental Social Psychology*, 21, 181–227. New York: Academic Press.

Thibodeau, R. and Aronson, E. (1992) Taking a closer look: Reasserting the role of the self-concept in dissonance theory. *Personality and Social Psychology Bulletin*, 18, 591–602.

Turner, J.C. and Hogg, M.A. (1987) *Rediscovering the Social Group*. Oxford: Blackwell.

Waterman, C.K. and Katkin, E.S. (1967) Energizing (dynamogenic) effect of cognitive dissonance on task performance. *Journal of Personality and Social Psychology*, 6, 126–131.

Weaver, K.D. and Cooper, J. (2002) Self-standard accessibility and cognitive dissonance reduction. Paper presented at the Meeting of the Society for Personality and Social Psychology. Palm Springs, CA, January, 2002.

Zanna, M.P. and Cooper, J. (1974) Dissonance and the pill: An attribution approach to studying the arousal properties of dissonance. *Journal of Personality and Social Psychology*, 29, 703–709.

Zanna, M.P. and Sande, G.N. (1987) The effects of collective actions on the attitudes of individual group members: A dissonance analysis. In M.P. Zanna, J.M. Olson, and C.P. Herman (eds), *The Ontario Symposium: Vol. 5. Social Influence*, pp. 151–163. Hillsdale, NJ: Lawrence Erlbaum Associates.

Terror Management Theory

Jeff Greenberg and Jamie Arndt

ABSTRACT

Terror management theory was developed to explain the motivational underpinnings of phenomena such as self-esteem defense and prejudice. The theory is rooted in a long tradition of thought regarding human awareness of death and its role in psychological functioning. The theory posits that to manage the potential for terror engendered by the awareness of mortality, humans sustain faith in worldviews which provide a sense that they are significant beings in an enduring, meaningful world rather than mere material animals fated only to obliteration upon death. The theory is supported by a wide range of studies showing that self-esteem and worldviews provide protection against anxiety and death-related cognition, reminders of mortality instigate worldview bolstering and self-esteem striving, and threats to the worldview and self-esteem increase the accessibility of death-related thought. The research has also led to a dual defense model of responses to conscious and unconscious death thoughts. We then focus on two of many topics informed by the theory: attitudes and behavior regarding physical health, and political preferences and intergroup conflict. We then consider factors that mitigate destructive forms of terror management. Finally, we briefly summarize the contribution of terror management work so far and where it's heading.

TERROR MANAGEMENT THEORY

Terror management theory was born out of a fundamental dissatisfaction with the field of social psychology circa 1980 shared by three graduate students at the University of Kansas: Sheldon Solomon, Tom Pyszczynski, and the senior author of this chapter. Firmly in the grip of social cognition, the field portrayed people as dispassionate information processors guided by schemas and heuristics, operating in a historical, cultural, motivational, and affective vacuum. Raised in working class families characterized by joy and anger, sibling rivalry and love, passion and sarcasm, in neighborhoods centered around churches and synagogues, bars and ballparks, full of people driven by regional, ethnic, and occupational pride and conflict, social psychology at the time struck us as describing some affectless androids the three of us had never met.

In viewing people as highly motivated creatures driven by their needs to protect their self-esteem and assert the superiority of

their own group over other groups, we were like young salmon swimming upstream against the cognitive revolution. We focused our research on how motivation influenced people's perceptions of themselves and others, particularly in the context of defending self-esteem. We contributed to research on self-serving biases and self-handicapping, which showed that people bias their cognitions to protect their self-esteem. But we completed our graduate studies believing that we had no idea what drove people toward their prides and prejudices.

Spread around the country by the whims of the difficult job market (some things never change), the three of us began searching outside the discipline's journals for answers. I recall getting a call from Sheldon, back in 1982, in which he proclaimed that he had found a guy with the answers. It was a deceased cultural anthropologist named Ernest Becker, and the answers were in his book, *The Denial of Death*, which had won the 1973 Pulitzer Prize for general nonfiction. I quickly read the book. I found it frightening, brutal, and revelatory. Building on a wide array of sources, and taking an existential psychoanalytic perspective, the book seemed to explain all the human trends we had observed growing up and which were so well documented by social psychological research: conformity, obedience, self-serving biases, aggression, and prejudice. It explained everything from the rise of Nazism in Germany to the complexities of sex.

Sheldon, Tom, and I began discussing the book's many insights, as well as those provided by his earlier book, *The Birth and Death of Meaning* (1971), and his last book, *Escape from Evil* (1975). We began using Becker's ideas to teach social psychology in a more integrated way, as if the creature being discussed when we covered self-esteem was the same one we discussed when we covered social influence, aggression, prejudice, and close relationships. We were invited by Roy Baumeister to participate in a symposium entitled "Public and Private Self" at the October 1984 Society of Experimental Social

Psychology (SESP) conference in Snowbird, Utah, and to contribute a chapter to an associated edited volume. We decided at SESP to introduce the core of Becker's analysis to our colleagues in social psychology, though Tom couldn't make it. On a snowy morn in front of an unlit fireplace in our room at the Snowbird Ski Lodge, Sheldon and I rather excitedly hacked out a simple summary of the theory, which we there dubbed "terror management theory" (TMT).

The symposium room was moderately full as Sheldon started his presentation, but as he began discussing Marx, Kierkegaard, Freud, and Otto Rank, a good half of the audience jostled their way out of the room. From the back of the room, I spotted a few remaining dignitaries such as John Darley, Ned Jones, who was visibly shaking his head throughout the talk, and our graduate school advisor and one fan, Jack Brehm. At talk's end, instead of the high fives we had envisioned, we received stunned silence, shock, and dismay. I was comforted at least that the practice of stoning had gone out of fashion.

Undeterred by the chilly initial reception, we forged ahead with the chapter, which became the first written presentation of TMT (Greenberg et al., 1986). Simultaneously, we worked on a paper more fully presenting the theory and explaining its potential value as a broad explanatory framework. We were hoping such a paper would be well received at *American Psychologist*, a publication which clamors for broad, integrative ideas. The paper was rejected unceremoniously with two reviews; one a single paragraph, the other a single sentence: "I have no doubt that this paper would be of no interest to any psychologist, living or dead."

Socialized into our profession very well, we did not accept these concise but argument-free bases for rejection. After a year or so of back and forth, editor Leonard Eron explained that the ideas may have had some validity but would not gain currency until backed by empirical research. Actually, up to that moment, we had not thought of TMT as something to study empirically. But it struck

us – duh – that this was precisely what we had been trained to do – to derive testable hypotheses from theories, and then test them.

We believe that TMT explains some fundamental things about human beings and their social behavior. But TMT was not like your standard variety social psychological theory. Most theories in social psychology are mini-theories focused on a specific set of processes pertinent to a particular topic within the field: stereotype threat (prejudice); the elaboration likelihood model (persuasion); the culture of honor (aggression); self-verification theory (self). TMT is about the role of the unconscious fear of death in just about everything we humans do. But soon we were able to derive hypotheses from this broad existential psychodynamic theory and strategies for testing them in collaboration with our graduate students.

Realizing resistance was going to be strong, we put together six studies, a larger set than was customary at that time, before submitting to the *Journal of Personality and Social Psychology*. The paper was accepted (Rosenblatt et al., 1989); the reviews were along the lines of, “This can’t be right, I don’t like it, but I can’t explain their findings, so take it.” We thought at the time, “Fair enough.”

Since then, the empirical support for TMT has blossomed into a literature consisting of over 400 studies and counting, conducted in 16 countries. This body of work, which has included numerous theoretical expansions and refinements, has benefited greatly from the second (and now third) generation of TMT researchers, former students of the original trio including the junior author of this chapter. In the last decade, researchers from independent labs throughout the world have also made valuable contributions to this continually expanding literature. We believe that the growth of this area of study reflects the generative value of a broad theory that integrates widely varied human endeavors by exploring deeply rooted forces that drives humans to behave the way they do. Indeed,

the junior author of this chapter became enthralled as an undergraduate at Skidmore College with the broad existential perspective TMT offered and the promise of subjecting such ideas to empirical scrutiny. In due course, we will provide a brief overview of the literature that has developed over the past 20 years with an emphasis on recent directions pertinent to contemporary concerns. But first we should step back to acknowledge the rather extensive roots of the theory.

THE DISTAL AND PROXIMAL ROOTS OF TMT

Although met with surprise and skepticism when introduced to social psychologists in the 1980s, a decent case can be made that TMT is an ancient theory that can be traced back to one of the first narrative texts from around 3000 bc, the Sumerian Epic of Gilgamesh. A central theme of this story, which influenced the major religions that later emerged out of the Middle East, is the main character’s deep concerns about death and his consequent search for immortality. Gilgamesh, devastated by the death of his friend Enkidu, realizes that he too will die. He roamed the desert, lamenting: “How can I rest, how can I be at peace? Despair is in my heart. What my brother is now, that shall I be when I am dead ... I am afraid of death...” He then embarked on a journey to find immortality.

From that early time on, the notion that we humans fear death and urgently desire to deny or transcend it in some way has been a prominent theme in literature, religious writings, and philosophy. Indeed, Schopenhauer declared death the muse of philosophy; we couldn’t even begin to review its role in philosophical thought here. But we should note that the first person to put the basic points of TMT together seems to have been the famed Greek historian Thucydides.

Around 400 bc, Thucydides was focused on understanding the vicious intergroup

conflicts that plagued Ancient Greece. He proposed that the fear of inevitable death drove people to seek immortality in three ways: through heroic, noble actions restoring justice, which qualified them for a divinely awarded afterlife; through memories of their heroic deeds; and through identification with death transcending group identifications. As Ahrensdorf (2000: 591) put it: "Thucydides contends that they will seek to escape by somehow overcoming their mortal condition, by living on after their death – either through their city, or through their glory, or in an afterlife – and by winning the gods' favor through the vehement affirmation of their own nobility or piety, or justice." Thucydides also noted that the increasing salience of mortality once conflicts begin intensifies the desire to heroically vanquish the enemy.

Jumping forward about 2000 years to the modern English literary tradition, poets from Shakespeare to Wordsworth, Keats, and Shelley, to Dickinson and Emerson, among hundreds of others, have acknowledged the role of the fear of death and desire to escape it in the human psyche. Similarly, novelists such as Swift, Dickens, Dostoevsky, and Tolstoy down to twentieth-century writers such as James Baldwin, Don DeLillo, James Joyce, Phillip Roth, Milan Kundera, and Kurt Vonnegut have explored how the fear of death drives various forms of human behavior. Here, Baldwin concisely captures the gist of TMT:

Life is tragic simply because the earth turns and the sun inexorably rises and sets, and one day, for each of us, the sun will go down for the last, last time. Perhaps the whole root of our trouble, the human trouble, is that we will sacrifice all the beauty of our lives, will imprison ourselves in totems, taboos, crosses, blood sacrifices, steeples, mosques, races, armies, flags, nations, in order to deny the fact of death, which is the only fact we have (James Baldwin, *The Fire Next Time*, 1963).

Awareness of the role of death in the human psyche has not been limited to the great philosophers, poets, and novelists, but can also be found extensively in visual art (e.g., Van Gogh, Klimt), music (e.g., Schubert and

Mahler), and film (e.g., Woody Allen and Ingmar Bergman). And it is not limited to representatives of "high culture," indeed, it seems increasingly prevalent in contemporary pop culture also. As we've been writing this chapter in early 2009, two films, David Fincher's Fitzgerald-inspired *The Curious Case of Benjamin Button*, and Frank Miller's comic-book-inspired *The Spirit*, as well as the second 2009 episode of the sitcom *Scrubs* have focused squarely on the psychological problem of mortality. Recognition of this problem in so many of the products of Western culture makes it all the more interesting that in the 1980s social psychology texts made no mention of death and TMT was treated by most people in the field as outlandish and irrelevant to understanding human social behavior.

As noted earlier, TMT was formalized out of three of Becker's books, *The Birth and Death of Meaning*, *The Denial of Death*, and *Escape from Evil*. These books combine insights from anthropology, evolutionary biology, philosophy, psychoanalysis, and sociology. We urge anyone interested in understanding human behavior to read them.

Around the time Becker was putting his synthesis together, evolutionary philosopher Susanne Langer and psychoanalytic historian Robert Jay Lifton were coming to the same view, and a bit later in 1980, so did existential psychotherapist Irvin Yalom. The primary influences on Becker's analysis were Kierkegaard, William James, Freud, Gregory Zilboorg, Erving Goffman, Norman Brown, and especially Otto Rank. Within psychology, Rank, Freud's disciple and an impressive interdisciplinary scholar back in the 1930s, first granted the fear of death and the urge toward immortality a central role in human culture and social behavior. By systematizing these existential psychodynamic ideas into a coherent explanatory framework, TMT integrates much of what we know about human behavior and provides a fertile basis of generating a wealth of testable hypotheses. The TMT research in turn has inspired useful expansions and refinements of the theory.

THE BASICS OF TMT

TMT starts with two simple observations. The first is that humans are animals with many systems designed to keep them alive, including a fight-or-flight response to impending threats to their continued existence. The second is that the cognitive abilities of humans have led to the awareness that death is inevitable and could come at any time due to a host of potential causes. The theory posits that this knowledge, in an animal built to avert death, creates an ever-present potential for intense anxiety, or terror, which must be managed continuously. The terror is managed by sustaining faith in a view of the world and oneself that denies the precarious and transient nature of one's own existence.

Cultural worldviews have been shaped to serve this function from ancient times to now. Presumably there was a point in time when the cerebral cortex of our species became sufficiently developed to provide self-awareness and the ability to think in terms of past, present, and future. These largely adaptive cognitive capacities also led to the awareness of mortality. Although fear in response to imminent threat is often adaptive, constant anxiety regarding the fragility of continued existence and its inevitable end would not be.

At this point, our ancestors constructed and gravitated toward shared conceptions of reality that most effectively quelled the potential fear that knowledge of our vulnerability and mortality engendered. These cultural worldviews imbue external reality with order, stability, meaning, and purpose, and offer ways in which people can believe they will endure beyond death either literally through an everlasting soul, symbolically through a death transcending identity, or for most worldviews, in both ways.

At the most basic level, all cultural worldviews allow people to live out their lives largely within a conception of reality in which they view themselves as symbolic or spiritual beings who exist in a meaningful world, rather than as mere transient animals

fated only to obliteration upon death. Indeed, the content of consciousness is structured by the cultural worldview in which a given individual is raised. We think in terms of names, dates, months, days of the week, hours, minutes, social roles, and categories. But these are ultimately elaborate window dressing that lends artificial and largely arbitrary structure to an ongoing subjective experience of unique sensations and perceptions that begins at birth and ends at death. Questions like "Who are you?" "Where are you?" "What time is it?" can only be answered with reference to culturally created constructs.

How do people become imbedded in a cultural worldview that provides them with their fundamental psychological security? Developmentally, the human newborn is the most helpless and dependent of all living creatures. It is also a highly distressed creature because, as Rank (1932 /1989) noted, it was conceived and developed attached to its mother and was suddenly, violently separated from its warm home, the womb, at birth. From exposure to the first words that flow out of their parents' mouths, the infant begins to be socialized into the prevailing cultural worldview. In early childhood, the only basis of psychological security for these helpless creatures is the care and love of the parents. The parents provide knowledge, comfort, nutrition, and protection.

As the parents begin to require particular actions from the child to sustain their love and approval, the child learns that when she does the right things, all is good in the world but when she does the wrong things, all hell could break lose. To sustain the psychological security of the parental love, the child internalizes the parental values of good and bad, and tries to live up to the good, despite the infringement on the child's natural desires and urges this entails. For the good little five-year-old sitting on mom's lap, all is right in the world. For the bad five-year-old who just punctured one of dad's new stereo speakers with a flying action figure, the world has just become a scary place. In this way, the child comes to associate being good and valued

with feeling secure, and being bad and worthless with dread. Self-esteem, that feeling of being valuable, thus comes to serve an anxiety-buffering function.

The plot thickens as the child's cognitive development leads her to realize that the things she fears – the dark, monsters, ghosts, big dogs – threaten her very existence and that the parents are limited and can't always protect them from everything. Concomitant with these dawning realizations, the child's basis of security gradually shifts to larger more powerful sources, deities, presidents, and the culture in general. This transference process is facilitated by the fact that the parents have been inculcating the child into the larger worldview, including the parent's own bases of psychological security, throughout the child's upbringing.

So over childhood, psychological security becomes predicated on being a good Christian, American, and so forth, on being valuable within the context of the individual's internalized version of the cultural worldview. This sense of enduring significance is not only linked implicitly with safety and security but explicitly with the literal immortality of heaven and the symbolic immortality of being part of death-transcending entities such as the family line and the nation, and lasting cultural achievements in science, politics, and the arts. In these ways, the significant person in a society can feel like he or she is an eternal part of an enduring reality.

Summary of TMT and its basic implications

In sum, the simple formula for effective terror management is faith in a meaning providing cultural worldview and the belief that one is a valuable contributor to that meaningful world (the TMT conceptualization of self-esteem). We initially derived two basic implications from the theory (Solomon et al., 1991). First, self-esteem and the worldview upon which it is predicated serve a critical anxiety-buffering function. So people strive

for and defend self-esteem in the service of psychological security and these efforts depend on the specific cultural worldview to which the individual subscribes. Second, because these constructs are ultimately fragile social constructions, people will react negatively to anyone or anything that undermines faith in their worldview or their self-worth. We believe that this provides a very basic insight into prejudice and intergroup conflict. People who criticize one's own worldview, or simply adhere to a worldview very different from one's own, call into question the validity of one's own basis of psychological security.

TMT therefore posited that such different people are inherently threatening and are reacted to with four defenses. The first, most widespread, is derogation. If these others are ignorant or evil, then their alternative beliefs can be dismissed. The second is assimilation. If these people are wrong, then I can help them see the light, which will make me all the more certain my worldview is the right one. Missionary activity provides a vivid example of this strategy. The third is accommodation, which involves incorporating appealing aspects of the alternative worldview into one's own. In this way, one can preserve faith in the central components of one's own worldview. Rock and rap music initially expressing anger and challenging mainstream American culture lose their threatening nature when converted to muzak or used to sell fast food burgers. Finally there is annihilation. If one's worldview is right and the others' is wrong, then one's own should prevail.

The core tests of hypotheses derived from TMT

TMT is consistent with a wide range of evidence from anthropology, archeology, and history, but our challenge was to deduce predictions that put the theory to the test. The central TMT hypotheses are based on its account of the psychological function of cultural worldviews and self-esteem. If these structures provide

protection against death-related fears, then: (a) reminding people of their mortality should increase bolstering of the worldview and striving for self-worth; (b) bolstering these structures should reduce anxiety in response to threat and defensive reactions to reminders of death; and (c) threatening the structures should arouse anxiety and bring death-related concerns closer to consciousness.

Hundreds of studies have supported these hypotheses. First, let's consider hypothesis (a). Reminding people of their mortality (mortality salience or MS) increases positive reactions to people who validate aspects of participant's worldviews and negative reactions to people who challenge aspects of their worldview or espouse a different one (see Greenberg et al., 2008 for a review). For example, MS leads to negative evaluations of Jews by Christians, of a critic of the US by Americans, and of foreign products by Germans. MS even leads to increased aggression against a critic of one's preferred political orientation, whether liberal or conservative. On the other hand, MS increases positive reactions to heroes and celebrities, one's national soccer team, and members of one's own religious group, and leads to more donations to valued charities. MS also increases striving for self-worth (see Greenberg et al., 2008). In people who value such things, MS increases bold driving, displays of physical strength, proenvironmental intentions, focus on one's appearance, interest in self-esteem enhancing dating partners, and ratings of the importance of and desire for fame and wealth.

A substantial number of studies have also supported hypothesis (b). Bolstering self-esteem reduces anxiety in response to threat (Greenberg et al., 1992b) and defensive reactions and death-thought accessibility following reminders of mortality (Harmon-Jones et al., 1997). Similarly, bolstering or defending one's worldview also reduces defensive reactions and death-thought accessibility following reminders of mortality (e.g., Arndt et al., 1997).

Recent research has also supported Hypothesis (c). Reminders that humans are

animals, criticism of participants' worldviews, and threats to participants' self-esteem all increase the accessibility of death-related thoughts and not the accessibility of other negative thoughts (e.g., Friedman and Rholes, 2007; Schimel et al., 2007).

This body of research has operationalized reminding people of their mortality in a variety of ways and has extensively compared reminders of mortality with reminders of other negative thoughts (see Greenberg et al., 2008). People have been reminded of death via two simple questionnaire items, writing one sentence about death, fear of death scales, proximity to cemeteries and funeral homes, and subliminal death primes. Thinking about mortality has been compared to thinking about being paralyzed, failing, feeling uncertain, general anxieties, giving a speech in public, intense pain, dental pain, disease, unpredictable, intense bouts of pain, life being meaningless, being socially excluded, upcoming exams, unexpected events, and worries after college.

Although some studies have found similar effects of other aversive thoughts in circumscribed contexts (e.g., McGregor et al., 2001; van den Bos, 2001) the vast majority have found quite different effects for reminders of mortality. TMT theorists (e.g., Greenberg et al., 2008) have argued that this is because death is the only inevitable future event, it is what human biological systems, including the fight or flight system, are most focused on avoiding, and it threatens to undermine all human desires, whether for belonging, cognition, control, or growth. One possibility is that when these other threats lead to defenses similar to those instigated by MS, they may undermine structures that serve terror management, thereby bringing thoughts of death closer to focal awareness. For example, Landau et al. (2004a) found that threatening individuals' belief in a just world led to elevated death-thought accessibility. These other aversive experiences may also, under some conditions, be threatening in their own right. Given the clear evidence that thoughts of mortality so often arouse different responses

than these other aversive cognitions, an important direction for future inquiry research is understanding when psychological defenses are serving terror management and when they are serving other concerns.

A theoretical refinement: the dual process model

After the Rosenblatt et al. (1989) paper was published, a German psychologist, Randolph Ochsmann, reported having difficulty replicating MS effects. In contrast to two brief items, Ochsmann was reminding people of death by leading them through an extensive guided imagery exercise of imagining their own death and interment. This discrepancy led to an important insight into the processes instigated by MS. Realizing that the initial terror management studies invariably included some kind of intervening experience between the MS manipulation and assessment of the dependent measures (e.g., mood scales, experimenter instructions) we posited that terror management defenses do not occur when death is prominent in consciousness but rather when death is highly accessible but no longer in focal attention.

Studies (see Pyszczynski et al., 1999 for a review) initially supported this idea by showing, for example, that MS only leads to worldview defense if there is a non-death-related task between the MS induction and the dependent measure. They also showed that immediately after an MS induction, the accessibility of death-related thought is low. After a delay, in contrast, death-related thought is no longer in focal attention but becomes high in accessibility. Given people's reluctance to consciously dwell on their own mortality, we hypothesized that, confronted with such thoughts, people initially suppress them. A series of studies supported this reasoning, finding for example that MS led to immediate increases in death-thought accessibility if participants were placed under high cognitive load. Another series of studies further clarified that role of consciousness in MS-induced worldview defenses by

showing that worldview defenses emerge in response to unconscious thoughts of death (e.g., when people are exposed to subliminal reminders of mortality).

A dual process model of terror management emerged from this line of research. Explicit thoughts of death instigate direct proximal defenses to remove death-related thoughts from current focal attention. Such "pseudo-rational" mechanisms make death seem like a far-off problem, thereby allowing the individual to stop thinking about it. However, after thoughts of death have been removed from focal attention, death-thought accessibility increases, heightening the potential to experience death-related anxiety. This in turn instigates symbolic distal defenses such as bolstering one's worldview or self-worth. These terror management defenses then bring death thought accessibility down to baseline levels.

Summary

TMT was originally developed to explain two facts of human experience, namely that people have difficulty getting along with those who are different and that people have a trenchant need to feel good about themselves. Research shows that both these propensities are strands in the fabric by which people insulate themselves from deeply rooted fears of mortality. Further, studies have supported a dual process model of the defenses instigated by reminders of mortality. Based on this foundation, the theory has guided investigation of a variety of aspects of human experience, many more so than were envisioned on that snowy day in Utah.

THE GRAVE MATTERS: APPLYING TMT TO CONTEMPORARY HUMAN CONCERNs

TMT is relevant to a wide range of domains of human behavior. Although we have been accused on occasion of trying to explain

everything with TMT, our position is not that everything stems from the need for terror management, but that most, if not all, significant domains of human behavior are influenced by terror management concerns. Indeed, the theory has provided insights into a wide range of phenomena, including human ambivalence about sex (e.g., Goldenberg et al., 2000), reactions to the handicapped (e.g., Hirschberger, 2006), the operation of stereotypes (e.g., Schimel et al., 1999), academic performance (Landau et al., 2009), altruism (e.g., Jonas et al., 2002), parenthood (e.g., Wisman and Goldenberg, 2005), attachment and close relationships (e.g., Mikulincer et al., 2003), stigmatization (e.g., Salzman, 2001), reactions to nature (e.g., Koole and van den Berg, 2005), attitudes toward women (e.g., Landau et al., 2006a), religion (Greenberg et al., in press), art (e.g., Landau et al., 2006b), film (Sullivan et al., 2010), and even human–android relations (see e.g., MacDorman, 2005).

The theory has also generated some useful insights, following Kurt Lewin, into affairs of particular practical significance, including legal affairs (Arndt et al., 2005a), consumer behavior (Arndt et al., 2004), and mental health (Arndt et al., 2005b). Given space limitations, we will focus on two heavily researched applications of TMT. The first concerns the decisions that people make with regard to the health, and the second concerns the political sphere and our seemingly escalating proclivities for intergroup violence. We will then conclude by considering factors that mitigate the more harmful defensive responses provoked by concerns about mortality.

Understanding the role of awareness of death in everyday health decisions

It seems rather obvious that death does not bode well for an individual's health. Despite this fact, theory and research had largely ignored the deeper psychological significance

of concerns about mortality for understanding the decisions people make regarding their physical health. This is surprising given that many health campaigns remind individuals that noncompliance with health recommendations can hasten their demise. What are the possible consequences of such reminders? Goldenberg and Arndt (2008) recently extended TMT with a terror management health model (TMHM) to address such questions.

The TMHM starts with the idea that health-relevant scenarios have varying potential to activate death-related cognition, and then builds from the proximal and distal terror management defenses previously described. When a health situation leads people to explicitly think about death, or thoughts of death are otherwise in focal attention, health decisions are influenced by a proximal motivation to jettison the threatening cognitions from focal attention. For example, an individual may take proactive (e.g., exercise more) or avoidant (e.g., deny perceived risk) steps to render death less of an immediate concern. When thoughts of death are activated implicitly (i.e., outside of focal attention), health-relevant decisions are driven more by a desire to affirm the symbolic value of the self; for example, by pursuing the standards upon which an individual's self-esteem is based, investing in worldview beliefs, or distancing oneself from the creatureliness of the body. Responses to both conscious and unconscious death-related thought can be beneficial or harmful to health, but responses to unconscious death thoughts will be moderated by the individual's worldview and bases of self-worth, whereas responses to conscious death thoughts will be moderated by the perceived efficacy of efforts directed toward reducing the threat of death (see Goldenberg and Arndt, 2008, for a review of the supportive research).

In one study illustrating this basic distinction, participants were reminded of their mortality (or not) and then immediately thereafter or after a delay answered questions

about their intentions to tan. Immediately after thinking about death, participants reported lower intentions to suntan, suggesting efforts to reduce vulnerability and protect health. However, after a delay (when thoughts of death were accessible but not conscious) participants increased their intentions to suntan, suggesting an effort to bolster esteem by enhancing attractiveness as defined in contemporary society. Thus, conscious thoughts of death led to a health over beauty decision, but unconscious thoughts of death led to a beauty over health decision.

In drawing attention to these distinctions, TMHM helps explain a number of consequences of fear appeals, such as a so-called “boomerang” effect whereby fear appeals (which often highlight the threat of death) can have consequences antithetical to those intended. Whereas the initial response to a fear appeal may reflect efforts to protect one’s physical self (which then remove death from consciousness), the delayed effect of unconscious thoughts of death may instigate self-esteem striving that entails more risky behavior.

The TMHM focuses on three general propositions. The first is that people’s health decisions will reflect efforts to remove conscious thoughts of death from focal attention. Accordingly, research has found that when confronted with explicit thoughts of death, people react by trying to avoid the threat or by trying to reduce its likelihood with health-promoting behaviors (such as increasing exercise). Of course, the critical question is when and for whom a threat avoidance or health-promoting response is most likely to emerge. When an individual perceives a health response as effective, maintains optimism, approaches health situations with active coping strategies, or construes vulnerability as low, they respond to conscious thoughts of death with health-oriented rather than threat-avoidance strategies, particularly in death-related health domains (e.g., screenings for breast and skin cancer).

The second direction of research has explored how the unconscious resonance of

death-related cognition promotes maintaining, not one’s health, but a sense of meaning and self-esteem. For example, thoughts of death outside of focal attention increase both health-defeating (e.g., intentions to suntan) and health-facilitating (e.g., exercising) outcomes depending on situational and dispositional self-esteem contingencies (see Goldenberg and Arndt, 2008). As this suggests, MS can lead to seemingly counterintuitive risky effects on health-relevant outcomes, both for self and others. In one study, after being reminded of mortality, Christian medical students were more likely to adopt a cautious triage approach to Christian patients complaining of chest pain but adopted a more cavalier approach to Muslim patients reporting identical symptoms (Arndt et al., 2009). This suggests that even people who are more routinely exposed to issues concerning death may still be vulnerable to terror management effects. But why would individuals make such decisions? Such responses help to bolster the symbolic buffers that offer protection from deeply rooted existential fear. Consistent with this analysis, Vess et al. (2009) found that not only did religious fundamentalists respond to reminders of death by endorsing faith-based over medical-based treatments for illness (both for themselves and others), but doing so helped to satiate their need for meaning in life.

This does not imply, however, that unconscious thoughts of death necessarily increase health-risk behavior. If individuals derive self-esteem from behaviors conducive to health, they display health-facilitating outcomes across different health domains when death thought accessibility is high. Moreover, in a number of areas (e.g., exercise, smoking cessation, tanning, breast cancer screening), when self-esteem contingencies are experimentally altered in health-conducive ways, people respond to increased death thought accessibility with more protective intentions and behavior. For example, shifting normative appearance standards toward pale skin led actual beachgoers, when distracted from

activated thoughts of death, to request higher SPF sun lotions when offered compensatory samples, especially among those who derive self-esteem from extrinsically defined standards (e.g., Arndt et al., 2009). Indeed, following MS, showing those who smoke to fit in with others a commercial about how smoking reduces one's popularity leads to increased smoking cessation intentions.

The third direction of TMHM research builds from Goldenberg and colleagues' (2000) work on how reminders of our physical, creaturely nature intensify terror management concerns. Confrontations with the physical body threaten the illusion of symbolic significance, and thus our psychological security. Research has shown both reminders of death and of the creaturely aspects of the body influence health-related attitudes and behavior. For example, they increase negative reactions to breastfeeding and pregnant women, increase perceptions of discomfort regarding mammograms, and reduce the thoroughness of breast self-exams. In sum, TMHM research has shown that the awareness of mortality plays a significant role in attitudes and behavior pertinent to cancer detection and prevention as well as to more general healthful living.

TMT, politics, and intergroup conflict

The awareness of mortality also contributes to intergroup strife, in two primary ways. First, as Becker (1971) noted, the mere presence of groups with worldviews very different from one's own threatens individuals' terror-assuaging faith in their worldview. We have reviewed research showing that reminders of mortality lead to derogation and even aggression against those who criticize the participant's worldview or simply subscribe to a different one.

However, in a chapter called "the Nexus of Unfreedom" from *The Denial of Death*, Becker (1973) elucidated an even more destructive consequence of the fear of death.

He argued that the most blood has been spilled not by those with evil motives but by those who are serving the great leader, God, and country. Our mortality requires us to find something bigger than ourselves for our salvation. When that individual or institution designates some other group as evil or that group is perceived as a threat to the terror-assuaging leader, deity, or entity, violent attempts to eradicate that evil threat commonly result.

In his final book, Becker (1975) also noted that no matter what our worldview is, residual death anxiety will exist and people will seek controllable sources of that anxiety to mask its true cause – scapegoats. So the greatest sense of death transcendence is provided by participating in a heroic triumph over evil. In other words, concerns about mortality will draw people toward worldviews and leaders who provide them most compellingly with a sense that they are part of something great and have a mission to heroically triumph over those who are evil; rigid worldviews and charismatic leaders that clearly delineate who is good and who is evil will do this best. Becker used this analysis to explain something that otherwise seems inexplicable: the rise of Nazism in Germany culminating in the Holocaust (as well as other similar historical phenomena). Indeed, much of the course of human history seems to have been guided by efforts to heroically triumph over evil in the service of the death-transcending ideologies and leaders in which people have invested.

TMT research has supported these ideas in a number of ways. Reminders of mortality draw people toward charismatic leaders and ideologies that sell a simple vision of the greatness of the ingroup and the need to rid the world of evildoers. Using descriptions of hypothetical gubernatorial candidates, Cohen et al. (2004) found that MS increased the appeal of a candidate with a charismatic style; that is, one who exuded self-confidence and emphasized the greatness of the state and nation. As the 2004 American presidential election approached, Americans were faced

with a choice that would allow a unique test of the TMT analysis of attraction to leaders. One was George W. Bush, a folksy, at-ease incumbent president who emphasized the greatness of America and the need to rid the world of evildoers, such as the “Axis of Evil.” The other was John Kerry, a Democratic challenger with a stiff speaking style who had a complex, sometimes hard to decipher view of issues and who was portrayed by the Republican spin machine as a waffling flip-flopper. Landau et al. (2004b) posited that MS, as well as reminders of terrorism by virtue of activating fears of death, would increase the appeal of Bush and reduce the appeal over Kerry. And in a series of studies prior to the 2004 election, that’s exactly what they found. Subsequently, with the help of a Bin Laden tape released days before the election, Bush won re-election even though he had invaded a country and embroiled the US in a protracted war based on false claims.

The Cohen study suggests that the Bush effect resulted from Bush’s charismatic style and simple good-versus-evil worldview. However, Bush could have also had an advantage for terror management purposes as the current leader of the nation. In addition, Jost and colleagues (2003) have suggested that conservative rightwing ideologies may serve terror management better than liberal or leftwing ones. However, in the Landau et al. (2004b) studies, despite increasing the preference for Bush, MS did not increase participants’ self-reported political conservatism. In addition, a recent study by Kosloff et al. (2010) building on Cohen et al. found that MS leads people to prefer a charismatic hypothetical leader only if that leader also espouses policies that match the individual’s pre-existing political orientation, whether conservative or liberal.

So MS seems to move people toward those who espouse more straightforward good-versus-evil ideologies as long as those ideologies fit with the individual’s pre-existing worldview. Based on these findings, Pyszczynski et al. (2006) wondered if MS would also increase the appeal of engaging in

violence against those designated by one’s culture as evil. In their first study, they exposed Iranian college students to MS or a control topic, and then asked them to react to a fellow student who espoused either suicide bombing against Americans or peaceful approaches to addressing issues with the US. In the control condition, the Iranian students preferred the peaceful fellow student. However, after MS, they preferred the supporter of suicide bombing, and expressed increased interest in joining the cause themselves.

Lest we’re tempted to confine such reactions to Iranians, Routledge and Arndt (2008) similarly found that British students after MS reported an increased willingness to die for mother England. Further, people don’t just show a proclivity to blow themselves up after MS, but even more so to simply blow up the enemy. Indeed, in a second study, Pyszczynski et al. (2006) manipulated MS and asked conservative and liberal Americans how much they supported extreme military actions, including the use of nuclear weapons, against perceived threats to the US in the Middle East and elsewhere. Compared with the control condition, MS increased conservative participants’ advocacy of such violent measures. Similarly, MS increased violent reactions among Israelis upset with the then upcoming 2005 pullout from the Gaza Strip and the northern West Bank (Hirschberger and Ein-Dor, 2006). Finally, supporting the terror management value of death to the designated evil other, Hayes et al. (2008) showed that after threatening Christians’ worldviews, informing them that 117 Muslims died in a plane crash actually reduced the accessibility of death-related thought and worldview defense.

Taken together, the TMT literature suggests that the threat of mortality increases negativity toward different others and fuels support for simple “we are good, they are evil” worldviews and desires to rid the world of those designated as evil. Thus, research supports the role of terror management needs in intergroup conflict around the world.

WHERE DO WE GO FROM HERE? INSIGHTS INTO BETTER WAYS TO MANAGE TERROR

Given the often destructive means that people use to manage the fear of mortality, how can we reduce such effects? Fortunately, a number of lines of research have addressed this question and suggest ways such responses can be reduced or even reversed. According to Becker, the key is to identify secure, enduring worldviews that provide widely accessible avenues for self-worth to their constituents, but do so in a way that minimizes the costs both to those inside and outside the culture. These insights inspired two broad approaches to mitigating harmful terror management that have been examined empirically.

Fortifying the shield: psychological buffers against thoughts of mortality

The earliest statements of TMT posited that, because self-esteem serves to buffer anxiety, boosting self-esteem or having high dispositional self-esteem should render people less vulnerability to the consequences of awareness of mortality. Harmon-Jones et al. (1997) and others have supported the specific hypothesis that bolstering an individual's self-worth would enable the person to contemplate mortality without experiencing subsequent increases in death thought accessibility and associated defensive responses. As long as the threat that other people pose does not undermine their own feelings of self-worth (Arndt and Greenberg, 1999), self-esteem is capable of offering people psychological protection against responding defensively to reminders of death.

People can also derive existential protection through their attachments with close others. Rank (1941/1958) originally proposed that interpersonal relationships have become an especially important security blanket as dominant religious doctrines have lost some of their luster. In accord with this insight,

Mikulincer and others have shown that people who have secure attachments to close others when reminded of death are less likely to respond defensively and advocate violence (e.g., Mikulincer et al., 2003; Weise et al., 2008).

While both relational attachment and self-esteem (and other buffers we don't have the space to touch on here) reduce consequences of MS, these critical bases of psychological security are often hard to sustain over the lifespan. When they falter, problems with anxiety, depression, and substance abuse are likely. Thus, the more durable the bases of self-worth and attachments, the less destructive an individual's efforts at terror management are likely to be.

Worldviews espousing tolerance, compassion, and open-mindedness

Another tack for mitigating harmful forms of terror management focuses not on fortifying the existential shield, but on funneling the routes that responses take in more constructive directions. This can be achieved by appealing to worldviews that espouse tolerance and open-mindedness. If an individual values tolerance and MS motivates effort to live up to important values, then MS should increase tolerance. Accordingly, Greenberg and colleagues (1992a) showed that when tolerance was primed or dispositionally important, participants did not respond to MS with increased derogation of those who threatened their beliefs. Such findings of course are hopeful, as people can socialize their children into valuing acceptance of others.

Many components of culture contain such prescriptions, though the message often gets lost. For instance, could religiously advocated compassion encourage more peaceful coexistence with those who are different? To find out, Rothschild et al. (2009) investigated how priming compassionate religious values might redirect the effects of MS on support for violent worldview defense among religious fundamentalists. In these studies, using

both American samples as well as a Shia population in Iran, participants were reminded of death or a control topic, and then were presented with either compassionate or neutral religious scripture, or compassionate or neutral secular adages, followed by a measure of support for extreme military responses (against those who challenge the US or toward the US in the case of the Shia). Religious fundamentalists in the control condition were supportive of violent worldview defense regardless of type of teaching. However, fundamentalists reminded of death and exposed to compassionate religious teachings demonstrated a significant decrease in their support for violence. Presumably the death reminder motivated an increased need to live up to compassionate values when they were linked to the worldview.

Can tolerance and compassion be accentuated without explicit priming? One possibility involves people reconceptualizing ingroup identifications to recognize our common humanity with people the world over. For example, Motyl and colleagues (submitted) reminded American participants of their death or a control topic, and then exposed them to pictures of common non-American families from diverse regions of the world, stereotypically American families, or a group of unrelated people. In the neutral and American families conditions, MS elicited an increase in anti-Arab prejudice; however, when reminded of the common humanity shared by all people, MS significantly reduced anti-Arab prejudice. In a second study, Americans were reminded either of death or a control topic, and then presented with brief stories of common childhood experiences ostensibly written by either Americans or people from diverse areas of the world. When presented with American childhood experiences, MS increased hostile prejudice against immigrants; however, when those same childhood experiences were attributed to a foreign author, this negative effect was eliminated.

Encouraging people to consider our shared humanity is a promising avenue to reducing destructive effects, particularly given the

increasing globalization of the world. Such goals may also be facilitated by more generally stimulating open-minded and flexible thinking. By virtue of its utilization of open-minded thinking, creative action may be one way to encourage more acceptance of different others.

The initial research examining creativity and terror management (see Greenberg et al., 2008, for a review) built from the theorizing of Otto Rank (e.g., 1932). Creative engagement, by separating the individual from conventional thinking, threatens the person's valued place within the security-providing worldview. Therefore, when reminded of mortality, creative action can lead to feelings of guilt (an emotion that reflects a desire for social reparation). But fortified with an enhanced sense of such social connection, people can engage in creativity after being reminded of death without experiencing guilt, and can then reap the more positive psychological effects of such activity. Accordingly, after being reminded of death, people are more creative if the product is directed toward communal benefit, but less creative if the product is directed toward individual gain.

This work also suggests that as people are faced with managing existential fears, creativity has the potential to facilitate more optimal engagement with life (Routledge and Arndt, 2009). One study found that socially validated creativity inspires a more open-minded orientation which reduces people's tendency to manage terror by derogating those with conflicting beliefs. Another set of studies showed that priming the cultural value of creativity after MS can even increase willingness to expose oneself to ideas that run counter to prevailing cultural beliefs. This and other research suggests that people can manage the awareness of death in ways that actually embrace the rich diversity of perspectives that the world offers, the possibility of authentic self-transcendence, and intrinsic goal pursuit (see, for example, Lykins et al., 2007).

Finally, individual differences in need for structure may play a role in how people cope

with their mortality concerns. Vess, Routledge et al. (2009) showed in a series of studies that while those high in need for structure prefer more rigid routes to epistemic knowledge (as in the work of Landau et al., 2004a, 2006a), those low in need for structure may utilize more explorative forms of discovery and integrative processing to extract meaning and significance in their life. Thus, again, much of the potential for reducing destructive forms of terror management seems to go back to the content of the worldview to which an individual subscribes. When one's worldview prescribes prosocial behavior, flexible thinking, or tolerance and compassion, constructive responses to the human existential predicament are likely.

LONG DAY'S JOURNEY INTO NIGHT

TMT has brought the problem of death into social psychology, and we hope that you agree that has been a valuable (albeit not pleasant to think about) contribution. The threat of death lurks behind most, if not all, the things people care about: health, economic wellbeing, the environment, terrorism, war, close relationships, aging, and spirituality. TMT work has helped clarify how the awareness of mortality contributes to attitudes and behavior regarding these topics and many more as well. In our defiance of our mortality, we all urgently defend our cherished beliefs and strive hard to make the greatest mark on the world we can. TMT research has shown that these urgings contribute to the most noble and ignoble forms of human behavior.

TMT work has also raised many questions currently being considered. How do defenses aroused by conscious and unconscious concerns relate to each other? What does TMT work imply about the basic nature of the unconscious? What brain regions contribute to the awareness of mortality and the defenses it instigates? What are the effects of sustained heightened mortality salience over

time; for example, for professionals such as oncologists and morticians? One study showed that workers who perform death ceremonies in India had stronger pro-India bias than other workers, independent of an MS induction, whereas the other workers were as strongly pro-India as the death ceremony workers only if they were first induced to consider their mortality (Fernandez et al., 2010). But much more research on this issue is needed. Are there particular components of cultural worldviews that are particularly important for terror management? Are there certain beliefs and values that are common to all worldviews because of the way evolutionary has shaped our basic motivations, ways of thinking, and emotions (e.g., empathy, desire for justice)? Are there other motivations besides terror management which contribute to the desires to preserve faith in a meaningful world and one's own significance?

We also like to think that TMT has contributed to a general broadening of the field, a greater awareness of the important roles of culture, of unconscious motivations and processes, of emotions, and of core human concerns in social behavior. One aspect of this has been the emergence of the subfield within social psychology known as experimental existential psychology (XXP; Greenberg et al., 2004; Pyszczynski et al., 2010). This subfield focuses on the impact on thought, feeling, and behavior of five major existential concerns: death, meaning, identity, isolation, and freedom. We are hopeful that XXP will continue to progress toward a comprehensive understanding of these core facets of the human condition and their roles in the social phenomena that matter to us most; such a prospect gives us a sense of fundamental satisfaction.

REFERENCES

Ahrensdorf, P.J. (2000) The fear of death and the longing for immortality: Hobbes and Thucydides on

human nature and the problem of anarchy. *The American Political Science Review*, 94, 579–593.

Arndt, J., Cox, C.R., Goldenberg, J.L., Vess, M., Routledge, C. and Cohen, F. (2009) Blowing in the (social) wind: Implications of extrinsic esteem contingencies for terror management and health. *Journal of Personality and Social Psychology*, 96, 1191–1205.

Arndt, J. and Greenberg, J. (1999) The effects of a self-esteem boost and mortality salience on responses to boost relevant and irrelevant worldview threats. *Personality and Social Psychological Bulletin*, 25, 1331–1341.

Arndt, J., Greenberg, J., Solomon, S., Pyszczynski, T. and Simon, L. (1997) Suppression, accessibility of death-related thoughts, and cultural worldview defense: Exploring the psychodynamics of terror management. *Journal of Personality and Social Psychology*, 73, 5–18.

Arndt, J., Lieberman, J.D., Cook, A. and Solomon, S. (2005a) Terror management in the courtroom: Exploring the effects of mortality salience on legal decision-making. *Psychology, Public Policy, and Law*, 11, 407–438.

Arndt, J., Routledge, C., Cox, C.R. and Goldenberg, J.L. (2005b) The worm at the core: A terror management perspective on the roots of psychological dysfunction. *Applied and Preventative Psychology*, 11, 191–213.

Arndt, J., Solomon, S., Kasser, T. and Sheldon, K.M. (2004) The urge to splurge: A terror management account of materialism and consumer behavior. *Journal of Consumer Psychology*, 14, 198–212.

Arndt, J., Vess, M., Cox, C.R., Goldenberg, J.L. and Lagle, S. (2009) The psychosocial effect of personal thoughts of mortality on cardiac risk assessment. *Medical Decision Making*, 29, 175–181.

Baldwin, J. (1963) *The Fire Next Time*. New York: Random House.

Becker, E. (1971) *The Birth and Death of Meaning*, 2nd Edition. New York: Free Press.

Becker, E. (1973) *The Denial of Death*. New York: Academic Press.

Becker, E. (1975) *Escape from Evil*. New York: Academic Press.

Cohen, F., Solomon, S., Maxfield, M., Pyszczynski, T. and Greenberg, J. (2004) Fatal attraction: The effects of mortality salience on evaluations of charismatic, task-oriented, and relationship-oriented leaders. *Psychological Science*, 15, 846–851.

Fernandez, S., Castano, E. and Singh, I. (2010) Managing death in the burning grounds of Varanasi, India: A terror management investigation. *Journal of Cross-Cultural Psychology*, 41, 182–194.

Friedman, M. and Rholes, W.S. (2007) Successfully challenging fundamentalist beliefs results in increased death awareness. *Journal of Experimental Social Psychology*, 43, 794–801.

Goldenberg, J.L. and Arndt, J. (2008) The implications of death for health: A terror management health model for behavioral health promotion. *Psychological Review*, 115, 1032–1053.

Goldenberg, J.L., McCoy, S.K., Pyszczynski, T., Greenberg, J. and Solomon, S. (2000) The body as a source of self-esteem: The effects of mortality salience on identification with one's body, interest in sex, and appearance monitoring. *Journal of Personality and Social Psychology*, 79, 118–130.

Greenberg, J., Koole, S. and Pyszczynski, T. (eds) (2004) *Handbook of Experimental Existential Psychology*. New York: Guilford Press.

Greenberg, J., Landau, M.J., Solomon, S. and Pyszczynski, T. (in press) The case for terror management as the primary psychological function of religion. In D. Wulff (ed.), *Handbook of the Psychology of Religion*. London: Oxford University Press.

Greenberg, J., Pyszczynski, T. and Solomon, S. (1986) The causes and consequences of a need for self-esteem: A terror management theory. In R.F. Baumeister (ed.), *Public Self and Private Self*, pp. 189–212. New York: Springer-Verlag.

Greenberg, J., Simon, L., Pyszczynski, T., Solomon, S. and Chatel, D. (1992a) Terror management and tolerance: Does mortality salience always intensify negative reactions to others who threaten one's worldview? *Journal of Personality and Social Psychology*, 63, 212–220.

Greenberg, J., Solomon, S. and Arndt, J. (2008) A uniquely human motivation: Terror management. In J. Shah and W. Gardner (eds), *Handbook of Motivation Science*, pp. 113–134. New York: Guilford.

Greenberg, J., Solomon, S., Pyszczynski, T., Rosenblatt, A., Burling, J., Lyon, D., Pinel, E. and Simon, L. (1992b) Assessing the terror management analysis of self-esteem: Converging evidence of an anxiety-buffering function. *Journal of Personality and Social Psychology*, 63, 913–922.

Harmon-Jones, E., Simon, L., Greenberg, J., Pyszczynski, T., Solomon, S. and McGregor, H. (1997) Terror management theory and self-esteem: Evidence that increased self-esteem reduces mortality salience effects. *Journal of Personality and Social Psychology*, 72, 24–36.

Hayes, J., Schimel, J. and Williams, T.J. (2008) Fighting death with death: The buffering effects of learning that worldview violators have died after worldview threat. *Psychological Science*, 19, 501–507.

Hirschberger, G. (2006) Terror management and attributions of blame to innocent victims: Reconciling compassionate and defensive responses. *Journal of Personality and Social Psychology*, 91, 832–844.

Hirschberger, G. and Ein-Dor, T. (2006) Defenders of a lost cause: Terror management and violent resistance to the disengagement plan. *Personality and Social Psychology Bulletin*, 32, 761–769.

Jonas, E., Schimel, J., Greenberg, J. and Pyszczynski, T. (2002) The Scrooge effect: Evidence that mortality salience increases prosocial attitudes and behavior. *Personality and Social Psychology Bulletin*, 28, 1342–1353.

Jost, J.T., Glaser, J., Kruglanski, A.W. and Sulloway, F. (2003) Political conservatism as motivated social cognition. *Psychological Bulletin*, 129, 339–375.

Koole, S.L. and van den Berg, A.E. (2005). Lost in the wilderness: Terror management, action orientation, and nature evaluation. *Journal of Personality and Social Psychology*, 88, 1014–1028.

Kosloff, S., Greenberg, J., Weise, D. and Solomon, S. (2010) The effects of reminders of death on political preferences: The roles of charisma and political orientation. *Journal of Experimental Social Psychology*, 46, 139–145.

Landau, M.J., Goldenberg, J., Greenberg, J., Gillath, O., Solomon, S., Cox, C., Martens, A. and Pyszczynski, T. (2006a) The siren's call: Terror management and the threat of men's sexual attraction to women. *Journal of Personality and Social Psychology*, 90, 129–146.

Landau, M.J., Greenberg, J. and Rothschild, Z. (2009) Motivated cultural worldview adherence and culturally loaded test performance. *Personality and Social Psychology Bulletin*, 35, 442–453.

Landau, M.J., Greenberg, J., Solomon, S., Pyszczynski, T. and Martens, A. (2006b) Windows into nothingness: Terror management, meaninglessness, and negative reactions to modern art. *Journal of Personality and Social Psychology*, 90, 879–892.

Landau, M.J., Johns, M., Greenberg, J., Pyszczynski, T., Solomon, S. and Martens, A. (2004a) A function of form: Terror management and structuring of the social world. *Journal of Personality and Social Psychology*, 87, 190–210.

Landau, M.J., Solomon, S., Greenberg, J., Cohen, F., Pyszczynski, T., Arndt, J., Miller, C.H., Ogilvie, D.M. and Cook, A. (2004b) Deliver us from evil: The effects of mortality salience and reminders of 9/11 on support for President George W. Bush. *Personality and Social Psychology Bulletin*, 30, 1136–1150.

Lykins, E.L.B., Segerstrom, S.C., Averill, A.J., Evans, D.R. and Kemeny, M.E. (2007) Goal shifts following reminders of mortality: Reconciling posttraumatic growth and terror management theory. *Personality and Social Psychology Bulletin*, 33, 1088–1099.

MacDorman, K.F. (2005) Mortality salience and the uncanny valley. *Proceedings of the 5th IEEE-RAS International Conference on Humanoid Robots*, Tsukuba, Japan, 399–405. Doi: 10.1109/ICHR.2005.1573600.

McGregor, I., Zanna, M.P., Holmes, J.G. and Spencer, S.J. (2001) Compensatory conviction in the face of personal uncertainty: Going to extremes and being oneself. *Journal of Personality and Social Psychology*, 80, 472–488.

Mikulincer, M., Florian, V. and Hirschberger, G. (2003) The existential function of close relationships: Introducing death into the science of love. *Personality and Social Psychology Review*, 7, 20–40.

Motyl, M., Pyszczynski, T., Cox, C., Siedel, A., Maxfield, M. and Weise, D.R. (submitted) We are family: The effects of mortality salience and priming a sense of common humanity on intergroup prejudice.

Pyszczynski, T., Abdollahi, A., Solomon, S., Greenberg, J., Cohen, F. and Weise, D. (2006) Mortality salience, martyrdom, and military might: The Great Satan versus the Axis of Evil. *Personality and Social Psychology Bulletin*, 32, 525–537.

Pyszczynski, T., Greenberg, J., Koole, S. and Solomon, S. (2010) Experimental existential psychology: Coping with the facts of life. In S. Fiske, D. Gilbert, and G. Lindzey (eds), *Handbook of Social Psychology*, Vol. 1, 5th Edition, pp. 724–757. London: John Wiley and Sons.

Pyszczynski, T., Greenberg, J. and Solomon, S. (1999) A dual-process model of defense against conscious and unconscious death-related thoughts: An extension of terror management theory. *Psychological Review*, 106, 835–845.

Rank, O. (1932/1989) *Art and Artist: Creative Urge and Personality Development*. New York: Knopf.

Rank, O. (1941/1958) *Beyond Psychology*. New York: Dover.

Rosenblatt, A., Greenberg, J., Solomon, S., Pyszczynski, T. and Lyon, D. (1989) Evidence for terror management theory I: The effects of mortality salience on reactions to those who violate or uphold cultural values. *Journal of Personality and Social Psychology*, 57, 681–690.

Rothschild, Z.K., Abdollahi, A. and Pyszczynski, T. (2009) Does peace have a prayer? The effect of mortality salience, compassionate values, and religious fundamentalism on hostility toward outgroups. *Journal of Experimental Social Psychology*, 45, 816–827.

Routledge, C. and Arndt, J. (2008) Self-sacrifice as self-defense: Mortality salience increases efforts to affirm a symbolic immortal self at the expense of the physical self. *European Journal of Social Psychology*, 38, 531–541.

Routledge, C. and Arndt, J. (2009) Creative terror management: Creativity as a facilitator of cultural exploration after mortality salience. *Personality and Social Psychology Bulletin*, 35, 493–505.

Salzman, M.B. (2001) Cultural trauma and recovery: Perspectives from terror management theory. *Trauma, Violence and Abuse*, 2, 172–191.

Schimel, J., Hayes, J., Williams, T.J. and Jahrig, J. (2007) Is death really the worm at the core? Converging evidence that worldview threat increases death-thought accessibility. *Journal of Personality and Social Psychology*, 92, 789–803.

Schimel, J., Simon, L., Greenberg, J., Pyszczynski, T., Solomon, S., Waxmonski, J. and Arndt, J. (1999) Support for a functional perspective on stereotypes: Evidence that mortality salience enhances stereotypic thinking and preferences. *Journal of Personality and Social Psychology*, 77, 905–926.

Solomon, S., Greenberg, J. and Pyszczynski, T. (1991) A terror management theory of social behavior: The psychological functions of self-esteem and cultural worldviews. In M.P. Zanna (ed.), *Advances in Experimental Social Psychology*, 24, 93–159. New York: Academic Press.

Sullivan, D., Greenberg, J. and Landau, M.J. (2010) Toward a new understanding of two films from the dark side: Terror management theory applied to *Rosemary's Baby* and *Straw Dogs*. *Journal of Popular Film and Television*, 37, 42–51.

van den Bos, K. (2001) Uncertainty management: The influence of uncertainty salience on reactions to perceived procedural fairness. *Journal of Personality and Social Psychology*, 80, 931–941.

Vess, M., Arndt, J., Cox, C.R., Routledge, C. and Goldenberg, J.L. (2009) The terror management of medical decisions: The effect of mortality salience and religious fundamentalism on support for faith-based medical intervention. *Journal of Personality and Social Psychology*, 97, 334–350.

Vess, M., Routledge, C., Landau, M. and Arndt, J. (2009) The dynamics of death and meaning: The effects of death-relevant cognitions and personal need for structure on perceptions of meaning in life. *Journal of Personality and Social Psychology*, 97, 728–744.

Weise, D., Pyszczynski, T., Cox, C., Arndt, J., Greenberg, J., Solomon, S. and Kosloff, S. (2008) Interpersonal politics: The role of terror management and attachment processes in political preferences. *Psychological Science*, 19, 448–455.

Wisman, A. and Goldenberg, J.L. (2005) From the grave to the cradle: Evidence that mortality salience engenders a desire for offspring. *Journal of Personality and Social Psychology*, 89, 46–61.

Self-Determination Theory

Edward L. Deci and Richard M. Ryan

ABSTRACT

Self-determination theory (SDT) is an empirically derived theory of human motivation and personality in social contexts that differentiates motivation in terms of being autonomous and controlled. Work leading to the theory began with experiments examining the effects of extrinsic rewards on intrinsic motivation. During more than thirty years since the initial studies, we have developed five mini-theories to address different, though related, issues: the effects of social environments on intrinsic motivation; the development of autonomous extrinsic motivation and self-regulation through internalization and integration; individual differences in general motivational orientations; the functioning of fundamental universal psychological needs that are essential for growth, integrity, and wellness; and the effects of different goal contents on well-being and performance. We have subsequently used SDT and its mini-theories to guide and interpret research on many new issues, including motivation and wellness across cultures, close relationships, enhancement and depletion of energy and vitality, and the roles of both mindful awareness and nonconscious processes in behavioral regulation. Although much of SDT was developed through laboratory experiments, it is also supported by a great deal of applied research

using both field studies and clinical trials to address significant social issues. We briefly mention some of that work, especially related to health behavior change, education, psychotherapy, work motivation, sport and exercise, and prosocial behaviors.

INTRODUCTION

The central focus of social psychology has long been the effects of social environments on people's attitudes, values, motivations, and behaviors, and there can be little doubt that environmental forces have an enormous impact on these outcomes. Many social psychological theories have tended, either implicitly or explicitly, to view learning – that is, acquiring the attitudes, values, motivations, and behaviors – primarily in terms of social environments teaching individuals what to think, value, need, and do. This view has been called the “standard social science model” of a relatively plastic human nature, molded by its social contexts (see Tooby and Cosmides, 1992). In developmental psychology, this

perspective is most clearly conveyed in social learning theory, where modeling and reinforcements are the principal mechanisms of learning and growth. In social psychology, social cognition, and cultural relativism, such models are frequently employed to explain situationally influenced cognitions and behaviors.

The social psychology of self-determination theory (SDT) has also focused on the influences of social environments on attitudes, values, motivations, and behaviors both developmentally and in current situations; however, it has taken a quite different approach to these issues. Specifically, SDT assumes that the human organism is evolved to be inherently active, intrinsically motivated, and oriented toward developing naturally through integrative processes. These qualities need not be learned; they are inherent in human nature. Still, they develop over time, play a central role in learning, and are affected by social environments.

For these natural, active processes of intrinsic motivation and integration to operate effectively toward healthy development and psychological well-being, human beings need particular nutriments – both biological and psychological (Ryan, 1995). In the relative absence of such nutriments, these natural processes will be impaired, resulting in experiences, development, and behaviors that are less than optimal. Within SDT, we focus primarily on psychological nutriments and their dynamics within social environments, although biological supports, as well as inherent individual differences, also play important roles.

According to SDT, based on decades of empirical work, there are at least three universal psychological needs – specifically, needs for competence, autonomy, and relatedness – that are essential for optimal development and functioning. Unlike some evolutionary perspectives, we understand these needs as underlying the adaptive organization of behavior and being supported by many individual adaptations, rather than themselves being functionally specific or

modular “add-ons” (Deci and Ryan, 2000). As will be apparent throughout the chapter, the starting point of natural activity, intrinsic motivation, the integrative tendency, and fundamental psychological needs has led to many instances where the predictions of SDT have diverged dramatically from those of other prominent social-psychological theories, and we will point out a few of them in this chapter.

THE DEVELOPMENT OF SDT

SDT evolved out of research on the effects of extrinsic rewards on intrinsic motivation. In the first published studies (Deci, 1971), college students were paid for working on intrinsically interesting puzzles and the monetary rewards undermined their intrinsic motivation for the activity. This was followed by more than 100 similar studies (see Deci et al., 1999) confirming this controversial idea that rewards do not always motivate subsequent persistence; indeed they can undermine intrinsic motivation. We have used the attributional concept *perceived locus of causality* (de Charms, 1968; Heider, 1958) as part of our account of that effect and of other changes in intrinsic motivation, but we have also linked intrinsic motivation, and the social-contextual effects on it, to the basic human needs for competence and self-determination (i.e., autonomy). Intrinsic motivation was considered an inherent characteristic of human beings and was viewed as the prototype of psychological freedom or self-determination. It could be either undermined or enhanced depending on whether the social environment supported or thwarted the needs for competence and self-determination. If a reward or other external event such as threat of punishment (Deci and Cascio, 1972), positive feedback (Deci, 1971), competition (Deci and Betley et al., 1981), or choice (Zuckerman et al., 1978) were expected to thwart these basic needs, it was predicted to prompt an external

perceived locus of causality and undermine intrinsic motivation; but if the event were expected to support these basic needs, it was predicted to prompt an internal perceived locus of causality and enhance intrinsic motivation. Monetary rewards, threats, and competition were predicted to thwart autonomy, and such events did typically undermine intrinsic motivation (e.g., Deci, Betley et al., 1981; Deci et al., 1999). In contrast, positive feedback and choice were predicted to enhance experiences of competence and self-determination, fostering greater intrinsic motivation, and results have confirmed this as well (Deci et al., 1999; Zuckerman et al., 1978).

Our theoretical postulate linking environmental factors to fundamental human needs as a basis for explaining effects of the social environment on intrinsic motivation stood in sharp contrast to the views of other psychologists examining intrinsic motivation. For example, Lepper et al. (1973) considered intrinsic motivation to be a postbehavior self-attribution, and they used Bem's (1972) self-perception theory to explain the undermining of intrinsic motivation by tangible rewards – that is, people simply attributed less intrinsic motivation to themselves because there was overjustification for doing the activity.

Cognitive evaluation theory

As the work progressed, it became increasingly necessary to consider the dynamic interplay of autonomy and competence needs within social contexts in order to explain the increasingly complex experimental phenomena such as the finding that performance-contingent rewards, which are given for doing well on a task (e.g., doing better than the 80th percentile), were less detrimental than task-contingent rewards, which are given for doing or completing a task (Ryan et al., 1983). Accordingly, we (Deci and Ryan, 1980) reviewed the extant literature and introduced a formal mini-theory, labeled *cognitive evaluation theory* (CET), to explain

the effects of extrinsic factors on intrinsic motivation.

CET specified two processes through which intrinsic motivation can be affected. First, to the extent that events such as rewards lead to an external perceived locus of causality and thwart the autonomy or self-determination need, the events will undermine intrinsic motivation; whereas, to the extent that events such as choice lead to an internal perceived locus of causality and support the autonomy need, they will enhance intrinsic motivation. The second process specifies that events such as positive feedback that lead to perceived competence by supporting the competence need will enhance intrinsic motivation, whereas events such as negative feedback that lead to perceived incompetence will undermine intrinsic motivation. However, the positive feedback must be for an autonomously motivated activity (Pritchard et al., 1977) or within an autonomy-supportive context (Ryan, 1982) in order for it to enhance intrinsic motivation.

Finally, CET specified that social-environmental events such as rewards or feedback have two aspects relevant to intrinsic motivation. The first is a *controlling* aspect that pressures people to think, feel, or behave in specific ways, thus prompting an external perceived locus of causality, thwarting autonomy, decreasing intrinsic motivation, and leaving motivation primarily controlled rather than autonomous. The second aspect is the *informational* aspect that conveys competence information within the context of some autonomy support. When this aspect affirms people's competence for an autonomous activity, it supports the competence need and to some extent the autonomy need, thus enhancing intrinsic motivation. However, when it prompts perceived incompetence and thwarts the competence need, it decreases intrinsic motivation. In fact, if the competence information is sufficiently negative, indicating that people are too incompetent to attain desired outcomes, it could undermine both intrinsic and extrinsic motivation, leaving individuals *amotivated* (i.e., without intention or

motivation toward the activity). The effect of an event on intrinsic motivation will depend on the relative salience of the two aspects. The CET proposition about the two aspects of rewards provided an explanation for such phenomena as tangible rewards undermining intrinsic motivation and verbal rewards (i.e., positive feedback) enhancing it. It also explained, for example, why performance-contingent rewards, although they significantly undermined intrinsic motivation, were not as detrimental as task-contingent rewards: because the controlling aspect of both types of rewards is similarly salient, but the informational aspect is more salient in performance-contingent rewards than in task-contingent rewards (Ryan et al., 1983).

Reward effects revisited

The research on reward effects was very controversial from its first appearance, and to some extent the controversy continued for decades. People not fond of the finding that tangible rewards tend to undermine intrinsic motivation argued that there were confounds in the methods (e.g., Calder and Staw, 1975), that there were behavioral explanations of the phenomena that were more valid than CET's cognitive-motivational ones (e.g., Scott, 1975), and that the findings provided no real reason to refrain from using rewards as a primary motivational strategy in education and elsewhere (Eisenberger and Cameron, 1996). These critiques led us to do a meta-analysis of 128 experiments on the effects of extrinsic rewards on intrinsic motivation (Deci et al., 1999). The meta-analysis strongly confirmed what we had long been saying, namely that: (1) positive feedback enhances intrinsic motivation; (2) tangible rewards undermine intrinsic motivation; and (3) both task-contingent rewards and performance-contingent rewards decrease intrinsic motivation, but that unexpected rewards and rewards that do not require doing the target task do not undermine intrinsic motivation for it. Further, CET provided a full account of the complex set of findings that emerged from the meta-analysis.

The social ambience

There was another important set of findings that emerged from the CET research, namely that the general interpersonal ambience of a situation such as a classroom or a work group can be characterized as either autonomy supportive or controlling. For example, Deci, Schwartz et al. (1981) found that when elementary-school teachers created an autonomy-supportive classroom ambience their students evidenced greater intrinsic motivation as well as greater competence need satisfaction, than when the teachers created a controlling ambience, because, in the former context, the students felt free to develop their own sense of competence. Further Deci et al. (1989) found that when managers were more autonomy supportive their subordinates were more satisfied and trusting than when the managers were more controlling.

In laboratory experiments, the social ambience was manipulated to examine the effects of various external events when administered in an autonomy-supportive versus controlling context. For example, Ryan (1982) found that when positive feedback was given within an autonomy-supportive ambience it tended to enhance intrinsic motivation, as had been found earlier, but when it was given within a controlling ambience it decreased intrinsic motivation, thus confirming that positive feedback enhances intrinsic motivation only if it is accompanied by some autonomy support. (In short, forced competence development does not enhance intrinsic motivation.) Similarly, Ryan et al. (1983) found that when performance-contingent monetary rewards were administered controllingly they undermined intrinsic motivation, as had been found earlier, but when they were administered in an autonomy-supportive context they enhanced intrinsic motivation relative to no rewards and no feedback, although these rewards still led to less intrinsic motivation than just positive feedback comparable to what was implicit in the performance-contingent rewards (e.g., "you did better than 80 percent of the other participants").

Koestner et al. (1984) found further that when limits were set on children in an autonomy-supportive context the limits had a positive effect on subsequent motivation, but when they were set in a controlling context their effect was negative.

Summary

The CET work was able to specify the types of external events that would, on average, enhance intrinsic motivation (e.g., positive feedback and choice) versus diminish it (e.g., tangible rewards and competition). The research was also able to characterize autonomy-supportive versus controlling social contexts and use those to predict the intrinsic motivation of people within them. Finally, it explained how the social context or ambience could interact with external events to moderate their results on intrinsic motivation.

Additional mini-theories

As we were writing the review that drew together the CET research (Deci and Ryan, 1980), we began thinking about two new questions. First, would it make sense to have an individual difference concept that paralleled the “state” ideas of autonomous motivation, controlled motivation, and amotivation? This seemed important because surely enduring person factors and not just concurrent social-contextual factors affect people’s motivation and behavior at any given time. The second question was whether extrinsically motivated behaviors, which the studies showed typically thwarted the autonomy need and undermined intrinsic motivation, could be performed autonomously, and, if so, how autonomous extrinsic motivation could be facilitated. This also seemed like a very important question, because daily behaviors involve for most of us a hefty portion of extrinsically motivated activities, some of which are enacted autonomously and others of which are clearly controlled and characterized by alienation.

Causality orientations theory

The first of these two new questions led to the concept of *causality orientations* – with the autonomous, controlled, and impersonal causality orientations – as individual differences (Deci and Ryan, 1985a). Everyone is said to have each of the orientations to some degree, so any or all of them can be used to predict outcomes. The autonomy orientation refers both to orienting toward internal and external cues in a way that gives them an autonomy-supportive or informational significance and also to being more autonomous in general across domains and time. The controlled orientation refers to interpreting cues as controls and demands and to being controlled in general at the person level. Finally, the impersonal orientation refers to orienting toward cues as indicators of incompetence and to being generally amotivated.

The causality orientations concept and its psychometric scale have worked effectively in predicting variance in a range of adult outcomes. For example, the autonomy orientation has been positively associated with self-actualization, self-esteem, more choiceful self-disclosure, and supporting autonomy in others; the controlled orientation has been positively associated with public self-consciousness, the type-A coronary prone behavior pattern, inconsistency in attitudes and behaviors, and greater defensiveness; and the impersonal orientation has been positively related to self-derogation, poorer self-regulation, and depression.

Consistent with SDT, both the autonomy and controlled orientations were positively related to an internal locus of control (Rotter, 1966). The concept of internal locus of control concerns only whether one believes that outcomes are contingent on behavior, but it does not differentiate between whether the resulting motivation is autonomous or controlled. An internal perceived locus of causality, in contrast, reflects only autonomous motivation.

The causality orientations refer to relatively stable motivational orientations that parallel the three concepts of autonomy-supportive,

controlling, and amotivating interpersonal contexts, and the causality orientations also parallel the state-like motivational concepts of autonomous motivation, controlled motivation, and amotivation. Importantly, research has shown that the causality orientations and the types of social contexts make parallel predictions for the three state-like motivations and also predict independent variance in the motivational states and in a range of other outcomes, such as work performance (Baard et al., 2004), and better maintained weightloss over two years (Williams, Ryan, and Deci, 1996). Causality orientations theory (COT) is more in the realm of personality than social psychology, but SDT concepts are all interrelated.

Organismic integration theory

The second question we were pondering in 1980, namely whether extrinsic motivation could become autonomous, led us to address the concept of *internalization* by differentiating between internalized extrinsic motivation that is controlled and internalized extrinsic motivation that is autonomous. The resulting conceptualization, with its manifold ramifications, was referred to as *organismic integration theory* (OIT; Deci and Ryan, 1985b; Ryan et al., 1985). It has at its core the assumption of an inherent integrative tendency, viewed as the fundamental developmental process, and this tendency, like intrinsic motivation, was predicted to be facilitated by support for the basic psychological needs and to be impaired by thwarting of the needs. To understand internalization, however, it became immediately clear that we needed the third basic psychological need, namely that of relatedness, in order to have a full understanding of internalization and integration. The *need for relatedness* – that is, the need to be close to, trusting of, caring for, and cared for by others – is similar to what Baumeister and Leary (1995) would later call the *need to belong*. Since 1985 we have worked with the concept of the three basic and universal psychological needs without having found a compelling reason to add a fourth.

Perhaps the most important and far-reaching element of OIT is its differentiation of the varied types of internalized extrinsic motivation. Whereas most internalization theories have considered a value or behavioral regulation to be either outside the person or inside the person (with inside being better), OIT specified varying degrees to which a behavioral regulation and its accompanying value could be internalized, resulting in different types of subsequent regulation. A relatively unstable form of internalization is represented by *introjection*, in which people adopt an ambient value or practice, and are motivated to maintain it, as they “should,” in order to maintain self-approval or avoid guilt. Self-esteem contingencies (Deci and Ryan, 1995) and ego-involvement (Ryan 1982) are both forms of introjected regulation. A second type of internalization is *identification*, which involves personally identifying with the value of a behavior and thus fully accepting it as their own. A final type of internalization is *integration*, in which people integrate identifications with other aspects of their core values and practices. Internalization is a manifestation of the organismic integration process, and integrated regulation results when that process has worked to fully internalize a behavioral regulation.

By specifying the degrees of internalization OIT provided an account for the troubling issue of people forcing themselves to behave using internal pressures such as contingent self-esteem or threats of guilt. These processes are internal to people but they are by no means optimally healthy ways for people to regulate themselves for they do not possess the qualities of autonomy – namely, flexibility, volition, and a sense of choice. It turns out, in fact, that introjection has correlates and consequences that are qualitatively closer to external control than to identification (e.g., Ryan et al., 1993). In contrast to introjection, identification and integration share many qualities with intrinsic motivation and represent relatively autonomous forms of extrinsic motivation. These styles of regulation, together with *external regulation*

(behavior controlled by external reward or punishment contingencies) and *intrinsic motivation*, represent five ways of regulating oneself, and we have proposed that the *reasons* people engage in a behavior can be aligned with the various types of motivation and regulation – namely, external reasons, introjected reasons, identified reasons, integrated reasons, and intrinsic reasons.

With this new conceptualization, the most salient and important distinction within SDT is neither “intrinsic versus extrinsic motivation” nor “internal versus external to the person,” but is rather *autonomous versus controlled motivation*. Autonomous motivation encompasses intrinsic motivation and identified/integrated extrinsic motivation, whereas controlled motivation comprises external control and introjected regulation. Moreover the issue of autonomy was considered relative, insofar as most behaviors represent a mixture of the various reasons for acting described by these five categories.

OIT has been vigorously researched, frequently using an assessment approach developed by Ryan and Connell (1989) that measures the degree to which individuals do particular behaviors for various autonomous and controlled reasons. Using this approach, for example, researchers have found that the types of regulation form a *simplex pattern* indicating that they fall along a relative-autonomy continuum anchored by external regulation on the controlling end and integrated regulation and intrinsic motivation on the autonomous end. Further, studies have shown that the more autonomous types of motivation are associated with such outcomes as wellness, engagement in work or schoolwork, perceived competence, and deeper conceptual learning (e.g., Grolnick and Ryan, 1987; Ryan and Connell, 1989; Vallerand, 1997).

Internalization is particularly important in childhood, but it is relevant across the lifespan. For example, in our research on health behavior change among adults, we view the change process as being based in internalization of the value for and regulation

of healthy behaviors (Williams et al., 1998). Research has shown that the more fully a regulation is internalized, such that the behavior is more autonomous, the more likely people are to change and maintain behaviors such as eating healthy diets and stopping smoking (e.g., Williams et al., 2006). Further, internalization plays a similarly important role in making positive changes within psychotherapy (Pelletier et al., 1997). This general formulation has been used in many life domains, as discussed later in the chapter.

Given the importance of well-internalized extrinsic motivation for effective functioning and well-being, we soon began examining the conditions most likely to promote full internalization. We hypothesized that social environments that support satisfaction of the basic psychological needs would promote fuller internalization of extrinsic motivation, and many studies have focused on the relation of autonomy support to internalization. *Autonomy support* refers to taking the others’ perspective, encouraging initiation and exploration, providing choice, and being responsive to the others. In an interview study with parents, Grolnick and Ryan (1989) found that when parents were more autonomy supportive rather than controlling in their child rearing, the children more fully internalized the regulation for doing their schoolwork and chores around the house. Further, a laboratory experiment by Deci et al. (1994) revealed that providing a meaningful rationale for engaging in an uninteresting task, acknowledging people’s feelings about the task, and making requests in ways that are more consistent with choice than with control all contributed to greater internalization and integration.

One common way for parents to be controlling rather than autonomy supportive is through *parental conditional regard*. Parents provide more attention and affection when the children do or be what the parents want, and they withdraw attention and love when the children do not live up to the parents’ expectations. Assor et al. (2004) showed that

when parents are conditionally regarding, their children tend to introject the demands – thus conditionally administering esteem to themselves as their parents had done to them. In the process, they experience internal conflict, shame and guilt when they fail, and anger and resentment toward their parents. Follow-up research compared parental conditional regard with parental autonomy support (Roth et al., 2009). Results showed that conditional regard predicted introjection, both suppression of and dysregulation of emotions, and both controlled motivation and amotivation in school; whereas autonomy support led to the experience of choice, integrated regulation of emotions, and interest-focused school engagement. In short, the consequences of autonomy support are far more adaptive than those of conditional regard.

Although much OIT research has focused on autonomy support for facilitating internalization, the theory proposes that it is satisfaction of all three needs that is essential for full internalization. It turns out that parents and other authorities who support autonomy also tend to support competence and relatedness, so it is often the case that, when autonomy is being supported, competence and relatedness are also being supported, although satisfaction of each need is associated with independent influences and resulting dynamic outcomes.

In sum, research guided by OIT has shown that extrinsic motivation can be internalized to differing degrees, leading to types of internal regulation that differ in the degree to which they represent autonomy. The types that are more autonomous (identified and integrated) have been associated with more positive outcomes in various domains including schoolwork and regulation of emotions, whereas the types of regulation that are controlling (external and introjection) are associated with poorer outcomes across domains. Finally, social contexts that support the basic needs for autonomy, competence, and relatedness also facilitate fuller internalization, whereas context that thwart need satisfaction,

such as using rewards and punishments or conditional regard promote only introjection and are accompanied by indicators of ill-being.

OIT is perhaps best viewed as a theory of personality development and self-regulation, although the idea that social contexts facilitate versus impair internalization by supporting versus thwarting the basic psychological needs is very much a social-psychological concept. In this regard we see that the more the social contexts are relationally accepting and inclusive, competence promoting, and autonomy supportive, the more fully people internalize ambient social values and norms. Hence, social-contextual conditions that facilitate internalization of extrinsic motivation share much with those that maintain or enhance intrinsic motivation.

By the late 1980s, SDT's three interrelated mini-theories (CET, COT, and OIT) were being validated by considerable research. Yet, as that research was accumulating, we also saw that the concept of basic psychological needs had additional utility. In contexts where the three basic needs were being satisfied, people were more likely to exhibit both intrinsic motivation and more integrated forms of extrinsic motivation. But just as importantly, across these studies we also observed that where needs were satisfied, participants were invariably reporting greater wellness; whereas when any of the needs was thwarted, various forms of defensiveness and ill-being were evident. This led us to formulate the fourth mini-theory focused on the core definition of basic psychological needs for autonomy, competence, and relatedness and their roles as essential nutriments for healthy development, well-being, and mature relationships. We now take a brief look at this fourth mini-theory and some of the research it organizes.

Basic psychological needs theory

The basic psychological needs theory (BPNT; Ryan et al., 1996) was built on the concept of universal psychological needs, which had been important since the earliest days of

our work. BPNT was formulated primarily to account for the well-being effects associated with satisfactions of autonomy, competence, and relatedness needs, and much research has shown both the necessity of basic need satisfaction for well-being and the importance of the basic needs as mediators of the effects of social contexts on wellness. For example, need satisfaction at the between-person level has predicted better performance and greater psychological health in the workplace (e.g., Baard et al., 2004), and increases in need satisfaction over time have mediated the longitudinal relations between autonomy support of law school professors and the well-being of law students (Sheldon and Krieger, 2007). Other BPNT research examined basic need satisfaction not only at the between-person level but also at the within-person level and found that general need satisfaction was associated with greater psychological health at the person level, and also that on a daily basis, people experienced greater positive affect and less negative affect on those days when they got more satisfaction of their basic needs (Reis et al., 2000; Ryan et al., 2010). As we will see later, much research has employed the concept of basic psychological needs across domains, contexts, and cultures.

As we were formulating BPNT, work by Kasser and Ryan (1993) had begun examining the importance of different goal contents. In that work, life goals were differentiated, based on factor analysis, as being either intrinsic (directly satisfying of the basic needs) or extrinsic (more distal from the needs and may be antagonistic to them). We initially interpreted the goal-content work with BPNT, but more recently we recognized that this research area had become so extensive and complex that it required its own mini-theory. The fifth mini-theory within SDT is thus referred to as goal content theory (GCT).

Goal content theory

Based on the factor analyses, intrinsic aspirations or life goals included personal growth,

affiliation, and community, whereas extrinsic goals included wealth, fame, and image (Kasser and Ryan, 1996). The studies showed that, when people rated the extrinsic aspirations as being strong relative to the intrinsic aspirations, they evidenced less self-actualization and vitality, and more depression, anxiety, and narcissism. Whereas the line of work initiated by Kasser and Ryan looked at aspirations or life goals as individual differences, other studies by Vansteenkiste and colleagues manipulated the salience of people's goals and found that focusing people on extrinsic goals led to poorer performance on learning activities (e.g., Vansteenkiste et al., 2004). In short, strongly pursuing extrinsic life goals, whether because of an individual difference or a prompt, has led to less well-being, more ill-being, and poorer performance, presumably because the extrinsic aspirations do not directly satisfy the basic needs, and indeed often crowd out or compromise their satisfaction. For example, as materialists spend effort and time accumulating, they often compromise autonomy and relatedness in the pursuit of more "stuff." Additional research showed that not only pursuing extrinsic aspirations, but also attaining them, could be problematic for psychological health (Niemiec et al., 2009). Specifically, whereas attainment of intrinsic aspirations was associated with greater well-being and less ill-being, mediated by basic psychological need satisfaction, attainment of extrinsic aspirations did not enhance well-being but did relate to greater ill-being.

Research further showed that people develop stronger intrinsic life goals when their parents are accepting, affirming, and autonomy supportive (Williams et al., 2000), but they develop stronger extrinsic goals when their parents are rejecting and controlling. Presumably, when parents are cold and pressuring, their children fail to experience adequate need satisfaction, thus experiencing insecurity and developing what we call need substitutes, such as the pursuit of wealth, fame, and image. These goals then guide subsequent actions, which result in additional

need thwarting and an ongoing vicious cycle.

The continued expansion of SDT

In recent years, SDT has continued to expand in order to incorporate a broad range of new research topics. We have not, however, specified new mini-theories beyond the five already discussed, but instead have used just the SDT macro-theory to guide and interpret this new work. In this section we briefly address four of the basic research topics that have been examined within SDT during the past decade or so.

Cross-cultural studies

SDT states strongly that the three basic psychological needs are universal such that their satisfaction versus thwarting affects the psychological well-being of all people. This proposition has two important implications. First, it requires that the proposition be compatible with an evolutionary perspective, and second that the relation between need satisfaction-versus-thwarting and well-being be confirmed across a variety of cultures with different economic and political systems and with different cultural values. Elsewhere (Deci and Ryan, 2000), we presented an argument that the proposition of universal psychological needs for competence, relatedness, and autonomy is indeed consistent with an evolutionary perspective. Further, there have now been several studies supporting the cross-cultural relevance of the needs proposition. Here we briefly discuss two of those studies.

Within psychology, the idea of a need for relatedness (or belongingness or love) is fairly widely accepted and there is little argument about it being relevant across cultures. In addition, the idea of a need for competence is consistent with several prominent theories, and its relevance to multiple cultures has not been contested. On the other hand, however, the concept of a basic, universal psychological need for autonomy is

very controversial, with cross-cultural theorists such as Markus et al. (1996) maintaining that people acquire their needs from their cultures and that East Asian cultures do not value autonomy and independence but instead value relatedness and interdependence. Hence, these authors maintain that the SDT concept of autonomy is relevant only to Western cultures where individualism is valued. This, of course, implies that autonomy is not a universal need.

From the SDT perspective, autonomy across cultures would be evident in two ways. First, the prototype of autonomy is the intrinsic motivation inherent to the nature of people, so it should be apparent regardless of culture. Second, autonomy should also be manifest in all cultures as behaviors that are motivated by well-internalized extrinsic motivations. Most of the cross-cultural research has focused on the second.

Still, anecdotally, it should be obvious to anyone who has observed young children in any culture that intrinsically motivated learning and play are ubiquitous. Of course, the degree of support for these activities and opportunities to enact intrinsically motivated behaviors may differ in different cultures, but the phenomenon of being self-motivated to move and laugh and play can be found in any place where it has not been stamped out. Further, there is some education research indicating that intrinsic motivation is affected by autonomy support versus control in the social context and that intrinsic motivation leads to more effective learning in Japan, China, Korea, and other collectivist contexts (e.g., Bao and Lam, 2008; Jang et al., 2009; Kage and Namiki, 1990), just as is the case in the US.

In one cross-cultural study of autonomy, Chirkov et al. (2003) pointed out that autonomy within Western or Eastern cultures can accrue from relevant values having been fully internalized. For example, persons in an Eastern culture could be autonomous when enacting a collectivist cultural value, just as persons in a Western culture could be autonomous when enacting an individualist cultural

value if they had fully internalized the target value. Chirkov et al. sampled students in Russia, Turkey, South Korea, and the US, who completed a self-regulation questionnaire assessing their reasons for engaging in various cultural practices. Results showed that the degree to which the participants had internalized (and expressed greater relative autonomy for enacting) the values and regulations for the various practices predicted their degree of psychological health and well-being. This confirmed that being autonomous was just as important for psychological wellness in Korea as in the US and each of the other countries in the study. Chirkov et al. found further that the relation between autonomy and well-being was not moderated by gender. This is important because theorists such as Jordan (1997) have argued that autonomy is a male characteristic and has little relevance to women's sensibilities. The Chirkov et al. study indicates that it was indeed relevant for both women and men in each culture studied. The researchers pointed out that part of the reason for the strongly disparate positions was that authors such as Markus et al. (1996) and Jordon tend to conflate autonomy (volition) and independence (nonreliance), rather than differentiating these important constructs as SDT research and theory have done (Deci and Ryan, 2000).

A study by Deci et al. (2001) examined working adults in the US, which has a capitalist economy, and Bulgaria, which at the time of data collection had a central planning economy for the majority of companies, which were still state-owned. The researchers found that the degree to which the workers in both countries perceived their managers to be autonomy supportive positively predicted satisfaction of the needs for autonomy, competence, and relatedness while at work. Need satisfaction, in turn, positively predicted engagement with work and psychological adjustment at work. Together, this study, the Chirkov et al. (2003) study, and many other more recent studies (e.g., Lynch et al., 2009) using varied methodologies have

supported the view that basic psychological need satisfaction, and in particular satisfaction of the need for autonomy, is important for psychological well-being across a range of cultures regardless of whether they value individualism or collectivism more strongly.

Close personal relationships

It is often stated that being in a meaningful relationship, such as with a romantic partner or best friend, requires that people relinquish autonomy to keep the relationship strong. The SDT perspective, however, is that feeling autonomous *within* the relationship is an essential element for the relationship to be strong and intimate. Consequently, SDT researchers have, over the past decade, engaged in research examining the importance of autonomy and autonomy support for the highest quality relationships.

Three studies by La Guardia et al. (2000) examined within-person variance in security of attachment among close relational partners (mother, father, best friend, and romantic partner). Attachment theorists suggest that children develop working models of attachment based largely on early interactions with primary caregivers, and these between-person individual differences strongly influence the security of attachment with all subsequent close attachment partners. The La Guardia et al. research indicated that there was between-person variance in security of attachment; however, there was also substantial variance explained at the within person level. The security of attachment that people had within their closest relationships varied substantially from partner to partner, and the felt security within each relationship was a function of the level of basic psychological need satisfaction, including autonomy, experienced within that relationship. Lynch et al. (2009) assessed autonomy support from close others in multiple cultural contexts and similarly found that relationship satisfaction as well as self-functioning were higher in relationships characterized by autonomy support. Patrick et al. (2007) found further that satisfaction of the basic psychological needs

within romantic relationships predicted personal well-being, relational well-being, and effective management of conflict within the relationship. Yet another examination of autonomy in relationships focused on best friends (Deci et al., 2006). The researchers found that receiving autonomy support from a best friend was associated with greater relationship quality and greater well-being in the person receiving it. Further, the giving of autonomy support by that person to the best friend was also associated with the person experiencing greater relationship quality and well-being. In other words, both receiving and giving autonomy support within a friendship accounted for significant independent variance in the person experiencing higher relationship quality and greater well-being.

From these and other studies, it seems clear that feeling a sense of autonomy and volition within close relationships is important for experiencing the relationships as satisfying (La Guardia and Patrick, 2008). Thus, feeling autonomy and relatedness are not inherently antagonistic but rather are mutually supportive, although it is possible to make these needs antagonistic, as when a relational partner provides conditional regard, or requires that the person give up his or her autonomy in order to receive affection or regard from the partner.

Vitality: Energy available to the self

Ryan and Frederick (1997) used the concept of subjective vitality to refer to the sense of aliveness and vigor that energizes volitional actions. They suggested that vitality results from satisfaction of the basic psychological needs, is an important indicator of health, and provides the necessary energy for effective self-regulation and coping with challenges. Ryan and Deci (2008) proposed that, whereas attempts to control oneself (i.e., to act in accord with one's introjects) can drain energy and diminish aliveness, autonomous self-regulation is not depleting but is instead vitalizing. Vitality and autonomous self-regulation are thus activating, but it is a type of activation involving positive affect and is

different from the energy people experience when they are angry or anxious. Vitality, in short, is energy associated with people's integrated sense of self, which thus invigorates the processes of choice, volition, and effective coping with challenges (e.g., Rozanski, 2005).

Baumeister et al. (1998) proposed that any form of self-regulation consumes psychological energy so they predicted that having choice, which has been found in many situations to enhance autonomous motivation, should be depleting of energy and vitality. They reported results that they interpreted as supporting this assertion. However, in line with the SDT viewpoint, Moller et al. (2006) argued that true choice should not be depleting, and they pointed out that the condition Baumeister et al. had referred to as "high choice" actually represented a controlled rather than autonomous condition (i.e., pressure toward a particular choice) and hence was not true choice. In contrast, Moller et al. (2006) found in three experiments, that when participants were given true choice with no pressure there was no depletion and it led to significantly more energy and vitality than the "controlled choice" condition used by Baumeister et al. (1998).

Other experiments have also found that controlled regulation (i.e., self-control) is depleting as SDT predicts but that autonomous regulation tended to be vitalizing (e.g., Muraven et al., 2008; Nix et al., 1999). In sum, SDT's distinction between controlled regulation (i.e., self-control) and autonomous regulation (i.e., true self-regulation) is critical for understanding both vitality and depletion. Self-control depletes energy and vitality, but self-regulation, which promotes psychological need satisfaction, is vitalizing.

Within SDT, vitality, which is the energy available to the self, has been linked to both physical health and satisfaction of psychological needs for competence, autonomy, and relatedness (Ryan and Frederick, 1997). Thus, although energy has diurnal patterns and biological underpinnings (e.g., Thayer, 2001) it also varies with the support versus thwarting

of psychological needs affecting a wide range of outcomes from motivation to mood.

Nonconscious processes and mindful awareness

Because we use “experience of choice” about one’s behavior as an aspect of autonomy’s definition, various writers have interpreted this to mean that autonomy necessitates conscious decision-making. However, a sense of choice does not require deliberate decision making, it merely requires endorsing one’s actions. Thus, SDT allows for nonconscious initiation of autonomous behavior, and numerous experiments have now examined the nonconscious prompting of autonomous and controlled regulation of behavior.

Levesque and Pelletier (2003) primed autonomous and controlled motivation with words relevant to each concept and then allowed participants to spend 15 minutes working on an interesting activity. Subsequent dependent measures indicated that those participants primed with autonomy-related words were more intrinsically motivated for the task than those primed with control-related words. Several experiments by Hodgins and colleagues have also shown that people primed with autonomy display less defensiveness than those primed with control. For example, Hodgins et al. (2006) used a word-priming procedure and then examined participants’ self-handicapping around performance on a physical activity. Self-handicapping is a defensive response in which people deliberately do something (e.g., stay up very late the night before an important activity) to have an excuse in case they perform badly. The researchers found that people primed with control showed more self-handicapping than those primed with autonomy. These results, which were prompted with a priming procedure, parallel those found by Knee and Zuckerman (1998) in which self-reported control assessed with the causality orientations scale also led to more self-handicapping than did self-reported autonomy.

In short, nonconscious initiation of actions can be consistent with either autonomous or

controlled functioning. Further, when people are lacking awareness they are more vulnerable to being controlled through primes; however, when controlled primes are present, people still have the capacity, through mindfulness, to be autonomous in spite of the primes. Since the early days of SDT (e.g. Deci and Ryan, 1980) we have proposed that *mindful awareness* facilitates more autonomous regulation of behavior, but beginning with Brown and Ryan (2003), the last few years have witnessed more vigorous research efforts linking mindfulness with more autonomous, self-endorsed regulation, less defensiveness, and thus greater need satisfaction and wellness.

SDT ROOTS AND RELATIONS

The inspirations and foundations for SDT have come from several psychological traditions. The concept of intrinsic motivation first appeared in experimental psychology when Hullian drive theory concepts were found to be inadequate for explaining the exploratory behaviors of rats and monkeys (e.g., Harlow, 1950). White (1959) drew together that work and introduced the concept of competence as a fundamental motivation and basic need. Using these starting points, SDT has operated wholly within the empirical tradition since its beginning, and Heider’s (1958) attribution theory provided the route into the empirical work on intrinsic motivation at a time when the field of motivation was essentially moribund within psychology.

Although SDT has empirical roots in experimental social psychology (Heider, 1958), its theoretical roots extend farther afield to the organismic (Goldstein, 1939), ego-psychology (Hartmann, 1958), and existential-phenomenological (Pfander, 1910/1967) traditions, which focus on the critical importance of human experience and meaning in the determination of action, and the biologically inherent tendencies

toward integrated functioning (see Ryan, 1995; Ryan and Deci, 2004). In other words, we share meta-theory and some aspects of theory with these latter traditions while working within the empirical tradition.

For example, when ego-psychologists within the psychoanalytic tradition abandoned psychosexual staging as its main theory of normal development, the idea of conflict-free ego energy emerged as being an inherent motivation rather than being a derivative of the id. This so-called independent ego energy (White, 1963) is evident in intrinsic motivation, and is a motivational underpinning of the developing ego and healthy self. Within SDT, intrinsic motivation, with its underlying psychological needs for autonomy, competence, and relatedness, energizes the operation of the organismic integration process. This integrative process, which we view as the inherent and natural developmental process, is energized by intrinsic motivation and involves the internalization and integration of attitudes, values, motivations, and emotional regulatory processes, and it has much in common with Loevinger's (1976) concept of ego-development, which is central to her structural stage theory of the ego and its regulation. Like Loevinger's theory, Piaget's (1971) theory of cognitive development is also an organismic theory for it too assumes an inherent developmental process toward assimilation and integration. SDT, although explicitly not a stage theory, has common threads with these and related organismic theories that assume a natural tendency toward development, which thus does not have to be "programmed" by the environment, although, as SDT has emphasized, environmental supports are necessary for effective functioning of the integrative process. Finally, humanistic theories (e.g., Rogers, 1963) also posit an inherent developmental process, which these theories tend to call self-actualization.

SDT does recognize age-related changes in motivation, but our focus is on: (1) the fundamental integrative process that is operative across the lifespan; (2) the basic psychological

needs for autonomy, competence, and relatedness that energize the natural developmental process; and (3) the different regulatory processes, which, although they differ in the degree to which they reflect mature regulation, do not develop sequentially through age-related stages. Further, unlike stage theories, we maintain that adults are regulated to some degree by intrinsic motivation and each type of extrinsic motivation – external, introjected, identified, and integrated – and that people's motivational profiles will vary over time but not necessarily unidirectionally. Finally, although SDT is a theory of human needs and their relations with integrated functioning, it rejects Maslow's (1971) hierarchy of needs, instead asserting that the three basic needs are implicated across development (Ryan et al., 2006).

Perhaps SDT's major point of departure from these theories, other than that it being empirically based, is its focus on motivation. Whereas each of these other theories assumes inherent activity and an integrative tendency underlying development, they simply propose that the tendency functions without addressing the interaction between inherent needs and social conditions that supports its functioning. One of the most important reasons for addressing this issue is that it provides a means for predicting the conditions under which the developmental process will function most effectively. Specifically, it will function effectively to the degree that the needs for competence, autonomy, and relatedness are satisfied. This allows a theoretically based examination of the conditions that promote, for example, healthy child development, effective therapeutic change, optimal learning, skilled performance, and prosocial behavior.

APPLICATIONS TO SOCIAL ISSUES

SDT, perhaps as much as, or more than, any other theory of social psychology, has been applied to a broad range of life domains and

social issues. Some of the studies have been longitudinal or cross-sectional field studies, some have been randomized trials, and some have been lab experiments. SDT, and specifically the CET mini-theory, has identified specific environmental events such as rewards, deadlines, threats of punishment, competition, and evaluations, as well as controlling interpersonal contexts, that tend to (1) undermine intrinsic motivation, (2) have a corrupting effect in which people take the shortest path to the outcome, sometimes even when the path is inappropriate or immoral (Ryan and Brown, 2005), and (3) relate to poorer heuristic performance and well-being. SDT research has also identified external events such as choice, positive feedback, and acknowledging feelings, as well as autonomy-supportive social contexts, that enhance intrinsic motivation, internalization, and psychological well-being.

Because SDT research has shown that social contexts and communication styles affect motivation, performance, and well-being, many studies dealing with real-world contexts and social issues have examined autonomy-supportive versus controlling styles as they affect outcomes such as learning, socialization, healthy behaving, job satisfaction, prosocial behaviors, therapeutic outcomes, and rehabilitation in various treatment settings. We address just a small portion of this work, offering illustrative examples rather than a comprehensive review.

Promoting healthy behaving

Behavioral choices made by individuals everyday are among the most serious threats to physical health. For example, tobacco use has serious consequences such as heart disease and cancer; unhealthy diets and a lack of physical activity promote obesity, diabetes, and cardiovascular disease; and medication nonadherence works against the amelioration of illness. SDT research has tested process models for smoking cessation (Williams et al.,

2002), weight loss (Williams et al., 1996), glucose management by diabetic patients (Williams et al., 2004), and medication adherence (Williams and Rodin et al., 1998), among other health issues, finding that autonomy support provided by physicians or other practitioners predicts patients' autonomous motivation and perceived competence, which in turn predicts maintained health behavior change, as well as concrete health indicators such as glycosylated hemoglobin or chemically verified cessation.

This consistent pattern of findings has led to clinical trials involving SDT-based interventions designed to be autonomy, competence, and relationally supportive. Randomized trials from our labs and others have thus far addressed tobacco cessation and abstinence, improved diet, exercise, and LDL cholesterol (Williams et al., 2006); increased physical activity (Fortier et al., 2007); and better oral health (Halvari and Halvari, 2006). The Williams et al. trial found, for example, that an autonomy-supportive intervention led to significantly greater smoking cessation at the end of six months relative to that of a community-care control group, a significant difference that was still evident at both 18 and 30 months. Importantly, these patients were of relatively low socioeconomic status and over 50 percent of them said at their initial clinic visit that they did not want to make a quit attempt within the subsequent 30 days, belying the idea that a stage of "readiness" is a prerequisite for treatment and change.

Promoting learning and adjustment in schools

SDT researchers have examined links among autonomy-support (relative to control) within classrooms and homes, students' autonomous motivation and perceived competence, and the outcomes of improved learning, achievement, and well-being. For example, Deci et al. (1981) found that autonomy-supportive classrooms led to increased intrinsic

motivation and satisfaction of the competence need; Grolnick and Ryan (1987, 1989) found links from autonomy support to internalization of extrinsic motivation and well-being; Chirkov and Ryan (2001) found these relations in Russia; Reeve et al. (2002) found that a rationale for behaviors given in an autonomy-supportive way led to fuller internalization and more engagement in learning; both Benware and Deci (1984) and Grolnick and Ryan (1987) linked autonomy support to enhanced deep learning and conceptual understanding; and Vansteenkiste et al. (2004) found that intrinsic rather than extrinsic learning goals led to greater learning.

A network of findings that includes those just mentioned has led to the development of reform interventions that have used SDT as the primary theoretical basis (Deci, 2009). The SDT approach stands in clear contrast to the recent focus on incentives, accountability, and high-stakes testing within this country, where the assumption is that administrators, teachers, and students need to be held accountable for their performance (typically assessed as student scores on state-administered exams) and that the pressured accountability and the use of various incentives (i.e., high stakes) will motivate individuals at each level of the educational hierarchies to perform more effectively. SDT's view is that the high stakes with the pressure surrounding them tend to undermine autonomous motivation for teaching and learning and promote various types of "gaming," perhaps the most extreme of which is "manipulating" tests scores and student records (see Ryan and Brown, 2005).

In contrast, school reform based on SDT has included work by Feinberg et al. (2007) in Israel in which they used an autonomy-supportive approach to teach the basic principles of SDT to administrators and teachers, and then facilitated a process of the school personnel creating and implementing strategies for improvement. In the US an approach to school reform developed by James Connell is a comprehensive structured approach with many SDT-related elements including

improved teacher-student relationships by making smaller units within the schools, facilitating greater choice for teachers and students, and making instruction more optimally challenging and engaging. Results from evaluations of this approach have been very promising for improving attendance, graduation rates, and achievement (e.g., Gambone et al., 2004).

Psychotherapy and behavior change

In most psychotherapies, clients are encouraged to address and often change maladaptive behaviors, troubled relationships, or other presenting problems. These clients, in turn, have the choice about whether or not to engage in self-reflection and change. Successful psychotherapy requires a genuine willingness on the part of people to engage a process of change, particularly if the effects are expected to last beyond treatment (Deci and Ryan, 1985b; Ryan and Deci, 2008). SDT specifically argues that the maintenance and transfer of therapeutic changes, be they behavioral or psychological, requires the support of internalization and autonomous motivation (Pelletier et al., 1997). From the SDT perspective, this in turn suggests the importance of therapists providing an autonomy-supportive context for change.

SDT provides a broad, treatment-relevant framework for conceptualizing various developmental influences on psychopathology (Ryan et al., 2006). It also specifies the major elements of an approach to the practice of psychotherapy, arguing that the processes of growth and motivation entailed in psychological and behavioral change require an understanding of the dynamics of basic psychological needs and their interpersonal support from practitioners (Ryan and Deci, 2008). SDT has been applied to problems from suicide (Britton et al., 2008) and depression (e.g., Zuroff et al., 2007), to drug and alcohol rehabilitation (e.g., Ryan et al., 1995; Zeldman et al., 2004). SDT, in fact, is not only an approach to psychotherapy but can

also be applied to any existing interventions whose approach to motivating and implementing change processes affects clients' engagement and volition.

Sport and physical activity

One of the most vigorous applied areas of research within SDT is sport and physical activity. Studies have shown the role of both autonomous self-regulation and autonomy support in promoting greater motivation and persistence in physical activities, and SDT has been applied in physical education, health promotion, and coaching settings around the globe, as indicated by two recent collections of papers (Hagger and Chatzisarantis, 2007; Vlachopoulos, 2009).

Other social issues

Although SDT has relevance for many other social issues across a range of life domains, we very briefly mention just three more. The first concerns people experiencing job satisfaction and well-being in the workplace while performing effectively. For example, studies by Baard et al. (2004) in the banking industry found that employees who experienced more autonomy support from managers also reported more satisfaction of their basic psychological needs for autonomy, competence, and relatedness, and in turn had higher performance evaluations, higher well-being, and lower ill-being. Indeed, many studies point to the relevance of basic need supports for a productive, well-functioning workplace (e.g., Deci et al., 2001).

The second of these issue concerns prosocial behaviors. Recent studies by Weinstein and Ryan (2010) revealed that when participants were given choice (relative to no choice) over acts of helping, they in turn experienced significantly greater need satisfaction and displayed greater psychological well-being. Just as importantly these and other studies have shown that recipients

of help benefit more, are less psychologically threatened, and are more grateful when the help they receive is autonomously motivated.

Yet another intriguing area of research is motivation in virtual worlds. SDT research has demonstrated the role of need satisfaction in motivating videogame play, as well as the relation of basic needs to issues such as overuse and aggression in videogames, among others (e.g., Przybylski et al., 2009; Ryan et al., 2006).

As we stated, we could at best provide a very partial review of the expanding applied work based on SDT. In fact, SDT research is voluminous in areas of sustainability (Pelletier and Sharp, 2008), parenting (Grolnick, 2003), religion (Ryan et al., 1993), and the nature of happiness (Ryan and Deci, 2001), among others. We believe this is a result of philosophically well-founded and empirically testable theoretical foundations, as well as a focus on central issues in human experience and their impact on motivation and well-being.

CONCLUSION

SDT is a psychological macro-theory that focuses to a substantial extent on the effects of social–contextual factors on human motivation, behavior, and personality. In this chapter we explored the theory and its development, emphasizing the importance of autonomy-supportive, relative to controlling, interpersonal contexts for optimal motivation, effective behavior, healthy development, and psychological well-being. We presented each of the five mini-theories underlying SDT, as well as some newer areas of work that are extending that basic framework. We pointed out that SDT is an empirically derived theory, but its meta-theory has elements drawn from organismic, phenomenological, ego-psychological, and humanistic traditions, thus leading to basic assumptions and theoretical elements that make the theory

very different from many other mainstream social psychological theories. Finally, we briefly reviewed applications of SDT that are relevant for addressing social issues, including health behavior change, education, psychotherapy, work motivation, virtual environments, and prosocial behavior.

REFERENCES

Assor, A., Roth, G. and Deci, E.L. (2004) The emotional costs of parents' conditional regard: A self-determination theory analysis. *Journal of Personality*, 72, 47–89.

Bao, X. and Lam, S. (2008) Who makes the choice? rethinking the role of autonomy and relatedness in Chinese children's motivation. *Child Development*, 79, 269–283.

Baard, P.P., Deci, E.L. and Ryan, R.M. (2004) Intrinsic need satisfaction: A motivational basis of performance and well-being in two work settings. *Journal of Applied Social Psychology*, 34, 2045–2068.

Baumeister, R.F., Bratslavsky, E., Muraven, M. and Tice, D.M. (1998) Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology*, 74, 1252–1265.

Baumeister, R. and Leary, M.R. (1995) The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, 117, 497–529.

Bem, D.J. (1972) Self-perception theory. In L. Berkowitz (ed.), *Advances in Experimental Social Psychology*, 6, 1–62. New York: Academic Press.

Benware, C. and Deci, E.L. (1984) The quality of learning with an active versus passive motivational set. *American Educational Research Journal*, 21, 755–766.

Britton, P.C., Williams, G.C. and Conner, K.R. (2008) Self-determination theory, motivational interviewing, and the treatment of clients with acute suicidal ideation. *Journal of Clinical Psychology*, 64, 52–66.

Brown, K.W. and Ryan, R.M. (2003) The benefits of being present: Mindfulness and its role in psychological well-being. *Journal of Personality and Social Psychology*, 84, 822–848.

Calder, B.J. and Staw, B.M. (1975) The interaction of intrinsic and extrinsic motivation: Some methodological notes. *Journal of Personality and Social Psychology*, 31, 76–80.

Chirkov, V.I. and Ryan, R.M. (2001) Parent and teacher autonomy-support in Russian and U.S. adolescents: Common effects on well-being and academic motivation. *Journal of Cross Cultural Psychology*, 32, 618–635.

Chirkov, V.I., Ryan, R.M., Kim, Y. and Kaplan, U. (2003) Differentiating autonomy from individualism and independence: A self-determination theory perspective on internalization of cultural orientations and well-being. *Journal of Personality and Social Psychology*, 84, 97–110.

de Charms, R. (1968) *Personal Causation*. New York: Academic Press.

Deci, E.L. (1971) Effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology*, 18, 105–115.

Deci, E.L. (2009) Large-scale school reform as viewed from the self-determination theory perspective. *Theory and Research in Education*, 7, 247–255.

Deci, E.L., Betley, G., Kahle, J., Abrams, L. and Porac, J. (1981) When trying to win: Competition and intrinsic motivation. *Personality and Social Psychology Bulletin*, 7, 79–83.

Deci, E.L. and Cascio, W.F. (1972) Changes in intrinsic motivation as a function of negative feedback and threats. Paper presented at the Eastern Psychological Association, Boston, April.

Deci, E.L., Connell, J.P. and Ryan, R.M. (1989) Self-determination in a work organization. *Journal of Applied Psychology*, 74, 580–590.

Deci, E.L., Eghrari, H., Patrick, B.C. and Leone, D.R. (1994) Facilitating internalization: The self-determination theory perspective. *Journal of Personality*, 62, 119–142.

Deci, E.L., Koestner, R. and Ryan, R.M. (1999) A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, 125, 627–668.

Deci, E.L., La Guardia, J.G., Moller, A.C., Scheiner, M.J. and Ryan, R.M. (2006) On the benefits of giving as well as receiving autonomy support: Mutuality in close friendships. *Personality and Social Psychology Bulletin*, 32, 313–327.

Deci, E.L. and Ryan, R.M. (1980) The empirical exploration of intrinsic motivational processes. In L. Berkowitz (ed.), *Advances in Experimental Social Psychology*, 13, 39–80. New York: Academic Press.

Deci, E.L. and Ryan, R.M. (1985a) The General Causality Orientations Scale: Self-determination in personality. *Journal of Research in Personality*, 19, 109–134.

Deci, E.L. and Ryan, R.M. (1985b) *Intrinsic Motivation and Self-determination in Human Behavior*. New York: Plenum.

Deci, E.L. and Ryan, R.M. (1995) Human autonomy: The basis for true self-esteem. In M. Kernis (ed.),

Efficacy, Agency, and Self-esteem, pp. 31–49. New York: Plenum.

Deci, E.L. and Ryan, R.M. (2000) The 'what' and 'why' of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, 11, 227–268.

Deci, E.L., Ryan, R.M., Gagné, M., Leone, D.R., Usunov, J. and Kornazheva, B.P. (2001) Need satisfaction, motivation, and well-being in the work organizations of a former Eastern Bloc country. *Personality and Social Psychology Bulletin*, 27, 930–942.

Deci, E.L., Schwartz, A.J., Sheinman, L. and Ryan, R.M. (1981) An instrument to assess adults' orientations toward control versus autonomy with children: Reflections on intrinsic motivation and perceived competence. *Journal of Educational Psychology*, 73, 642–650.

Eisenberger, R. and Cameron, J. (1996) Detrimental effects of reward: Reality or myth? *American Psychologist*, 51, 1153–1166.

Feinberg, O., Kaplan, H., Assor, A. and Kanat-Maymon, Y. (2007) The concept of 'internalization' (based on SDT) as a guide for a school reform program. Paper presented at the Third International Conference on Self-Determination Theory, Toronto, May.

Fortier, M.S., Sweet, S.N., O'Sullivan, T.L. and Williams, G.C. (2007) A self-determination process model of physical activity adoption in the context of a randomized controlled trial. *Psychology of Sport and Exercise*, 8, 741–757.

Gambone, M.A., Klem, A.M., Summers, J.A., Akey, T.A. and Sipe, C.L. (2004) *Turning the Tide: The Achievements of the First Things First Education Reform in the Kansas City, Kansas Public School District*. Philadelphia, PA: Youth Development Strategies, Inc.

Goldstein, K. (1939) *The Organism*. New York: American Book Co.

Grolnick, W.S. (2003) The psychology of parental control. Mahwah, NJ: Erlbaum.

Grolnick, W.S. and Ryan, R.M. (1987) Autonomy in children's learning: An experimental and individual difference investigation. *Journal of Personality and Social Psychology*, 52, 890–898.

Grolnick, W.S. and Ryan, R.M. (1989) Parent styles associated with children's self-regulation and competence in school. *Journal of Educational Psychology*, 81, 143–154.

Hagger, M.S. and Chatzisarantis, N.L. (2007) *Intrinsic Motivation and Self-determination in Exercise and Sport*. Champaign, IL: Human Kinetics.

Halvari, A.M. and Halvari, H. (2006) Motivational predictors of change in oral health: An experimental test of self-determination theory. *Motivation and Emotion*, 30, 295–306.

Harlow, H.F. (1950) Learning and satiation of response in intrinsically motivated complex puzzle performance by monkeys. *Journal of Comparative and Physiological Psychology*, 43, 289–294.

Hartmann, H. (1958) *Ego Psychology and the Problem of Adaptation*. New York: International Universities Press. (Originally published 1939.)

Heider, F. (1958) *The Psychology of Interpersonal Relations*. New York: Wiley.

Hodgins, H.S., Yacko, H.A. and Gottlieb, E. (2006) Autonomy and nondefensiveness. *Motivation and Emotion*, 30, 283–293.

Jang, H., Reeve, J., Ryan, R.M. and Kim, A. (2009) Can self-determination theory explain what underlies the productive, satisfying learning experiences of collectivistically oriented Korean students? *Journal of Educational Psychology*, 101, 644–661.

Jordan, J.V. (1997) Do you believe that the concepts of self and autonomy are useful in understanding women? In J.V. Jordan (ed.), *Women's Growth in Diversity: More Writings from the Stone Center*, pp. 29–32. New York: Guilford Press.

Kage, M. and Namiki, H. (1990) The effects of evaluation structure on children's intrinsic motivation and learning. *Japanese Journal of Educational Psychology*, 38, 36–45.

Kasser, T. and Ryan, R.M. (1993) A dark side of the American dream: Correlates of financial success as a central life aspiration. *Journal of Personality and Social Psychology*, 65, 410–422.

Kasser, T. and Ryan, R.M. (1996) Further examining the American dream: Differential correlates of intrinsic and extrinsic goals. *Personality and Social Psychology Bulletin* 22, 80–87.

Knee, C.R. and Zuckerman, M. (1998) A nondefensive personality: Autonomy and control as moderators of defensive coping and self-handicapping. *Journal of Research in Personality*, 32, 115–130.

Koestner, R., Ryan, R.M., Bernieri, F. and Holt, K. (1984) Setting limits on children's behavior: The differential effects of controlling versus informational styles on intrinsic motivation and creativity. *Journal of Personality*, 52, 233–248.

La Guardia, J.G. and Patrick, H. (2008) Self-determination theory as a fundamental theory of close relationships. *Canadian Psychology*, 49, 201–209.

La Guardia, J.G., Ryan, R.M., Couchman, C. E. and Deci, E.L. (2000) Within-person variation in security of attachment: A self-determination theory perspective on attachment, need fulfillment, and well-being.

Journal of Personality and Social Psychology, 79, 367–384.

Lepper, M.R., Greene, D. and Nisbett, R.E. (1973) Undermining children's intrinsic interest with extrinsic rewards: A test of the 'overjustification' hypothesis. *Journal of Personality and Social Psychology*, 28, 129–137.

Levesque, C. and Pelletier, L.G. (2003) On the investigation of primed and chronic autonomous and heteronomous motivational orientations. *Personality and Social Psychology Bulletin*, 29, 1570–1584.

Loevinger, J. (1976) *Ego Development*. San Francisco: Jossey-Bass.

Lynch, M.F., La Guardia, J.G. and Ryan, R.M. (2009) On being yourself in different cultures: Ideal and actual self-concept, autonomy support, and well-being in China, Russia, and the United States. *Journal of Positive Psychology*, 4, 290–394.

Markus, H.R., Kitayama, S. and Heiman, R.J. (1996) Culture and basic psychological principles. In E.T. Higgins and A.W. Kruglanski (eds), *Social Psychology: Handbook of Basic Principles*, pp. 857–913. New York: Guilford Press.

Maslow, A.H. (1971) *The Farther Reaches of Human Nature*. New York: Viking Press.

Moller, A.C., Deci, E.L. and Ryan, R.M. (2006) Choice and ego-depletion: The moderating role of autonomy. *Personality and Social Psychology Bulletin*, 32, 1024–1036.

Muraven, M., Gagné, M. and Rosman, H. (2008) Helpful self-control: Autonomy support, vitality, and depletion. *Journal of Experimental Social Psychology*, 44, 573–585.

Niemiec, C.P., Ryan, R.M. and Deci, E.L. (2009) The path taken: Consequences of attaining intrinsic and extrinsic aspirations in post-college life. *Journal of Research in Personality*, 43, 291–306.

Nix, G.A., Ryan, R.M., Manly, J. B. and Deci, E.L. (1999) Revitalization through self-regulation: The effects of autonomous and controlled motivation on happiness and vitality. *Journal of Experimental Social Psychology*, 35, 266–284.

Patrick, H., Knee, C.R., Canevello, A. and Lonsbary, C. (2007) The role of need fulfillment in relationship functioning and well-being: A self-determination theory perspective. *Journal of Personality and Social Psychology*, 92, 434–457.

Pelletier, L.G. and Sharp, E. (2008) Persuasive communication and proenvironmental behaviours: How message tailoring and message framing can improve integration of behaviors through self-determined motivation. *Canadian Psychology*, 49, 210–217.

Pelletier, L.G., Tuson, K.M. and Haddad, N.K. (1997) Client motivation for therapy scale: A measure of intrinsic motivation, extrinsic motivation, and amotivation for therapy. *Journal of Personality Assessment*, 68, 414–435.

Pfander, A. (1910/1967) *Phenomenology of Willing and Motivation*. Chicago: Northwestern.

Piaget, J. (1971) *Biology and Knowledge*. Chicago: University of Chicago Press.

Pritchard, R.D., Campbell, K.M. and Campbell, D.J. (1977) Effects of extrinsic financial rewards on intrinsic motivation. *Journal of Applied Psychology*, 62, 9–15.

Przybylski, A.K., Ryan, R.M. and Rigby, C.S. (2009) The motivating role of violence in video games. *Personality and Social Psychology Bulletin*, 35, 243–259.

Reeve, J., Jang, H., Hardre, P. and Omura, M. (2002) Providing a rationale in an autonomy-supportive way as a strategy to motivate others during an uninteresting activity. *Motivation and Emotion*, 26, 183–207.

Reis, H.T., Sheldon, K.M., Gable, S.L., Roscoe, J. and Ryan, R.M. (2000) Daily well-being: The role of autonomy, competence, and relatedness. *Personality and Social Psychology Bulletin*, 26, 419–435.

Rogers, C. (1963) The actualizing tendency in relation to 'motives' and to consciousness. In M.R. Jones (ed.), *Nebraska Symposium on Motivation*, 11, 1–24. Lincoln, NE: University of Nebraska Press.

Roth, G., Assor, A., Niemiec, C.P., Ryan, R.M. and Deci, E.L. (2009) The emotional and academic consequences of parental conditional regard: Comparing conditional positive regard conditional negative regard and autonomy support as parenting practices. *Developmental Psychology*, 45, 1119–1142.

Rotter, J. (1966) Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs*, 80 (1, Whole No. 609), 1–28.

Rozanski, A. (2005) Integrating psychologic approaches into the behavioral management of cardiac patients. *Psychosomatic Medicine*, 67, S67–S73.

Ryan, R.M. (1982) Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory. *Journal of Personality and Social Psychology*, 43, 450–461.

Ryan, R.M. (1995) Psychological needs and the facilitation of integrative processes. *Journal of Personality*, 63, 397–427.

Ryan, R.M., Bernstein, J.H. and Brown, K.W. (2010) Weekends, work, and well-being: Psychological need satisfactions and day of the week effects on mood, vitality, and physical symptoms. *Journal of Social and Clinical Psychology*, 29, 95–122.

Ryan, R.M. and Brown, K.W. (2005) Legislating competence: The motivational impact of high stakes testing as an educational reform. In C. Dweck and A.E. Elliot (eds), *Handbook of Competence*, pp. 354–374. New York: Guilford Press.

Ryan, R.M. and Connell, J.P. (1989) Perceived locus of causality and internalization: Examining reasons for acting in two domains. *Journal of Personality and Social Psychology*, 57, 749–761.

Ryan, R.M., Connell, J.P. and Deci, E.L. (1985) A motivational analysis of self-determination and self-regulation in education. In C. Ames and R.E. Ames (eds), *Research on Motivation in Education: The Classroom Milieu*, pp. 13–51. New York: Academic Press.

Ryan, R.M. and Deci, E.L. (2001) On happiness and human potentials: A review of research on hedonic and eudaimonic well-being. In S. Fiske (ed.), *Annual Review of Psychology*, 52, 141–166. Palo Alto, CA: Annual Reviews, Inc.

Ryan, R.M. and Deci, E.L. (2004) Autonomy is no illusion: Self-determination theory and the empirical study of authenticity, awareness, and will. In J. Greenberg, S.L. Koole, and T. Pyszczynski (eds), *Handbook of Experimental Existential Psychology*, pp. 449–479. New York: Guilford Press.

Ryan, R.M. and Deci, E.L. (2008) From ego depletion to vitality: Theory and findings concerning the facilitation of energy available to the self. *Social and Personality Psychology Compass*, 2, 702–717.

Ryan, R.M., Deci, E.L., Grolnick, W.S. and La Guardia, J.G. (2006) The significance of autonomy and autonomy support in psychological development and psychopathology. In D. Cicchetti and D.J. Cohen (eds), *Developmental Psychopathology, Vol 1: Theory and Method*, 2nd Edition, pp. 795–849. New Jersey: Wiley.

Ryan, R.M. and Frederick, C.M. (1997) On energy, personality, and health: Subjective vitality as a dynamic reflection of well-being. *Journal of Personality*, 65, 529–565.

Ryan, R.M., Mims, V. and Koestner, R. (1983) Relation of reward contingency and interpersonal context to intrinsic motivation: A review and test using cognitive evaluation theory. *Journal of Personality and Social Psychology*, 45, 736–750.

Ryan, R.M., Plant, R.W. and O'Malley, S. (1995) Initial motivations for alcohol treatment: Relations with patient characteristics, treatment involvement and dropout. *Addictive Behaviors*, 20, 279–297.

Ryan, R.M., Rigby, S. and King, K. (1993) Two types of religious internalization and their relations to religious orientations and mental health. *Journal of Personality and Social Psychology*, 65, 586–596.

Ryan, R.M., Rigby, C.S. and Przybylski, A. (2006) Motivation pull of video games: A self-determination theory approach. *Motivation and Emotion*, 30, 347–365.

Ryan, R.M., Sheldon, K.M., Kasser, T. and Deci, E.L. (1996) All goals are not created equal: An organismic perspective on the nature of goals and their regulation. In P.M. Gollwitzer and J.A. Bargh (eds), *The Psychology of Action: Linking Cognition and Motivation to Behavior*, pp. 7–26. New York: Guilford.

Scott, W.E., Jr. (1975) The effects of extrinsic rewards on 'intrinsic motivation': A critique. *Organizational Behavior and Human Performance*, 15, 117–129.

Sheldon, K.M. and Krieger, L.S. (2007) Understanding the negative effects of legal education on law students: A longitudinal test of self-determination theory. *Personality and Social Psychology Bulletin*, 33, 883–897.

Thayer, R.E. (2001) *Calm Energy: How People Regulate Mood with Food and Exercise*. New York: Oxford University Press.

Tooby, J. and Cosmides, L. (1992) The psychological foundations of culture. In J.H. Barkow, L. Cosmides, and J. Tooby (eds), *The Adapted Mind*, pp. 19–136. New York: Oxford University Press.

Vallerand, R.J. (1997) Toward a hierarchical model of intrinsic and extrinsic motivation. In M.P. Zanna (ed.), *Advances in Experimental Social Psychology*, 29, 271–360. San Diego: Academic Press.

Vansteenkiste, M., Simons, J., Lens, W., Sheldon, K.M. and Deci, E.L. (2004) Motivating learning, performance, and persistence: The synergistic effects of intrinsic goal contents and autonomy-supportive contexts. *Journal of Personality and Social Psychology*, 87, 246–260.

Vlachopoulos, S.P. (2009) Prologue. *Hellenic Journal of Psychology*, 6, vii–x.

Weinstein, N. and Ryan, R.M. (2010) When helping helps: Autonomous motivation for prosocial behavior and its influence on well-being for the helper and recipient. *Journal of Personality and Social Psychology*, 98, 222–244.

White, R.W. (1959) Motivation reconsidered: The concept of competence. *Psychological Review*, 66, 297–333.

White, R.W. (1963) *Ego and Reality in Psychoanalytic Theory*. New York: International Universities Press.

Williams, G.C., Cox, E.M., Hedberg, V. and Deci, E.L. (2000) Extrinsic life goals and health risk behaviors in adolescents. *Journal of Applied Social Psychology*, 30, 1756–1771.

Williams, G.C., Deci, E.L. and Ryan, R.M. (1998) Building health-care partnerships by supporting

autonomy. In A.L. Suchman, P. Hinton-Walker, and R. Botelho (eds), *Partnerships in Healthcare*, pp. 67–87. Rochester, NY: University of Rochester Press.

Williams, G.C., Gagné, M., Ryan, R.M. and Deci, E.L. (2002) Facilitating autonomous motivation for smoking cessation. *Health Psychology*, 21, 40–50.

Williams, G.C., Grow, V.M., Freedman, Z., Ryan, R.M. and Deci, E.L. (1996) Motivational predictors of weight loss and weight-loss maintenance. *Journal of Personality and Social Psychology*, 70, 115–126.

Williams, G.C., McGregor, H.A., Sharp, D., Kouides, R.W., Levesque, C., Ryan, R.M. and Deci, E.L. (2006) A self-determination multiple risk intervention trial to improve smokers' health. *Journal of General Internal Medicine*, 21, 1288–1294.

Williams, G.C., McGregor, H.A., Zeldman, A., Freedman, Z.R. and Deci, E.L. (2004) Testing a self-determination theory process model for promoting glycemic control through diabetes self-management. *Health Psychology*, 23, 58–66.

Williams, G.C., Rodin, G.C., Ryan, R.M., Grolnick, W.S. and Deci, E.L. (1998) Autonomous regulation and adherence to long-term medical regimens in adult outpatients. *Health Psychology*, 17, 269–276.

Zeldman, A., Ryan, R.M. and Fiscella, K. (2004) Client motivation, autonomy support and entity beliefs: Their role in methadone maintenance treatment. *Journal of Social and Clinical Psychology*, 23, 675–696.

Zuckerman, M., Porac, J., Lathin, D., Smith, R. and Deci, E.L. (1978) On the importance of self-determination for intrinsically motivated behavior. *Personality and Social Psychology Bulletin*, 4, 443–446.

Zuroff, D.C., Koestner, R., Moskowitz, D.S., McBride, C., Marshall, M. and Bagby, M. (2007) Autonomous motivation for therapy: A new common factor in brief treatments for depression. *Psychotherapy Research*, 17, 137–147.

The Theory of Planned Behavior

Icek Ajzen

ABSTRACT

This chapter describes the theory of planned behavior (TPB), a prominent reasoned action model, its conceptual foundation, its intellectual history, and the research it has generated. From its roots in propositional control and expectancy theory, the TPB emerged as a major framework for understanding, predicting, and changing human social behavior. According to the theory, intention is the immediate antecedent of behavior and is itself a function of attitude toward the behavior, subjective norm, and perceived behavioral control; and these determinants follow, respectively, from beliefs about the behavior's likely consequences, about normative expectations of important others, and about the presence of factors that control behavioral performance. Empirical support for the theory comes from a host of correlational studies demonstrating its ability to predict intentions and behavior as well as from interventions showing that changes in behavioral, normative, and control beliefs can produce changes in intentions, and that these changes in intentions are reflected in subsequent behavior. The chapter also considers the TPB's reasoned action approach in the context of recent work on automatic, nonconscious processes in human social behavior. It is argued that insight into automaticity can complement the

understanding of behavior provided by a reasoned action approach.

INTRODUCTION

The tenets of a reasoned action approach to human behavior strike a familiar chord. We are introspectively aware of the thoughts and feelings that lead up to our decisions and we find in these processes a convincing explanation for our behavior. By this account, the immediate causes of human social behavior are neither mysterious nor outside conscious awareness. Behavior is performed not automatically or mindlessly but follows reasonably and consistently from the behavior-relevant information available to us. Ever since I entered graduate school at the University of Illinois in 1966 and began to work with Martin Fishbein, this reasoned action assumption has guided my theoretical approach and empirical research.

Many contemporary models of human social behavior bear the hallmarks of a reasoned action approach. Among these models are Bandura's (1986, 1997) social cognitive theory, Triandis's (1972) theory of subjective culture and interpersonal relations, the health belief model (Rosenstock et al., 1994), goal setting theory (Locke and Latham, 1994), the information-motivation-behavioral skills model (Fisher and Fisher, 1992), and the technology acceptance model (Davis et al., 1989). The discussion in this chapter, however, focuses on what is perhaps the most influential reasoned action approach, the theory of reasoned action (TRA), which I developed in close collaboration with Martin Fishbein (Ajzen and Fishbein, 1980; Fishbein and Ajzen, 1975) and my extension of this model, the TPB (Ajzen, 1985, 1991a, 2005a).

The roots of the TRA can be traced, at least in part, to a confrontation with radical behaviorism and its law of effect. According to operant conditioning principles, behaviors followed by rewarding events are reinforced and behaviors followed by punishing events are weakened. This process is assumed to be automatic, requiring neither conscious awareness of behavior-outcome contingencies nor any other higher-order cognitive mediation. Yet there is scant evidence in human adults for operant conditioning of this kind or, for that matter, for classical conditioning without awareness (see Brewer, 1974).

At the University of Illinois I was exposed to, and greatly influenced by, the ideas of Don Dulany who was working in the area of verbal learning. Dulany (1962, 1968) was struck by the stark contrast between the behaviorist explanation of behavior and by our intuitive understanding of its determinants. In a series of laboratory experiments, Dulany set out to test the contrasting views. Participants in these experiments were shown pairs of sentences, one pair on each trial, and were asked to choose one sentence and read it aloud. They performed this task while seated in a chamber that was kept at a

constant temperature of 110 degrees Fahrenheit and 35 percent humidity. Sentences containing certain words were followed, with predetermined frequencies, by either a hot stream of air across the participant's face, a cool stream of air, or a stream of air at chamber temperature.

Now, according to a strict interpretation of operant conditioning principles, the hot blast of air – being a punishing event – should reduce the likelihood of the verbal responses that preceded it; that is, a hot blast of air should reduce the likelihood that participants will, on subsequent trials, choose sentences containing the stimulus words that were followed by the punishing event. Conversely, the cool blast of air should serve to reinforce emission of the preceding type of sentences. Dulany, however, believed that the affectively positive or negative events did not strengthen or weaken response tendencies directly. Instead, he proposed that their effects were mediated by higher mental processes, specifically by the participant's interpretation of the events' significance. To test this idea, he manipulated – orthogonally to the cool, neutral, or hot stream of air – instructions about the event's significance. Some participants were told that the stream of air indicated a correct response, some that it indicated neither a correct nor an incorrect response, and some that it indicated an incorrect response.

Dulany assumed that, in this situation, participants could form two hypotheses. First, they could come to believe that certain verbal responses (sentence choices) were followed by a certain event (a cool, neutral, or hot stream of air) with some degree of probability. He termed this belief the *hypothesis of the distribution of reinforcement*. Second, participants could form the belief that the stream of air meant that they had just done what they were supposed to do, what they were not supposed to do, or what they were neither supposed to do nor avoid doing. He termed this belief the *hypothesis of the significance of a reinforcer*. These two hypotheses provide the foundation for

Dulany's *theory of propositional control*. According to the theory, people form a conscious intention to select a certain response, and it is this behavioral intention (*BI*) that determines the response actually emitted. The intention itself is a function of two factors. The first is the hypothesis of the distribution of reinforcement (*RHd*) weighted by the subjective value of the reinforcer (*RSv*). The more strongly people believe that a certain response will lead to a certain outcome, and the more positively they value that outcome, the stronger their intention to produce the response in question. The second factor influencing intentions is the hypothesis of the distribution of reinforcement weighted by the hypothesis of the reinforcer's significance. This compound variable is termed the *behavioral hypothesis (BH)*. The behavioral hypothesis is a function of the extent to which a response is perceived to produce a certain outcome and the extent to which this outcome is believed to indicate a correct response, that is, a response that is expected of them. Intentions to produce the expected response are strengthened to the extent that participants are *motivated to comply (MC)* with what they think they are supposed to do. Dulany's (1968) theory of propositional control in relation to behavioral intentions is shown symbolically in Equation (21.1); the actual response is assumed to be a direct function of the behavioral intention.

$$BI = (RHd) (RSv) + (BH) (MC) \quad (21.1)$$

Dulany's theory does not, of course, imply that reinforcement is irrelevant; it only suggests that the effect of reinforcement on behavior is mediated by the beliefs people form about reinforcement contingencies and what these contingencies signify. Consistent with the theory, intentions to select sentences of certain kinds had a correlation of 0.94 with the actual sentences selected, and behavioral intentions were predicted with a multiple correlation of 0.88 from (*RHd*) (*RSv*) and (*BH*) (*MC*). Interestingly, instructions regarding the significance of the

reinforcement had a much stronger effect on intentions and actual response selection than had the affective value of the reinforcement. In other words, participants were guided in their response selection primarily by what sentences they thought the experimenter wanted them to read aloud, rather than by the affective consequences of their choices. Thus, they were willing to endure a hot stream of air to their faces while seated in a hot environment if they thought that this was expected of them in the experiment (and if they were motivated to comply).

THEORY OF REASONED ACTION

To me, the results of Dulany's experiments were quite compelling and difficult to reconcile with automatic strengthening of responses by response-contingent events. His work suggested instead that human behavior is mediated and indeed controlled by higher mental processes. Dulany (1968) attributed a causal or instrumental role to conscious, volitional processes, although he recognized that, with practice, voluntary responses will tend to habituate and conscious rules may become unconscious. The causal role of conscious awareness and volition in human social behavior remains controversial however (see Bargh and Chartrand, 1999; Wegner and Wheatley, 1999); I will return to this topic later in this chapter.

Whether always consciously experienced or not, the cognitive processes identified in Dulany's theory of propositional control have their counterparts in the TRA (Ajzen and Fishbein, 1980; Fishbein, 1967a; Fishbein and Ajzen, 1975), the forerunner of the TPB (Ajzen, 1991b). Consider, first, the hypothesis of the distribution of reinforcement. In the TRA, this hypothesis is termed a *behavioral belief*. It is defined as a person's subjective probability that performing a certain behavior will produce a particular outcome, and the subjective value of the reinforcer is designated the person's evaluation of that outcome.

Of course, most behaviors of interest to social psychologists are capable of producing more than one outcome and therefore, in the TRA, people are assumed to hold multiple behavioral beliefs each of which links performance of the behavior to a different outcome. A study on alcohol and drug use among college students (Armitage et al., 1999: 306, Table 2) provides a concrete example. In this study, the following behavioral beliefs about using alcohol and marijuana were held with relatively high frequency: “Makes me more sociable,” “Leads to me having poorer physical health,” “Will result in my becoming dependent on it,” “Will result in me getting into trouble with authority,” and “Makes me feel good.”

The second component in Dulany’s model, the behavioral hypothesis, is termed a *normative belief* in the TRA. It is defined as a person’s subjective probability that a particular normative referent (the experimenter in Dulany’s case) wants the person to perform a given behavior. As in Dulany’s model, this normative belief is weighted (multiplied) by the person’s motivation to comply with the referent’s perceived expectation. However, in the TRA it is assumed that people can hold normative beliefs with respect to more than one referent individual or group. Commonly identified referents are a person’s spouse or partner, close family, friends, and, depending on the behavior under consideration, coworkers, health professionals, and law enforcement authorities.

Attitude toward a behavior: the expectancy-value model

As noted above, people generally hold a number of behavioral beliefs in relation to any given behavior. Each of these beliefs links the behavior to an outcome, and each outcome has a certain subjective value. In the TRA it is assumed that these behavioral beliefs and outcome evaluations combine to produce an overall positive or negative *attitude toward the behavior*. Specifically, the

subjective value or evaluation of each outcome contributes to the attitude in direct proportion to the person’s subjective probability that the behavior produces the outcome in question. This *expectancy-value* model of attitude is shown in Equation (21.2), where A represents attitude toward a behavior, b_i is the subjective probability or belief that the behavior produces outcome i , e_i is the evaluation of outcome i , and the sum is over the total number of behavioral beliefs.

$$A \propto \sum b_i e_i \quad (21.2)$$

The multiplicative combination of subjective probabilities (beliefs) and values (evaluations) in the TRA’s attitude model can be traced to general theorizing about the formation and structure of social attitudes. In a paper on attitudes and motivation, Peak hypothesized that the attitude toward any object “is related to the ends which it serves, that is, to its consequences” (1955: 153) and that the attitude therefore is “some function of (1) the judged probability that the object leads to good or bad consequences, and (2) the intensity of the affect expected from those consequences” (1955: 154). She proposed that the judged probability of a given consequence be multiplied by its expected valence and that the products be summed across all consequences to provide an estimate of the affect or evaluation associated with the object, that is, to provide an estimate of attitude (see also Carlson, 1956; Rosenberg, 1956).

Whereas Peak arrived at her attitude model by a consideration of motivation and its implication for attitude structure, Fishbein (1963) proposed a very similar model but supplied it with a concept formation and learning theory foundation (Fishbein, 1967b). His *summation model* of attitudes inspired the expectancy-value model in the TRA. In an initial test of his model, Fishbein (1963) examined the relation between beliefs about and attitudes toward African Americans. In contrast to the approach taken by Rosenberg (1956), who constructed a predetermined set

of end states with respect to which beliefs were assessed, Fishbein elicited people's own beliefs in a free-response format. The participants were read, five times, a list of different groups of people, including "Negroes," and were asked to respond with a word they believed to be characteristic of the group in question. The ten attributes mentioned most frequently in relation to Negroes were selected for further investigation. Among these attributes were dark skin, curly hair, musical, athletic, uneducated, and hard workers.

In the second part of the study, evaluations of the ten attributes were assessed by means of five bipolar evaluative scales, including *good–bad*, *clean–dirty*, and *wise–foolish*. To measure belief strength, participants were asked to rate, for each of the ten attributes, the likelihood that blacks have the attribute in question on five probabilistic scales (e.g., *unlikely–likely*, *probably–improbably*, *false–true*). Finally, attitudes toward African Americans were assessed by means of the same five evaluative scales that had been used to measure attribute evaluations. Belief strength was multiplied by attribute evaluation and the products were summed over the ten attributes. This expectancy-value composite was found to correlate 0.80 with the direct attitude measure.

The cognitive foundation of attitudes

In sum, several lines of theorizing in the 1950s and 1960s converged on the expectancy-value model of attitudes (for reviews, see Dabholkar, 1999; Feather, 1982).¹ However, the expectancy-value model of attitude in the TRA has certain features that are not necessarily shared by other expectancy-value approaches to attitude formation and structure.

Causal effects of beliefs on attitudes

Perhaps most important from a theoretical perspective are differences between the

TRA's attitude construct and other attitudinal approaches in the assumptions they make about the nature of the relations between beliefs and attitudes. Following Rosenberg and Hovland (1960), many investigators consider beliefs (or cognitions) and evaluations (or affect) to be two components of attitude, together with behavioral inclinations (or conation), the third component. According to this approach, attitudes can be inferred from cognitive, affective, or conative responses to the attitude object, but there is no assumption that one component causally precedes another. The tripartite model merely stipulates that there will be pressure for the three components to be evaluatively consistent with each other (Rosenberg, 1965). By way of contrast, in the TRA, beliefs that performing a behavior will lead to certain outcomes, together with the evaluations of these outcomes, are assumed to produce a favorable or unfavorable attitude toward the behavior in question. And, as we shall see below, this attitude is further assumed to have a causal effect on intentions to engage in the behavior, that is, on conation.

It is important to note, however, that not all potential outcomes of a behavior are expected to influence attitudes. According to the TRA's expectancy-value model, only beliefs that are *readily accessible* in memory determine the prevailing attitude. This limits the number of beliefs that provide the basis for an observed attitude toward a behavior and it also implies that appropriate means must be employed to identify the readily accessible beliefs. It is not sufficient, for example, to simply provide participants in a study with a list of belief statements constructed by the investigator. Many of these statements may not represent beliefs that are readily accessible, and some accessible beliefs may be missing. Although responses to a priori set of belief statements can be used to infer underlying attitudes, it would be a mistake to assume that these responses necessarily provide information about accessible beliefs that provide the causal basis for the attitude.

Subjective norm

As noted earlier, in the TRA, Dulany's (1968) behavioral hypothesis is termed normative belief, that is, a belief that a particular referent other wants us to perform a given behavior. It is assumed that the normative beliefs regarding different social referents combine to produce an overall perceived social pressure or *subjective norm*. Drawing an analogy to the expectancy-value model of attitude toward a behavior, the prevailing subjective norm (*SN*) is determined by the total set of readily accessible normative beliefs concerning the expectations of important referents. Specifically, the strength of each normative belief (n_i) is weighted by motivation to comply (m_i) with referent i , and the products are aggregated across all accessible referents, as shown in Equation (21.3).

$$SN \propto \sum n_i m_i \quad (21.3)$$

We can form beliefs as to what is expected of us by being told or by inferring what important others want us to do (*injunctive norms*), or based on the observed or inferred actions of those important social referents (*descriptive norms*) (see Cialdini et al., 1990; Fishbein and Ajzen, 2010). Subjective norms are conceptually independent of attitudes toward the behavior. People can, in principle, hold favorable attitudes toward a given behavior, yet perceive social pressure not to perform it; they can hold negative attitudes toward the behavior and favorable subjective norms; or their attitudes and subjective norms may coincide. However, in practice, personal attitudes and subjective norms are rarely, if ever, completely orthogonal to each other. This is due to the fact that many events are likely to lead to the formation of parallel behavioral and normative beliefs. Consider, for example, the publication in the popular press of the results of new medical research indicating that a low-calorie diet prolonged the life of laboratory mice by 35 percent. People exposed to this information may form the behavioral belief that eating a low-calorie

diet is likely to prolong their own lives and, at the same time, also form the normative beliefs that their partners and doctors would want them to eat a low-calorie diet. As a result, attitudes toward eating a low-fat diet and subjective norms with respect to this behavior are likely to correlate with each other.

The cognitive foundation of subjective norms

The product term in the subjective norm model (see Equation [21.3]) implies that the effect of normative beliefs on subjective norms is moderated by motivation to comply. The belief that an important social referent wants us to perform a particular behavior increases perceived social pressure to do so only to the extent that we are motivated to comply with the referent in question. Similar to tests of the expectancy-value model of attitudes, tests of the subjective norm model usually involve correlating the summed products of normative belief strength times motivation to comply with a direct measure of subjective norm. Direct measures are obtained by means of items that ask participants how likely it is that important others think they should perform a behavior of interest, how likely it is that important others themselves perform or would perform the behavior, and so forth. Similar items are formulated to assess normative beliefs with respect to particular social referents. That is, participants indicate how likely they think it is that certain persons or groups of individuals (e.g., spouse, coworkers) want them to perform the behavior or themselves would perform the behavior. (The normative referents that readily come to mind in a given research population are identified by means of elicitation in a free-response format.) Finally, participants are asked to rate how motivated they are to comply with each of the normative referents. The measures of normative belief strength are multiplied by the corresponding measures of motivation to

comply, and the products are summed to produce a normative belief aggregate which is then correlated with the direct subjective norm measure (for illustrations, see Ajzen and Driver, 1991; Conner et al., 1998).

Empirical evidence is supportive of a correlation between an aggregate of normative beliefs on one hand and perceived social pressure or subjective norm on the other. The strength of this correlation is conveyed in a meta-analysis of research with the TPB (Armitage and Conner, 2001). Across 34 sets of data dealing with diverse kinds of behavior, the mean correlation between normative beliefs and subjective norms was 0.50. However, a number of investigators have reported that this correlation is attributable to the measure of normative belief strength and that taking motivation to comply into account does little to improve the correlation or may even lower it slightly (e.g., Ajzen and Driver, 1991; Budd et al., 1984).² One possible explanation of these findings is that people generally tend to be motivated to comply with their social referents and there is therefore relatively little meaningful variance in motivation to comply measures. Under these circumstances, multiplying normative beliefs by motivation to comply can do little to improve prediction of subjective norms.

Historical and theoretical context

Soon after its publication, the TRA began to stimulate a great deal of empirical research designed to predict and explain behaviors in various domains (see Sheppard et al., 1988 for an early review of this literature). The theory's broad appeal must be understood in the context of previous failures to find a strong link between verbal attitudes and actual behavior. Prior to development of the TRA, much attitude theory and research dealt with attitudes toward such general concepts as institutions, policies, racial or ethnic groups, and other broad objects. Investigators assumed that such attitudes

would be predictive of any behavior toward the object of the attitude, but empirical research challenged this assumption. Attitudes toward African Americans were related neither to conformity with the judgments made by African Americans (Himelstein and Moore, 1963) nor to willingness to have a picture taken with an African American (De Fleur and Westie, 1958; Linn, 1965); job satisfaction attitudes failed to predict job performance, absenteeism, and turnover (e.g., Bernberg, 1952; Vroom, 1964); attitudes toward labor unions failed to predict attendance at labor union meetings (Dean, 1958), and so forth. In a highly influential review of this literature, Wicker (1969) called attention to the inconsistency between verbal attitudes and overt behaviors and, like several theorists before him (e.g., Blumer, 1955; Deutscher, 1966; Festinger, 1964), questioned the utility of the attitude concept.

Compatibility

When I began my work with Martin Fishbein in the 1960s, we faced the challenge of explaining why verbal attitudes failed to predict actual behavior. In our work on the TRA, we (Ajzen, 1982; Ajzen and Fishbein, 1980; Fishbein and Ajzen, 1975) drew a distinction between two kinds of attitudes: general attitudes toward physical objects, institutions, groups, policies, and events – attitudes of the kind studied in most prior research; and attitudes toward performing particular behaviors, whether related to matters of health and safety (exercising, using contraception, getting a cancer screening, wearing a safety helmet, eating a healthy diet), race relations (hiring a member of a minority group, inviting an outgroup member to a party), politics (participating in an election, donating money to a political candidate, voting for a candidate), the environment (using public transit, recycling, conserving energy), or any other domain. We formulated the principle of correspondence or compatibility (Ajzen, 1987; Ajzen and Fishbein, 1977) to help clarify the nature of the relation

between verbal attitudes and overt actions. According to this principle, attitudes and behavior correlate with each other to the extent that they are compatible in terms of their action, target, context, and time elements. Measures of behavior typically involve a specific action (e.g., making friends) and target (e.g., a gay person), and often also a specific context (e.g., at school) and time frame (e.g., in the next six months). By way of contrast, general attitudes (e.g., toward gays) identify only the target; they do not specify any particular action, context, or time element. We proposed that this lack of compatibility, especially in the action element, was responsible for the reported low and often nonsignificant relations between general attitudes and specific behaviors.

This is not to say, however, that general attitudes toward targets are irrelevant when it comes to the prediction of behavior. According to the principle of compatibility, general attitudes predict broad patterns or aggregates of behavior (Ajzen, 2005a; Ajzen and Fishbein, 1977). When we aggregate different behaviors directed at a given target, we generalize across actions, contexts, and time elements, thus assuring compatibility with equally broad attitudes toward the target in question. Thus, attitudes toward religion and the church, though largely unrelated to individual behaviors in this domain, were shown to correlate strongly with broad patterns of religious behavior (Fishbein and Ajzen, 1974); and attitudes toward protection of the environment predicted an aggregate of individual behaviors protective of the environment (Weigel and Newman, 1976).

The TRA succeeded where general attitudes had failed: it offered a way to use attitudes as a means to predict and explain individual behaviors. According to the theory, and in line with the principle of compatibility, individual behaviors can be predicted from attitudes toward the particular behavior of interest, and this is indeed how the attitudinal component in the TRA is defined. In addition, the theory went beyond the

impact of personal attitudes by considering the role of perceived social norms, again in relation to the particular behavior of interest. Thus, the theory stipulated that the intention to perform a particular behavior is a joint function of a favorable or unfavorable attitude toward the behavior and of a subjective norm that encourages or discourages its performance.

THE THEORY OF PLANNED BEHAVIOR

When we initially formulated the TRA we explicitly confined it to behaviors over which people have complete volitional control under the assumption that this category includes most behaviors of interest to social psychologists (see Ajzen and Fishbein, 1980). However, I soon came to realize that this formulation imposed too severe a limitation on a theory designed to predict and explain all manner of socially significant behavior. Many behaviors, even if in principle under volitional control, can pose serious difficulties of execution. Consider, for example, a study on physical exercise among cancer patients (Courneya et al., 2000). The investigators reported that medical complications following high-dose chemotherapy and bone marrow transplantation present significant challenges to patients' ability to adhere to a recommended exercise regimen. I decided that, to accommodate behaviors over which people may have limited volitional control, the TRA model had to be expanded by taking degree of control over the behavior into account (Ajzen, 1985). The TPB (Ajzen, 1987, 1991b, 2005a) was designed to accomplish this goal.

Behavioral control

Many factors, internal and external, can impair (or facilitate) performance of a given behavior: the extent to which people possess

the requisite information, mental and physical skills and abilities, the availability of social support, emotions and compulsions, and absence or presence of external barriers and impediments (see Ajzen, 2005a: Chapter 5). People should be able to act on their intentions to the extent that they have the information, intelligence, skills, abilities, and other internal factors required to perform the behavior and to the extent that they can overcome any external obstacles that may interfere with behavioral performance. The degree of actual behavioral control is thus expected to moderate the effect of intentions on behavior. When control is uniformly high such that virtually everybody can perform the behavior if so inclined, intentions alone should be sufficient to predict behavior, but when degree of control varies among individuals, intentions and control should interact to jointly affect behavioral performance. Individuals who intend to perform the behavior and who have a high degree of control over it should be most likely to perform it.

Perceived behavioral control

Perhaps less self-evident than the importance of actual control, but more interesting from a psychological perspective, is the role of *perceived* behavioral control – the extent to which people believe that they can perform a given behavior if they are inclined to do so. The conceptualization of perceived behavioral control in the TPB owes much to Albert Bandura's work on self-efficacy (Bandura, 1977, 1986, 1997). In Bandura's social cognitive theory, people's beliefs about their capabilities to exercise control over events that affect their lives function as proximal determinants of human motivation and action. Bandura emphasized that self-efficacy is not a context-free global disposition but that, instead, it "refers to beliefs in one's capabilities to organize and execute the courses of action required to produce given attainments" (Bandura, 1997: 3). Clearly, the

concept of perceived behavioral control in the TPB, though focusing on the extent to which people believe that they are capable of, or have control over, performing a given behavior, is very similar to Bandura's conception of self-efficacy.

A considerable body of research attests to the powerful effects of self-efficacy beliefs on motivation and performance (see Bandura and Locke, 2003 for a review). The strongest evidence comes from studies in which level of self-efficacy was experimentally manipulated to observe the effect on perseverance at a task and/or on task performance. Much of this research has been conducted in situations where intentions to perform the behavior of interest can be taken as given. Under these conditions, perseverance and task performance are found to increase with perceived self-efficacy. For example, Cervone and Peake (1986) had participants work on a series of intellectual problems (anagrams or cyclical graphs) that had no solution. Prior to this task, they manipulated self-efficacy beliefs by means of the anchoring and adjustment heuristic (Tversky and Kahneman, 1974). After drawing, ostensibly at random, either a relatively high number (18) or a relatively low number (4), participants were asked to indicate whether they thought they would be able to solve more, an equal number, or fewer problems than the number they had drawn, and – as a measure of self-efficacy – how many problems they thought they would be able to solve. The high anchor was found to produce a significantly higher level of perceived self-efficacy than the low anchor. The investigators then recorded how many times participants attempted to solve problems of a given type before switching to the second task. The results showed that participants in the high anchor condition persevered significantly longer on the unsolvable task than did participants in the low anchor condition, and this effect was completely mediated by measured self-efficacy.

Because the problems in this experiment had no solution, it was impossible to assess

task performance. Other experiments have shown that manipulating self-efficacy influences not only perseverance but also actual task performance. Bandura and Adams (1977), for example, had ophidiophobics (people with a fear of snakes) undergo symbolic desensitization by visualizing threatening snake scenes while deeply relaxed. This manipulation increased the participants' sense of self-efficacy in relation to handling snakes and it also increased their subsequent ability to perform various snake-handling tasks. Self-efficacy manipulations by means of bogus performance feedback were even found to increase pain tolerance on a cold-pressor test (Litt, 1988) and performance on a physical endurance task (Weinberg et al., 1981).

Clearly, then, one way in which self-efficacy or perceived behavioral control can influence performance of difficult behaviors is by its effect on perseverance. The more people believe that they have the capacity to perform an intended behavior, the more likely they are to persevere and, therefore, to succeed. However, in the TPB, the role of perceived behavioral control goes beyond its effect on perseverance in at least two ways. First, the TPB is a general model designed to be applicable to any behavior, not only behaviors that individuals are motivated to perform. In fact, for most behaviors of interest to social psychologists, people vary greatly in their intentions. Some individuals intend to exercise, others do not; some intend to smoke marijuana, others do not; some intend to get a cancer screening, others have no intention to do so. In the TRA, intentions of this kind were said to be a function of attitudes and of subjective norms with respect to the behavior of interest. In the TPB, perceived behavioral control is added as a third determinant of behavioral intentions. Specifically, the more favorable people's attitudes and subjective norms, and the more they believe that they are capable of performing the behavior, the stronger should be their behavioral intentions. Conversely, people who do not believe that they are capable of performing the

behavior under investigation will be unlikely to form an intention to do so.

Perceived behavioral control can thus influence behavioral performance indirectly by its effects on intentions to engage in the behavior and on perseverance in the face of difficulties encountered during execution. In addition, perceived behavioral control can potentially serve as a proxy for actual control. Recall that actual control is expected to moderate the effect of intentions on behavior. However, in most TPB applications, measures of actual control are unavailable. Indeed, with respect to many behaviors, it would be difficult to identify, let alone measure, the various internal and external factors that may facilitate or inhibit behavioral performance. It is perhaps for this reason that many investigators rely on measures of perceived behavioral control. This of course assumes that perceptions of behavioral control accurately reflect the person's actual control in the situation. To the extent that perceptions of control are veridical, they can serve as a proxy for actual control and contribute to the prediction of behavior.

The cognitive foundation of perceived behavioral control

Like attitudes and subjective norms, perceptions of behavioral control are assumed to follow consistently from readily accessible beliefs, in this case beliefs about resources and obstacles that can facilitate or interfere with performance of a given behavior. Analogous to the expectancy-value model of attitudes, the power of each control factor to facilitate or inhibit behavioral performance is expected to contribute to perceived behavioral control in direct proportion to the person's subjective probability that the control factor is present. This model is shown in Equation (21.4), where PBC is perceived behavioral control, c_i is the subjective probability or belief that control factor i is present, p_i is the power of control factor i to facilitate or inhibit performance of the behavior, and

the sum is over the total number of accessible control beliefs.

$$PBC \approx \sum c_i p_i \quad (21.4)$$

In support of this model, empirical evidence shows a strong correlation between a direct measure of perceived behavioral control and a composite of control beliefs. Direct measures of perceived behavioral control are typically obtained by asking people whether they believe that they are capable of performing the behavior of interest, whether they believe that doing so is completely under their control, and so forth. Readily accessible control factors are elicited in a free-response format. To illustrate, in a study on eating a low-fat diet (Armitage and Conner, 1999), the seven most frequently mentioned control factors dealt largely with obstacles to maintaining a low-fat diet: that doing so is time-consuming, expensive, and inconvenient; that it requires strong motivation and knowledge of the fat contents of various foods; that low-fat foods must be readily available; and that high-fat foods pose temptation.

Although investigators have frequently measured control belief strength; that is, the subjective probabilities that certain control factors will be present, in only a few studies have they also secured measures of the power of these control factors to facilitate or inhibit behavioral performance. Nevertheless, the results of empirical research provide support for the proposition that perceived behavioral control can be predicted from control beliefs. For example, in an analysis of 16 of their own studies in the health domain, Gagné and Godin (2000) found a median correlation of 0.57 between control belief composites and direct measures of perceived behavioral control, and in a meta-analysis of 18 studies on a variety of different behaviors, Armitage and Conner (2001) reported a mean correlation of 0.52.

Predicting intentions and behavior

To summarize briefly, according to the TPB, human action is guided by three kinds

of considerations: readily accessible beliefs about the likely outcomes of the behavior and the evaluations of these outcomes (behavioral beliefs), readily accessible beliefs about the normative expectations and actions of important referents and motivation to comply with these referents (normative beliefs), and readily accessible beliefs about the presence of factors that may facilitate or impede performance of the behavior and the perceived power of these factors (control beliefs). In their respective aggregates, behavioral beliefs produce a favorable or unfavorable attitude toward the behavior; normative beliefs result in perceived social pressure or subjective norm; and control beliefs give rise to perceived behavioral control. In combination, attitude toward the behavior, subjective norm, and perception of behavioral control lead to the formation of a behavioral intention. As a general rule, the more favorable the attitude and subjective norm, and the greater the perceived control, the stronger the person's intention to perform the behavior in question. Finally, given a sufficient degree of actual control over the behavior, people are expected to carry out their intentions when the opportunity arises. Intention is thus assumed to be the immediate antecedent of behavior. However, because many behaviors pose difficulties of execution, it is useful to consider perceived behavioral control in addition to intention. Perceived control influences perseverance in the face of difficulties and, to the extent that it is veridical, it can serve as a proxy for actual control and contribute to the prediction of behavior.

Accessibility of beliefs

I have emphasized the idea that it is the currently accessible behavioral, normative, and control beliefs that provide the cognitive foundation for attitudes, subjective norms, and perceived control, respectively. Although belief accessibility was an important feature of my theoretical framework from the outset, I became fully aware of its various

implications only over time. In terms of explaining attitudes and behavior, the importance of belief accessibility is readily apparent. When we identify people's readily accessible beliefs we obtain a snapshot of the kinds of considerations that guide their attitudes, subjective norms, and perceptions of control and that therefore influence their intentions and actions at a given point in time. What may be less obvious is that the beliefs that are readily accessible in memory can change over time. This possibility can help explain frequently observed gaps between intentions and actions. Intentions measured at time 1 will be influenced by the beliefs that are accessible at that time. Behavior, however, is performed at a later point in time, and at that time different beliefs may have become accessible, producing different intentions. In short, intentions measured at time 1 can be expected to predict behavior at time 2 only to the extent that the same beliefs (or beliefs of equivalent valence) are readily accessible at the two points in time (Ajzen and Sexton, 1999).

Predicting behavior

Fundamental to the TPB and other reasoned action models is the idea that behavior is guided by intentions. This idea implies, first, a strong relation between intentions and behavior, though this relation can be moderated by the degree of control over performance of the behavior. Second, it implies that changes in intentions will be followed by changes in behavior. There is ample evidence for both of these propositions.

Many studies have shown that behavioral intentions account for a considerable proportion of variance in behavior. To give just one example from our own research program, in an application of the TPB to outdoor recreation activities (Hrubes et al., 2001), a correlation of 0.62 was observed between intentions to go hunting and self-reported hunting behavior. Meta-analyses of research in diverse behavioral domains – from

physical activity, health screening, and illicit drug use to playing videogames, donating blood, and smoking cigarettes – have reported mean intention-behavior correlations ranging from 0.44 to 0.62 (e.g., Armitage and Conner, 2001; Notani, 1998; Randall and Wolff, 1994; Sheppard et al., 1988). In a meta-analysis of such meta-analyses, Sheeran (2002) reported an overall mean correlation of 0.53 between intention and behavior.

As noted previously, perceived behavioral control is expected to moderate the relation between intentions and behavior such that intentions will predict behavior better when perceived control is high rather than low. Tests of this hypothesis, when undertaken at all, typically rely on multiple regression analyses in which intentions and perceived control are entered on the first step and the product of these variables on the second step. In many of these tests the interaction term does not reach conventional levels of significance and even when it does have a significant regression coefficient, it tends to account for relatively little additional variance in the prediction of behavior (see Ajzen, 1991b; Armitage and Conner, 2001; Yang-Wallentin et al., 2004).

A likely explanation for such findings is the fact that there is relatively little variance in perceived behavioral control for many behaviors studied by social psychologists. Although people vary greatly in their intentions to engage in physical activity, eat a healthy diet, donate blood, recycle glass and paper, attend church, vote in an upcoming election, drink alcohol, attend class, and so forth, most believe that they can, if they so desire, engage in these kinds of activities. Under these conditions, intentions will have good predictive validity, but we cannot expect that perceived behavioral control will exert a strong moderating effect on the correlation between intentions and behavior.

With respect to some behaviors, the situation seems to be reversed such that people generally intend to perform the behavior in question but they vary greatly in their perceived control. We noted examples of this

situation earlier in the discussion of research on self-efficacy. Participants in this research are asked to perform an intellectual or physical task, to overcome certain phobias, to tolerate pain, and so forth. It can be assumed that they try their best in these situations, that is, that they intend to perform the task to the best of their abilities. When this is the case, behavioral achievement is found to covary with perceived self-efficacy, that is, with perceived control over the behavior. Although intentions are usually not assessed in these studies, it stands to reason that measures of intention would show relatively little variance and hence would not make much of a contribution to the prediction of behavior, either as a main effect or in interaction with perceived behavioral control.

In short, we cannot expect to find a strong interaction between intentions and perceived behavioral control when there is relatively little variance in either of these factors. Only when, in the population under investigation, people vary greatly in their intentions to perform the behavior of interest and vary in their perceptions of control over the behavior, can we expect a strong moderating effect. (For a discussion of a parallel problem in relation to the interaction of beliefs and evaluations in the expectancy-value model of attitudes, see Ajzen and Fishbein, 2008.)

Causal effect of intentions on behavior

In reasoned action models, such as the TPB, intentions are assumed to be causal antecedents of corresponding behavior. The correlational nature of most empirical evidence shows that intentions can indeed be used to predict behavior, but such evidence is not definitive proof of their causal impact. There is growing evidence, however, for a causal effect of intentions on actions coming mainly from intervention studies. In a meta-analysis of 47 studies in which an intervention was shown to have had a significant effect on intentions (Webb and Sheeran, 2006), this effect was also shown to promote a change in actual behavior. On average, the interventions reviewed produced medium to large

changes in intentions (mean $d = 0.66$), followed by small to medium changes in behavior (mean $d = 0.36$).

Predicting and explaining intentions

There is also ample evidence to show that, consistent with the TPB, intentions can be predicted from attitudes, subjective norms, and perceptions of behavioral control. Two examples will serve to illustrate successful applications of the theory. In the study by Hrubes et al. (2001) mentioned earlier, the multiple correlation for the prediction of hunting intentions was 0.92, showing that attitudes, subjective norms, and perceived control accounted for 86 percent of the variance in intentions. Each of the three antecedents of intentions made a significant contribution to the prediction although attitudes were found to be most important ($\beta = 0.58$), followed by subjective norms ($\beta = 0.37$) and perceived control ($\beta = 0.07$). A different pattern of influence was observed in a study on leisure time physical activity among individuals with spinal cord injury (Latimer and Martin Ginis, 2005). The multiple correlation for the prediction of intentions was 0.78, indicating that attitudes, subjective norms, and perceptions of behavioral control accounted for 61 percent of the variance in intentions. Again, the regression coefficients were statistically significant for all three predictors. However, perhaps not surprisingly given the difficulties individuals with spinal cord injuries are likely to face, perceived control made a larger independent contribution to the prediction of exercise intentions ($\beta = 0.46$) than did either attitudes ($\beta = 0.29$) or subjective norms ($\beta = 0.27$).

It is beyond the scope of this chapter to review the large body of research attesting to the proposition that intentions can be predicted from attitudes toward the behavior, subjective norms, and perceived behavioral control. In meta-analytic syntheses covering

varied behaviors (Armitage and Conner, 2001; Cheung and Chan, 2000; Notani, 1998; Rivas and Sheeran, 2003; Schulze and Wittmann, 2003), the mean multiple correlations for the prediction of intentions ranged from 0.59 to 0.66. Meta-analyses in specific behavioral domains show similar results. In two meta-analytic reviews of research on condom use, the mean multiple correlations were found to be 0.71 (Albarracín et al., 2001) and 0.65 (Sheeran and Taylor, 1999), and in two meta-analyses of research on physical activity, the mean multiple correlations were 0.55 (Downs and Hausenblas, 2005) and 0.67 (Hagger et al., 2002). An extensive review of this literature can be found in Fishbein and Ajzen (2010, chapter 6).

REASONED ACTION

The TPB emphasizes the controlled aspects of human information processing and decision making. Its concern is primarily with behaviors that are goal-directed and steered by conscious self-regulatory processes. According to the TPB, intentions and behavior are guided by expected consequences of performing the behavior, by perceived normative pressures, and by anticipated difficulties. This focus has often been misinterpreted to mean that the theory posits an impassionate, rational actor who reviews all available information in an unbiased fashion to arrive at a behavioral decision. In reality, the theory draws a much more complex and nuanced picture.

First, there is no assumption in the TPB that behavioral, normative, and control beliefs are formed in a rational, unbiased fashion or that they accurately represent reality. Beliefs reflect the information people have in relation to the performance of a given behavior, but this information is often inaccurate and incomplete; it may rest on faulty or irrational premises, be biased by self-serving motives, or otherwise fail to reflect reality. Clearly, this is a far cry from a rational actor.

However, no matter how people arrive at their behavioral, normative, and control beliefs, their attitudes toward the behavior, their subjective norms, and their perceptions of behavioral control follow automatically and consistently from their beliefs. It is only in this sense that behavior is said to be reasoned or planned. Even if inaccurate, biased, or otherwise irrational, our beliefs produce attitudes, intentions, and behaviors consistent with these beliefs (see, e.g., Geraerts et al., 2008).

Second, there is no assumption in the TPB that people carefully and systematically review all their beliefs every time they are about to perform a behavior. On the contrary, the theory recognizes that most behaviors in everyday life are performed without much cognitive effort. Consistent with contemporary theorizing in social psychology (see Carver and Scheier, 1998; Chaiken and Trope, 1999; Petty and Cacioppo, 1986), it is assumed that the amount of information processing people engage in prior to performing a behavior varies along a continuum, from shallow to deep (Ajzen and Sexton, 1999). In-depth processing is reserved for important decisions and behaviors in novel situations that demand careful consideration of the behavior's likely consequences, the normative expectations of significant others, and the obstacles that may be encountered. When it comes to routine, everyday behaviors like eating breakfast, taking one's vitamin supplements, going to work, watching the news on TV, and so forth, no careful deliberation is required or postulated. Attitudes, subjective norms, and perceptions of control as well as intentions in relation to these kinds of behaviors are assumed to guide behavior implicitly without cognitive effort and often below conscious awareness (see Ajzen and Fishbein, 2000 for a discussion of these issues).

Habituation and automaticity in social behavior

Notwithstanding the above qualifications, the reasoned action approach represented in the

TPB stands in contrast to recent trends in social psychology that view much of human social behavior as habitual, automatic, and driven by nonconscious goal pursuit (Bargh, 1990; Bargh and Barndollar, 1996; Bargh et al., 2001; Hassin et al., 2009; Kruglanski et al., 2002; Ouellette and Wood, 1998). In this section, I briefly consider the issues raised and how they can be reconciled with a reasoned action perspective.

Habituation

With repeated performance, behavior is said to habituate such that it comes under direct control of stimulus cues, bypassing intentions as a determinant of behavior. This argument implies that behavioral intentions lose their predictive validity once a strong habit has been established (e.g., Aarts et al., 1998; Neal et al., 2006; Ouellette and Wood, 1998). Empirical findings lend little support to this hypothesis. In a meta-analysis of 15 data sets, Ouellette and Wood (1998) classified each set as dealing with a behavior that can be performed frequently and hence can habituate (e.g., seatbelt use, coffee drinking, class attendance) or infrequently and not likely to habituate (e.g., flu shots, blood donation, nuclear protest). Contrary to the habit hypothesis, prediction of behavior from intentions was found to be quite accurate for both types of behavior (mean $r = 0.59$ and $r = 0.67$ for high- and low-opportunity behaviors, respectively; difference not significant). The same conclusion comes from a more extensive meta-analysis based on 51 data sets (Sheeran and Sutton, unpublished data). For behaviors that could be performed infrequently (once or twice a year) the intention-behavior correlation was 0.51, and it was 0.53 for high-opportunity behaviors that could be performed daily or at least once a week. This meta-analysis also compared behaviors typically performed in the same context and thus amenable to habit formation to behaviors performed in variable contexts. Again, there was little difference in the predictive validity of intentions. If anything, the pattern of results was contrary to

what would be predicted by the habit hypothesis. The mean intention-behavior correlation was 0.40 with respect to behaviors performed in unstable contexts (where intentions should be most relevant) compared with a mean intention-behavior correlation of 0.56 for behaviors performed in stable contexts. Nor were the results more supportive of the habit hypothesis in a direct test with primary data (Ouellette and Wood, 1998; see Ajzen, 2002 for a discussion of these issues).

In sum, although behavior can become routine with repeated performance, no longer requiring much if any conscious deliberation, there is no evidence to suggest that intentions become irrelevant when behavior is routine. On the contrary, empirical evidence demonstrates that intentions predict routine as well as relatively novel behaviors. Moreover, this conclusion is not necessarily inconsistent with a habit perspective. "Within current theorizing, habits are automated response *dispositions* that are cued by aspects of the performance context" (Neal et al., 2006: 198; emphasis added). It may thus be argued that it is not the routinized behavior itself that is automatically initiated by the supporting context but a disposition to perform the behavior, such as an implicit intention. Consistent with this idea, the TPB assumes that, in the case of routine behaviors, implicit intentions are activated automatically and are then available to guide performance of the behavior.

Automaticity

Priming research has demonstrated that a large array of psychological concepts and processes can be activated automatically, below conscious awareness (see Bargh, 2006). Initial studies showed that activation of such knowledge structures as trait concepts (kindness, hostility) or ethnic stereotypes can influence encoding, comprehension, and judgments of ambiguous social behavior (e.g., Higgins et al., 1977; Srull and Wyer, 1979). In recent years, research has turned to nonconscious goal pursuit, demonstrating that desired

outcomes can be primed subconsciously and influence pursuit of the activated goal without the person's awareness (Hassin et al., 2009; Kruglanski et al., 2002). Thus, for example, primed activation of an achievement goal was found to improve level of performance on a word-search task (Bargh et al., 2001).

More relevant for present purposes, however, is the proposition that automatic activation of knowledge structures or goals can influence not only judgments or achievements but can also have a direct effect on behavior. Bargh et al. (1996), for example, showed that participants primed with the stereotype of the elderly walked more slowly down the hallway when leaving the experiment than did control participants, and when primed with the concept of rudeness, they interrupted the experimenter more frequently and quickly than when primed with the concept of politeness. Similarly, Aarts and Dijksterhuis (2003) found that when primed with the concept of silence by exposure to a picture of a library, participants spoke more softly, and following exposure to a picture of an exclusive restaurant, they were more likely to remove crumbs after eating a biscuit. These effects have usually been attributed to automatic enactment of a response made ready accessible by priming a certain construct (ideo-motor expression).

Traditionally, goal-directed behavior has been conceptualized as a controlled process that involves some measure of conscious deliberation and awareness, a view inherent in the TPB and other reasoned action models (e.g., Bandura, 1986, 1997; Deci and Ryan, 1985; Locke and Latham, 1990; Triandis, 1977). Although consistent with our intuitive sense that the pursuit of nontrivial goals is a controlled, conscious process, this perspective has, in the past 20 years, given way to theorizing that denies the importance of consciousness as a causal agent (Wegner, 2002; Wegner and Wheatley, 1999) and views much human social behavior as driven by implicit attitudes (Greenwald and Banaji, 1995) and other unconscious or subconscious mental processes (Aarts and Dijksterhuis, 2000; Bargh, 1989,

1996; Bargh and Chartrand, 1999; Brandstätter et al., 2001; Uhlmann and Swanson, 2004).

In one sense, the importance of nonconscious processes is undeniable. Our ability to exercise conscious, intentional control is constrained by limited information processing capacity such that most moment-to-moment mental processes must occur below conscious awareness (Bargh and Chartrand, 1999). The question of how much of our day-to-day behaviors is subject to automatic versus controlled processes is complicated by the fact that behavior involves a complex sequence of events. Many attributes of behavioral performance are outside conscious awareness. Thus, we don't pay much attention to how we move our legs and arms as we walk or how we produce sentences as we speak, nor do we ordinarily consciously monitor our facial expressions, tone of voice, or body posture. Even more complex behaviors can become automatic with sufficient practice. When we learn to drive, for example, we initially pay close attention to various aspects of this behavior, but once we have become skilled at the task, we can perform it more or less automatically as evidenced by the fact that we can at the same time engage in conversation or other activities that occupy our cognitive resources.

The research reviewed above has demonstrated that priming of constructs, knowledge structures, and goals can initiate these kinds of automatic processes and routine action sequences, although questions can be raised as to whether important decisions, such as buying a car, are ever completely automatic. More importantly, it should be noted that the observed automaticity in behavior is also consistent with a reasoned action perspective if we assume that attitudes and intentions with respect to common behaviors can become implicit and exert their influence below conscious awareness. Recent empirical research (e.g., Cesario et al., 2006; see also Förster et al., 2005) has provided support for this idea by showing that the priming of a category activates implicit preparatory responses, such as an implicit attitude toward

the primed category, and that these implicit responses determine the effects of the prime on behavior. For example, Cesario et al. (2006, Study 2) assessed implicit attitudes toward elderly and youth and, in a second session, subliminally presented either elderly picture primes or youth picture primes (or no primes in the control condition). Following this manipulation, the participants' speed of walking was recorded. As in previous research, priming the elderly construct slowed walking speed whereas priming the youth construct increased it. Of greater interest, however, was the role of implicit attitudes. The more positive participants' implicit attitudes toward the elderly, the slower they walked; and the more negative their implicit attitudes toward the elderly, the faster they walked. Comparable results were obtained with respect to activation of implicit attitudes toward youth. Speed of walking was therefore not a simple automatic response directly produced by priming of the elderly or youth category. Instead, in a proximal sense, it was the result of implicit preparatory responses, that is, of implicit attitudes toward the elderly or toward youth that were activated by priming these categories.

Of course, theorists interested in automaticity do not deny the importance of controlled processes in social attitudes and behavior (Devine and Monteith, 1999; Wegner and Bargh, 1998). Most endorse a dual-mode processing perspective that has room for automatic as well as controlled processes (Chaiken and Trope, 1999), but in recent years the pendulum may have swung too much in the direction of automaticity. If history is a guide, the pendulum is eventually bound to swing in the opposite direction and perhaps the time has come for social psychologists to rediscover reasoned action.

to be a useful framework for understanding, predicting, and changing human social behavior. Judging by the sheer number of investigations it has stimulated, the TPB is perhaps the most popular of the reasoned action models (for a list of publications, see bibliography in Ajzen, 2005b). Its application in varied domains has allowed investigators to identify important psychological determinants of socially significant behaviors. Armed with the conceptual framework and methodologies provided by the TPB, investigators have collected information about the behavioral, normative, and control-related determinants of many different behaviors, from exercising, eating a healthy diet, donating blood, and using illicit drugs to conserving energy, using public transportation, and practicing safer sex. Such knowledge can, of course, also provide the basis for effective interventions designed to modify social behavior in a desirable direction. Although the number of actual intervention studies to date is relatively small, especially in comparison with the large number of prediction studies, the theory has demonstrated its utility as a basis for designing and evaluating the effectiveness of interventions of various kinds, including interventions to discourage car use (Bamberg and Schmidt, 2001), limit infant sugar intake (Beale and Manstead, 1991), promote effective job search behaviors (Van Ryn and Vinokur, 1992), and encourage testicular self-examination (Brubaker and Fowler, 1990) and condom use (Fishbein et al., 1997) (see Ajzen, in press, for a review). I am hopeful that, in this fashion, the theory will continue to make a valuable contribution to the solutions of critical social problems.

NOTES

APPLICATIONS OF THE THEORY

Looking back over the past 30 years, I am very gratified to see that the TPB has proven

1 A popular approach to the analysis of decisions under uncertainty, the subjective expected utility model of behavioral decision theory (Coombs and Beardslee, 1954; Edwards, 1954), was developed at about the same time. According to this model, the

expected utility of a choice alternative is a function of the subjective probability that the alternative possesses certain attributes, multiplied by the subjective values or utilities of those attributes. It is assumed that a subjective expected utility is produced for each choice alternative and that decision makers choose the alternative with the highest SEU. For a comparison of the expectancy-value and SEU models, see Ajzen (1996).

2 Some investigators (e.g., Fekadu and Kraft, 2002; Rimal and Real, 2003) have assessed identification with the social referent, instead of motivation to comply, and have examined the moderating effect of identification. These investigations have shown that identification with social referents also does little to moderate the effect of normative beliefs.

REFERENCES

Aarts, H. and Dijksterhuis, A. (2000) Habits as knowledge structures: Automaticity in goal-directed behavior. *Journal of Personality and Social Psychology*, 78, 53–63.

Aarts, H. and Dijksterhuis, A. (2003) The silence of the library: Environment, situational norm, and social behavior. *Journal of Personality and Social Psychology*, 84, 18–28.

Aarts, H., Verplanken, B. and van Knippenberg, A. (1998) Predicting behavior from actions in the past: Repeated decision making or a matter of habit? *Journal of Applied Social Psychology*, 28, 1355–1374.

Ajzen, I. (1982) On behaving in accordance with one's attitudes. In M.P. Zanna, E.T. Higgins and C.P. Herman (eds), *Consistency in Social Behavior: The Ontario Symposium*, 2, 3–15. Hillsdale, NJ: Erlbaum.

Ajzen, I. (1985) From intentions to actions: A theory of planned behavior. In J. Kuhl and J. Beckman (eds), *Action-control: From Cognition to Behavior*, pp. 11–39. Heidelberg: Springer.

Ajzen, I. (1987) Attitudes, traits, and actions: Dispositional prediction of behavior in personality and social psychology. In L. Berkowitz (ed.), *Advances in Experimental Social Psychology*, Vol. 20, pp. 1–63. San Diego: Academic Press.

Ajzen, I. (1991a) Benefits of leisure: A social psychological perspective. In B.L. Driver and P.J. Brown (eds), *Benefits of Leisure*, pp. 411–417. State College: Venture Publishing.

Ajzen, I. (1991b) The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50, 179–211.

Ajzen, I. (1996) The social psychology of decision making. In E.T. Higgins and A.W. Kruglanski (eds), *Social Psychology: Handbook of Basic Principles*, pp. 297–325. New York: Guilford Press.

Ajzen, I. (2002) Residual effects of past on later behavior: Habituation and reasoned action perspectives. *Personality and Social Psychology Review*, 6, 107–122.

Ajzen, I. (2005a) *Attitudes, Personality, and Behavior*, 2nd Edition. Maidenhead: Open University Press.

Ajzen, I. (2005b) The Theory of Planned Behavior. Available at: <http://www.people.umass.edu/aizen/tptbrefs.html>, accessed 3 March 2005.

Ajzen, I. (in press) Behavioral interventions: Design and evaluation guided by the theory of planned behavior. In M.M. Mark, S.I. Donaldson and B. Campbell (eds), *Social Psychology and Program/Policy Evaluation*. New York: Guilford Press.

Ajzen, I. and Driver, B.L. (1991) Prediction of leisure participation from behavioral, normative, and control beliefs: An application of the theory of planned behavior. *Leisure Sciences*, 13, 185–204.

Ajzen, I. and Fishbein, M. (1977) Attitude-behavior relations: A theoretical analysis and review of empirical research. *Psychological Bulletin*, 84, 888–918.

Ajzen, I. and Fishbein, M. (1980) *Understanding Attitudes and Predicting Social Behavior*. Englewood-Cliffs, NJ: Prentice-Hall.

Ajzen, I. and Fishbein, M. (2000) Attitudes and the attitude-behavior relation: Reasoned and automatic processes. In W. Stroebe and M. Hewstone (eds), *European Review of Social Psychology*, 11, 1–33. Chichester: Wiley.

Ajzen, I. and Fishbein, M. (2008) Scaling and testing multiplicative combinations in the expectancy-value model of attitudes. *Journal of Applied Social Psychology*, 33, 2222–2247.

Ajzen, I. and Sexton, J. (1999) Depth of processing, belief congruence, and attitude-behavior correspondence. In S. Chaiken and Y. Trope (eds), *Dual-process Theories in Social Psychology*, pp. 117–138. New York: Guilford.

Albaracín, D., Johnson, B.T., Fishbein, M. and Muellerleile, P.A. (2001) Theories of reasoned action and planned behavior as models of condom use: A meta-analysis. *Psychological Bulletin*, 127, 142–161.

Armitage, C.J. and Conner, M. (1999) Distinguishing perceptions of control from self-efficacy: Predicting consumption of a low-fat diet using the theory of planned behavior. *Journal of Applied Social Psychology*, 29, 72–90.

Armitage, C.J. and Conner, M. (2001) Efficacy of the theory of planned behavior: A meta-analytic review. *British Journal of Social Psychology*, 40, 471–499.

Armitage, C.J., Conner, M., Loach, J. and Willets, D. (1999) Different perceptions of control: Applying an extended theory of planned behavior to legal and illegal drug use. *Basic and Applied Social Psychology*, 21, 301–316.

Bamberg, S. and Schmidt, P. (2001) Theory-driven subgroup-specific evaluation of an intervention to reduce private car use. *Journal of Applied Social Psychology*, 31, 1300–1329.

Bandura, A. (1977) Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84, 191–215.

Bandura, A. (1986) *Social Foundations of Thought and Action: A Social Cognitive Theory*. Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1997) *Self-efficacy: The Exercise of Control*. New York: Freeman.

Bandura, A. and Adams, N.E. (1977) Analysis of self-efficacy theory of behavioral change. *Cognitive Therapy and Research*, 1, 287–310.

Bandura, A. and Locke, E.A. (2003) Negative self-efficacy and goal effects revisited. *Journal of Applied Psychology*, 88, 87–99.

Bargh, J.A. (1989) Conditional automaticity: Varieties of automatic influence in social perception and cognition. In J.S. Uleman and J.A. Bargh (eds), *Unintended Thought*, pp. 3–51. New York: Guilford Press.

Bargh, J.A. (1990) Auto-motives: Preconscious determinants of social interaction. In E.T. Higgins and R.M. Sorrentino (eds), *Handbook of Motivation and Cognition: Foundations of Social Behavior*, pp. 93–130. New York: Guilford Press.

Bargh, J.A. (1996) Automaticity in social psychology. In E.T. Higgins and A.W. Kruglanski (eds), *Social Psychology: Handbook of Basic Principles*, pp. 169–183. New York: Guilford Press.

Bargh, J.A. (2006) Agenda 2006: What have we been priming all these years? On the development, mechanisms, and ecology of nonconscious social behavior. *European Journal of Social Psychology*, 36, 147–168.

Bargh, J.A. and Barndollar, K. (1996) Automaticity in action: The unconscious as repository of chronic goals and motives. In P.M. Gollwitzer and J.A. Bargh (eds), *The Psychology of Action: Linking Cognition and Motivation to Behavior*, pp. 457–481. New York: Guilford Press.

Bargh, J.A. and Chartrand, T.L. (1999) The unbearable automaticity of being. *American Psychologist*, 54, 462–479.

Bargh, J.A., Chen, M. and Burrows, L. (1996) Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology*, 71, 230–244.

Bargh, J.A., Gollwitzer, P.M., Lee-Chai, A., Barndollar, K. and Troetschel, R. (2001) The automated will: Nonconscious activation and pursuit of behavioral goals. *Journal of Personality and Social Psychology*, 81, 1014–1027.

Beale, D.A. and Manstead, A.S.R. (1991) Predicting mothers' intentions to limit frequency of infants' sugar intake: Testing the theory of planned behavior. *Journal of Applied Social Psychology*, 21, 409–431.

Bernberg, R.E. (1952) Socio-psychological factors in industrial morale: I. The prediction of specific indicators. *Journal of Social Psychology*, 36, 73–82.

Blumer, H. (1955) Attitudes and the social act. *Social Problems*, 3, 59–65.

Brandstätter, V., Lengfelder, A. and Gollwitzer, P.M. (2001) Implementation intentions and efficient action initiation. *Journal of Personality and Social Psychology*, 81, 946–960.

Brewer, W.F. (1974) There is no convincing evidence for operant or classical conditioning in adult humans. In W.B. Weimer and D.S. Palermo (eds), *Cognition and the Symbolic Processes*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Brubaker, R.G. and Fowler, C. (1990) Encouraging college males to perform testicular self-examination: Evaluation of a persuasive message based on the revised theory of reasoned action. *Journal of Applied Social Psychology*, 20, 1411–1422.

Budd, R.J., North, D. and Spencer, C. (1984) Understanding seat-belt use: A test of Bentler and Speckart's extension of the 'theory of reasoned action'. *European Journal of Social Psychology*, 14, 69–78.

Carlson, E.R. (1956) Attitude change through modification of attitude structure. *Journal of Abnormal and Social Psychology*, 52, 256–261.

Carver, C.S. and Scheier, M.F. (1998) *On the Self-regulation of Behavior*. Cambridge: Cambridge University Press.

Cervone, D. and Peake, P.K. (1986) Anchoring, efficacy, and action: The influence of judgmental heuristics on self-efficacy judgments and behavior. *Journal of Personality and Social Psychology*, 50, 492–501.

Cesario, J., Plaks, J.E. and Higgins, E.T. (2006) Automatic social behavior as motivated preparation to interact. *Journal of Personality and Social Psychology*, 90, 893–910.

Chaiken, S. and Trope, Y. (eds) (1999) *Dual-process Theories in Social Psychology*. New York: Guilford Press.

Cheung, S.-F. and Chan, D.K.-S. (2000) The role of perceived behavioral control in predicting human behavior: A meta-analytic review of studies on the theory of planned behavior. Unpublished manuscript, Chinese University of Hong Kong.

Cialdini, R.B., Reno, R.R. and Kallgren, C.A. (1990) A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of Personality and Social Psychology*, 58, 1015–1026.

Conner, M., Sherlock, K. and Orbell, S. (1998) Psychosocial determinants of ecstasy use in young people in the UK. *British Journal of Health Psychology*, 3, 295–317.

Coombs, C.H. and Beardslee, D. (1954) On decision-making under uncertainty. In R.M. Thrall, C.H. Coombs and R.L. Davis (eds), *Decision Processes*, pp. 255–285. New York: Wiley.

Courneya, K.S., Keats, M.R. and Turner, R. (2000) Social cognitive determinants of hospital-based exercise in cancer patients following high-dose chemotherapy and bone marrow transplantation. *International Journal of Behavioral Medicine*, 7, 189–203.

Dabholkar, P.A. (1999) Expectancy-value models. In P.E. Earl and S. Kemp (eds), *The Elgar Companion to Consumer Research and Economic Psychology*, pp. 200–208. Cheltenham: Edward Elgar.

Davis, F.D., Bagozzi, R.P. and Warshaw, P.R. (1989) User acceptance of computer technology: A comparison of two theoretical models. *Management Science*, 35, 982–1003.

De Fleur, M.L. and Westie, F.R. (1958) Verbal attitudes and overt acts: An experiment on the salience of attitudes. *American Sociological Review*, 23, 667–673.

Dean, L.R. (1958) Interaction, reported and observed: The case of one local union. *Human Organization*, 17, 36–44.

Deci, E.L. and Ryan, R.M. (1985) The general causality orientations scale: Self-determination in personality. *Journal of Research in Personality*, 19, 109–134.

Deutscher, I. (1966) Words and deeds: Social science and social policy. *Social Problems*, 13, 235–254.

Devine, P.G. and Monteith, M.J. (1999) Automaticity and control in stereotyping. In S. Chaiken and Y. Trope (eds), *Dual-process Theories in Social Psychology*, pp. 339–360. New York: Guilford Press.

Downs, D.S. and Hausenblas, H.A. (2005) The theories of reasoned action and planned behavior applied to exercise: A meta-analytic update. *Journal of Physical Activity and Health*, 2, 76–97.

Dulany, D.E. (1962) The place of hypotheses and intentions: An analysis of verbal control in verbal conditioning. In C. Eriksen (ed.), *Behavior and Awareness*, pp. 102–129. Durham, NC: Duke University Press.

Dulany, D.E. (1968) Awareness, rules, and propositional control: A confrontation with S-R behavior theory. In D. Hornton and T. Dixon (eds), *Verbal Behavior and S-R Behavior Theory*, pp. 340–387. Englewood Cliffs, NJ: Prentice-Hall.

Edwards, W. (1954) The theory of decision making. *Psychological Bulletin*, 51, 380–417.

Feather, N.T. (ed.) (1982) *Expectations and Actions: Expectancy-value Models in Psychology*. Hillsdale, NJ: Erlbaum.

Fekadu, Z. and Kraft, P. (2002) Expanding the theory of planned behaviour: The role of social norms and group identification. *Journal of Health Psychology*, 7, 33–43.

Festinger, L. (1964) Behavioral support for opinion change. *The Public Opinion Quarterly*, 28, 404–417.

Fishbein, M. (1963) An investigation of the relationships between beliefs about an object and the attitude toward that object. *Human Relations*, 16, 233–240.

Fishbein, M. (1967a) Attitude and the prediction of behavior. In M. Fishbein (ed.), *Readings in Attitude Theory and Measurement*, pp. 477–492. New York: Wiley.

Fishbein, M. (1967b) A behavior theory approach to the relations between beliefs about an object and the attitude toward the object. In M. Fishbein (ed.), *Readings in Attitude Theory and Measurement*, pp. 389–400. New York: Wiley.

Fishbein, M. and Ajzen, I. (1974) Attitudes towards objects as predictors of single and multiple behavioral criteria. *Psychological Review*, 81, 59–74.

Fishbein, M. and Ajzen, I. (1975) *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research*. Reading, MA: Addison-Wesley.

Fishbein, M. and Ajzen, I. (2010) *Prediction and Change of Behavior: The Reasoned Action Approach*. New York: Psychology Press.

Fishbein, M., Guenther-Grey, C., Johnson, W., Wolitski, R.J., McAlister, A., Rietmeijer, C.A. and O'Reilly, K. (1997) Using a theory-based community intervention to reduce AIDS risk behaviors: The CDC's AIDS community demonstration projects. In M.E. Goldberg, M. Fishbein and S. Middlestadt (eds),

Social Marketing: Theoretical and Practical Perspectives, pp. 123–146. Mahwah, NJ: Lawrence Erlbaum Associates.

Fisher, J.D. and Fisher, W.A. (1992) Changing AIDS-risk behavior. *Psychological Bulletin*, 111, 455–474.

Förster, J., Liberman, N. and Higgins, E.T. (2005) Accessibility from active and fulfilled goals. *Journal of Experimental Social Psychology*, 41, 220–239.

Gagné, C. and Godin, G. (2000) The theory of planned behavior: Some measurement issues concerning belief-based variables. *Journal of Applied Social Psychology*, 30, 2173–2193.

Geraerts, E., Bernstein, D.M., Merckelbach, H., Linders, C., Raymaeckers, L. and Loftus, E.F. (2008) Lasting false beliefs and their behavioral consequences. *Psychological Science*, 19, 749–753.

Greenwald, A.G. and Banaji, M.R. (1995) Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, 102, 4–27.

Hagger, M.S., Chatzisarantis, N.L.D. and Biddle, S.J.H. (2002) A meta-analytic review of the theories of reasoned action and planned behavior in physical activity: Predictive validity and the contribution of additional variables. *Journal of Sport and Exercise Psychology*, 24, 3–32.

Hassin, R.R., Aarts, H., Eitam, B., Custers, R. and Kleiman, T. (2009) Non-conscious goal pursuit and the effortful control of behavior. In E. Morsella, P.M. Gollwitzer and J.A. Bargh (eds), *Oxford Handbook of Human Action*, pp. 549–568. New York: Oxford University Press.

Higgins, E.T., Rholes, W.S. and Jones, C.R. (1977) Category accessibility and impression formation. *Journal of Experimental Social Psychology*, 13, 141–154.

Himelstein, P. and Moore, J. (1963) Racial attitudes and the action of Negro and white background figures as factors in petition-signing. *Journal of Social Psychology*, 61, 267–272.

Hrubes, D., Ajzen, I. and Daigle, J. (2001) Predicting hunting intentions and behavior: An application of the theory of planned behavior. *Leisure Sciences*, 23, 165–178.

Kruglanski, A.W., Shah, J.Y., Fisbach, A., Friedman, R.S., Chun, W.Y. and Sleeth-Keppler, D. (2002) A theory of goal systems. In M.P. Zanna (ed.), *Advances in Experimental Social Psychology*, pp. 331–378. San Diego: Academic Press.

Latimer, A.E. and Martin Ginis, K.A. (2005) The theory of planned behavior in prediction of leisure time physical activity among individuals with spinal cord injury. *Rehabilitation Psychology*, 50, 389–396.

Linn, L.S. (1965) Verbal attitudes and overt behavior: A study of racial discrimination. *Social Forces*, 43, 353–364.

Litt, M.D. (1988) Self-efficacy and perceived control: Cognitive mediators of pain tolerance. *Journal of Personality and Social Psychology*, 54, 149–160.

Locke, E.A. and Latham, G.P. (1990) *A Theory of Goal Setting and Task Performance*. Englewood Cliffs, NJ: Prentice-Hall.

Locke, E.A. and Latham, G.P. (1994) Goal setting theory. In H.F.J. O'Neil and M. Drillings (eds), *Motivation: Theory and Research*, pp. 13–29. Hillsdale, NJ: Lawrence Erlbaum Associates.

Neal, D.T., Wood, W. and Quinn, J.M. (2006) Habits – a repeat performance. *Current Directions in Psychological Science*, 15, 198–202.

Notani, A.S. (1998) Moderators of perceived behavioral control's predictiveness in the theory of planned behavior: A meta-analysis. *Journal of Consumer Psychology*, 7, 247–271.

Ouellette, J.A. and Wood, W. (1998) Habit and intention in everyday life: The multiple processes by which past behavior predicts future behavior. *Psychological Bulletin*, 124, 54–74.

Peak, H. (1955) Attitude and motivation. In M.R. Jones (ed.), *Nebraska Symposium on Motivation*, 3, 149–189. Lincoln, NE: University of Nebraska Press.

Petty, R.E. and Cacioppo, J.T. (1986) *Communication and Persuasion: Central and Peripheral Routes to Attitude Change*. New York: Springer Verlag.

Randall, D.M. and Wolff, J.A. (1994) The time interval in the intention-behaviour relationship: Meta-analysis. *British Journal of Social Psychology*, 33, 405–418.

Rimal, R.N. and Real, K. (2003) Understanding the influence of perceived norms on behaviors. *Communication Theory*, 13, 184–203.

Rivis, A. and Sheeran, P. (2003) Social influences and the theory of planned behaviour: Evidence for a direct relationship between prototypes and young people's exercise behaviour. *Psychology and Health*, 18, 567–583.

Rosenberg, M.J. (1956) Cognitive structure and attitudinal affect. *Journal of Abnormal and Social Psychology*, 53, 367–372.

Rosenberg, M.J. (1965) Inconsistency arousal and reduction in attitude change. In I.D. Steiner and M. Fishbein (eds), *Current Studies in Social Psychology*, pp. 121–134. New York: Holt, Reinhart, and Winston.

Rosenberg, M.J. and Hovland, C.I. (1960) Cognitive, affective, and behavioral components of attitudes.

In C.I. Hovland and M.J. Rosenberg (eds), *Attitude Organization and Change: An Analysis of Consistency Among Attitude Components*, pp. 1–14. New Haven, CT: Yale University Press.

Rosenstock, I.M., Strecher, V.J. and Becker, M.H. (1994) The health belief model and HIV risk behavior change. In R.J. DiClemente and J.L. Peterson (eds), *Preventing AIDS: Theories and Methods of Behavioral Interventions. AIDS Prevention and Mental Health*, pp. 5–24. New York: Plenum Press.

Schulze, R. and Wittmann, W.W. (2003) A meta-analysis of the theory of reasoned action and the theory of planned behavior: The principle of compatibility and multidimensionality of beliefs as moderators. In R. Schulze, H. Holling and D. Böhning (eds), *Meta-analysis: New Developments and Applications in Medical and Social Sciences*, pp. 219–250. Cambridge, MA: Hogrefe & Huber.

Sheeran, P. (2002) Intention-behavior relations: A conceptual and empirical review. In W. Stroebe and M. Hewstone (eds), *European Review of Social Psychology*, 12, 1–36. Chichester: Wiley.

Sheeran, P. and Taylor, S. (1999) Predicting intentions to use condoms: A meta-analysis and comparison of the theories of reasoned action and planned behavior. *Journal of Applied Social Psychology*, 29, 1624–1675.

Sheppard, B.H., Hartwick, J. and Warshaw, P.R. (1988) The theory of reasoned action: A meta-analysis of past research with recommendations for modifications and future research. *Journal of Consumer Research*, 15, 325–342.

Srull, T.K. and Wyer, R.S. (1979) The role of category accessibility in the interpretation of information about persons: Some determinants and implications. *Journal of Personality and Social Psychology*, 37, 1660–1672.

Triandis, H.C. (1972). *The Analysis of Subjective Culture*. New York: Wiley.

Triandis, H.C. (1977) *Interpersonal Behavior*. Monterey, CA: Brooks/Cole.

Tversky, A. and Kahneman, D. (1974) Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124–1131.

Uhlmann, E. and Swanson, J. (2004) Exposure to violent video games increases automatic aggressiveness. *Journal of Adolescence*, 27, 41–52.

Van Ryn, M. and Vinokur, A.D. (1992) How did it work? An examination of the mechanisms through which an intervention for the unemployed promoted job-search behavior. *American Journal of Community Psychology*, 20, 577–597.

Vroom, V.H. (1964) *Work and Motivation*. New York: Wiley.

Webb, T.L. and Sheeran, P. (2006) Does changing behavioral intentions engender behavior change? A meta-analysis of the experimental evidence. *Psychological Bulletin*, 132, 249–268.

Wegner, D.M. (2002) *The Illusion of Conscious Will*. Cambridge, MA: MIT Press.

Wegner, D.M. and Bargh, J.A. (1998) Control and automaticity in social life. In D.T. Gilbert, S.T. Fiske and L. Gardner (eds), *The Handbook of Social Psychology*, Vol. 1, 4th Edition, pp. 446–496. Boston, MA: McGraw-Hill.

Wegner, D.M. and Wheatley, T. (1999) Apparent mental causation: Sources of the experience of will. *American Psychologist*, 54, 480–492.

Weigel, R.H. and Newman, L.S. (1976) Increasing attitude-behavior correspondence by broadening the scope of the behavioral measure. *Journal of Personality and Social Psychology*, 33, 793–802.

Weinberg, R.S., Gould, D., Yukelson, D. and Jackson, A. (1981) The effect of preexisting and manipulated self-efficacy on a competitive muscular endurance task. *Journal of Sport Psychology*, 3, 345–354.

Wicker, A.W. (1969) Attitudes versus actions: The relationship of verbal and overt behavioral responses to attitude objects. *Journal of Social Issues*, 25, 41–78.

Yang-Wallentin, F., Schmidt, P., Davidov, E. and Bamberg, S. (2004) Is there any interaction effect between intention and perceived behavioral control? *Methods of Psychological Research Online*, 8, 127–157.

Social Comparison Theory

Jerry Suls and Ladd Wheeler

ABSTRACT

This chapter reviews the history and development of the study of social comparison processes – the search for and utilization of information about other persons' standings and opinions for the purpose of accurate self-assessment or for self-esteem enhancement/protection. We describe the origins of Festinger's classic comparison theory in the group dynamics tradition, ambiguities in the classic theory, the later attributional reformulation of the theory, the research shift from self-evaluation to self-enhancement as a dominant motive for comparison (i.e., downward comparison theory [DCT]) to more recent social cognitive approaches which clarify what Festinger really meant (the proxy model) and emphasize the importance of knowledge accessibility (selective accessibility model) and social judgment (interpretation–comparison model) for self-evaluation. The essay concludes with some unresolved questions and illustrative applications of social comparison for education, health and subjective wellbeing.

INTRODUCTION

Although the statement, "Everything is relative," is an exaggeration, many things are

relative, thereby complicating judgment and action. Apropos of a joke:

A snail was mugged by two turtles. When the police asked him what happened, he said: I don't know. It all happened so fast.

(Cathcart and Klein, 2007: 273)

Social philosophers and early social scientists recognized that people often rely on how they stand relative to other people to assess their opinions and potential. Leon Festinger (1954a) was the first to systematically consider this topic in "A theory of social comparison processes." "Social comparison" referred to the search for and utilization of information about other persons' standings and opinions for the purpose of self-assessment – judging the correctness of one's opinions, beliefs, and capabilities (Wood, 1996). Later researchers also placed emphasis on additional motives for comparison.

In this chapter, we survey the path of social comparison research from the 1950s to the current day. Although proceeding in chronological fashion, readers will see that the path was not straight, with some wrong

turns, rest stops, and blind canyons. One of us (Ladd Wheeler) started on the trail in the early 1960s; the other (Jerry Suls) joined in the early 1970s. For two decades, we were “fellow travelers,” but in the mid 90s we joined forces, a collaboration that has happily continued to the present. (Lefty and Dusty, the two cowboys, from *Prairie Home Companion* come to mind.) Because we are contributors to this research area, we may be biased in some conclusions so we will try to label our opinions as such. However, one lesson from social comparison theory is that distinguishing facts from opinions is no simple matter.

GROUP DYNAMICS

In the 1940s, Festinger was busy with research on informal communication in small groups at the Research Center for Group Dynamics at MIT and later at the University of Michigan. That work culminated in his theory of informal social communication which posited that people desire to attain uniformity of opinion either because group consensus provides confidence in one’s opinion or because agreement was needed to achieve group goals (Festinger, 1950). The research with his students showed that patterns of communication and rejection of opinion deviates followed the propositions of the theory and provided a foundation for further study in conformity and group performance (e.g., Allen, 1965; Turner, 1991).

This group research prompted the Ford Foundation to award Festinger a grant to summarize and integrate empirical studies on social influence. In the 1950s, there was a keen interest to find ways to use social science research to inform policy and societal change. Festinger also was invited to give a talk at the second annual symposium on motivation held at the University of Nebraska. The talk was published as a chapter in an edited collection and “A theory of social comparison processes” was published in the

journal, *Human Relations*, along with some related empirical papers, in the same year, 1954.

Both papers extended the arguments of informal communication theory, but the emphasis changed from the power of the group over the individual to how individuals use groups to evaluate the need to evaluate opinions and abilities. We now turn to details of the theory.

THE CLASSIC 1954 FORMULATION

Festinger began with the premise that people have the need to evaluate their abilities and opinions to be able to act in the world. He noted that they prefer to evaluate themselves relative to objective or physical standards, but these are not always available and in some cases never are. We can verify the Lone Ranger’s (Clayton Moore) date of birth or whether the ground is muddy, but there are no objective standards for opinions about gun control or whether a cowboy will win a contest at the rodeo. But in many cases, we must make a judgment in the absence of objective information. In those instances, Festinger proposed that uncertainty induced a drive-like state that could be satisfied by social comparison (Recall in the 1950s drive theory was still popular in behavioral psychology).

A third key element in the theory was that comparison was important for the evaluation of opinions *and* abilities. The latter were not part of informal communication theory, but he realized, as noted above, that often there is no objective standard to assess our abilities, especially about future performance. Since abilities, unlike opinions, cannot be changed through communication, Festinger replaced “comparison” for “communication” – the prominent element in the earlier theory.

Festinger also recognized a distinctive feature of abilities. People want to be slightly better than everyone else because the desire to be better or to improve is emphasized in

Western cultures. This tendency was referred to as the *unidirectional drive upward*.

These ideas were well received. The remainder of his *Human Relations* paper (1954a), however, focused on “the need for similarity” which might well have been the subtitle of the paper. We can evaluate our abilities and opinions *accurately* (1954a: 120; italics in original) only by comparison with other people of similar abilities and opinions. Consequently, we choose to compare to similar others, we try to change others to become more similar to ourselves, we change ourselves to be more similar to others, and failing all else, we simply cease comparing with dissimilar others. Why? Festinger claimed because we will not be able to make a “subjectively precise evaluation.” The startling aspect of the paper is that it never explains why similarity is necessary for accurate evaluation. However, this gap was not apparent for some time (see Deutsch and Krauss, 1965). Perhaps this is because the general idea that social comparison is a core element of social life seemed both evident and powerful.

Let’s look at opinions and abilities separately to try to determine why similarity might be so important. In the case of opinions, Festinger wrote: “. . . , a person who believes that Negroes are the intellectual equals of whites does not evaluate his opinion by comparison with the opinion of a person who belongs to some very anti-Negro group” (1954a: 120–121). That comparison would lead to disagreement, and the equality believer would still not know if his opinion was correct. Instead we should compare with someone who is similar. If such a person agreed about the intellectual equality of Negroes and whites, the equality believer would be somewhat more confident in the correctness of the opinion. We are not told in which ways the comparison other should be similar beyond the fact that the person should not belong to a very anti-Negro group. Presumably, the person should be similar in values and world outlook, and we should have an expectation of agreement.

Turning to abilities, Festinger wrote: “. . . , a college student does not compare himself to inmates of an institution for the feeble minded to evaluate his own intelligence. Nor does a person who is just beginning to learn the game of chess compare himself to the recognized masters of the game” (1954a: 120). It is not clear how comparing to another chess beginner would give one an accurate appraisal of one’s chess ability. In fact, it would seem that comparing to a chess master would be more informative; one would at least have an idea of the level of ability that could be achieved. Bumbling around the board with another bumbler would say little about one’s ability.

In short, Festinger’s *Human Relations* article makes some degree of sense about the importance of similarity in feeling that our opinions are correct, but the argument regarding ability evaluation is completely unsatisfactory. There is no reason given about why we would be more accurate in ability appraisal as a result of comparing to someone of similar ability. Fortunately, the *Nebraska Symposium on Motivation* paper (1954b), written after the *Human Relations* article, is less formal, and offers a relatively clear explanation for the importance of similarity in ability evaluation. The explanation stresses action possibilities in the real world. We should be able to do the same things as others of similar ability can do. Below is a quote from the Nebraska paper illustrating this point.

Let us take as an illustration a person who tries to find out precisely how intelligent he is. Let us remember that his motivation is to know what his ability does and does not permit him to do in the real world in which he lives and acts. Simply knowing his score on some intelligence test does not tell him this. It does, however, allow him to compare himself with others. Suppose the others with whom he compares himself are all very divergent from himself in their score. He then knows that his possibilities for action in the world are very different from theirs. But this is negative knowledge and he still does not know precisely what he himself can do. Consider, however, the case where the others with whom he compares himself have scores very close

to his own. He then knows that his own possibilities for action in the environment are identical or very similar to those for these other persons. This gives him the subjective feeling of knowing what he can or cannot do in the same way that having other agree with one's opinion gives the person a subjective feeling that his own opinion is correct.

1954b: 196–197

There are numerous other statements in the *Nebraska* chapter consistent with the one just given. For example:

[L]et us imagine a high school student who wants to know whether his intellectual ability is such as to enable him to go through college. He ... cannot adequately evaluate his ability in the real world. Clearly, going to college and seeing what happens would give him an evaluation of his ability for that purpose, but it is not possible to 'reality test' his ability before going to college. In the case of abilities it is particularly true that, even when a clear and unambiguous 'performance score' is available, it may not provide a satisfactory evaluation because of the large variety of situations for which the ability is relevant and the large number of purposes which make the ability important. Also important is the frequent desire to evaluate the ability *before* engaging in the action which would test it.

(1954b: 195; italics in original)

Thus, by emphasizing the importance of action possibilities in the *Nebraska* paper, Festinger sent a different message than in the *Human Relations* paper. The only part of the *Human Relations* paper that suggested that similarity is necessary to predicting what we can and cannot do is the Corollary to Derivation E that says:

An increase in the importance of an ability or an opinion, or an increase in its relevance to immediate behavior, will increase the drive toward reducing the discrepancies concerning that opinion or ability.

(1954a: 130¹)

We do know that almost everyone who subsequently worked in social comparison theory in the late 1950s to the mid 1970s used the *Human Relations* paper rather than the *Nebraska* chapter as their guide, and thus they simply accepted that similarity was

important without knowing why it was important. The only paper that investigated the action possibility argument was by Steve Jones and Dennis Regan (1974) – a paper we will return to later.

COMPARISON EXTENDED TO AFFILIATION AND EMOTION

During the time Festinger was working on comparison theory, he already had formulated his theory of cognitive dissonance which was soon to become his dominant interest. We don't know whether Festinger recognized the ambiguity in comparison theory because he never returned to the topic.

Comparison theory might have languished, but for the use his ex-student Stanley Schachter made of it in his research on affiliation. In a series of clever experiments, Schachter (1959) showed that inducing fear in experimental subjects prompts them to want to wait with others, particular others who were also awaiting the same fearful stimulus (i.e., painful electric shocks) – inspiring the memorable, "Misery doesn't love just any company; misery loves miserable company." Schachter thought that subjects preferred to affiliate to learn whether their emotional responses were appropriate for the circumstances. The fact that their preferred affiliates were people in the *same* circumstances (also waiting to receive the shocks) seemed to support Festinger's similarity hypothesis. (Subjects didn't want to wait with people who were there for some other purpose, such as waiting to see their advisors.)

The research had an enormous influence on the psychology of emotion, but gave little attention to the process of emotion comparison. As far as comparison studies were concerned, the biggest contribution was the recognition that comparisons apply to emotions as well as to abilities and opinions. This opened the door to apply social comparison to other attributes of the person in addition to abilities and opinions.

SCHACHTER'S INDIRECT EFFECT AND THE JESP SUPPLEMENT

While Schachter pursued a different research direction, he taught a graduate course in group dynamics at the University of Minnesota that spread the gospel about informal communication and social comparison. Many of the graduate students, including Radloff, Hakmiller, Singer, Latané, Arrowood, and one of the present authors (Ladd Wheeler), conducted experiments on social comparison. In 1966, these were eventually published in a special supplement to the *Journal of Experimental Social Psychology, Studies in Social Comparison*, edited by Bibb Latané.

Space prevents us from a complete description of the supplement's contents, but the empirical findings showed support for some of the theory's propositions and extended its scope. People under stress don't just prefer to affiliate with others; they also grow to like others who are present, presumably because they satisfy the need for comparison (Latané et al., 1966). People prefer to affiliate with others who agree with them when uncertain about the correctness of their opinions (Gordon, 1966). Lacking comparison information, people are unstable and feel uncertain in their personal performance evaluations (Radloff, 1966). This study was what initially piqued JS's interest in comparison theory. Radloff extrapolated that exceptionally talented children might actually underperform academically because they lacked any similar peers with whom to compare. Similar others in these experiments were operationalized as "people in the same circumstances" or in agreement. There was little acknowledgment of the aforementioned ambiguity of the similarity concept.

This is most apparent, with hindsight, in the two most influential experiments in the supplement. Wheeler (1966) developed a procedure – subsequently known as the rank-order paradigm – to test the unidirectional drive upward idea which Festinger proposed

would motivate the person to be slightly better than others (Only *slightly better*, according to Festinger, because the desire to be better was counteracted by the presumed desire to be similar). In Wheeler's experimental paradigm, subjects received a score and their rank on a bogus trait inventory. They were then offered the choice to see the actual score of someone else of a different rank to assess how they had done. When subjects were motivated to think they were good, the likelihood of choosing to see the score of someone ranked above them (i.e., upward comparison) was greater. It is worth noting that those who chose to see the score of a higher ranked subject – in essence made an *upward comparison* – felt they were more similar to the person above them in the rank order than to the person below them. "The comparer is attempting to prove to himself that he is almost as good as the very good ones" (Wheeler, 1966: 30).

The second influential experiment was conducted by Hakmiller (1966), who was a graduate student with Wheeler and who adapted the rank order paradigm to introduce the idea of downward comparison. In the experiment, subjects were led to believe through a cover story that they received a high score on a supposed measure of "hostility toward one's parents." In one condition, subjects were led to believe the high score represented a bad thing (high threat); in the other condition, they were led to believe it represented something relatively positive (low threat). Then all subjects were offered the opportunity to see someone else's score in the rank ordering. Threatened subjects were more interested in seeing the score of someone higher in hostility and therefore presumably worse off. This suggested to Hakmiller (1966) that threat to self-esteem increases the motivation to protect or enhance the self and thereby prompts downward comparison. The idea was a major source of inspiration for Wills' (1981) subsequent downward comparison theory that was influential a decade later.

The Wheeler (1966) and Hakmiller (1966) experiments had interesting implications – the former was suggestive of assimilative-type processes that later became central in social cognition research; the latter provided evidence for the motive to enhance or protect self-esteem. But, at the time, their most salient implication was that similarity should be conceived in terms of relative standing on the dimension under evaluation. In this way, similarity referred to people who shared the same opinions and performed in the same way. There is circularity in this, however, because in real life choosing to compare with a similar other presupposes that an implicit social comparison has already occurred (i.e., how does one know they are similar in the first place?). The rank-order paradigm solved this experimentally by providing rank information prior to the comparison choice. It took some time before the full extent of this problem was recognized.

RESTING AROUND THE CAMPFIRE

Between the publication of the *JESP* supplement and the late 1970s, comparison was not a dominant interest of social psychology; the hot areas were cognitive dissonance and attribution theory. However, there were a few developments brewing. Morse and Gergen (1970) conducted an experiment to show that casual exposure to another person may be sufficient to produce an impact on a person's self-esteem. Job applicants encountered an accomplice whose personal appearance was either highly desirable (Mr. Clean) or highly undesirable (Mr. Dirty). Under the guise of collecting information about the applicants, subjects completed self-esteem scales both before and after exposure to Mr. Clean/ Mr. Dirty. The results showed that exposure to Mr. Clean produced a decrease in self-esteem and exposure to Mr. Dirty produced an increase in self-esteem. "As a result of others' characteristics appearing more desirable or less desirable than his own, a person's

generalized self-estimate is displaced downward or upward" (1970: 154). The Morse and Gergen experiment was important because it showed that social comparison was relevant to broader domains than specific abilities, opinions or emotions. Also, prior to this time, social comparison was considered to fall under the umbrella of "social influence," consistent with its origins in group dynamics. Morse and Gergen's research placed comparison under the umbrella of the psychology of the self.

COMPARISON'S FIRST BOOK

Comparison theory seemed to be at rest. There was research activity but studies were published piecemeal, there was no "big picture," and not even much of an invisible comparison college. (A candidate for the "big picture" was Tom Pettigrew's chapter, "Self-evaluation theories," in the 1967 *Nebraska Symposium on Motivation*, where he described the common themes and implications of social comparison, equity, and relative deprivation theories. But his chapter appeared in the same volume with Harold Kelley's "Attribution principles in social psychology," which clearly took most social psychologists' hearts and minds at that time.)

But primed by chutzpah, Radloff's experiments, Pettigrew's essay and the threat of unemployment more than any editing experience, JS and a colleague, Rick Miller, floated the idea of an edited collection of original essays on social comparison. To our surprise, we learned there were many social psychologists, some well known, who were conducting comparison research and who were eager to contribute. Persuading a publisher to take on this project was much harder than getting commitments from authors. At least one anonymous reviewer, described to us only as "an eminent social psychologist," told a prospective publisher that nothing new in social comparison had happened since 1959

so there was no need for a new book (JS thinks this was Schachter). But although publisher, Larry Erlbaum, did not offer us a contract, he persuaded another publisher (Hemisphere) to do so – a favor for which we remain grateful.

That volume, *Social Comparison Processes: Theoretical and Empirical Perspectives* (Suls and Miller, 1977), had 13 chapters on a broad range of comparison-related topics by both veteran and relatively junior social psychologists and a concluding commentary. Two chapters proved to be very influential. In “Pleasure and pain of social comparison,” Brickman and Bulman proposed, counter to Festinger, that sometimes people want to vigorously avoid social comparison because of the threats it can pose. Their essay emphasized self-enhancement and self-protection rather self-evaluation. Brickman and Bulman’s (1977) essay was a direct precursor of Thomas Ashby Wills’ (1981) later downward comparison principles article. As a historical aside, Brickman was approached to write an update of his essay “Hedonic relativism and planning the good society,” coauthored with Donald Campbell (1971). But in our conversations, Phil wanted us to reprint the original essay which he felt was little read (it is now considered a classic essay). JS recalls insisting that we could not reprint; we needed a new essay. Phil reluctantly agreed. Months later, he sent something entirely different that included six experiments specifically designed and conducted for the chapter.

The other chapter by Goethals and Darley (1977) received widespread acceptance when the volume was published. Ladd Wheeler, who coauthored a commentary with Miron Zuckerman for the book, thought this was the flagship chapter in the volume. (The commentary represented a change in plans. Originally, Jerry Suls and Miller were supposed to coauthor the commentary chapter, but Miller relocated to Germany and Jerry Suls had [realistic] doubts he could do it justice by himself. Ladd Wheeler originally

had been contracted to summarize the rank-order paradigm experiments, but he discovered it had too much overlap with another contributing author so he was invited to do the commentary. LW asked Miron Zuckerman to help him because LW had broken his writing arm sliding into second base during a softball game [not falling off a horse].

Goethals and Darley proposed to resolve the ambiguity of the similarity hypothesis one had to consider the important role of related attributes. Related attributes are related to, and predictive of, an ability or an opinion. Festinger’s Hypothesis VIII was: ‘If persons who are very divergent from one’s own opinion or ability are perceived as different from oneself on *attributes consistent with the divergence*, the tendency to narrow the range of comparability becomes stronger’ (1954a: 133). Festinger noted that in level of aspiration studies, college students do not compete with high-scoring graduate students, because the college students see themselves as inferior on attributes consistent with the difference in scores. Goethals and Darley (1977) married Hypothesis VIII to attribution theory (Kelley, 1967, 1973) and argued that people will discount differences in performance if they are consistent with differences in related attributes. According to Kelley, the discounting principle is as follows: the role of a given cause in producing an effect is discounted if other plausible causes are also present.

If we want to evaluate an ability, we must observe a performance and make inferences about the underlying ability. There are many determinants of performance that are not indicative of the underlying ability – effort, luck, difficulty, age, practice, and so forth. Only if we compare our performance to that of someone similar on all these related attributes can we make a reasonable inference about our ability relative to the ability of the other person. This, according to Goethals and Darley, is what Festinger meant when he wrote that we must compare ourselves to similar others in order to make a subjectively precise self-evaluation.

The argument for the evaluation of opinions is somewhat different. Goethals and Darley distinguished between beliefs (a potentially verifiable assertion about the true nature of an entity) and values (liking or disliking of an entity). Only beliefs can be correct or incorrect, and they seem to correspond better than values to what Festinger meant by opinions. For that reason, we will discuss only beliefs. According to Kelley's attribution theory, a person needs to determine if his belief is entity-caused (which means that he is correct) or person-caused (which means that his belief is a biased view based on his needs, values, wishes, etc.). Goethals and Darley argued that comparison with others similar on related attributes would not be very useful unless they disagreed with the individual, in which case his belief in the correctness of his opinion would decrease. If similar others agreed, the individual would not know if their agreement was entity-caused or person-caused. If similar others disagreed, the individual would conclude that the disagreement was entity-caused. The same argument applies, *mutatis mutandis*, to comparison with others who are dissimilar on related attributes. Their disagreement would not be informative, but their agreement should increase one's confidence in the belief. Goethals and Darley concluded that they couldn't make a prediction about who would be chosen for comparison on the basis of related attributes.

We now view the related attributes hypothesis with less enthusiasm than when it appeared in 1977, but we still think it is a useful formulation for some purposes. If we compare our ability with someone who is perfectly similar to us on all related attributes, we can conclude that we are about as good as we ought to be, or better or worse than we ought to be, relative to that comparison person. That knowledge does not tell us, however, where we are on the ability scale, and it does not tell us what we can accomplish in the world. It is useful in helping us to evaluate a particular performance, which is often an important question. And if we have additional

information about what the comparison person has accomplished in the world, we may be able to predict our own accomplishments. However, it was two decades before we fully appreciated these things and presented them in the "proxy model," to be described later (Suls et al., 2002).

SELF-ENHANCEMENT AND DOWNWARD COMPARISON

In the 1980s, Goethals and Darley were perceived to have resolved the ambiguity in Festinger's theory (Arrowood, 1986; Suls, 1986; Wood, 1989) and a series of experiments provided empirical support for the attributional reformulation (Gastorf and Suls, 1978; Suls et al., 1978; Wheeler and Koestner, 1984; Wheeler et al., 1982; Zanna et al., 1975). With self-evaluation apparently covered, the field moved on to self-enhancement. Inspired in part by Hakmiller's experiment and Brickman and Bulmann (1977), Tom Wills (1981) presented an integrative review of prior research and original theory about how comparisons operate when the self is threatened. His paper, "Downward comparison principles in social psychology," covered a range of evidence from gossip and aggression to projection and self-esteem, arguing that downward comparison processes had a broad reach. For the most part, the studies Wills reviewed had been conducted to test other theories, but his analysis was persuasive. His review seemed to support two central ideas: threatened people (and those of low self-esteem) are more likely to compare with others who are worse off than themselves than do others who are better off (or of high self-esteem); and exposure to a less fortunate other (i.e., a downward target) boosts subjective wellbeing. DCT was broad, bold, and focused on "hot cognition" and it is not surprising that researchers found it appealing. For some time, a motive for self-enhancement, rather than self-evaluation, occupied the minds (and ambitions) of social psychological researchers.

DCT really took off due to a field study published by Joanne Wood, Shelley Taylor and Rosemary Lichtman (1985). They interviewed breast cancer survivors to learn how they adjusted to their condition. To the surprise of the researchers, a large majority of the women spontaneously reported coping better than other cancer patients (e.g., "I only had a lumpectomy, but those other women lost a breast.") In other words, the patients compared with others who were less fortunate. While Wills (1981) had relied on already-published studies, Wood et al. (1985) presented a new interesting result that reinforced his general thesis. Together, the papers showed that (downward) social comparison could serve as a means of coping, added to the novelty and applicability of DCT by making connections with the growing literature on stress and coping, health psychology, and the psychology of the self (Crocker and Major, 1989).

A stream of experiments followed showing how downward comparisons might aid the adjustment of medical patients and other stressed populations (see Gibbons and Gerrard, 1991; Tennen et al., 1991). Other researchers looked at psychological factors that moderated the motivation to compare downward and responses to such comparisons (Major et al., 1991; Tesser, 1998; Testa and Major, 1990). An entire edited volume by Bram Buunk and Rick Gibbons (1997) was devoted to coping and social comparison, mostly through downward comparison. DCT was a solid pillar of social psychology. Furthermore, by the 1990s a coherent narrative had emerged – people seek out and are influenced by similar others on related attributes when they want to accurately assess their opinions and abilities; but they seek out downward comparisons when they want to reduce threat or enhance the self (Wood, 1989).²

UP AND DOWN

There were researchers, including ourselves, not as sanguine about the ubiquity or benefits

of downward comparison. Hakmiller's (1966) threatened subjects did show more interest in seeing the score of someone higher in hostility, but most of the relevant evidence showed avoidance of more positive scorers (e.g., Smith and Insko, 1987). Reluctance to compare with the most fortunate is not the same as seeking someone worse-off (Wheeler, 2000; Wheeler and Miyake, 1992). Jerry Suls was not convinced because in several earlier experiments, he and Miller found that subjects who thought they had performed poorly still preferred to affiliate with people who had more ability (Suls and Miller, 1978). Of course, affiliation should not be equated with social comparison because wanting company can result from other motives. But Suls and Tesch (1978) found that even students who failed an exam preferred to know the high performance scores – scarcely evidence for downward comparison. However, we felt like party-poopers at the campfire – disconcerted about all of the fuss made about downward comparison. Truth be told, Dusty (Ladd Wheeler) was a confirmed nonbeliever, while Lefty (Jerry Suls) was more of an agnostic. We didn't doubt that downward comparison could make a person feel better, but would a threatened person be able to find and use downward targets?

The picture began to change as more evidence accumulated. First, there was the recognition that the prediction that downward comparison could protect threatened self-esteem was better supported by correlational survey and interview studies than experimental evidence. Also, contrary to downward comparison's prediction, people with high self-esteem seemed to benefit more from downward comparisons (Crocker et al., 1987). Another problem was that experiments that actually manipulated upward versus downward comparison often failed to include no-comparison control groups so it was unclear whether downward comparison increased well-being or upward comparison decreased it (Wheeler 2000). When LW measured spontaneous social comparisons in

daily life, he found people made downward comparison when they felt happy rather than unhappy – directly contrary to the DCT prediction (Wheeler and Miyake, 1992).

The solidity of downward comparison also was questioned by Taylor and Lobel (1989) who proposed that persons under threat might be served by both upward *and* downward comparisons. Perceiving oneself as better off than someone less fortunate can enhance wellbeing while having contact with someone more fortunate can provide inspiration and provide information about how to improve. The authors did not actually have any direct empirical evidence for these ideas, but as one of the authors, Shelley Taylor, was so strongly associated with the earlier breast cancer survivor downward comparison results and the paper was published in *Psychological Review*, the major theoretical journal in the field, the proposal received much notice.

The other article also included Taylor as a coauthor. Buunk et al. (1990) reported the results from two surveys that inquired about the degree to which cancer patients and college students experienced both positive and negative effects of making upward and downward comparisons. The results showed that direction of comparison was not intrinsically associated with positive or negative affect. In some cases, subjects reported positive reactions to upward comparison and negative reactions to downward comparisons. Affective responses seemed to depend less on the direction and more on the salient implication of the comparison – “Will I get better or will I get worse?” At the time, we had two opposing reactions to Buunk et al.’s article. On the positive side, we hoped the paper would give downward comparison researchers pause, but the argument rested on a correlational methodology and small effects.

Fortunately, experimental results eventually emerged showing that exposure to superstars produced positive effects on self-evaluations – a result directly in contradiction to downward comparison’s basic theses. Lockwood and Kunda (1997) had first- and fourth-year students read a newspaper article

about an outstanding fourth-year student of matching major and gender. This student was outstanding in academics, student government, sports, and leadership. Subjects in the control condition were not exposed to a target. Afterward, subjects completed self-ratings. For fourth-year students, the superstar had no effect on self-ratings, but for the first-year students, exposure to the superstar was associated with higher self-ratings. The researchers reasoned that the superstar could be inspiring to first-year students because such success was still attainable for them; this was, of course, not the case for fourth-year students. The idea that upward comparisons could be inspiring and could elicit positive affect led comparison researchers in a new direction. The analysis of the role of upward and downward comparison required a more nuanced treatment than offered heretofore.

THE RETURN OF SELF-EVALUATION: PROXY RIDES INTO TOWN

During the heyday of downward comparison theory/research, one of the cowboys (Jerry Suls) still was obsessing about the attributional reformulation and whether it really resolved the original ambiguity in Festinger’s theory. He had already begun to discuss this at some length with Renny Martin, a sharpshooter (the “Annie Oakley” of this story) at the University of Iowa. He also shared this worry with Ladd Wheeler at a Nags Head conference over a plate of barbecue and beans. Did comparison with someone similar in related attributes provide “an accurate evaluation” of one’s ability, as Goethals and Darley (1977) claimed? If I outride someone who has the same experience on horseback, the same age, and so on (i.e., related attributes), what have I learned? It means I have more than met my potential. If the other person outrides me, then it means I have not met my potential. If both of us ride at the same level, then “I am as good as

I ought to be.” But this doesn’t address whether I am a “good,” “average,” or “poor” rider. Goethals and Darley’s analysis answered a perfectly fine question, but it wasn’t the one they emphasized in their essay.

It was then that Lefty, Dusty and Annie reread Festinger’s *Nebraska Symposium* chapter and realized that Festinger’s question was “Can I do X?” and not “Do I perform X as well as I should?” We also looked at a paper published by Jones and Regan back in 1974, which was the only paper that investigated Festinger’s action possibility argument. They showed, first, that people are most interested in comparative information about an ability when they anticipate making a decision about an action based on that ability. Second, they found that the preference for comparison with similar others is strongest when these similar others have experience using that ability in situations relevant to the participants’ decisions.

Jones and Regan noted that

[w]hat appears to be a hidden assumption in the theory is that by comparison the individual learns not only where he stands relative to others but also something about what it means for his decisions, actions, and outcomes to stand at that level. In other words, the critical knowledge the individual seeks to gain concerns not so much the level of his ability as what he can and cannot accomplish with whatever level of ability he has.

(1974: 140–141)

We believe that Jones and Regan (1974) completely nailed social comparison theory. However, we don’t recall reading the paper when it appeared in the *Journal of Experimental Social Psychology* (we would have read it), and the paper had minimal impact (four citations in the 70s and 80s). It is a nice illustration of scientists not being ready to hear something. When we finally came to accept the Jones and Regan (1974) argument, we extended it and developed what we called the “proxy model of social comparison” (Wheeler et al., 1997). Our basic assumption was that by comparing ourselves with a proxy who has attempted to

perform a behavior X, we can determine our own likelihood of success at X. In order for this to be true, the proxy must be similar to us on the underlying ability. However, we can’t directly observe ability, but only performance, and we need some way of knowing if the performance we observe is truly indicative of the proxy’s ability.

Proxy extended the Jones and Regan (1974) argument by including the concept of “related attributes” from Goethals and Darley (1977) and recognizing the importance of maximal effort. A proxy’s prior success on a novel task (“X”) should be a good index of one’s likely future performance on “X,” if both self and proxy performed similarly on a prior related task and the proxy is known to have exerted maximal effort on that occasion. If it is unclear whether proxy made a maximum effort then proxy may be an inappropriate comparison (e.g., if proxy was fatigued on the first performance, it may be an underestimate of what they can do and a poor prognosticator for one’s own future success).

Another element of proxy theory is that if information about proxy’s maximum effort is unavailable, then proxy’s success/failure can still be informative if the individual and proxy share related attribute standing. Finally, knowing proxy made a maximum effort and performed the same as self on a prior (related) task means similarity on related attributes is not critical. This means that related attribute similarity is not always required for a proxy to be an appropriate and useful comparison other.

The “proxy” research was an intellectual satisfying collaborative effort for the three of us, resulting in a theoretical article (Wheeler et al., 1997) in the first issue of *Personality and Social Psychology Review*. The theory article was followed by a series of four experiments supporting the predictions described above (Martin et al., 2002) and publication of two chapters in edited collections (Martin, 2000; Suls et al., 2000). (We should note around the same time William Smith of Vanderbilt also recognized

the limitations of the Goethals and Darley formulation and, in fact, was the first to use the term “proxy”; see Smith and Sachs, 1997).

The work on proxy was satisfying to us, but like Jones and Regan (1974), nobody cites it. Most writers continue to refer to Goethals and Darley (1977) without recognition that, like Festinger (1954a), their approach was ambiguous about what self-evaluative question it answers. A cynical cowboy might conclude that perhaps the unexamined life *is* worth living (or at least being cited for). More likely, the lack of interest is a result of current social psychological excitement about automatic processes, heuristics, and hot cognitions. Proxy is about deliberative processing and will probably remain uncited until the social psychology pendulum swings back, as we trust and hope it will (prior to our move to the Cowboy Retirement Home).

In the meanwhile, we continue to maintain that the questions subsumed under the “self-evaluation” umbrella need to carefully distinguished. On some occasions, people want to know, “Do I play golf as well as I ought to?” On other occasions, especially where there are significant costs associated with failing, they want to know, “Can I win this match?” There are probably also other kinds of self-evaluative questions waiting for consideration (for our “spin” on opinion comparison, see Suls et al., 2000).

SOCIAL COGNITION’S TURN

Although social comparison researchers had ingested attribution theory and the psychology of self in the 1970s to 1990s (Suls & Wheeler, 2000), the area showed little recognition or application of basic cognitive process of judgment and perception as Kruglanski and Mayseless (1990) and Wills and Suls (1991) noted. Social comparison should be relevant to and informed by what psychologists have learned about psychophysical processes (Herr et al., 1983;

Parducci, 1974), judgments of similarity and dissimilarity (Tversky, 1977), automatic processing, and social cognition. This began to be recognized in the mid 1990s and happily continues to the present.

In a *Psychological Bulletin* review, Collins (1996) resurrected the Wheeler (1966) results for upward comparison. Her basic idea was that upward comparison could lead either to contrast or to assimilation, depending on whether the comparison is construed as indicating similarity to, or difference from, the comparison person. A major factor in this determination is the comparer’s expectation, which acts as a top-down cognitive influence on perception and judgment. People are more likely to perceive similarity when they expect to find it (Manis et al., 1991). Since people want to believe they have positive characteristics, they perceive similarity with upward targets and conclude, “[T]hey are among the better ones” (Collins, 2000: 170). By this time, sufficient evidence was available demonstrating that social comparison could lead to contrast or assimilation. The other major lesson was that comparison outcomes are not intrinsically linked to the direction of the comparison (Buunk et al., 1990). It remained to be determined what were the critical factors leading an assimilative or contrastive outcome.

Two contenders at the Comparison Corral

Once assimilation was advanced as a potential outcome of social comparison, connections to the priming literature in social cognition became apparent. The basic idea of priming is that increasing a category’s cognitive accessibility, by prior use or subliminal presentation, increases the probability for a new ambiguous stimulus to be interpreted consistently with the accessible category (Bruner, 1957; Neely, 1976, 1991). Priming a concept, for example, by having subjects unscramble four-word sequences to create three-word

sentences that described behaviors related to a trait construct such as hostility, would cognitively activate the concept of hostility. Subjects who subsequently read a passage describing a target's activities ambiguous with respect to hostility (e.g., complained to a store clerk) would then rate the target's behaviors as more hostile (Srull and Wyer, 1979; also see Higgins, 1996; for a meta-analysis see DeCoster and Claypool, 2004). In essence, priming moves the evaluation of the target stimulus closer to the primed standard – assimilation. Subsequent research (Lombardi et al., 1987) showed that under certain conditions, priming also could displace the judgment of the stimulus away from the primed standard – contrast. Extension from other-evaluation to social comparison and self-evaluation was a logical development. Two major theories have emerged.

Selective accessibility model (SAM)

In Mussweiler's (2003; see also Mussweiler and Strack, 2000) approach, whether the response to social comparison elicits self-evaluations toward or away from the comparison target depends on what information is cognitively accessible. Almost at the moment of exposure, an initial holistic assessment of the similarity between the target and the self is made. This is consistent with cognitive psychology, which finds that people rapidly consider a small number of salient features to determine whether an object and a target are generally similar or dissimilar. In the case of social comparison, this might be salient characteristics, such as gender or age. Then, the holistic impression prompts subsequent information retrieval that focuses on hypothesis-consistent evidence (Klayman and Ha, 1987). A general impression of similarity between the comparer and the target sets in motion a process of "similarity testing." Because there are many facets of the self, people can construe self-knowledge in such a way that accessible knowledge is consistent with the initial holistic impression. The consequence

is that the self-evaluations should draw closer to the target after selective search for similarity, leading to assimilation.

There are cases where the initial holistic impression is one of dissimilarity, which should prompt selective retrieval of target-inconsistent knowledge about the self, or construal consistent with the initial impression of difference. After dissimilarity search, self-evaluations should be displaced from the comparison target, leading to contrast.

Empirical support for SAM is good (e.g., Mussweiler et al., 2005). For example, priming social standards makes certain kinds of standard-consistent or standard-inconsistent information more cognitively accessible and hence faster to access in lexical priming tasks (Mussweiler and Strack, 2000). In addition, if subjects are initially led to focus on similarities versus differences; for example, between two pictorial scenes, this primes subsequent similarity versus dissimilarity testing with respect to a comparison target, such as another college student described as adjusting well or poorly. When the subjects subsequently evaluated their own college adjustment, those who had been primed to focus on similarities listed more social activities and friends after comparison with a very adjusted target than after a poorly adjusted target. Those subjects primed to focus on differences reported their adjustment to college was worse after comparison with a well-adjusted target than a poorly adjusted target. Assimilation (versus contrast) with a comparison target occurred depending on whether subjects had been primed to focus on similarities or differences.

SAM has the dual virtues of being parsimonious and explaining why several factors are moderators of assimilation/contrast. These factors include attainability (Lockwood and Kunda, 1997), perceived control (Major et al., 2001), and psychological closeness to the comparison target (Brown et al., 1992; Mussweiler and Bodenhausen, 2002). They are important because each should lead to a holistic impression of similarity, promoting further testing for similarity. If information

about the self is retrieved that is consistent with the target then assimilation should result. Conversely, unattainability, low control, the absence of psychological closeness, or extreme comparison targets (Stapel and Koomen, 2000) should produce an initial impression of dissimilarity, thereby setting in motion a search for differences and eventuating in contrastive self-evaluations.

A distinctive feature of SAM is its contention that selective accessibility operates directly on the mental representations of self-evaluation. This contrasts with psychophysicists (Biernat et al., 1991) who claim that a comparison target may serve as an anchor that changes the meaning of, and interpretation of, the judgmental scale. In the company of a professional basketball player, two cowboys might rate themselves as “very short,” though objectively, the cowboys might be of average height. The “pro” might serve as a reference point that influences the interpretation of the points on the rating scale so that “average” and “short” take on a different meaning.

To avoid the possibility that it is merely the language or ratings people use to make comparative judgments, Mussweiler (2003) prefers to have subjects make absolute judgments rather than subjective ratings. Thus, instead of asking subjects how tall they are on a seven-point scale with (“very tall” to “very short”), he would ask how tall they are in feet and inches. The latter should represent the comparer’s underlying mental representation.

A final feature of SAM is that similarity testing is assumed to operate as a default. This is consistent with the literature on judgment and decision-making where people initially tend to focus on similarities rather than differences in comparisons (Chapman and Johnson, 1999). Of course, in SAM, the search for similarity is short-circuited when a target is initially perceived as extreme or distinct, does not share related attributes, or has some salient attribute suggesting the self and target are probably quite different. The holistic impression of dissimilarity should

set in motion a search for differences that leads to contrast.

Interpretation-comparison model (ICOM)

The second approach (Stapel, 2007; Stapel and Koomen, 2000) also depends on knowledge accessibility but emphasizes the way social comparison information is *used*. ICOM posits that social comparison can instigate two different processes with opposing effects. The standing of a comparison target may provide an interpretative framework to define the self and be incorporated in the self-definition with assimilation (as priming hostility prompts the interpretation of a person’s ambiguous behaviors as aggressive). Alternatively, the same comparison may serve as an extreme standard against which the self is evaluated such that the information is excluded from the self-concept – producing contrast. Stapel (2007) thinks both assimilative and contrastive processes can occur simultaneously, but when the interpretative pull is stronger, assimilation is the consequence. When the comparative push is stronger, contrast results.

The degree of push versus pull depends on several factors. One example is the *distinctness* of the comparison. When the behavior of a comparison target activates distinct actor-trait links (“Gene Autry is a skilled horseman”), self-evaluation is likely to be contrastive (“I am a poor rider”); if the target activates indistinct trait information (“skilled horseman”), the outcome is assimilative (“I am a skilled horseman”). The latter outcome is especially likely when the comparer’s self-concept is mutable or unclear, such that there is room for inclusion and/or a need for filling in the gaps in the self-concept (Stapel and Koomen, 2000).

Another consideration is whether people are *explicitly* asked to compare themselves with a comparison standard (“Do you sing as well as Hank Williams?”) or whether the comparison is *implicit* (“Think about Hank Williams”). Stapel and Suls (2004) reasoned that intentionally comparing will make self-relevant information accessible that is

consistent with the comparison target because a common ground is needed on which the target and comparer can be compared (Gentner and Markman, 1994). This similarity focus should be less likely to be a feature of spontaneous, effortless reactions to implicit social comparisons. Results showed that explicit social comparisons more often lead to assimilation with a superior or inferior comparison target; implicit comparisons were associated with contrastive self-evaluations. Explicit comparisons are also subject to more boundary conditions. For example, if the self is not perceived a mutable or there is too much delay between comparison and self-evaluation then contrast “wins” over assimilation. Contrast appears to be a more typical outcome. With hindsight, this might account for why the social comparison field took longer than expected to appreciate the possibility of assimilative outcomes.

It's a draw for now

Currently, Mussweiler's and Stapel's theories are the most popular and well-supported approaches for understanding responses to upward and comparison. While they share an underlying premise about comparison-induced knowledge accessibility, they differ in important aspects. For SAM, everything depends on whether similarity testing versus difference testing is initiated – an “either/or” process. ICOM assumes outcomes are driven by interpretation or comparison with a standard, but both may be simultaneously operating. The evaluative outcome depends on whether interpretation or comparison “wins out.” SAM posits that similarity testing tends to be a cognitive default, while ICOM does not. SAM posits that absolute judgments and behaviors are the best proof of comparison-induced changes in self-evaluations. ICOM finds consistent evidence for subjective rating, absolute judgments, and behavioral outcomes and assumes the same underlying processes are responsible. SAM assumes the same processes and outcomes result whether comparisons are implicit or explicit; ICOM argues they have

different effects. SAM seems more parsimonious to us, but only strong inference testing between the two theories will decide.

Definitions and control groups

Comparison research inspired by social cognition has focused on forced comparisons subject to, for the most part, automatic processes. We perceive progress although there are disparate results that could lead to the impression that there are no strong conclusions to be drawn. We don't share that impression, but we are keenly aware of sources of confusion and the need for better research.

One source of confusion is definitional. In a recent survey of the literature on assimilation in social comparison (Wheeler and Suls, 2007), we identified several kinds of outcomes measured in recent studies of assimilation-contrast: (single) attribute evaluations, global evaluation, mood, and behavior. Throughout this chapter, we have used the terms “contrast” and “assimilation” carefully to refer to self-evaluations displaced toward or away from the comparison target. The literature is not as consistent; there is a tendency for researchers to lump together as evidence different kinds of outcomes and not fully acknowledge nonparallel findings. For example, Buunk and Ybema (2003) exposed Dutch married women to a description by another married woman of her either happy or unhappy marriage. Affect was more positive following an upward comparison, but subjective evaluation of the participant's relationship was lower. The authors concluded that this demonstrated assimilation on affect and contrast on self-evaluation. According to our definition, assimilation can occur only on self-evaluation; it cannot occur on affect. There are various reasons why affect might increase following exposure to a successful other, but assimilation is not one of them.

In addition, some writers equate assimilation with Tesser's (1988) reflection

or basking-in-reflected-glory. As we observed elsewhere, “feeling good because one knows Mick Jagger’s friend is not the same as believing one can perform before crowds in Madison Square Garden,” (Suls and Wheeler, 2007: 35). Assimilation simply does not exist in Tesser’s model; all comparison results in contrast. The reflection process leads to positive affect but not to increased self-evaluation on the comparison dimension.

A second problem is the missing control group problem (Wheeler, 2000; Wheeler and Suls, 2007). Researchers rarely include no-comparison control groups in their designs so it is not possible to determine whether upward target or downward target is attracting or repelling or both are acting. This is a critical issue because inclusion of a control group is the only way to know whether the cognitive processes posited by Mussweiler and Stapel are actually operating as they claim. Also, both cognitive models imply that upward assimilation and downward assimilation are just as likely given the right circumstances. In our reviews of the literature, however, while we find that upward assimilation appears to be robust, downward assimilation is rare: people do not assimilate to those who are worse off. This is consistent with Collins’ (1996) assertion that people generally think they are good and assume similarity upward, not downward.

Self-enhancement

The latter observation provides a transition to the third issue – the role of the self-enhancement motive. It is a bit ironic that responses to upward and downward comparison were originally connected to self-enhancement, but currently popular theories, such as Stapel’s and Mussweiler’s, emphasize cognitive processes and cognitive moderators. To address this gap, our recent work has involved the development of an approach/avoidance model of upward–downward comparison effects (Suls and Wheeler, 2008). Our framework borrows heavily from SAM but

recognizes that self-enhancement lumps together two very different motives – the need to excel and the need to avoid failure. Whether our approach-avoidance model properly integrates cognitive and affective processes only more time on the research trail will tell.

APPLICABILITY TO SOCIAL ISSUES

This chapter has emphasized the progression of theory and evidence, but social comparison also has offered insights into a wide variety of social problems and applications. In the interests of space, our examples will be selective.

Education

Radloff (1966) found that extreme scorers, who, by definition, would have difficulty finding similar comparison others, are very uncertain about their abilities and fluctuate in their self-evaluations. He suggested this is a good reason to institute ability tracking in the schools – to provide both gifted and untalented classmates with appropriate peers with whom to compare. It has been documented, however, that tracking is associated with negative academic self-concept, especially in higher-ability classrooms. Students with the same ability (as measured by standardized tests) typically have *lower* academic self-concept when they attend *higher*-ability schools than when they attend lower-ability schools, a finding known as the *big-fish-little-pond effect* (“BFLPE”; Marsh and Hau, 2003). The BFLPE is explained as a negative contrast effect created by invidious classroom comparisons, but heretofore little direct evidence for comparisons mediating the negative contrast effect. The BFLPE also seems to be contradicted by consistent findings that students’ performance improves if they report comparing their exam grades with a classmate who performed (slightly) *better* than

themselves (Blanton et al., 1999; Huguet et al., 2001).

We (Wheeler and Suls, 2005) proposed that forced upward social comparisons, characteristic of BFLPE, and strategic deliberate upward comparison choices can coexist. Students in low- and high-ability schools may deliberately select classmates with *slightly better* grades (and therefore attainable accomplishments) as comparison targets, but students in high-ability schools are also involuntarily exposed to “superstars” (whose accomplishments might be seen as unattainable), and thus suffer a decline in self-concept. The net result would be a lower academic self-concept in the high-ability schools or classes.

With an international team of collaborators including Lefty and Dusty, Pascal Huguet and Florence Dumas tested the role of deliberate and forced comparisons and their effect on academic self-concept in a sample of more than 2,000 French secondary school students. Measures of academic self-concept and perceived classroom standing and nominations of the classmate with whom they preferred to compare their grades were collected. Students’ aptitude test scores were also available. The negative contrast effect was replicated in this sample, but it disappeared after statistically controlling for students’ invidious comparisons in the classroom, providing the first direct evidence that social comparison mediates the BFLPE. But the BFLPE coexisted with the tendency for students to nominate a better-performing classmate to compare their grades. Moreover, this tendency was positively related to academic self-concept, suggestive of an assimilation effect. After statistically controlling for the comparison choice effect, the magnitude of BFLPE was reduced, but not eliminated. Huguet et al. (2009) concluded that BFLPE is the net effect of counterbalancing influences: stronger negative contrast effects associated with forced exposure to invidious comparisons at the classroom level and weaker assimilation effects associated with upward social comparison choices.

For educational policy makers, this work suggests that academic selective schools do not automatically benefit the students who attend them, contrary to a largely uncritical belief.

Health

The original application of downward comparison to the adaptation of the chronically ill medical patients was already described. This was mainly correlational-descriptive research (Buunk and Gibbons, 1997; Wood et al., 1985) and has not caught up with recent lab findings on forced automatic processes. An exception is Stanton et al. (1999) who assigned breast cancer survivors to listen to an audiotape interview of a (supposed) patient whose comments reflected good, poor, or unspecified psychological and physical health status. While patients who had listened to the poorly adjusted patient rated their own adjustment as better than those exposed to the well-adjusted patient, even those who heard a poor-functioning survivor reported feeling better than the patient in the interview. In short, and reminiscent of Buunk et al. (1990), patients were able to find positive meaning in either direction. Their findings also reinforce our belief that downward assimilation rarely occurs, even when a downward comparison is forced. In any case, the kind of research conducted by Stanton et al. has the potential to inform medical professionals about what kinds of patient models and instructions are appropriate for psycho-educational materials, such as videos used in medical rehabilitation.

The relevance of proxy and upward comparison can be seen in Kulik and Mahler’s (1987) demonstration that coronary-bypass surgery patients experience better adjustment post-bypass if prior to the procedure they are assigned to share a room with a patient who already undergone bypass (versus a patient who also is waiting). Assigning patients to different roommates is feasible and has implications for cardiology practice.

Support groups allowing for open disclosure and discussion are a standard psychosocial referral for cancer patients. These groups tend to include all patients with a particular cancer, regardless of the need for psychosocial treatment, which means the groups are heterogeneous in terms of distress level. One concern is that nondistressed members may actually suffer deficits in such groups, leading to suggestions that groups should be homogenous in composition. This idea has been resisted because nondistressed patients are thought to serve as important role models. Several reviews find that there are overall improvements in depression, anxiety, and self-efficacy as a function of group interventions (e.g. Spiegel et al., 1981), but the biggest gains are typically seen in those patients with the fewest psychosocial resources.

Support groups fulfill needs for affiliation, receipt of information, emotional ventilation, and social comparison. A recent social comparison analysis of the social support group process (Carmack Taylor et al., 2007) drew on findings from the upward and downward comparison literature. On this basis, they concluded that distressed patients need to be in heterogeneous-composition groups so they are exposed to better functioning patients from whom they can be inspired, learn, and assimilate. There is insufficient evidence to determine whether nondistressed patients may receive no benefits or even experience negative effects from associating with low-functioning patients. Thus far, only a single study has shown that patients with ample psychosocial resources may be harmed by the group experience (Helgeson et al., 2000). Based on our previous comments, we are doubtful that nondistressed patients experience downward assimilation, but they may find the support experience irrelevant for them. Carmack Taylor et al. correctly call for more research, but propose that well-functioning patients might be recruited to serve as “group volunteers” and trained to meet the needs of distressed patients for affiliation and information. Framed in

this way, nondistressed patients may desire participation and volunteer their time in the spirit of altruism.

Economics, subjective wellbeing, and greed

We are writing this chapter during serious economic times. The stock market lost millions of dollars, housing mortgages are in default, and the federal government has given record amounts of money to banks to keep them afloat. Social comparisons are prevalent and irritating. One large company, which has already received billions of dollars in bailout money from the federal government and has asked for more, gave top company administrators million dollar bonuses. Letters to the editor from “average wage-earners” complained about the unfairness of the huge bonuses while more people are losing their jobs. It is obvious these highly paid CEO’s are creating invidious comparisons that effect short-term changes in subjective wellbeing among the general public.

The current economic recession/depression has been explained by faulty regulation in the last two presidencies and increasing levels of greed. The reason for a lack of regulatory oversight is clear, but why should greed have increased? It has many sources, social comparison being one. This is very nicely demonstrated by an archival analysis tracking the causal impact of the introduction of television from 1951 to 1955 on FBI indicators of crime, burglary, auto theft, and larceny in the US (Hennigan et al., 1982). The researchers hypothesized that as the possessions and lifestyles of the advantaged become increasingly visible to the less advantaged, opportunities for comparison and frustration increase. Television allows everyone to learn about both the availability of goods and the conspicuous consumption of the “rich and famous.” In a time-series design, the introduction of television in different geographical regions had no effect on violent crimes, burglary, or auto theft.

However, larceny increased as television was introduced. If one feels one deserves desirable goods, but lack the legal means, then theft becomes an option. In recent years, exposure to luxury goods, vacations, and wealth has become even more prevalent – leading to invidious social comparisons and perhaps increasing greed. The current economic predicament provides a veritable comparison laboratory.

IS THIS THE END OF THE TRAIL?

This chapter has covered more than 50 years of theorizing and research to tell the story of social comparison. In 40 pages, this could not be the complete story. There are more yarns and other cowboys and cowgirls to tell them. One thing we are sure of is that comparison, after a slow start, has shown sustained growth – we think because it is a core element of social life. The lesson may be to never kick what you think is a dead horse; it may only be getting a few winks of sleep while waiting for the sun to come-up and the cattle drive to continue.

NOTES

1 The prose in this and other sections of the *Human Relations* article is tedious. Anyone who is familiar with Festinger's other publications knows he was capable of very strong writing. Someone once speculated that Festinger wrote the *Human Relations* paper in a style intended as a sly parody of Clark Hull and Kenneth Spence's theoretical papers which were filled with dense hypotheses and corollaries. (Festinger went to graduate school at the University of Iowa and for a time was Spence's research assistant.) This is such a good story and so consistent with Festinger's sense of fun that we hope it is true.

2 In this chapter, we focus on self-evaluation and self-enhancement as comparison motives. Wood (1989) suggested a third motive – self-improvement. This refers to studying the way experts do things (e.g., swing a golf club, play a piano) in order to imitate them or by an upward comparison. We are reluctant to accept a need for a third motive,

however. First, how can we distinguish inspiration provided by a role model and self-evaluation (with unidirectional drive)? Second, self-improvement simply may be a form of self-enhancement because by improving, self-concept should be enhanced.

REFERENCES

Allen, V. (1965) Situational factors in conformity. In L. Berkowitz (ed.), *Advances in Experimental Social Psychology*, 2, 133–175. New York: Academic Press.

Arrowood, A.J. (1986) Comments on 'Social comparison theory: Psychology from the lost and found'. *Personality and Social Psychology Bulletin*, 12, 279–281.

Blanton, H., Buunk, B.P., Gibbons, F.X. and Kuyper, H. (1999) When better-than-others compare upward: Choice of comparison and comparative evaluation as independent predictors of academic performance. *Journal of Personality and Social Psychology*, 76, 420–430.

Brickman, P. and Bulman, R.J. (1977) Pleasure and pain in social comparison. In J. Suls and R.L. Miller (eds), *Social Comparison Processes: Theoretical and Empirical Perspectives*, pp. 149–186. Washington, DC: Hemisphere.

Brickman, P. and Campbell, D.T. (1971) Hedonic relativism and planning the good society. In M.H. Appley (ed.), *Adaptation-level Theory: A Symposium*, pp. 287–302. New York: Academic Press.

Brown, J.D., Novick, N.J., Lord, K.A. and Richards, J.M. (1992) When Gulliver travels: Social context, psychological closeness, and self-appraisals. *Journal of Personality and Social Psychology*, 62, 717–727.

Bruner, J.S. (1957) On perceptual readiness. *Psychological Review*, 64, 123–152.

Buunk, B.P., Collins, R.L., Taylor, S.E., VanYperen, N.W. and Dakof, G.A. (1990) The affective consequences of social comparison: Either direction has its ups and downs. *Journal of Personality and Social Psychology*, 59, 1238–1249.

Buunk, B. and Gibbons, F.X. (eds) (1997) *Health, Coping and Well-being: Perspectives from Social Comparison Theory*. Mahwah, NJ: Lawrence Erlbaum Associates.

Buunk, B.P. and Ybema, J.F. (2003) Feeling bad, but satisfied: The effect of upward and downward comparison upon mood and marital satisfaction. *British Journal of Social Psychology*, 42, 613–628.

Carmack Taylor, C.L., Kulik, J., Badr, H., Smith, M., Basen-Engquist, K., Penedo, F. and Gritz, E.R. (2007) A social comparison theory analysis of group composition and efficacy of cancer support group programs. *Social Science and Medicine*, 65, 262–273.

Cathcart, T. and Klein, D. (2007) *Plato and a Platypus Walk into a Bar*. New York: Penguin Books.

Collins, R.L. (1996) For better or for worse: The impact of upward social comparisons on self-evaluations. *Psychological Bulletin*, 119, 51–69.

Collins, R.L. (2000) Among the better ones: Upward assimilation in social comparison. In J. Suls and L. Wheeler (eds), *Handbook of Social Comparison: Theory and Research*, pp. 159–171. New York: Kluwer Academic/Plenum.

Crocker, J. and Major, B. (1989) Social stigma and self-esteem: The self-protective properties of stigma. *Psychological Review*, 96, 608–630.

Crocker, J., Thompson, L., McGraw, K. and Ingerman, C. (1987) Downward comparison prejudice and evaluations of others: Effects of self-esteem and threat. *Journal of Personality and Social Psychology*, 52, 907–916.

DeCoster, J. and Claypool, H.M. (2004) A meta-analysis of priming effects on impression formation supporting a general model of information biases. *Personality and Social Psychology Review*, 8, 2–27.

Deutsch, M. and Krauss, R.M. (1965) *Theories in Social Psychology*. New York: Basic Books.

Festinger, L. (1950) Informal social communication theory. *Psychological Review*, 57, 271–282.

Festinger, L. (1954a) A theory of social comparison processes. *Human Relations*, 7, 117–140.

Festinger, L. (1954b) Motivation leading to social behavior. In M.R. Jones (ed.), *Nebraska Symposium on Motivation*, pp. 191–218. Lincoln, NE: University of Nebraska Press.

Gastorf, J.W. and Suls, J. (1978) Performance evaluation via social comparison: Related attribute similarity vs. performance similarity. *Social Psychology Quarterly*, 41, 297–305.

Gibbons, F.X. and Gerrard, M. (1991). Downward comparison and coping with threat. In J. Suls and T.A. Wills (eds), *Social Comparison: Contemporary Theory and Research*, pp. 317–346. Hillsdale, NJ: Lawrence Erlbaum.

Goethals, G. and Darley, J. (1977) Social comparison theory: An attributional approach. In J. Suls, and R. Miller (eds), *Social Comparison Processes: Theoretical and Empirical Perspectives*, pp. 259–278. Washington, DC: Hemisphere.

Gordon, B.F. (1966) Influence and social comparison as motives for affiliation. *Journal of Experimental Social Psychology, Suppl.* 1, 2, 55–65.

Helgeson, V.S., Cohen, S., Schulz, R. and Yasko, J. (2000) Group support interventions for women with breast cancer: Who benefits from what? *Health Psychology*, 19, 107–114.

Hennigan, K., Rosario, M., Heath, L., Cook, T.D., Wharton, J. and Calder, B. (1982) Impact of the introduction of television on crime in the United States: Empirical findings and theoretical implications. *Journal of Personality and Social Psychology*, 42, 461–477.

Herr, P.M., Sherman, S.J. and Fazio, R.H. (1983) On the consequences of priming: Assimilation and contrast effects. *Journal of Experimental Social Psychology*, 19, 323–340.

Higgins, E.T. (1996) Knowledge activation: Accessibility, applicability and salience. In E.T. Higgins and A. Kruglanski (eds), *Social Psychology: Handbook of Basic Principles*, pp. 133–168. New York: Guilford Press.

Huguet, P., Dumas, F., Marsh, H., Seaton, M., Wheeler, L., Suls, J., Regner, I. and Nezlek, J. (2009) Clarifying the role of social comparison in the big-fish-little-pond effect (BFLPE): An integrative study. *Journal of Personality and Social Psychology: Interpersonal Relationships and Group Processes*, 97, 156–170.

Huguet, P., Dumas, F., Monteil, J.M. and Genestoux, N. (2001) Social comparison choices in the classroom: Further evidence for students' upward comparison tendency and its beneficial impact on performance. *European Journal of Social Psychology*, 31, 557–578.

Jones, S. and Regan, D. (1974) Ability evaluation through social comparison. *Journal of Experimental Social Psychology*, 10, 133–146.

Kelley, H. (1967) Attribution theory in social psychology. In D. Levine (ed.), *Nebraska Symposium on Motivation*, pp. 192–238. Lincoln, NE: University of Nebraska Press.

Kelley, H. (1973) The process of causal attribution. *American Psychologist*, 28, 107–128.

Klayman, J. and Ha, Y.W. (1987) Confirmation, disconfirmation and information in hypothesis-testing. *Psychological Review*, 94, 211–228.

Kruglanski, A. and Myslinski, O. (1990) Classic and current social comparison research: Expanding the perspective. *Psychological Bulletin*, 108, 195–208.

Kulik, J.A. and Mahler, H.I.M. (1987) Effects of preoperative roommate assignment on preoperative anxiety and postoperative recovery from coronary bypass surgery. *Health Psychology*, 6, 525–543.

Latané, B. (ed.) (1966) Studies in social comparison. *Journal of Experimental Social Psychology, Suppl. 1*.

Latané, B., Eckman, J. and Joy, V. (1966) Shared stress and interpersonal attraction. *Journal of Experimental Social Psychology, Suppl. 1, 2*, 92–102.

Lockwood, P. and Kunda, Z. (1997) Superstars and me: Predicting the impact of role models on the self. *Journal of Personality and Social Psychology, 73*, 91–103.

Lombardi, W.J., Higgins, E.T. and Bargh, J.A. (1987) The role of consciousness in priming effects on categorization: Assimilation versus contrast as a function of awareness of the priming task. *Personality and Social Psychology Bulletin, 13*, 411–429.

Major, B., Testa, M. and Blysm, W. (1991) Responses to upward and downward social comparisons: The impact of esteem-relevance and perceived control. In J. Suls and T.A. Wills (eds), *Social Comparison: Contemporary Theory and Research*, pp. 237–260. Hillsdale, NJ: Lawrence Erlbaum Associates.

Manis, M., Biernat, M. and Nelson, T.F. (1991) Comparison and expectancy processes in human judgment. *Journal of Personality and Social Psychology, 61*, 203–211.

Marsh, H.W. and Hau, K. (2003) Big-fish-little-pond-effect on academic self-concept. A cross-cultural (26 country) test of the negative effects of academically selective schools. *American Psychologist, 58*, 364–376.

Martin, R. (2000). 'Can I do X?' Using the proxy comparison model to predict performance. In J. Suls and L. Wheeler (eds), *Handbook of Social Comparison: Theory and Research*, pp. 67–80. New York: Kluwer/Plenum.

Martin, R., Suls, J. and Wheeler, L. (2002) Ability evaluation by proxy: The role of maximum performance and related attributes in social comparison. *Journal of Personality and Social Psychology, 82*, 781–791.

Morse, S. and Gergen, K.J. (1970) Social comparison, self-consistency, and the concept of the self. *Journal of Personality and Social Psychology, 16*, 148–156.

Mussweiler, T. (2003) Comparison processes in social judgment: Mechanisms and consequences. *Psychological Review, 110*, 472–489.

Mussweiler, T. and Bodenhausen, G. (2002) I know you are, but what am I? Self-evaluative consequences of judging in-group and out-group members. *Journal of Personality and Social Psychology, 82*, 19–32.

Mussweiler, T., Epstrude, K. and Ruter, K. (2005) The knife that cuts both ways: Comparison processes in social perception. In M. Alicke, D. Dunning, and J. Krueger (eds), *The Self in Social Judgment*, pp. 109–130. New York: Psychology Press.

Mussweiler, T. and Strack, F. (2000) The 'Relative Self': Informational and judgmental consequences of comparative self-evaluation. *Journal of Personality and Social Psychology, 79*, 23–38.

Neely, J.H. (1976) Semantic priming and retrieval from lexical memory: Evidence for facilitory and inhibitory processes. *Memory and Cognition, 4*, 648–654.

Neely, J.H. (1991) Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner and G.W. Humphreys (eds), *Basic Processes in Reading: Visual Word Recognition*, pp. 264–336. Hillsdale, NJ: Lawrence Erlbaum.

Parducci, A. (1974) Contextual effects: A range-frequency analysis. In E.C. Carterette and M.P. Friedman (eds), *Handbook of Perception, 7*, 127–141. New York: Academic Press.

Pettigrew, T. (1967) Social evaluation theory: Convergences and applications. In D. Levine (ed.), *Nebraska Symposium on Motivation*, pp. 241–318. Lincoln, NE: University of Nebraska Press.

Radloff, R. (1966) Social comparison and ability evaluation. *Journal of Experimental Social Psychology, Suppl. 1, 2*, 6–26.

Schachter, S. (1959) *The Psychology of Affiliation*. Stanford, CA: Stanford University of Press.

Smith, W. and Sachs, P. (1997) Social comparison and task prediction: Ability similarity and the use of a proxy. *British Journal of Social Psychology, 36*, 587–602.

Spiegel, D., Bloom, J. and Yalom, I. (1981) Group support for patients with metastatic cancer: A randomized prospective outcome study. *Archives of General Psychiatry, 38*, 527–533.

Srull, T.K. and Wyer, R.S. (1979) The role of category accessibility in the interpretation of information about persons: Some determinants and implications. *Journal of Personality and Social Psychology, 37*, 1660–1672.

Stanton, A., Danoff-Burg, S., Cameron, C., Snider, P. and Kirk, S. (1999) Social comparison and adjustment to breast cancer: An experimental examination of upward affiliation and downward evaluation. *Health Psychology, 18*, 151–158.

Stapel, D.A. (2007) In the mind of the beholder: The Interpretation Comparison Model of accessibility effects. In D.A. Stapel and J. Suls (eds), *Assimilation and Contrast in Social Psychology*, pp. 143–164. New York: Psychology Press.

Stapel, D.A. and Koomen, W. (2000) Distinctiveness of others, mutability of selves: Their impact on self-evaluations. *Journal of Personality and Social Psychology*, 79, 1068–1087.

Stapel, D.A. and Koomen, W. (2001b) The impact of interpretation versus comparison goals on knowledge accessibility effects. *Journal of Experimental Social Psychology*, 37, 134–149.

Stapel, D.A. and Suls, J. (2004) Method matters: Effects of explicit versus implicit social comparisons on activation, behavior, and self-views. *Journal of Personality and Social Psychology*, 87, 860–875.

Suls, J. (1986) Notes on the occasion of Social Comparison Theory's thirtieth birthday. *Personality and Social Psychology Bulletin*, 12, 289–296.

Suls, J., Gastorf, J. and Lawhorn, J. (1978) Social comparison choices for evaluating a sex- and age-related ability. *Personality and Social Psychology Bulletin*, 4, 102–105.

Suls, J., Martin, R. and Wheeler, L. (2000) Three kinds of opinion comparison: The Triadic Model. *Personality and Social Psychology Review*, 4, 219–237.

Suls, J., Martin, R. and Wheeler, L. (2002) Social comparison: Why, with whom and with what effect? *Current Directions in Psychological Science*, 11, 159–163.

Suls, J. and Miller, R.L. (1977) *Social Comparison Processes: Theoretical and Empirical Perspectives*. Washington, DC: Hemisphere.

Suls, J. and Miller, R.L. (1978) Ability comparison and its effects on affiliation preferences. *Human Relations*, 31, 267–282.

Suls, J. and Tesch, F. (1978) Students' preferences for information about their test performance: A social comparison study. *Journal of Applied Social Psychology*, 8, 189–197.

Suls, J. and Wheeler, L. (eds) (2000) *Handbook of Social Comparison*. New York: Kluwer/Plenum.

Suls, J. and Wheeler, L. (2007) Psychological magnetism: A brief history of assimilation and contrast in psychology. In D. Stapel and J. Suls (eds), *Assimilation and Contrast in Social Psychology*, pp. 9–44. New York: Psychology Press.

Suls, J. and Wheeler, L. (2008) A reunion for approach/avoidance motivation and social comparison. In A. Elliot (ed.), *Handbook of Approach and Avoidance Motivation*. Mahwah, NJ: Lawrence Erlbaum Associates.

Taylor, S. and Lobel, M. (1989) Social comparison under threat: Downward evaluation and upward contacts. *Psychological Review*, 96, 569–575.

Tennen, H., McKee, T.E. and Affleck, G. (1991) Social comparison processes in health and illness. In J. Suls and L. Wheeler (eds), *Handbook of Social Comparison: Theory and Research*, pp. 443–486. New York: Kluwer Academic/Plenum.

Tesser, A. (1988) Toward a self-evaluation maintenance model of social behavior. In L. Berkowitz (ed.), *Advances in Experimental Social Psychology*, 21, 181–227. New York: Academic Press.

Testa, M. and Major, B. (1990) The impact of social comparisons after failure: The moderating effects of perceived control. *Basic and Applied Social Psychology*, 11, 205–218.

Turner, J. (1991) *Social Influence*. Belmont, CA: Brooks/Cole.

Wheeler, L. (1966) Motivation as a determinant of upward comparison. *Journal of Experimental Social Psychology, Suppl.* 1, 2, 27–31.

Wheeler, L. (2000) Individual differences in social comparison. In J. Suls and L. Wheeler (eds), *Handbook of Social Comparison*, pp. 141–158. New York: Kluwer Academic/Plenum.

Wheeler, L. and Koestner, R. (1984) Performance evaluation: On choosing to know the related attributes of others when we know their performance. *Journal of Experimental Social Psychology*, 20, 263–271.

Wheeler, L., Koestner, R. and Driver, R. (1982) Related attributes in the choice of comparison others: It's there, but it isn't all there is. *Journal of Experimental Social Psychology*, 18, 489–500.

Wheeler, L., Martin, R. and Suls, J. (1997) The proxy social comparison model for self-assessment of ability. *Personality and Social Psychology Review*, 1, 54–61.

Wheeler, L. and Miyake, K. (1992) Social comparison in everyday life. *Journal of Personality and Social Psychology*, 62, 760–733.

Wheeler, L. and Suls, J. (2005) Social comparison and self-evaluations of competence. In A. Elliot and C. Dweck (eds), *Handbook of Competence and Motivation*. New York: Guilford Press.

Wheeler, L. and Suls, J. (2007) Assimilation in social comparison: Can we agree on what it is? *Revue Internationale de Psychologie Sociale*, 20, 31–51.

Wills, T.A. (1981) Downward comparison principles in social psychology. *Psychological Bulletin*, 90, 245–271.

Wills, T.A. and Suls, J. (1991) Commentary: neo-social comparison and beyond. In J. Suls and T.A. Wills (eds), *Social Comparison: Contemporary Theory and*

Research, pp. 395–412. Hillsdale, NJ: Lawrence Erlbaum.

Wood, J.V. (1989) Theory and research concerning social comparison of personal attributes. *Psychological Bulletin*, 106, 231–248.

Wood, J.V. (1996) What is social comparison and how should we study it? *Personality and Social Psychology Bulletin*, 22, 520–537.

Wood, J.V., Taylor, S. and Lichtman, R. (1985) Social comparison in adjustment to breast cancer. *Journal of Personality and Social Psychology*, 49, 1169–1183.

Zanna, M., Goethals, G.R. and Hill S. (1975) Evaluating a sex related ability: Social comparison with similar others and standard setters. *Journal of Experimental Social Psychology*, 11, 86–93.

Regulatory Focus Theory

E. Tory Higgins

ABSTRACT

Regulatory focus theory was the child of self-discrepancy theory and the parent of regulatory fit theory. As the child of self-discrepancy theory, it distinguishes between self-regulation in relation to hopes and aspirations (promotion ideals) versus self-regulation in relation to duties and obligations (prevention oughts). But in regulatory focus theory, promotion and prevention orientations are states that vary not only predispositionally across individuals but also can be situationally induced. And the emotional and motivational effects of success and failure can be on a current task rather than just chronic congruencies and discrepancies to ideal and ought self-guides. Most importantly, the emphasis is on strategic differences between promotion and prevention in how desired states are attained, with promotion preferring eager means of advancement and prevention preferring vigilant means of maintenance. This emphasis on strategic differences also distinguishes regulatory focus theory from control system theories' concern with approach and avoidance at the system level rather than the strategic level. The asymmetry between promotion and prevention in strategic preferences and in the motivational effects of success and failure, with success strengthening motivation in promotion but failure strengthening motivation in

prevention, gave birth to the regulatory fit idea that the manner of goal pursuit can sustain or disrupt a self-regulatory orientation.

INTRODUCTION

I moved from New York University (NYU) to Columbia University in 1989. The decision was very difficult, mostly because of the wonderful qualities of both options – an approach–approach conflict. The conflict would have been even worse if I had recognized and fully considered all of the consequences of each option. But I didn't. I had other things on my mind. My wife and I were about to be blessed with our first child and this was what I was thinking about. And, not surprisingly, our child was what I continued to think about for several months after she was born. It was not until my second term at Columbia (Spring, 1990) that I finally began thinking about what research I would do at Columbia. And when I did begin thinking about it, I was thunderstruck. I suddenly

realized that I could no longer do the kind of research that I had done at NYU for years. I could no longer work on self-discrepancy theory. Like an irresponsible parent, I had thoughtlessly abandoned my theory by moving to Columbia. What was I going to do?

What struck me at that painful moment was that the research conditions at Columbia were completely different from those at NYU. At NYU, where I had developed self-discrepancy theory, there was a very large subject pool – thousands of potential participants per year. In contrast, the subject pool at Columbia was just a few hundred. And there was another difference that was at least as important. At NYU all students in the subject pool filled out a battery booklet that contained not only demographic questions but also personality questionnaires. You guessed it – one of those questionnaires was the Selves Questionnaire that my lab used to measure individuals' self-discrepancies. This meant that even before my lab contacted potential participants for a study, we already had information about their self-discrepancies. We could select participants as a function of whether they were high or low in different kinds of self-discrepancies. At Columbia, on the other hand, there was no such battery booklet. My lab would have to do a formal study with hundreds of participants just to obtain the background information on their self-discrepancies. By the time we had done that, we would have no more subject hours for research. It was hopeless!

So there I was, mourning over the lost of my theory and wondering what kind of research I would do instead. My first thought was, "Well, I will just have to go back to doing research on priming and accessibility full-time." It was a comforting thought, but not comforting enough. I still missed self-discrepancy theory. After all, whereas construct accessibility theory was an independent teenager by this time (Higgins et al., 1977; see Chapter 4, this volume), self-discrepancy theory was still only a child (Higgins, 1987).

It was during this period, under these trying circumstances, that I began to find

a solution to the problem of what to do about self-discrepancy theory – a solution that ultimately became regulatory focus theory. Without question, it was trying to continue parenting self-discrepancy theory that gave birth to regulatory focus theory (see Higgins, 2004, 2006a). I realized that what I needed was to find a way to test self-discrepancy theory without having to measure people's chronic self-discrepancies. In thinking about this problem, and ultimately finding a solution, regulatory focus theory was born, although I didn't appreciate that immediately. It took me some time to recognize what had happened (a separate theory construction issue I will discuss later). But now, I need to begin at the beginning. For you to appreciate the dilemma I faced in the spring of 1990, I need to provide some basic background information about self-discrepancy theory and describe the kinds of studies we had conducted to test it.

SELF-DISCREPANCY THEORY: THE PARENT OF REGULATORY FOCUS THEORY

Why do people emotionally react so differently to the same tragic event? More specifically, why is it that when people are emotionally overwhelmed by a serious setback in their life, such as the death of their child, the loss of their job, or the breakup of their marriage, some suffer from depression whereas others suffer from anxiety? Self-discrepancy theory was developed to answer this question. Self-discrepancy theory proposed that even when people have the same specific goals, such as seniors in high school wanting to go to a good college or older adults wanting a good marriage, they often vary in how they represent these goals. The goals or standards that direct or guide our self-regulation are called *self-guides* in self-discrepancy theory. Some of us represent our self-guides as hopes or aspirations, the kind of person we ideally want to be – *ideal*

self-guides. Others of us represent our self-guides as duties or obligations, the kind of person we believe we ought to be – *ought* self-guides.

According to self-discrepancy theory (Higgins, 1987), it is the difference between failing to meet our ideals versus failing to meet our oughts that provides the key to unlocking the mystery of why we have different emotional reactions to the same negative life event. Self-discrepancy theory proposes that when a negative life event happens to us, it is represented as saying something about how we are doing. We compare our current, actual self to our self-guides: “Compared to the kind of person I want to be (e.g., going to a good college; having a good marriage), how am I doing?” We suffer emotionally when there is a discrepancy between our actual self and a self-guide – a *self-discrepancy*. When our actual self is discrepant from an ideal self-guide, we feel sad, disappointed, discouraged – dejection-related emotions that relate clinically to depression. When our actual self is discrepant from an ought self-guide, we feel nervous, tense, and worried – agitation-related emotions that relate clinically to anxiety disorders. According to self-discrepancy theory, then, our vulnerabilities to different kinds of emotional suffering depend on which type of self-guide is emphasized in our self-regulation – dejection/depression suffering when ideals are emphasized, and agitation/anxiety suffering when oughts are emphasized.

Research with clinically depressed and clinically anxious patients has found support for these proposals about emotional vulnerabilities. Discrepancies between patients’ actual selves and their ideal self-guides predicts their suffering from depression more than it predicts their suffering from anxiety disorders, whereas discrepancies between patients’ actual selves and their ought self-guides predicts their suffering from anxiety disorders more than it predicts their suffering from depression (Strauman, 1989). Because some individuals have

actual-self discrepancies from both their ideal and ought self-guides, they can suffer from both depression and anxiety disorders.

At any one time, however, either individuals’ ideal self-guides or their ought self-guides can be more accessible, and whichever is more accessible will determine which emotional syndrome they experience. This means that momentary situations can determine which syndrome is experienced by priming or activating either ideal or ought self-guides. For example, there is evidence that either actual-ideal discrepancies or actual-ought discrepancies can be made temporarily more accessible by exposing individuals either to words which relate to an ideal they possess or to an ought they possess. When such priming of either an ideal or an ought occurs in an experiment, participants whose actual-ideal discrepancy is activated suddenly feel sad and disappointed and fall into a depression-related state of low activity (e.g., talk slower). In contrast, participants whose actual-ought discrepancy is activated suddenly feel nervous and worried and fall into an anxiety-related state of high activity (e.g., talk quicker). And these kinds of effects have been found with both clinical samples (Strauman, 1989) and non-clinical samples (Strauman and Higgins, 1987).

What is the psychological mechanism that underlies these effects? Self-discrepancy theory proposes that different emotions relate to different psychological situations that people experience. That is, the psychological situations produced by success or failure to meet our *ideals* are different from the psychological situations produced by success or failure to meet our *oughts*. Specifically, when events are related to our ideal self-guides (i.e., to our hopes and aspirations), we experience success as the presence of a positive outcome (a gain), which is a happy experience, and we experience failure as the absence of positive outcome (a nongain), which is a sad experience. In contrast, when events are related to our ought self-guides (i.e., our beliefs about our duties and obligations), we experience success as the absence of

a negative outcome (a nonloss), which is a relaxing experience, and we experience failure as the presence of a negative outcome (a loss), which is a worrying experience. Consistent with this proposed underlying mechanism, research has shown (e.g., Higgins and Tykocinski, 1992) that individuals with strong ideals remember better events that reflect the absence or the presence of positive outcomes (gains and nongains), whereas individuals with strong oughts remember better events that reflect the presence or absence of negative outcomes (nonlosses and losses). People also remember better those events in their own lives that relate to whichever type of self-guide is more accessible for them (Strauman, 1992).

What kind of parenting is likely to result in children having either strong ideal self-guides or strong ought self-guides? In answering these questions, self-discrepancy theory relies on the basic idea that self-regulation in relation to ideal self-guides versus ought self-guides involves experiencing different psychological situations. When children interact with their parents (or other caretakers), the parents respond to their child in ways that make the child experience one of the different kinds of psychological situations. Over time, the children respond to themselves like their parents respond to them, producing the same specific kinds of psychological situations, and this develops into the kind of self-guide (ideal or ought) that is associated with those psychological situations (see Higgins, 1991).

What pattern of parenting, then, predicts the development of strong ideal self-guides in children? It is when parents combine bolstering (when managing success) and love withdrawal (when disciplining failure). Bolstering occurs, for instance, when parents encourage the child to overcome difficulties, hug and kiss them when they succeed, or set up opportunities for the child to engage in success activities – it creates an experience of the presence of positive outcomes in the child. Love withdrawal occurs, for instance, when parents end a meal when the child

throws some food, take away a toy when the child refuses to share it, stop a story when the child is not paying attention – it creates an experience of the absence of positive outcomes in the child.

What pattern of parenting predicts the development of strong ought self-guides in children? It is when parents combine prudence (when managing success) and punitive/critical (when disciplining failure). Prudence occurs, for instance, when parents “child-proof” the house, train the child to be alert to potential dangers, or teach the child to “mind your manners” – it creates an experience of the absence of negative outcomes in the child. Punitive/critical occurs, for instance, when parents play roughly with the child to get his or her attention, yell at the child when he or she doesn’t listen, criticize the child when he or she makes a mistake – it creates an experience of the presence of negative outcomes. Indeed, consistent with these self-discrepancy theory predictions, there is recent evidence of positive associations between the critical and punitive parenting style and prevention-focused self-regulation and between the bolstering parenting style and promotion-focused self-regulation (see Higgins and Silberman, 1998; Keller, 2008; Manian et al., 2006).

In addition to distinguishing between ideal and ought self-guides, self-discrepancy theory distinguishes between different standpoints that can be taken in self-regulation (Higgins, 1987) – between self-regulation from our *own* independent standpoint (“What are my own goals and standards for myself?”) and self-regulation from the standpoint of a *significant other* person in our lives (e.g., “What are my mother’s goals and standards for me?”). For example, in North America at least, there is evidence that discrepancies from independent self-guides are a more important determinant of emotional vulnerabilities for males than for females. In contrast, discrepancies from significant other self-guides are more important for females than for males (Moretti and Higgins, 1999a, 1999b).

NOW BACK TO THE STORY: THE BIRTH OF REGULATORY FOCUS THEORY

Research testing self-discrepancy theory had always used the participants' self-discrepancies as part of the study. Most studies had tested the effects of self-discrepancies on emotional responses – the relation between self and affect – but occasionally we took advantage of previous work on priming and accessibility to make either actual-ideal or actual-ought discrepancies temporarily more accessible by priming either ideal or ought self-guides, respectively. For example, in one study by Higgins et al. (1986) participants were selected who were either high in *both* actual-ideal and actual-ought discrepancies or low in both, and when they arrived for the study they were asked either to discuss their own and their parents' hopes and aspirations for them (ideal priming) or to discuss their own and their parents' beliefs concerning their duties and obligations (ought priming). The study found that for "high in both" participants, but not "low in both" participants, ideal priming produced dejection-related emotions whereas ought priming produced agitation-related emotions.

The idea of using priming to make temporarily more accessible either ideal or ought self-guides (i.e., ideal or ought goals) was itself new to us. In the beginning, our research was based solely on using the Selves Questionnaire measure of self-discrepancies, which ideo-graphically measured individuals' stable self-discrepancies (e.g., Higgins et al., 1985). It took a while for us to realize that we could take advantage of what we knew about accessibility and priming (see Chapter 4, this volume) to activate different kinds of self-discrepancies, thereby having more experimental control over our tests of self-discrepancy theory by activating momentarily either actual-ideal or actual-ought discrepancies. This experimental method discovery turned out to be a critical turning point for the birth of regulatory focus theory, but we did not know that at the time.

At the time, we still thought of self-discrepancy theory in strictly personality terms. We were using priming to make one kind of chronic self-discrepancy more accessible than another; that is, make *it* the more currently active self-discrepancy. We knew from research that was going on during the same period that temporary accessibility from situational priming could trump, at least for a while, chronic accessibility from established individual differences (Bargh et al., 1988). But self-discrepancy theory was still about chronic individual differences in stored self-discrepancies. For individuals with both ideal and ought discrepancies, as in the Higgins et al. (1986) study, we were simply using the priming to make one or the other discrepancy more active at the moment.

So here I was in 1990 realizing that I could not conduct studies at Columbia which required measuring participants' self-discrepancies beforehand. To me that meant I could no longer do research on self-discrepancy theory. What was I to do? As I mentioned earlier, one thought was to return to priming and accessibility research full-time. But I did not want to abandon self-discrepancy theory so early in its development. Perhaps I was thinking about priming and accessibility while wanting to continue working on ideals and oughts that was critical. I don't know. What I do know is that I suddenly realized that there was a solution to my problem. What had fascinated me about accessibility for a long time was the fact that accessibility was a state, and individuals did not know the source of their accessibility state – it could be from chronic accessibility, priming, or both (see Chapter 4, this volume). I had thought about accessibility being a *common language* for variability across persons (chronic accessibility) and variability across situations (priming), which provided a different perspective on the classic "person-situation" debate (Higgins, 1990). I had also thought that standards provided another common language for variability across persons (personal standards) and variability across situations (contextual standards)

(Higgins, 1990). But I had *not* fully appreciated what the notion of a common language implied nor had I thought about how accessibility and standards could be combined.

I now realized that the distinction between ideal and ought self-regulation was being unnecessarily restricted by self-discrepancy theory. Rather than this distinction being about chronic discrepancies (or congruencies) between the actual self and ideal or ought self-guides, it was more generally about *two different systems of self-regulation*. At any moment, people could be in a state of regulating in relation to hopes or wishes (ideals) or they could be in a state of regulating in relation to duties or responsibilities (oughts). And, importantly, this could be true regardless of whether they did or did not possess chronic ideal or ought discrepancies.

From this broader, two-distinct-systems perspective, the emotional and motivational implications of individuals being in a *state* of ideal self-regulation versus a *state* of ought self-regulation could be studied *without any need to measure individuals' self-discrepancies*. What mattered now was the distinction between two systems of self-regulation, and one or the other *system* could be activated through priming or other experimental manipulations. It was *not* necessary to measure self-discrepancies in order to study the implications of ideal versus ought self-regulation. My problem was solved!

An issue did remain, however. Specifically, would I still be testing self-discrepancy theory – that is, *the theory itself* – if I was no longer measuring self-discrepancies? For many years I have taught a course on theory construction in psychology (see Higgins, 2004). I was well aware in 1990 that it was important to distinguish between an *extension* of a theory versus a *new* theory. Failure to do so can be very unfair to both the “old” theory and the “new” theory. The question, then, was whether there were two theories – the old self-discrepancy theory and the new “X” theory – or simply earlier and later versions of self-discrepancy theory. Why treat what might be just an extension or

elaboration of the extant self-discrepancy theory as a new theory? I have noted elsewhere that this is a gray area (Higgins, 2004), and it comes down to whether one believes that the “new” theory, in fact, really adds something that is fundamentally new.

In brief, I did believe that something fundamentally new had been added by the “new,” and as-yet-unnamed, theory. The “new” theory – what became called *regulatory focus theory* – was concerned with distinct self-regulatory *states* that varied across *both* persons *and* situations. Unlike self-discrepancy theory, it was *not* a personality theory. Whereas ideal and ought self-guides varied chronically across persons in self-discrepancy theory, ideal and ought self-regulation in regulatory focus theory varied across situations as well as persons. And this difference really mattered. It inspired research that would not have been generated by self-discrepancy theory, such as framing effects on problem solving and decision making (see Higgins, 1998).

I believed then, as I believe now, that what distinguishes a theory are the discoveries that it generates (Higgins, 2004). And it was clear even then that this new theory would generate studies that could make discoveries that self-discrepancy theory would not. What made the most sense, and was most fair, was to let self-discrepancy theory continue to develop in its same basic form – after all it was still just a child – while treating its offspring as something deserving its own separate path of development. And self-discrepancy theory did continue to develop (see Moretti and Higgins, 1999a, 199b), such as considering the overlap between own standpoint and significant other standpoint on someone’s goals in order to distinguish among independent self-regulation (just own standpoint), identified self-regulation (shared reality from own and significant other overlap), and introjected self-regulation (just significant other “felt presence of other”). Moreover, when regulatory focus theory began to mature, it was used to enrich self-discrepancy theory and broaden its implications, as evident in

Tim Strauman's construction of *self-system therapy* (Strauman et al., 2006). Like all good children, then, regulatory focus theory had a positive impact on its parent. Parent and child have each benefited enormously from one another – a nice family story. It is time, then, to describe what this new regulatory focus theory was all about and discuss the new research directions that it inspired.

PROMOTION AND PREVENTION SYSTEMS OF SELF-REGULATION

From the ancient Greeks, through seventeenth- and eighteenth-century British philosophers, to twentieth-century psychologists (see Kahneman et al., 1999), the hedonic principle that people are motivated to approach pleasure and avoid pain has dominated our understanding of people's motivation. It is the basic motivational assumption of theories across all areas of psychology, including theories of emotion in psychobiology (e.g., Gray, 1982), conditioning in animal learning (e.g., Mowrer, 1960; Thorndike, 1935), decision-making in cognitive psychology (e.g., Edwards, 1955; Kahneman and Tversky, 1979), consistency in social psychology (e.g., Festinger, 1957; Heider, 1958), and achievement motivation in personality (e.g., Atkinson, 1964). Even when Freud, in *Beyond the Pleasure Principle* (Freud, 1950/1920), assigned a motivational role to the ego's reality principle, he made it clear that the reality principle "at bottom also seeks pleasure although a delayed and diminished pleasure" (Freud, 1952/1920: 365). Perhaps the clearest statement on the importance of hedonic experiences to motivation was given by Jeremy Bentham (1781/1988: 1): "Nature has placed mankind under the governance of two sovereign masters, *pain* and *pleasure*. It is for them alone to point out what we ought to do, as well as to determine what we shall do. On the one hand the standard of right and wrong, on the other the chain of causes and effects, are fastened to their throne."

Within this historical context, the contribution of self-discrepancy theory and, especially, regulatory focus theory has been to emphasize the significance of the different motivational systems that underlie pleasure and pain. It is not enough to know that people approach pleasure and avoid pain. It is critical to know *how* they do so. The starting assumption of regulatory focus theory is that the *different ways* that people approach pleasure and avoid pain can be *more* significant for understanding motivation and emotion than the hedonic principle per se.

To illustrate, for Gray (1982) and Mowrer (1960), as well as Carver and Scheier (1981, 1990 a,b), the important motivational distinction was between the approach system (BAS) and the avoidance or inhibition system (BIS). Gray (1982) and Mowrer (1960) explicitly included both approaching reward (the presence of a positive outcome) and approaching safety (the absence of a negative outcome) as *equivalent* cases of approaching a desired end-state. In contrast, self-discrepancy theory (Higgins, 1987) and regulatory focus theory (Higgins, 1996) explicitly distinguished between the promotion-ideal system versus the prevention-ought system as *two different ways of approaching desired end-states*. Within approach, and within avoidance, there were *distinct systems of self-regulation*, and, according to self-discrepancy theory and regulatory focus theory, it was the difference between the nature of these systems that mattered emotionally and motivationally rather than the hedonic principle per se.

THE SIGNIFICANCE OF REGULATORY FOCUS THEORY FOR EMOTIONAL EXPERIENCES

As I discussed earlier, self-discrepancy theory distinguished between different types of pleasures and different types of pains as a function of which self-guide was involved in an actual self-congruency or discrepancy,

cheerful-related and dejected-related feelings for ideal congruencies and discrepancies, and quiescent-related and agitated-related feelings for ought congruencies and discrepancies. Regulatory focus theory moved beyond this personality-based distinction by considering momentary successes and failures for people who happened to be in either a promotion focus state or a prevention focus state.

According to regulatory focus theory, the distinct emotional experiences from having a promotion focus versus a prevention focus are *not* restricted to personality. Anyone at a particular moment can be pursuing a goal with a promotion focus or a prevention focus. If a person in a promotion state succeeds in their goal pursuit, they will experience cheerful-related feelings (e.g., feel happy), and if they fail they will experience dejected-related feelings (e.g., feel sad). If a person in a prevention state succeeds, they will experience quiescent-related feelings (e.g., feel calm), and if they fail they will experience agitated-related feelings (e.g., feel tense) (see Idson et al., 2000).

The emotional significance of the two distinct self-regulatory systems was no longer restricted to actual-self relations to ideal and ought self-guides. It extended to any case of success or failure when self-regulating with a promotion focus or a prevention focus. One intriguing study which illustrates this included participants who did or did not believe that their father would strongly hope they would do well on the current task (promotion father) and other participants who did or did not believe that their father would view it as their obligation to do well on the task (prevention father). For “promotion father”-primed participants, success or failure on the current task produced feelings along the cheerful-dejected dimension, whereas for “prevention father”-primed participants success or failure produced feelings along the quiescent-agitated dimension (see Shah, 2003).

Although priming or activating a personal ideal or a personal ought is one way to induce

a promotion or prevention state, it is not the only way. Other momentary situations can induce promotion or prevention states. In an early study by Roney et al. (1995), for example, success and failure feedback was framed in a promotion or prevention manner. In the promotion-framing condition, success and failure feedback was “you got that one” and “you didn’t get that one,” respectively. In the prevention-framing condition, success and failure feedback was “you didn’t miss that one” and “you missed that one,” respectively. By the end of the task, everyone had failed. As predicted by regulatory focus theory, agitation-related emotions increased from pretest to post-test more for prevention-framed participants than for promotion-framed participants, whereas the opposite was true for dejection-related emotions.

Even when individuals’ promotion focus and prevention focus are chronic predispositions, it does not require that there be chronic self-discrepancies or self-congruencies in order for distinct emotions to be produced. Distinct emotions will be produced by success or failure on a current task. Our discovery of this fact involves another story with a message about theory construction and testing. It is remarkable, to me at least, how our knowledge in one scientific area is not readily applied to theory construction and theory testing in another scientific area. To illustrate, James Shah and I wanted to find a way to measure the chronic strength of individuals’ promotion focus and prevention focus independent of whether they had self-discrepancies or self-congruencies. We felt that this was important because it could predict how they would feel from success or failure in a momentary situation independent of their past history of success and failure. We believed, for example, that whether individuals had actual-ideal discrepancies or not, if they pursued a goal with a promotion focus they would feel sad after failure and happy after success; and whether they had actual-ought discrepancies or not, if they pursued a goal with a prevention focus they

would feel nervous after failure and relaxed after success.

For many months we tried various ways to measure the chronic strength of individuals' promotion focus and prevention strength that would predict the likelihood that they would pursue a goal in a promotion state or a prevention state. Time after time we developed what we thought was a very clever measure, but time after time the measure failed to work consistently. Then in 1993 I attended, as usual, the Social Psychology Winter Conference in Utah where I heard John Bassili describing his recent work on attitude strength, work that was inspired by the earlier work of Russ Fazio (see Bassili, 1996; Fazio, 1986). This work took advantage of the notion of chronic accessibility.

By the time of that conference, I had worked on chronic accessibility for many years (e.g., Higgins et al., 1982; see Chapter 4, this volume). But somehow it had never occurred to me to take advantage of what I knew about chronic accessibility to devise a measure of promotion and prevention strength. But now, after returning from the conference, James Shah and I did exactly that. In brief, we used individuals' response latencies when describing their personal ideals and oughts as a measure of the chronic accessibility of their ideals and oughts. Our assumption about ideals and oughts, like that of Bassili and Fazio about attitudes, was that chronic accessibility related positively to self-regulatory strength. Specifically, the more that a person's ideals had higher chronic accessibility than their oughts, the more that person's promotion system predominated; and the more that a person's oughts had higher chronic accessibility than their ideals, the more that person's prevention system predominated.

This implicit measure of promotion and prevention strength worked out very well. It turned out that promotion and prevention strength were independent of the extent to which individuals had ideal or ought discrepancies. Moreover, promotion and prevention strength moderated self-discrepancy effects,

such that an actual-ideal discrepancy predicted suffering from dejected-related emotions more when promotion strength was also high, and an actual-ought discrepancy predicted suffering from agitated-related emotions more when prevention strength was also high (Higgins et al., 1997). This was a very nice parent and child collaboration that benefited both family members. Equally nice was now we had what we always wanted – an implicit measure of the likelihood that a person would pursue a goal with either a promotion focus (predominant promotion strength) or a prevention focus (predominant prevention strength), independent of the extent to which that person had actual-ideal or actual-ought discrepancies.

Individuals with predominant promotion strength – as measured by greater chronic accessibility of ideals than oughts – should emotionally experience success and failure along the cheerful-dejected dimension, and should emotionally appraise objects in the world along the cheerful-dejected dimension. In contrast, individuals with predominant prevention focus – as measured by greater chronic accessibility of oughts than ideals – should emotionally experience success and failure along the quiescent-agitated dimension, and should emotionally appraise objects in the world along the quiescent-agitated dimension. All of these predictions were supported (see Idson et al., 2000; Shah and Higgins, 2001).

THE SIGNIFICANCE OF REGULATORY FOCUS THEORY FOR STRATEGIC GOAL PURSUIT

The bread and butter of self-discrepancy theory was to distinguish among different kinds of emotional experiences – different kinds of pleasure and different kinds of pain – and predict when they would occur. But even here, as I described above, regulatory focus theory had something new and important to add by predicting emotional

effects of success and failure on a current task that varied as a function of situationally induced promotion or prevention states, and by predicting emotional effects of chronic promotion or prevention strength independent of chronic self-discrepancies. The major difference between self-discrepancy theory and regulatory focus theory, however, was regulatory focus theory's concern with understanding how the two distinct systems of self-regulation worked; how exactly the promotion and prevention systems *worked differently* to approach pleasure and avoid pain. Addressing this *how* issue would really move us beyond the hedonic principle and beyond self-discrepancy theory. This issue became the major concern of our lab in the early 1990s.

At the time that we began working on regulatory focus theory, Carver and Scheier had also developed a self-regulatory model that was concerned with motivation and emotion (Carver and Scheier, 1990a, 1990b). In their model there were also two self-regulatory systems – an *approach* system of reducing discrepancies to desired states as reference points, and an *avoidance* system of amplifying discrepancies from undesired states as reference points. As I mentioned earlier, this general distinction between an approach system and an avoidance system was a classic and important way of thinking about self-regulatory systems, and it was elaborated and developed by Carver and Scheier in significant and innovative directions (see Carver and Scheier, Chapter 24, this handbook). What we needed to do conceptually and empirically was to clarify and demonstrate how the distinct systems identified by regulatory focus theory were *not* the same as the distinction between an approach system versus an avoidance system as described in Carver and Scheier's model or in other approach-avoidance models, such as Mowrer's (1960) model, Gray's (1982) model, Atkinson's (1964) model, Lopes' (1987) model, and so on.

Carver and Scheier's distinction between an approach system and an avoidance system involves a distinction between two different

reference points – a desired end-state as reference point versus an undesired end-state as reference point. I have called this a *regulatory reference* distinction (Higgins, 1997). (The distinction of Atkinson and Lopes between hope and fear is a *regulatory anticipation* distinction – see Higgins [1997].) In contrast, the regulatory focus theory distinction between a promotion focus system and a prevention focus system involves a distinction *within* a desired end-state as reference point and *within* an undesired end-state as reference point. Within a desired end-state as reference point, for example, individuals can have a promotion focus on ideals and accomplishments as the desired end-state or they can have a prevention focus on oughts and safety as the desired end-state. Regulatory focus and regulatory reference are orthogonal distinctions.

Continuing with a desired end-state as the reference point, how does self-regulation with a promotion focus differ from self-regulation with a prevention focus? Early on we decided that the difference between promotion versus prevention concerns would translate into a difference in *strategic* preferences for how to pursue goals. With a promotion focus, individuals would prefer the strategy of approaching self-states that were matches to a desired end-state. With a prevention focus, individuals would prefer the strategy of avoiding self-states that were mismatches to a desired end-state. This was a distinction at the *strategic* level instead of Carver and Scheier's distinction at the level of approach versus avoidance systems. Within Carver and Scheier's *approach* system, this was a distinction between a promotion strategy of approaching matches versus a prevention strategy of avoiding mismatches. Within Carver and Scheier's *avoidance* system, regulatory focus theory distinguished between a promotion strategy of approaching mismatches to an undesired end-state versus a prevention system of avoiding matches to an undesired end-state.

To test this new strategic distinction, an early study had participants read about

many different episodes in the life of an individual that occurred over several days (Higgins et al., 1994). In each of the episodes, the target was trying either to experience a desired end-state or not experience an undesired end-state. The episodes were of the following kinds:

- 1 *Approaching* matches to a *desired end-state*: "Because I wanted to be at school for the beginning of my 8:30 psychology class which is usually excellent, I woke up early this morning."
- 2 *Avoiding* mismatches to a *desired end-state*: "I wanted to take a class in photography at the community center, so I didn't register for a class in Spanish that was scheduled at the same time."
- 3 *Approaching* mismatches to an *undesired end-state*: "I dislike eating in crowded places, so at noon I picked up a sandwich from a local deli and ate outside."
- 4 *Avoiding* matches to an *undesired end-state*: "I didn't want to feel tired during my very long morning of classes, so I skipped the most strenuous part of my morning workout."

Using an "unrelated studies paradigm," the participants first described either their personal ideals or their personal oughts to experimentally induce either a promotion focus or a prevention focus. Then they read the story and, afterward, tried to remember it. Across the desired and undesired end-states as reference points, the promotion-focused participants better remembered the episodes that involved strategic approach than the episodes that involved strategic avoidance, whereas the opposite was true for the prevention-focused participants.

Another study by Higgins et al. (1994) had participants make their own strategic choices regarding the desired end-state of friendship. The first phase of the study identified different friendship tactics. There were three tactics for the strategy of *approaching* matches: (a) "Be generous and willing to give of yourself"; (b) "Be supportive to your friends. Be emotionally supportive"; and

(c) "Be loving and attentive." There were also three tactics for the strategy of *avoiding* mismatches: (a) "Stay in touch. Don't lose contact with friends"; (b) "Try to make time for your friends and not neglect them"; and (c) "Keep the secrets friends have told you and don't gossip about friends." In a later phase of the study, new participants varying in chronic strength of regulatory focus were given all six tactics and were asked the *same* general question about friendship: "When you think about strategies for *friendship*, which THREE of the following strategies would you choose?" Across all participants, strategic approach tactics were chosen more than strategic avoidance tactics. But the study also found that predominant promotion participants chose more strategic approach tactics than predominant prevention participants, and predominant prevention participants chose more strategic avoidance tactics than predominant promotion participants.

The results of these two initial studies conducted in the early 1990s provided the first support for the unique prediction of regulatory focus theory that the promotion and prevention systems differed in the *strategic* manner of their goal pursuit. The distinction between the promotion and prevention systems was *not* a distinction between approach versus avoidance systems like Carver and Scheier's model and other earlier models. *Within* the approach system, and *within* the avoidance system, promotion and prevention differed in *how* goals were pursued. This strategic difference between the two regulatory focus systems was a critical distinguishing feature. It would have implications for self-regulation that were not fully foreseen at the time that these studies were done – implications that later led to the birth of regulatory fit theory (Higgins, 2000). I believe that the implications were not foreseen in part because the *terminology* of regulatory focus theory at that point in its development was more of a hindrance than a help. It is time to turn to that part of the story of regulatory focus theory.

HOW “PROMOTION” AND “PREVENTION” BECAME THE LABELS FOR THE TWO SYSTEMS

I have already talked about how theory development is a family affair in the sense of one theory giving birth to another, as self-discrepancy theory gave birth to regulatory focus theory. There are other ways in which theory development is a family affair (see Higgins, 2006a). For example, friends and colleagues provide feedback and suggestions at key stages in the development of a theory that have different kinds of positive effects in its development. An illustration of this concerns the labeling of “regulatory focus theory” and its “promotion” and “prevention” systems. These were *not* the original labels.

In the summer of 1993, I attended a conference at Ringberg Castle that was organized by Peter Gollwitzer and John Bargh – a conference with the now classic titles, “For Whom the Ring Bergs” and “Four Days at Ringberg: Four Days at Ringberg” (see Gollwitzer and Bargh, 1996). At that conference, I discussed my new theory, called “regulatory outcome focus” that concerned two distinct self-regulatory systems, called “positive outcome focus” and “negative outcome focus” (Higgins, 1996; see also Roney et al., 1995). A couple of years later (1995), I attended another conference in Italy that was organized by Arie Kruglanski and Augusto Palmonari. I presented a similar talk to the one that I had presented at the Ringberg Conference. Sitting in the audience was Marilynn Brewer.

Marilynn came up to me after my talk and asked me why I was using the labels “positive outcome focus” and “negative outcome focus” in my new theory. As I remember our conversation, she said something like the following:

Don’t those labels contradict the very point you are trying to make? Aren’t you arguing that both of your systems have pain and pleasure, *both* have positive and negative outcomes? Isn’t your point *not* to confuse *your* two distinct systems with

approaching positive outcomes versus avoiding negative outcomes? But your labels, “positive outcome focus” and “negative outcome focus,” make it sound like the distinction *is* about positive versus negative outcomes.

It was hard to argue her point. Actually, a good life lesson with Marilynn is to listen carefully to what she has to say because you will learn something useful. When I arrived home I was convinced that I had to find new labels. But which labels? I knew that, developmentally, having a “positive outcome focus” was about “bolstering” and *nurturing advancements*. It was about fulfilling wishes and aspirations. For desired end-states, it was about strategically *approaching matches*. It did not take too long – with my trusty tool, *Roget’s Thesaurus* to help me – to find a label that worked: *promotion*. That definitely sounded right; it had all the right connotations.

Now I needed a new label for “negative outcome focus.” I realized at this point that to follow Marilynn’s wise counsel I needed a label that not only captured the psychology of a “negative outcome focus” but also sounded positive. I needed two labels that were both positive and, at the system level, both involved approaching a desired end-state. It would be nice if the labels also implied a difference at the strategic level, but it was critical that both have positive valence at the system level.

With 20/20 hindsight, the answer may seem obvious to you now, but it was *not* obvious to me then. This was in part because, thanks to the self-discrepancy theory parentage, our research emphasis regarding “negative outcome focus” was on *ought* self-guides (i.e., duties and obligations). When I began searching for a new label, it was *ought* self-regulation that I had in mind – not safety and security. But, eventually, I came across the concept of *prevention*. Not only did “prevention” capture the association with *security* concerns and strategically *avoiding mismatches* (for desired end-states), but it was also a three-syllable word with “p” as the initial consonant – just like “promotion”!

It was perfection (another three-syllable word with “p” as the initial consonant).

The theory was now about the distinction between self-regulation with a promotion focus and self-regulation with a prevention focus. To keep it simple, the theory should simply be called “regulatory focus theory.” To this day, I should note, papers appear that refer to the theory as “self-regulatory focus theory” – perhaps a more accurate phrase but too many words. When possible, try to find labels for your concepts, like “ideal” and “ought” or “promotion” and “prevention,” that are simple and everyday terms. (Donald Campbell’s [1958] label, “entitativity,” famously broke this rule but prospered anyway.)

I have always believed that words matter, and that exploring the different meanings of a word, its different denotations and connotations, is a great tool for discovering the psychological underpinnings of the concept to which the word refers. Indeed, for this reason, there was a kind of positive externality, an unexpected benefit, of changing the labels in the theory from “positive outcome focus” and “negative outcome focus” to “promotion focus” and “prevention focus.” For the reasons that Marilyn Brewer suggested, the original labels were more confusing than helpful because they differed from each other only by the words “positive” and “negative.” But this valence distinction is precisely *not* what the theory is about. In contrast, the difference between “promotion” and “prevention” *is* precisely what the theory is about.

The more I thought about the words “promotion” and “prevention” – once again with the help of my trusty *Roget’s Thesaurus* plus the *Webster* and *Oxford* English dictionaries – the more the differences between them became apparent. Perhaps most important initially, it occurred to me that the way to fulfill a promotion focus was by being *eager* and enthusiastic, whereas the way to fulfill a prevention focus was by being *vigilant* and careful. Up to this point I had emphasized the strategic difference between approaching matches to desired

ideals and avoiding mismatches to desired oughts. This strategic distinction, I believe, is still accurate, but it is less generative than thinking of the difference between being *eager* in the service of promotion and being *vigilant* in the service of prevention. Indeed, this new way of thinking ultimately gave birth to another new child – *regulatory fit theory* (Higgins, 2000).

REGULATORY FOCUS THEORY: THE PARENT OF REGULATORY FIT THEORY

I suppose that it would have been possible to think about approaching matches to a desired end-state as being a strategy that sustained (versus disrupted) a positive outcome focus and that avoiding mismatches to a desired end-state as being a strategy that sustained a negative outcome focus. But surely it is easier to think of an *eager* strategy as sustaining promotion and a *vigilant* strategy as sustaining prevention. And in the late 1990s, after the label changes, we began to do studies where we varied which strategies people used during goal pursuit (e.g., Förster et al., 1998). In one study by Shah et al. (1998), for example, participants solved green-colored anagrams in order to gain points versus red-colored anagrams in order not to lose points. We found that promotion-focused participants performed better on the (*eager*) green anagrams than the (*vigilant*) red anagrams, but the opposite was true for prevention-focused participants. Another study by Shah et al. (1998) found that predominant promotion participants performed better on an anagram task when the suggested strategy was to try and find 90 percent or more of the words (*eager*) versus when the suggested strategy was to try not to miss more than 10 percent of the words (*vigilant*).

Over time we began to think of these different strategies as being *eager* versus *vigilant* strategies, and we began to label them as such. Importantly, we thought of

strategic eagerness as an approach strategy and strategic vigilance as an avoidance strategy. Because this strategic difference was a kind of *approach versus avoidance* distinction, it became all the more important to distinguish regulatory focus theory from control theories like Carver and Scheier's. Whereas Carver and Scheier's approach-avoidance distinction was at the system level (i.e., approaching desired end-states versus avoiding undesired end-states), our approach-avoidance distinction was at the strategic level. The different conceptions of avoidance at the system level versus the strategic level provided a critical test between these theories.

At Carver and Scheier's system level of avoidance, moving away from an undesired end-state should *decrease* avoidance intensity over time (see Miller, 1959). At regulatory focus theory's strategic level of avoidance, moving toward a desired end-state with prevention-focused vigilant avoidance should *increase* motivational intensity over time according to a "goal looms larger effect" (Lewin, 1935; Miller, 1959). These different predictions were tested in several studies using alternative measures of change in motivational intensity over time (i.e., arm pressure; persistence) and the results supported the prediction of regulatory focus theory (see Förster et al. 1998, 2001).

During the same general period we were conducting our "goal looms larger effect" studies, we were conducting other studies which suggested that the combination of promotion plus eager and prevention plus vigilance involved some sort of compatibility that had its own motivational significance. Earlier, Lorraine Idson, Nira Liberman, and I (Idson et al., 2000) had conducted studies in which participants imagined buying a book and choosing between paying with cash or paying with a credit card, with the book's price being lower if you paid with cash (see Thaler, 1980). The participants reported either how good they would feel if they paid with cash (the positive "success" outcome) or how bad they would feel if they paid with

a credit card (the negative "failure" outcome). These earlier studies found that the pleasant feelings from the positive "success" outcome were more intense for promotion than prevention, that is, feeling cheerful versus feeling quiescent, whereas the painful feelings from the negative "failure" outcome were more intense for prevention than promotion, that is, feeling agitated versus feeling dejected.

In addition to measuring how good or bad participants felt about the outcome, our new studies included separate measures of pleasure/pain intensity and strength of motivational force (see Idson et al., 2004). For example, in one study that induced either a promotion focus or a prevention focus by priming either ideals or oughts, *pleasure-pain intensity* was measured by asking the participants how pleasant the positive outcome would be or how painful the negative outcome would be; and *strength of motivational force* was measured by asking the participants how motivated they would be to make the positive outcome happen (in the positive outcome condition) or how motivated they would be to make the negative outcome not happen (in the negative outcome condition).

We found that pleasure/pain intensity and strength of motivational force *each* made significant *independent* contributions to the perceived value of the imagined outcome (i.e., its goodness/badness). We also found that for the positive success outcome, strength of motivational force was higher in promotion than prevention; but for the negative failure outcome, strength of motivational force was higher in prevention than promotion. What these studies discovered was that there is an *asymmetry* between promotion and prevention with respect to whether success or failure yields a stronger motivational force. Other studies conducted around the same time also found this asymmetry. For promotion, there was better performance under conditions of success than failure. For prevention, there was better performance under conditions of failure than success

(Förster et al., 2001; Idson and Higgins, 2000). This asymmetry between promotion and prevention regarding the effects of success versus failure was another major feature distinguishing the two systems. It was not a feature that had been identified in models distinguishing approach versus avoidance systems. It was something new about the difference between the promotion and prevention systems, and I knew it was important.

The earlier performance studies had found that goal pursuit in a promotion focus yielded better performance when an eager than a vigilant strategy was used, and the opposite was true for goal pursuit in a prevention focus. Now there was this new asymmetry for the effects on motivational strength of success versus failure. What was going on? The key to the solution was to recognize that the effects for intensity of emotions that we had found earlier (Idson et al., 2000) reflected differences in motivational strength. That is, when promotion-focused individuals feel cheerful after success their motivation is high, but when they feel dejected after failure their motivation is low. In contrast, when prevention-focused individuals feel quiescent after success their motivation is low, but when they feel agitated after failure their motivation is high. And these differences in motivational strength are related to the same eagerness and vigilance that were involved in the earlier performance studies. That is, when promotion-focused individuals feel cheerful after success they are eager (strong motivation), but when they feel dejected after failure they are *not* eager (weak motivation). In contrast, when prevention-focused individuals feel agitated after failure they *are* vigilant (strong motivation), but when they feel quiescent after success they are *not* vigilant (weak motivation).

What united all the findings of these studies in the late 1990s was that being eager sustains motivational strength for individuals in a promotion focus, whereas being vigilant sustains motivational strength for individuals in a prevention focus. This solution to the mystery of what was going on in our recent

regulatory focus studies gave birth to regulatory fit theory. What became clear was that there was another self-regulatory principle which contributed to the effects we were finding – the principle of *regulatory fit*.

Within this same period, a serendipitous event occurred. In 1999 I learned that I would receive the APA's Award for Distinguished Scientific Contributions, which meant that I would be giving a talk in 2000 that would then appear as an article in the *American Psychologist*. I wasn't sure whether to give a talk about past research testing regulatory focus theory or to give a talk on something new. I preferred the latter. But what would a new talk be about? I was quite excited about the findings from our recent studies and the solution to the mystery of what might underlie them. So I decided that my speech and paper for the APA award would present the new principle of regulatory fit. But this meant that I had to develop the new theory in short order.

The major proposal of regulatory fit theory was that people experience regulatory fit when the *manner* of their engagement in an activity *sustains* (versus disrupts) their current regulatory orientation (Higgins, 2000). An eager manner sustains promotion and a vigilant manner sustains prevention. But regulatory fit was *not* the same as regulatory focus because it concerned the relation between *any* goal pursuit orientation and the strategic manner in which the goal is pursued. Indeed, when I reconsidered other research we had done during the same period, for example, research on how "fun" versus "important" task instructions impact performance (Bianco et al., 2003), I began to realize that regulatory fit was a very general principle that applied to other orientations and strategies. Nonetheless, it was the specific work testing regulatory focus theory that led to the discovery of regulatory fit. Regulatory focus theory was the parent of regulatory fit theory, and as this new child developed it began to generate its own separate studies and findings (see Higgins, 2008a, 2009). At the end of this chapter, I will

discuss how parent and child have benefited from one another.

APPLICABILITY TO SOCIAL ISSUES

As I noted earlier, it was self-discrepancy theory that gave birth to regulatory focus theory. And my inspiration for self-discrepancy theory was wanting to understand the distinct psychological underpinnings of depression and anxiety disorders. My first collaborator on self-discrepancy theory, Tim Strauman, had the same inspiration as me. Indeed, while a graduate student at NYU, he earned two PhDs – one in social psychology and one in clinical psychology. After leaving NYU, he continued to wear both hats, including being a director of clinical training and helping clinically depressed and anxious clients as a therapist. He began to develop a new form of clinical psychotherapy based on self-discrepancy theory. After the birth of regulatory focus theory, he expanded and modified this therapy to take advantage of the new insights provided by regulatory focus theory.

Conceptual advances and empirical discoveries from developing regulatory focus theory since 1990 have increased psychologists' understanding of the differences between depression-related promotion failure and anxiety disorder-related prevention failure. This has led to Strauman and his collaborators developing and testing, in clinical trials, a new-generation psychotherapy called "self-system therapy" (Vieth et al., 2003). Interventions specifically designed to reduce the actual-ideal discrepancies of depressed patients have proven to be effective. Indeed, for a theory-specified subset of depressed patients, it has been shown to be even more effective than cognitive therapy (Strauman et al., 2006).

The emotional and motivational significance of the difference between a failure in the promotion system and a failure in the prevention system sheds new light regarding

other clinical phenomenon. There is evidence, for example, that among women who become a mother for the first time, having an actual-ideal discrepancy prior to the birth of their child predicts increased vulnerability to post-partum depression, whereas having an actual-ought discrepancy predicts *decreased* vulnerability to post-partum anxiety (Alexander and Higgins, 1993). There is also evidence that possessing an actual-ideal discrepancy is a vulnerability factor for *bulimic* eating disorders whereas possessing an actual-ought discrepancy is a vulnerability factor for *anorexic* eating disorders (Higgins et al., 1992; Strauman et al., 1991).

In addition to its implications for clinical phenomenon, regulatory focus theory has implications for interpersonal relations and intergroup relations as well. There is evidence that people are more willing to forgive another person who apologizes for hurting them and are more empathic concerning another person's suffering when the promotion or prevention nature of the forgiveness message or the other person's suffering fits the promotion or prevention focus of the recipient or perceiver (Houston, 1990; Santelli et al., 2009). This illustrates an interpersonal benefit from *similarity* in regulatory focus. There is evidence as well that *complementarity* in regulatory focus can also have interpersonal benefits. Recent research has found that long-term married partners with complementary regulatory focus orientations have higher relationship wellbeing (Bohns et al., 2009). What appears to be critical for such complementarity effects is for each partner to be able to assume a separate role on shared tasks (i.e., division of labor) so that each can use the goal strategy that fits their regulatory focus orientation, such as the promotion partner taking on the eager parts of the task and the prevention partner taking on the vigilant parts of the task. There is also evidence that responses to social exclusion vary in a manner that relates to regulatory focus, with individuals responding in a prevention manner when they are rejected but responding in a promotion

manner when they are ignored (Molden et al., 2009).

Intergroup relations are also influenced by regulatory focus. Specifically, there is evidence that the classic phenomenon of *ingroup favoritism* (see Levine and Moreland, 1998) varies by regulatory focus. Two separate research programs have shown that favoritism that rewards and embraces ingroup members is driven largely by promotion concerns, whereas favoritism that punishes and rejects outgroup members is driven largely by prevention concerns (Sassenberg et al., 2003; Shah et al., 2004) – *promoting us* versus *protecting them*. And this effect is evident even for subtle measures of motivation for intergroup contact. For example, in a study by Shah et al. (2004), participants chose where to sit in a waiting room that had a backpack on a chair which supposedly was owned either by their future partner in an upcoming task or their future opponent. Participants with a stronger promotion focus chose to sit *closer to their teammate*, whereas a stronger prevention focus had no relation to sitting closer to one's teammate. In contrast, participants with a stronger prevention focus chose to sit *further away from their opponent*, whereas a stronger promotion focus had no relation to sitting further away from one's opponent.

Regulatory focus has other implications for intergroup relations as well. Being discriminated against is painful. But the nature of this pain and reactions to it can depend on perceivers' regulatory focus. When discrimination is perceived as blocking opportunities for advancement, the pain would involve dejection which has low motivational intensity. In contrast, when discrimination is perceived as a threat to one's security, the pain would involve agitation which has high motivational intensity. There is evidence, for example, that a prevention focus leads to more anger and agitation after social discrimination than a promotion focus, and especially when social discrimination is based on losses rather than on nongains (Sassenberg and Hansen, 2007).

These differences in emotions and motivation could translate into people responding differently when they are discriminated against. Consistent with this, Quinn and Olson (2004) demonstrated that, compared to promotion-focused women, prevention-focused women report stronger intentions to engage in future behaviors that are aimed at reducing discrimination toward women, such as participating in protests on women discrimination issues, as well as reporting that they have performed such actions more frequently in the past. Interestingly, when behaviors that protest discrimination are explicitly framed in terms of removing obstacles to advancement, that is, removing a barrier to accomplishing progress, then promotion-focused women report stronger intentions to engage in such behaviors than do prevention-focused women.

There is also evidence that regulatory focus is relevant for reducing the negative impact of stereotype threat on performance (Steele et al., 2002). Keller (2007) has shown that if a promotion focus, rather than a prevention focus, can be induced under stereotype threat conditions, the negative impact of stereotype threat can be reduced. Keller (2007) argues that when individuals are in a promotion focus, stereotype threat is more likely to be experienced as a challenge rather than as a threat, which in turn creates greater eagerness and engagement in maximal goals that enhance performance. Research by Seibt and Förster (2004) suggests that negative stereotypes induce a prevention focus that, in turn, motivates people to use vigilant strategies on a task (cf. Förster et al., 2004). If the task is one in which a vigilant strategy is useful, such as an analytical task, then this will not be a problem. But if the task requires the use of an eager strategy, or a mix of vigilant and eager strategies, then the prevention focus induced by negative stereotypes will hurt performance. On such tasks, Keller's (2007) intervention of inducing a promotion focus under stereotype threat conditions could be especially important.

One final way in which regulatory focus theory has been applied to social issues should be mentioned. Over the last several years, regulatory focus theory – often combined with the principle of regulatory fit – has been used to increase the effectiveness of persuasive messages (for reviews, see Cesario et al., 2008; Lee and Higgins, 2009). This application can be used, and *has* been used, to enhance the effectiveness of health messages. For instance, several studies have demonstrated that when recipients who are either promotion-focused or prevention-focused are given messages that are framed, respectively, in promotion-eager terms or in prevention-vigilant terms, the recipients are more persuaded to increase their consumption of fruits and vegetables (Cesario et al., 2004; Latimer et al., 2007; Spiegel et al., 2004), to use sunscreen (Keller, 2006; Lee and Aaker, 2004), increase physical activity (Latimer et al., 2008), and reduce intentions to smoke (Kim, 2006; Zhao and Pechmann, 2007).

In an early demonstration of this persuasion technique, Spiegel et al. (2004) gave participants health messages that advocated pursuit of the same desired end-state – eating more fruits and vegetables. The key manipulations took place as part of the messages that participants received. Although all participants received the same message advocacy (“eat more fruits and vegetables”), a promotion versus prevention focus was manipulated through the concerns that were highlighted within the messages – accomplishments for promotion and safety for prevention. Within each regulatory focus condition, participants were asked either to imagine the benefits they would get if they complied with the health message (eager strategy) or the costs they would incur if they didn’t comply with the health message (vigilant strategy). The participants in the fit conditions (promotion recipients/eager message; prevention recipients/vigilant message) ate more fruits and vegetables in the week following the first session than participants in the nonfit conditions

(prevention recipients/eager message; promotion recipients/vigilant message). Latimer et al. (2008) extended these findings to show that a single message framed to fit individuals’ chronic regulatory focus led to greater fruit and vegetable consumption even four months after message delivery.

CONCLUSION

From self-discrepancy theory to regulatory focus theory to regulatory fit theory, this *family of theories* has been developed and applied for over 20 years now. I should note that when a parent has a child, the parent does not stop developing. The child develops, but so too does the parent. Regulatory focus theory, for example, continued to develop on its own after giving birth to regulatory fit theory, as illustrated by the recent regulatory focus theory distinction between eager and vigilant *strategies* versus risky and conservative *tactics* (e.g., Scholer et al., 2010; see Scholer and Higgins, in press).

Importantly, it is not only the parent that affects the development of the child. The child affects the development of the parent as well. For example, regulatory fit theory taught regulatory focus theory that a particular strategy can have a consistent and stable association with a specific regulatory concern, such as eager with promotion and vigilant with prevention, because the strategy is *in the service of* sustaining that regulatory concern; a perspective that has proven useful when reconsidering the relation between culture and personality (Higgins, 2008b; Higgins, et al., in press). And in addition to a child affecting a parent, a grandchild can affect a grandparent. Regulatory fit theory, for example, provided new insights for self-discrepancy theory’s understanding of anhedonia, that is, the inability to gain pleasure from normally pleasurable activities, which is a central symptom of the depression that is associated with severe actual-ideal discrepancies. Because an actual-ideal discrepancy

is a promotion failure that reduces eagerness, and low eagerness is a nonfit for promotion, engagement in positive activities is weakened when people have severe actual-ideal discrepancies, which in turn deintensifies their attractiveness (see Higgins, 2006b). In this way, positive activities in general lose their attractiveness.

In my family of theories, grandparent, parent, child, and grandchild have all enriched one another. It has been an exciting journey, and I look forward to observing and participating in further developments. I am certain there will be new discoveries and new surprises. Let me conclude as I have in the past:

Children teach parents to appreciate life in new ways. Children help parents to discover new things about the world. Theories can too. Always remember to love your theory, enjoy your theory, and help it develop. It is what makes life as a scientist a joy

(Higgins, 2006a).

REFERENCES

Alexander, M.J. and Higgins, E.T. (1993) Emotional trade-offs of becoming a parent: How social roles influence self-discrepancy effects. *Journal of Personality and Social Psychology*, 65, 1259–1269.

Atkinson, J.W. (1964) *An Introduction to Motivation*. Princeton: D. Van Nostrand.

Bargh, J.A., Lombardi, W.J. and Higgins, E.T. (1988) Automaticity of chronically accessible constructs in person \times situation effects on person perception: It's just a matter of time. *Journal of Personality and Social Psychology*, 55, 599–605.

Bassili, J.N. (1996) Meta-judgmental versus operative indexes of psychological attributes: The case of measures of attitude strength. *Journal of Personality and Social Psychology*, 71, 637–653.

Bentham, J. (1888) *The Principles of Morals and Legislation*. Amherst, NY: Prometheus Books. (Original work published 1781).

Bianco, A.T., Higgins, E.T. and Klem, A. (2003) How 'fun/importance' fit impacts performance: Relating implicit theories to instructions. *Personality and Social Psychology Bulletin*, 29, 1091–1103.

Bohns, V.K., Lucas, G., Molden, D.C., Finkel, E.J., Coolsen, M.K., Kumashiro, M., Rusbult, C.E. and Higgins, E.T. (2009) When opposites fit: Increased relationship strength from partner complementarity in regulatory focus. Unpublished manuscript, Columbia University.

Carver, C.S. and Scheier, M.F. (1981) *Attention and Self-regulation: A Control-theory Approach to Human Behavior*. New York: Springer-Verlag.

Carver, C.S. and Scheier, M.F. (1990a) Principles of self-regulation: Action and emotion. In E.T. Higgins and R.M. Sorrentino (eds), *Handbook of Motivation and Cognition: Foundations of Social Behavior*, 2, 3–52. New York: Guilford Press.

Carver, C.S. and Scheier, M.F. (1990b) Origins and functions of positive and negative affect: A control-process view. *Psychological Review*, 97, 19–35.

Cesario, J., Grant, H. and Higgins, E.T. (2004) Regulatory fit and persuasion: Transfer from 'feeling right'. *Journal of Personality and Social Psychology*, 86, 388–404.

Cesario, J., Higgins, E.T. and Scholer, A.A. (2008) Regulatory fit and persuasion: Basic principles and remaining questions. *Social and Personality Psychology Compass*, 2, 444–463.

Edwards, W. (1955) The prediction of decisions among bets. *Journal of Experimental Psychology*, 51, 201–214.

Fazio, R.H. (1986) How do attitudes guide behavior? In R.M. Sorrentino and E.T. Higgins (eds), *Handbook of Motivation and Cognition: Foundations of Social Behavior*, pp. 204–243. New York: Guilford Press.

Festinger, L. (1957) *A Theory of Cognitive Dissonance*. Evanston: Row, Peterson.

Förster, J., Higgins, E.T. and Idson, C.L. (1998) Approach and avoidance strength as a function of regulatory focus: Revisiting the 'goal looms larger' effect. *Journal of Personality and Social Psychology*, 75, 1115–1131.

Förster, J., Higgins, E.T. and Werth, L. (2004) How threat from stereotype disconfirmation triggers self-defense. *Social Cognition*, 22, 54–74.

Förster, J., Grant, H., Idson, L.C. and Higgins, E.T. (2001) Success/failure feedback, expectancies, and approach/avoidance motivation: How regulatory focus moderates classic relations. *Journal of Experimental Social Psychology*, 37, 253–260.

Freud, S. (1950) *Beyond the Pleasure Principle*. New York: Liveright. (Original work published 1920).

Freud, S. (1952) *A General Introduction to Psychoanalysis*. New York: Washington Square Press. (Original work published 1920).

Gollwitzer, P.M. and Bargh, J.A. (eds) (1996) *The Psychology of Action: Linking Cognition and Motivation to Behavior*. New York: Guilford Press.

Gray, J.A. (1982) *The Neuropsychology of Anxiety: An Enquiry into the Functions of the Septo-hippocampal System*. New York: Oxford University Press.

Heider, F. (1958) *The Psychology of Interpersonal Relations*. New York: Wiley.

Higgins, E.T. (1987) Self-discrepancy: A theory relating self and affect. *Psychological Review*, 94, 319–340.

Higgins, E.T. (1990) Personality, social psychology, and person-situation relations: Standards and knowledge activation as a common language. In L.A. Pervin (ed.), *Handbook of Personality*, pp. 301–338. New York: Guilford Press.

Higgins, E.T. (1991) Development of self-regulatory and self-evaluative processes: Costs, benefits, and tradeoffs. In M.R. Gunnar and L.A. Sroufe (eds), *Self Processes and Development: The Minnesota Symposia on Child Psychology*, 23, 125–165. Hillsdale, NJ: Erlbaum.

Higgins, E.T. (1996) Ideals, oughts, and regulatory focus: Affect and motivation from distinct pains and pleasures. In P.M. Gollwitzer and J.A. Bargh (eds), *The Psychology of Action: Linking Cognition and Motivation to Behavior*, pp. 91–114. New York: Guilford Press.

Higgins, E.T. (2000) Making a good decision: Value from fit. *American Psychologist*, 55, 1217–1230.

Higgins, E.T. (2001) Promotion and prevention experiences: Relating emotions to nonemotional motivational states. In J.P. Forgas (ed.), *Handbook of Affect and Social Cognition*, pp. 186–211. Mahwah, NJ: Lawrence Erlbaum Associates.

Higgins, E.T. (2004) Making a theory useful: Lessons handed down. *Personality and Social Psychology Review*, 8, 138–145.

Higgins, E.T. (2006a) Theory development as a family affair. *Journal of Experimental Social Psychology*, 42, 549–552.

Higgins, E.T. (2006b) Value from hedonic experience and engagement. *Psychological Review*, 113, 439–460.

Higgins, E.T. (2008a) Regulatory fit. In J.Y. Shah and W.L. Gardner (eds), *Handbook of Motivation Science*, pp. 356–372. New York: Guilford Press.

Higgins, E.T. (2008b) Culture and personality: Variability across universal motives as the missing link. *Social and Personality Psychology Compass*, 2.

Higgins, E.T. (2009) Regulatory fit in the goal-pursuit process. In G.B. Moskowitz and H. Grant (eds), *The Psychology of Goals*. New York: Guilford Press.

Higgins, E.T., Bond, R.N., Klein, R. and Strauman, T. (1986) Self-discrepancies and emotional vulnerability: How magnitude, accessibility, and type of discrepancy influence affect. *Journal of Personality and Social Psychology*, 51, 5–15.

Higgins, E.T., King, G.A. and Marvin, G.H. (1982) Individual construct accessibility and subjective impressions and recall. *Journal of Personality and Social Psychology*, 43, 35–47.

Higgins, E.T., Klein, R. and Strauman, T. (1985) Self-concept discrepancy theory: A psychological model for distinguishing among different aspects of depression and anxiety. *Social Cognition*, 3, 51–76.

Higgins, E.T., Pierro, A. and Kruglanski, A.W. (in press) Re-thinking culture and personality: How self-regulatory universals create cross-cultural differences. In R.M. Sorrentino and S. Yamaguchi (eds), *Handbook of Motivation and Cognition Across Cultures*. New York: Academic Press.

Higgins, E.T., Rholes, W.S. and Jones, C.R. (1977) Category accessibility and impression formation. *Journal of Experimental Social Psychology*, 13, 141–154.

Higgins, E.T., Roney, C., Crowe, E. and Hymes, C. (1994) Ideal versus ought predilections for approach and avoidance: Distinct self-regulatory systems. *Journal of Personality and Social Psychology*, 66, 276–286.

Higgins, E.T., Shah, J. and Friedman, R. (1997) Emotional responses to goal attainment: Strength of regulatory focus as moderator. *Journal of Personality and Social Psychology*, 72, 515–525.

Higgins, E.T. and Silberman, I. (1998) Development of regulatory focus: Promotion and prevention as ways of living. In J. Heckhausen and C.S. Dweck (eds), *Motivation and Self-regulation Across the Life Span*, pp. 78–113. New York: Cambridge University Press.

Higgins, E.T. and Tykocinski, O. (1992) Self-discrepancies and biographical memory: Personality and cognition at the level of psychological situation. *Personality and Social Psychology Bulletin*, 18, 527–535.

Higgins, E.T., Vookles, J. and Tykocinski, O. (1992) Self and health: How 'patterns' of self-beliefs predict types of emotional and physical problems. *Social Cognition*, 10, 125–150.

Houston, D.A. (1990) Empathy and the self: Cognitive and emotional influences on the evaluation of negative affect in others. *Journal of Personality and Social Psychology*, 59, 859–868.

Idson, L.C. and Higgins, E.T. (2000) How current feedback and chronic effectiveness influence motivation: Everything to gain versus everything to lose. *European Journal of Social Psychology*, 30, 583–592.

Idson, L.C., Liberman, N. and Higgins, E.T. (2000) Distinguishing gains from nonlosses and losses from nongains: A regulatory focus perspective on

hedonic intensity. *Journal of Experimental Social Psychology*, 36, 252–274.

Idson, L.C., Liberman, N. and Higgins, E.T. (2004) Imagining how you'd feel: The role of motivational experiences from regulatory fit. *Personality and Social Psychology Bulletin*, 30, 926–937.

Kahneman, D., Diener, E. and Schwarz, N. (eds) (1999) *Well-being: The Foundations of Hedonic Psychology*. New York: Russell Sage Foundation.

Kahneman, D. and Tversky, A. (1979) Prospect theory: An analysis of decision under risk. *Econometrica*, 47, 263–291.

Keller, J. (2007) When negative stereotypic expectancies turn into challenge or threat: The moderating role of regulatory focus. *Swiss Journal of Psychology*, 66, 163–168.

Keller, J. (2008) On the development of regulatory focus: The role of parenting styles. *European Journal of Social Psychology*, 38, 354–364.

Keller, P.A. (2006) Regulatory focus and efficacy of health messages. *Journal of Consumer Research*, 33, 109–114.

Kim, Y. (2006) The role of regulatory focus in message framing for antismoking advertisements for adolescents. *Journal of Advertising*, 35, 143–151.

Latimer, A.E., Rivers, S.E., Rench, T.A., Katulak, N.A., Hicks, A., Hodorowski, J.K., Higgins, E.T. and Salovey, P. (2008) A field experiment testing the utility of regulatory fit messages for promoting physical activity. *Journal of Experimental Social Psychology*, 44, 826–832.

Latimer, A.E., Williams-Piehota, P., Katulak, N.A., Cox, A., Mowad, L.Z., Higgins, E.T. and Salovey, P. (in press) Promoting fruit and vegetable intake through messages tailored to individual differences in regulatory focus. *Annals of Behavioral Medicine*.

Lee, A.Y. and Aaker, J.L. (2004) Bringing the frame into focus: The influence of regulatory fit on processing fluency and persuasion. *Journal of Personality and Social Psychology*, 86, 205–218.

Lee, A.Y. and Higgins, E.T. (2009) The persuasive power of regulatory fit. In M. Wänke (ed.), *Social Psychology of Consumer Behavior*, pp. 319–333. New York: Psychology Press.

Levine, J.M. and Moreland, R.L. (1998) Small groups. In D.T. Gilbert, S.T. Fiske, and G. Lindzey (eds), *The Handbook of Social Psychology*, 4th Edition, pp. 415–469. Boston: McGraw-Hill.

Lewin, K. (1935) *A Dynamic Theory of Personality*. New York: McGraw-Hill.

Lopes, L.L. (1987) Between hope and fear: The psychology of risk. In L. Berkowitz (ed.), *Advances in Experimental Social Psychology*, 20, 255–295. New York: Academic Press.

Manian, N., Papadakis, A.A., Strauman, T.J. and Essex, M.J. (2006) The development of children's ideal and ought Self-Guides: Parenting, temperament, and individual differences in guide strength. *Journal of Personality*, 74, 1619–1645.

Miller, N.E. (1959) Liberalization of basic S-R concepts: Extensions to conflict behavior, motivation, and social learning. In S. Koch (ed.), *Psychology: A Study of a Science*, Vol. 2: *General Systematic Formulations, Learning, and Special Processes*, pp. 196–292. New York: McGraw-Hill.

Molden, D.C., Lucas, G.M., Gardner, W.L., Dean, K. and Knowles, M.L. (2009) Motivations for prevention or promotion following social exclusion: Being rejected versus being ignored. *Journal of Personality and Social Psychology*, 96, 415–431.

Moretti, M.M. and Higgins, E.T. (1999a) Internal representations of others in self-regulation: A new look at a classic issue. *Social Cognition*, 17, 186–208.

Moretti, M.M. and Higgins, E.T. (1999b) Own versus other standpoints in self-regulation: Developmental antecedents and functional consequences. *Review of General Psychology*, 3, 188–223.

Mowrer, O.H. (1960) *Learning Theory and Behavior*. New York: John Wiley.

Quinn, K.A. and Olson, J.M. (2004) Preventing prejudice or promoting progress? Regulatory framing and collective action. Unpublished manuscript.

Roney, C.J.R., Higgins, E.T. and Shah, J. (1995) Goals and framing: How outcome focus influences motivation and emotion. *Personality and Social Psychology Bulletin*, 21, 1151–1160.

Santelli, A.G., Struthers, C.W. and Eaton, J. (2009) Fit to forgive: Exploring the interaction between regulatory focus, repentance, and forgiveness. *Journal of Personality and Social Psychology*, 96, 381–394.

Sassenberg, K. and Hansen, N. (2007) The impact of regulatory focus on affective responses to social discrimination. *European Journal of Social Psychology*, 37, 421–444.

Sassenberg, K., Kessler, T. and Mummendy, A. (2003) Less negative = more positive? Social discrimination as avoidance or approach. *Journal of Experimental Social Psychology*, 39, 48–58.

Scholer, A.A. and Higgins, E.T. (in press) Distinguishing levels of approach and avoidance: An illustration using regulatory focus theory. In A.J. Elliot (ed.), *Handbook of Approach and Avoidance Motivation*.

Scholer, A.A., Zou, X., Fujita, K., Stroessner, S.J. and Higgins, E.T. (2010) When risk-seeking becomes a

motivational necessity. *Journal of Personality and Social Psychology*, 99, 215–231.

Seibt, B. and Förster, J. (2004) Stereotype threat and performance: How self-stereotypes influence processing by inducing regulatory foci. *Journal of Personality and Social Psychology*, 87, 38–56.

Shah, J. (2003) The motivational looking glass: How significant others implicitly affect goal appraisals. *Journal of Personality and Social Psychology*, 85, 424–439.

Shah, J. and Higgins, E.T. (2001) Regulatory concerns and appraisal efficiency: The general impact of promotion and prevention. *Journal of Personality and Social Psychology*, 80, 693–705.

Shah, J., Higgins, E.T. and Friedman, R. (1998) Performance incentives and means: How regulatory focus influences goal attainment. *Journal of Personality and Social Psychology*, 74, 285–293.

Shah, J.Y., Brazy, P.C. and Higgins, E.T. (2004) Promoting us or preventing them: Regulatory focus and manifestations of intergroup bias. *Personality and Social Psychology Bulletin*, 30, 433–446.

Spiegel, S., Grant-Pillow, H. and Higgins, E.T. (2004) How regulatory fit enhances motivational strength during goal pursuit. *European Journal of Social Psychology*, 34, 39–54.

Steele, C.M., Spencer, S. and Aronson, J. (2002) Contending with group image: The psychology of stereotype and social identity threat. In M.P. Zanna (ed.), *Advances in Experimental Social Psychology*, 34, 379–440. New York: Academic Press.

Strauman, T.J. (1989) Self-discrepancies in clinical depression and social phobia: Cognitive structures that underlie emotional disorders? *Journal of Abnormal Psychology*, 98, 14–22.

Strauman, T.J. (1992) Self-guides, autobiographical memory, and anxiety and dysphoria: Toward a cognitive model of vulnerability to emotional distress. *Journal of Abnormal Psychology*, 101, 87–95.

Strauman, T.J. and Higgins, E.T. (1987) Automatic activation of self-discrepancies and emotional syndromes: When cognitive structures influence affect. *Journal of Personality and Social Psychology*, 53, 1004–1014.

Strauman, T.J., Vieth, A.Z., Merrill, K.A., Kolden, G.G., Woods, T.E., Klein, M.H., Papadakis, A.A., Schneider, K.L. and Kwapil, L. (2006) Self-system therapy as an intervention for self-regulatory dysfunction in depression: A randomized comparison with cognitive therapy. *Journal of Consulting and Clinical Psychology*, 74, 367–376.

Strauman, T.J., Vookles, J., Berenstein, V., Chaiken, S. and Higgins, E.T. (1991) Self-discrepancies and vulnerability to body dissatisfaction and disordered eating. *Journal of Personality and Social Psychology*, 61, 946–956.

Thaler, R.H. (1980) Toward a positive theory of consumer choice. *Journal of Economic Behavior and Organization*, 1, 39–60.

Thorndike, E.L. (1935) *The Psychology of Wants, Interests, and Attitudes*. New York: Appleton-Century-Crofts.

Zhao, G. and Pechmann, C. (2007) The impact of regulatory focus on adolescents' response to antismoking advertising campaigns. *Journal of Marketing Research*, 44, 671–687.

A Model of Behavioral Self-regulation

Charles S. Carver and Michael F. Scheier

ABSTRACT

This chapter describes the evolution (or perhaps accretion) of a set of ideas bearing on the self-regulation of action and emotion. The ideas were drawn from many sources, eventually yielding a viewpoint in which goal-directed action is seen as reflecting a hierarchical set of feedback control processes, and the creation and reduction of affect are seen as reflecting another set of feedback processes. Also embedded in the model is the idea that confidence and doubt influence whether the person continues to struggle against adversity or gives up the goal that the adversity is threatening. The portion of the model devoted to affect is of particular interest in that it generates two positions that differ substantially from those deriving from other theories. The first is that both approach and avoidance give rise to both positive and negative feelings; the second is that positive affect leads to coasting, reduction in effort regarding the goal under pursuit. The recent interest in dual-process models, which distinguish between top-down goal pursuit and reflexive responses to cues of the moment, has caused us to re-examine some of our previous assumptions, considering the possibility that behavior is triggered in two distinct ways.

INTRODUCTION

This chapter outlines several aspects of a perspective we have adopted over an extended

period of time concerning the self-regulation of behavior and emotion. This perspective is more about the structure of behavior than it is about the content of behavior. It represents a viewpoint on the metaphorical bones and tendons that underlie very diverse sorts of action. We believe it is a viewpoint that is generally compatible with many other theories that are described in this book, standing beside them rather than in place of them.

There are two respects in which the ideas described here differ from those described elsewhere in this book. First, these ideas may be less a “theory” than a “meta-theory,” a very general way of conceptualizing interwoven functions. It is a declaration of belief about certain aspects of how complex systems are organized. Second, we actually developed on our own very little of the viewpoint we are about to describe. With a few exceptions, most of what we have done is to bring together ideas that had been developed by other people for their own reasons, and applied them to phenomena that are of interest to personality and social psychologists.

The viewpoint outlined here has long been identified with the term *self-regulation* (Carver and Scheier, 1981), a term that means different things to different people. We use it to imply purpose, with self-corrective

adjustments taking place as needed, which originate from within the person. These elements converge in the view that human behavior is a continuing process of moving toward (and sometimes away from) goal values. Some people ascribe to *self-regulation* the additional quality of restraining or overriding impulses (e.g., Baumeister and Vohs, 2004). We generally do not. We address issues of overriding impulses later in the chapter. When we do, we use the more restrictive term *self-control*.

INTELLECTUAL HISTORY

The intellectual history of this viewpoint on behavior is vast. It extends through the development of ideas about mechanical governors and computing machines (e.g., Ashby, 1940; Rosenblueth et al., 1943; Wiener, 1948), and homeostatic mechanisms within the body (Cannon, 1932). Its roots include the literature of expectancy-value models of motivation (e.g., Bandura, 1986; Feather, 1982; Rotter, 1954), and general systems theory (Ford, 1987; von Bertalanffy, 1968) – the idea that mechanisms with similar structural and functional properties operate at many levels of abstraction. A full depiction of this history would overwhelm the page limits of this chapter and is well beyond our reach.

What we can do, however, and what we have done in the pages that follow, is to situate these ideas in the historical context in which we came to them and began to work with them. It is nearly impossible to do this without writing about our own professional histories, because pursuit of this viewpoint has occupied much of our professional attention over a period of over 30 years. That, then, is the form taken by this chapter.

SELF-AWARENESS AND CYBERNETICS

We were graduate students together at the University of Texas in the early 1970s.

Although we were in personality psychology, we (along with several others in our program) were nearly as interested in social psychology. At about that time, Robert Wicklund, a young faculty member in the social psychology program, and Shelley Duval, a graduate student there, developed a theory that caught our attention. It was called objective self-awareness theory (Duval and Wicklund, 1972).

Self-awareness and conformity to standards

This theory had several facets, but what was most interesting to us at the time was its analysis of some of the circumstances leading a person's behavior to conform to situationally salient standards. The theory held that when a person's attention is directed inward to the self, it gravitates to self-aspects that can be evaluated by comparison with some standard of correctness. The theory's authors held that such comparison would generally reveal a discrepancy between the actual self-state and the salient standard. This would yield negative self-evaluation and negative affect. This, in turn, was presumed to prompt the person to seek ways to move out of the state of being either aware of the discrepancy or aware of the self.

There are two potential ways to escape this negative state. One is to avoid stimuli that induce self-awareness. No awareness of the discrepancy, no aversiveness. The other is to change the present state so that it conforms to the standard. Without a discrepancy, an awareness of the self is not aversive. It was the second of these responses – adjustment of behavior to match the standard – that stuck in our minds, and which led to some of our first explorations in the lab.

There we found that increased self-awareness could cause research participants to become both less punitive and more punitive to others, depending on circumstances. In one study, it led young men to behave more chivalrously toward young women,

punishing them less intensely for errors in a task (Scheier et al., 1974). In another case, it led participants to conform more closely to the “hint” dropped by the experimenter that stronger punishment would produce faster learning, the stated goal of the task (Carver, 1974).

These were by no means the first studies of how behavior shifts when attention is self-directed (see Duval and Wicklund, 1972), but they were formative for us. They helped lead us into the laboratory in a more serious way, and they helped solidify our impression that self-awareness could be a powerful and systematic influence on behavior. Of importance for the future direction of our thinking, we had also found that self-awareness does not just dampen behavior. We had found that self-awareness could lead in either of opposing directions: to more aggression in some circumstances, to less aggression in other circumstances. It induced shifting of behavior toward whatever standard was salient.

Duval and Wicklund’s (1972) theory lit our first real path into human research. It highlighted several phenomena to which we returned repeatedly. The one just described is that self-focus can cause behavior to conform more closely to salient standards. Another (also described above, in less detail) is the idea that people would sometimes act to avoid or escape from self-focus when confronting a discrepancy. Noted only briefly by Duval and Wicklund was the idea that some standards are prescriptive and others are proscriptive, with the latter being less directive than the former, requiring “only that particular points on personally relevant dimensions be avoided” (1972: 14). We eventually looked at all these phenomena, but ultimately did so with a different conceptual framing than Duval and Wicklund had used.

Cybernetics and systems

Duval and Wicklund’s (1972) theory was developed toward the end of a period in which many social psychological theories

were heavily influenced by more general drive-based models of motivation. Examples are cognitive dissonance theory (Festinger, 1957) and reactance theory (Brehm, 1966). Such theories relied on aversive drives to serve as the motivational engine of behavior. In classic drive theory, a need state develops over time, leading to an internal tension state – drive – which is aversive. The person thus is motivated to do something to reduce that aversiveness. Without the aversiveness, there is no need for action.

In social psychological applications of this idea, specific drives were ascribed to specific kinds of internal mental states and drive-specific classes of responses were posited. In dissonance theory, the conflict between two incompatible cognitions was said to create an aversive drive that could be reduced by reconciling or obscuring the conflict. In reactance theory, experiencing a loss of freedom was said to create an aversive drive that could be reduced by reasserting the freedom. In the same way, self-awareness theory held that awareness of a discrepancy from a salient standard produced an aversive drive state that could be reduced either by acting to reduce the discrepancy or by avoiding the self-aware state.

Drive theories are not, however, the only ways of thinking about motivation or the forces that guide behavior. Not long after encountering self-awareness theory, one of us was exposed for the first time to the concepts of cybernetics, a viewpoint that is quite different from drive theory. Cybernetics had already had a (relatively brief) heyday in the 1950s and 1960s. Probably the best-known example of this viewpoint was an engaging book by Miller et al. (1960). This book brought into the psychological lexicon the acronym TOTE, standing for test-operate-test-exit. These functions depict the action of a feedback loop, though they do so in a way that emphasizes a sequence of discrete steps rather than simultaneous occurrence of all the processes.

The logic of the TOTE unit – test (compare present state with a standard), operate (make

a change in the present state), test (compare again to ensure the change had the effect intended), and exit (move to the next thing that needs doing) – has an inescapable resemblance to core processes in self-awareness theory: that when attention is directed to the self, there is a tendency to compare one's present state with a salient standard of appropriate behavior, and to shift the present state so that it more closely approximates the standard. There was a real difference in connotation, though. To Duval and Wicklund (1972), this set of functions concerns *self-evaluation*. To Miller et al. (1960), this set of functions describes *the structure of goal-directed behavior*. It describes the way all human goals are attained.

We were rather taken by the idea that a small set of functions described the structure of behavior. We also resonated to the idea that a great deal of human behavior could be planned and conducted without the need for aversive internal states. It felt intuitive to us that a great deal of behavior was done because it was fun, or because the actions led to interesting outcomes. The idea that self-awareness processes paralleled the structure of a feedback loop, a construct that had been useful in a variety of places other than personality and social psychology, was also exciting.

Hierarchical organization

The ideas put forward by Miller et al. (1960) opened up a new avenue for exploration in literatures that were largely unfamiliar to us. It turned out that a number of people had proposed ideas with a similar character, over a fairly extended period (e.g., MacKay, 1956, 1966; Powers, 1973; for further review see Miller et al., 1960). Of particular interest, and of particular impact on our thinking, was a provocative statement published as we were becoming interested in feedback concepts. It was a book by William Powers (1973), in which he forcefully put forward a model of how human behavior may reflect a

hierarchy of feedback processes in simultaneous operation.

His goal was very ambitious. He tried to account for how the nervous system created physical movements by which intentions and even abstract human values become expressed behaviorally. The construct that took center stage in this account was the feedback loop. Powers tried to map several layers of feedback processes to aspects of the nervous system. Knowledge of the nervous system has progressed substantially since 1973, of course, and parts of the picture that Powers created are doubtlessly contradicted by later evidence. However, viability of the core idea that feedback processes underlie organized action need not depend entirely on details.

In Powers (1973) we found several strong themes. First, perhaps even more than Miller et al. (1960), Powers made a compelling case for the idea that the feedback construct was fully up to the challenge of accounting for the complexity of behavior. Not as one loop, of course, but as an interwoven network of loops, dealing with regulation of diverse properties simultaneously.

Second, he argued more specifically that the feedback processes underlying behavior form a hierarchy of varying levels of abstraction, which could be characterized by their properties. He started from the lowest level – regulation of muscle-fiber tensions – and worked his way upward. Each level deals with discrepancies at its own level; the reference value for a given level is the output of the level immediately above it. When a person does a relatively abstract behavior (e.g., expressing kindness to an elderly neighbor by shoveling the snow off her walks), all levels below that level of abstraction are at work simultaneously. This was (and remains) a very interesting idea with a great many implications.

Whether this model accurately portrays the control of action or only provides a comfortable illusion, we found it useful in two ways. First, it provides the sense that it is plausible to posit a way for the kinds of intentions that social and personality

psychologists are interested in to find their way into action. Second, it suggests a reason to attend to the literature of motor control (e.g., Rosenbaum et al., 2001). Some may see that literature as relevant only to exercise science and industrial psychology. We disagree. We believe that this literature also has things to tell personality and social psychologists. Nature is a miser and a recycler. It is very likely that principles embodied in movement control have more than just a little in common with principles that are embodied in higher mental functions (Rosenbaum et al., 2001).

A third contribution made by Powers was his particular view of how higher levels of control might be construed. Although he devoted very little attention (15 pages) to the three highest levels of control he argued for – those most relevant to the subject matter of personality and social psychology – he chose labels and wrote descriptions of these levels and their relationships to one another that are intuitively on-target and evocative of important concepts in our field.

Programs are organizations of behavior with choice points. They are clearly sequential and orderly (though the order can be quite flexible). They seem to require attention. If they are sufficiently well learned to have acquired an automatic runoff character, they are not programs, but a lower level of control that Powers called *sequences*. Programs are the level of the Powers hierarchy that most closely resembles Miller et al.'s TOTE construct, because of the sequencing of steps and subroutines that programs imply.

Principles, the level above programs, are roughly akin to values (Schwartz and Bilsky, 1990; Schwartz and Rubel, 2005). They are a basis for making the choices that programs entail, and they suggest certain programs to enter and to avoid. What Powers called *system concepts* are the coalesced essences of entities that imply certain principles and not others. One might think of the overall sense of ideal self as one example, the sense of an ideal relationship as another.

We have used these upper levels of the Powers hierarchy for decades as a conceptual heuristic for thinking about the organization of behavior. We never found the time to study the ideas empirically. However, these ideas have a certain amount in common with those of Vallacher and Wegner's (1987) action identification theory. This theory posits that people can identify any action at varying levels of abstraction. In identifying the action at that level, presumably they are also regulating it at that level. Within this framework, two complementary tendencies play out over time and circumstances: As the person becomes more adept at a behavior, less notice must be taken of its lower-level elements, and the person drifts to a higher-level construal of it. If the person encounters difficulties with the action at the level at which it is presently identified, the person is pulled to a lower-level construal of the action, to sort out the bits at the lower level.

The notion of hierarchical organization has many implications. A full treatment of those implications is unfortunately beyond the scope of this chapter. However, interested readers may find broader discussions elsewhere (e.g., Carver and Scheier, 1998, 1999a, 1999b; Powers, 1973).

CONFIDENCE AND DOUBT, EFFORT AND DISENGAGEMENT

The preceding section described our shift toward control-theory principles as a meta-theoretical grounding for interpreting self-awareness effects, and some of the complexity that ultimately came along with that shift. In this section we return to self-awareness theory and to another consequence of self-focused attention. As noted earlier, self-awareness was held to cause the person either to change the present state to fit the standard or to avoid self-focus, generally by withdrawing from or avoiding the situation in which self-focus was being created. The feedback model addresses the reducing

of the discrepancy. What about the other effect?

Several studies had, in fact, already found that when discrepancies were created, people acted to avoid self-awareness manipulations or to leave the situation (for reviews, see Carver and Scheier, 1981; Duval and Wicklund, 1972). On the other hand, evidence also suggested that these effects require more than just a discrepancy. In one study, for example, a situation was set up in which a discrepancy was portrayed as being either flexible or inflexible (Steenbarger and Aderman, 1979). Only when the discrepancy was seen as inflexible was self-focus experienced as aversive and only in that case did subjects avoid the self-focusing situation.

Our view was that whether self-focus led to increased efforts to reduce discrepancies or to disengagement of effort and avoidance of self-focusing stimuli depended on whether the person was confident or doubtful about eventual success at reducing the discrepancy. Confidence should lead to renewed (or greater) effort; sufficient doubt should lead to giving up and withdrawal.

We suggested that there is a sort of psychological watershed on the confidence dimension, and that the character of subsequent behavior flows either to renewed effort or to disengagement. Subsequent studies (reviewed in Carver and Scheier, 1981; see also Carver, 2003b) supported this view. Self-awareness enhanced the efforts of persons who had favorable expectations of being able to attain goals and cut short the efforts of persons who had unfavorable expectations. The opposing effects of self-focus as a function of differences in confidence suggest that there is indeed a breakpoint on the confidence dimension, analogous to a watershed on a mountain ridge.

Expectancy theories in the zeitgeist

We were not, of course, the only ones to have observed the importance of expectancies.

Many theorists have emphasized that theme over an extended period (e.g., Atkinson, 1964; Bandura, 1986; Feather, 1982; Rotter, 1954; Tolman, 1938; Vroom, 1964). Adoption of the expectancy construct as a way of addressing discrepancy reduction versus withdrawal linked our view of self-awareness phenomena to this tradition of expectancy-value models of motivation. Expectancy models vary in other ways, but in at least one respect their core argument is similar: confidence of success keeps the person engaged in the effort to succeed, and greater efforts tend to foster success; expectation of failure leads to not trying, and not trying often leads to failure.

Two models proposed at about that time resonated particularly well with this thinking: those of Klinger (1975) and Wortman and Brehm (1975). Both held that two regions of the range of expectancies form a dichotomy, with the resultant behavior falling into two categories: effort versus disengagement or withdrawal. Klinger (1975) contrasted commitment with disengagement. Wortman and Brehm (1975) contrasted the reassertion of control tied to reactance with the giving-up response tied to helplessness. Brehm's thinking would later evolve into a view in which people exert as much effort as needed to successfully complete a task, up to the point where success no longer seems worth the effort or no longer seems possible, at which point effort stops (Brehm and Self, 1989). These statements all imply that giving up is not merely lower effort; it is a shift from one class of response to another.

Giving up and moving on

The fact that a tendency to disengage plays a role in human self-regulation is hard to deny. However, it is also hard to capture all the roles this idea plays in one or two statements. For example, is disengagement good or bad? On the one hand, disengagement is often a maladaptive or dysfunctional response. A person who gives up whenever encountering

difficulty will never accomplish anything. Without continued struggle, it is often impossible to overcome obstacles. Some goals in life should not be given up easily, even if the struggle is hard and painful.

On the other hand, disengagement (at some level, at least) is also a necessity, a natural and indispensable part of self-regulation. If people are ever to turn away from efforts at unattainable goals, if they are ever to back out of blind alleys, they must be able to disengage, to give up and start over somewhere else. The importance of disengagement is particularly obvious with regard to concrete, low-level goals: people must be able to remove themselves from literal blind alleys and wrong streets, give up plans that have become disrupted by unexpected events, and spend the night in the wrong city if they have missed the last plane home.

Disengagement is also important, however, with regard to higher-level goals (Wrosch et al., 2003a). It is important to disengage and move on with life after the loss of a close relationship (e.g., Cleiren, 1993; Weiss, 1988). People sometimes must give up goals that are deeply embedded in the self if those values create too much conflict and distress in their lives (Pyszczynski and Greenberg, 1992). There often are goals of childhood that must be given up as it becomes apparent that they will never be realized (Baltes et al., 1979; Heckhausen and Schulz, 1995). Giving up thus is a double-edged sword. One of the thorniest questions in life is how to decide when to hang on and when to let go (Pyszczynski and Greenberg, 1992).

Another issue here is that certain kinds of decisions have simultaneous overtones of both continued effort and disengagement. Consider the scaling back of aspirations. Sometimes progress toward a goal is going poorly, expectancies of success are dim, and you want to quit. Rather than quit altogether, though, you abandon the more demanding goal for a less demanding one (e.g., a struggling student stops thinking in terms of an A and starts thinking in terms of a C).

This is a kind of limited disengagement. The first goal is being given up at the same time as the lesser one is being adopted. Limited disengagement has an important positive consequence: it keeps you engaged in the general domain in which you'd wanted to quit. By scaling back the goal (giving up in a small way), you keep trying to move ahead (thus *not* giving up, in a larger way).

A potential problem with the limited-disengagement strategy stems from the fact that goals are often interrelated. It may be fine in principle to lower your grade aspiration from an A to a C. But if a high grade in this course is a prerequisite to another goal – say, admission to medical school – the limited disengagement works only temporarily. The same issue will likely recur later on, with respect to the broader goal to which this one leads. In some cases, this bind is not easily resolved. If medical school is your ultimate educational goal and your grades are bad, some rearrangement of the ultimate goal is going to be necessary.

In the broader scheme, giving up on unattainable goals has multiple positive consequences. It conserves energy rather than waste it in futile pursuit of the unattainable (Nesse, 2000). It also eventually readies the person to take up alternative goals (Klinger, 1975). Finally, emotional pain from lost goals seems to reflect the combination of remaining committed to them and yet being unable to move forward (Carver and Scheier, 1998; Pyszczynski and Greenberg, 1992; Wrosch et al., 2003b). Disengaging fully from them thus removes a source of negative feelings.

Expectancies and feedback loops

The incorporation of the expectancy construct made fairly good sense with respect to the analysis of self-awareness processes. It provided a reason why self-focus would lead in some cases to discrepancy reduction efforts and in other cases to behavioral withdrawal. It is a little harder, though, to fit the expectancy construct into the model of

feedback loops. In some ways, these assumptions about expectancies and effort versus disengagement seem very ad hoc.

If one steps back from the hierarchy as a whole, however, the idea becomes more plausible. Powers (1973) had argued that a higher-level control loop operates by resetting the reference values of loops at the next lower level. Some kinds of resetting can be thought of as adjustments in level of aspiration (which we characterized above as a limited disengagement). Other kinds of resetting may be more complex, involving changes in entire programs of action that are being considered for enactment. If this program of action is not creating desired results at the higher level, resetting of lower-level goal may require abandoning that strategy altogether and trying a different one (there often being many ways to skin a cat).

DISCREPANCY ENLARGEMENT

Thus far we have talked about only those feedback processes that reduce discrepancies. There also exist feedback processes that enlarge discrepancies. These feedback loops are unstable. Unless overridden, they enlarge discrepancies without end. Some people believe that this kind of feedback is always problematic and dysfunctional (Powers, 1973). Others believe that positive loops are an important part of complex systems (DeAngelis et al., 1986; Maruyama, 1963; McFarland, 1971), but that in living systems (and other cases in which positive feedback is adaptive), the effect of this loop is limited in some way or other. There may be a natural endpoint (e.g., sexual arousal increases to the point of orgasm, thus ending the increase), or the discrepancy enlarging function may be constrained by a discrepancy reducing function.

One might view some discrepancy enlarging loops as avoidance processes. Examples of potential reference values for discrepancy-enlarging loops in social-personality psychology would include feared

or disliked possible selves (Markus and Nurius, 1986; Ogilvie, 1987) and negative reference groups. These are values to be avoided. If a positive standard can be viewed as a goal (Miller et al., 1960), these standards might be thought of as anti-goals (Carver and Scheier, 1998). If comparison of the present state with this standard suggests that the discrepancy is small, an effort to enlarge the discrepancy may follow.

As Duval and Wicklund (1972) noted, such standards are generally not as directive as positive standards. If a prohibitive standard is at one end of a dimension of variability, however, direction is thereby provided. As an example, Carver and Humphries (1981) recruited Cuban American students, and ascribed opinions to a group that is a negative reference group to them: the Castro government of Cuba. When asked to report their own opinions on the same issues, they took pains to differ from the opinions attributed to Cuban officials. Further, the tendency to do so was increased by self-focus. Thus, self-focus can increase discrepancy enlargement as well as discrepancy reduction (for more examples see Carver, 2003b; Carver and Scheier, 1998).

We have also suggested that social and personality psychology has examples of discrepancy-enlarging loops being constrained by discrepancy-reducing loops. This pattern seems represented in Higgins's (1996) concept of the ought self (Carver et al., 1999) and in Ryan and Deci's (2000) concept of introjected values. In both constructs, the initial impetus to behavior is the desire to avoid social sanction; a good way to avoid social sanction is to locate a socially approved value that is different from (or opposite to) the disapproved value, and move toward it. By homing in on the positive value, one simultaneously escapes the feared or disliked value.

Approach and avoidance

As just suggested, the dual concepts of discrepancy reducing and discrepancy enlarging

loops map onto the general form of approach and avoidance processes. Incentives are approached by systems that close discrepancies between present conditions and the incentives. Threats are avoided by systems that enlarge discrepancies between present conditions and the threats. The logic of feedback processes thus provides a way to construe this dichotomy among motivations.

The idea that behavior reduces to approach and avoidance tendencies is not new (e.g., Miller, 1944; Miller and Dollard, 1941), but it has re-emerged in recent years (e.g., Davidson 1998; Elliot, 2008). The idea that two sorts of feedback functions map onto these classes of motivations has led us to be more attentive to differences between approach and avoidance behavior (e.g., Carver and White, 1994). That has particular relevance for the next topic.

AFFECT

As described earlier, the view we had adopted on self-awareness effects did not include any assumption about aversive drive states. Yet it was clear that people do sometimes experience negative affect when experiencing self-awareness. This was most likely when the discrepancy between state and standard was relatively fixed – when there was doubt about being able to move forward. Further thought about these issues helped lead us to an elaboration of the model with which we were working. It led us to hazard a guess about the source of affect.

Origins

What is affect? Affect is positive or negative feelings. In many ways affect is the heart of emotion, though the term emotion often incorporates connotations of physiological changes that frequently accompany hedonic experiences. Affect pertains to one's desires and whether they are being met (Clore, 1994;

Frijda, 1986; Ortony et al., 1988). But what is the internal mechanism by which feelings arise?

Many different kinds of answers to this question have been offered, ranging from neurobiological (e.g., Davidson 1992) to cognitive (Ortony et al., 1988). We proposed an answer that focused on what appear to be some of the functional properties of affect (Carver and Scheier, 1990, 1998, 1999a, 1999b). In suggesting this answer, we used feedback control once more as an organizing principle. Now, however, the control bears on a different quality.

We suggested that feelings arise as a consequence of a feedback loop that operates simultaneously with the behavior-guiding process and in parallel to it. We regard its operation as automatic. The easiest characterization of what this second process is doing is that it is checking on how well the first process (the behavior loop) is doing. The input for this second loop thus is the *rate of discrepancy reduction in the action system over time*. (We focus first on discrepancy-reducing loops, then consider enlarging loops.)

Consider a physical analogy. Action implies change between states. Difference between states is distance. The action loop thus controls the psychological analog of distance. If the affect loop assesses the action loop's progress, then the affect loop is dealing with the psychological analog of velocity, the first derivative of distance over time. To the degree that this analogy is meaningful, the input to the affect loop should be the first derivative over time of the input used by the action loop.

Input (how well you are doing) does not by itself create affect; a given rate of progress has different affective consequences in different contexts. We argued that this input is compared to a reference value (cf. Frijda, 1986, 1988), just as in other feedback loops. In this case, the value is an acceptable or expected rate of behavioral discrepancy reduction. As in other feedback loops, the comparison checks for deviation from

the standard. If there is a discrepancy, an error is sensed and the output function changes.

We think the error signal in this loop is manifest in experience as affect, a sense of positive or negative valence. A rate of progress below the criterion yields negative affect. A rate high enough to exceed the criterion yields positive affect. If the rate is not distinguishable from the criterion, there is no valence. In essence, the argument is that feelings with positive valence mean you are doing better at something than you need to, and feelings with negative valence mean you are doing worse than you need to (for detail, including supporting evidence, see Carver and Scheier, 1998, Chapters 8 and 9). The absence of affect means being neither ahead nor behind.

We are not arguing for a deliberative thinking through of whether rate conforms to the criterion rate. We assume that the testing is continuous and automatic. Nor are we arguing for a deliberative thinking about what the affective valence means. We assume that the meaning (i.e., being ahead versus behind) is intrinsic to the affect's valence, which itself arises automatically.

One implication of this line of thought is that affects that might potentially exist regarding any given action should fall on a bipolar dimension. That is, it should be the case that affect can be positive, neutral, or negative for any given goal-directed action, depending on how well or poorly the action seems to be attaining the goal.

Reference criterion

What determines the criterion? There doubtlessly are many influences. Further, the orientation that a person takes to an action can induce a different framing that may change the criterion (Brendl and Higgins, 1996). What is used as a criterion is probably quite flexible when the activity is unfamiliar. If the activity is very familiar, the criterion is likely to reflect the person's accumulated

experience, in the form of an expected rate (the more experience you have, the more you know what is reasonable to expect). Whether "desired" or "expected" or "needed" is most accurate as a depiction of the criterion rate may depend greatly on the context.

The criterion can also change, sometimes readily, sometimes less so. The less experience the person has in a domain, the easier it is to substitute one criterion for another. We believe, however, that change in rate criterion in a relatively familiar domain occurs relatively slowly. Continuing overshoots result automatically in an upward drift of the criterion, continuing undershoots result in a downward drift (see Carver and Scheier, 2000). Thus, the system recalibrates over repeated events. A (somewhat ironic) consequence of such recalibration would be to keep the balance of a person's affective experiences (positive to negative, across a span of time) relatively similar, even if the rate criterion changes considerably.

Two kinds of action loops, two dimensions of affect

So far we have addressed only approach loops. The view just outlined was that positive feeling exists when a behavioral system is making more than adequate progress *doing what it is organized to do*. The systems addressed so far are organized to reduce discrepancies. Yet there seems no obvious reason why the principle should not apply to systems that enlarge discrepancies. If such a system is making rapid enough progress attaining its ends, there should be positive affect. If it is doing poorly, there should be negative affect.

That affects of both valences are possible seems applicable to both approach and avoidance. That is, both approach and avoidance have the potential to induce positive feelings (by doing well), and both have the potential to induce negative feelings (by doing poorly). But doing well at *approaching an incentive* is not quite the same experience as doing well

at *moving away from a threat*. Thus there may be differences between the two positives, and between the two negatives.

Drawing on the work of Higgins (e.g., 1987, 1996), we argue for two bipolar dimensions of affect, one bearing on approach, the other on avoidance (Carver, 2001; Carver and Scheier, 1998). Approach-related affect includes such positive affects as elation, eagerness, and excitement, and also such negative affects as frustration, anger, and sadness (Carver, 2004; Carver and Harmon-Jones, 2009). Avoidance-related affect includes such positive affects as relief, serenity, and contentment (Carver, 2009) and such negative affects as fear, guilt, and anxiety.

Affect and action: two facets of one event in time

This two-layered viewpoint implies a natural connection between affect and action. That is, if the input function of the affect loop is a sensed rate of progress in action, the output function of the affect loop must be a change in the rate of progress in that action. Thus, the affect loop has a direct influence on what occurs in the action loop.

Some changes in rate output are straightforward. If you are lagging behind, you try harder. Some changes are less straightforward. The rates of many “behaviors” are defined not by pace of physical action but in terms of choices among potential actions, or entire programs of action. For example, increasing your rate of progress on a project at work may mean choosing to spend a weekend working rather than playing with family and friends. Increasing your rate of being kind means choosing to do an act that reflects kindness when an opportunity arises. Thus, change in rate must often be translated into other terms, such as concentration, or allocation of time and effort.

The idea of two feedback systems functioning jointly is something we stumbled into. As it happens, however, this idea is quite common in control engineering

(e.g., Clark, 1996). Engineers have long recognized that having two systems functioning together – one controlling position, one controlling velocity – permits the device they control to respond in a way that is both quick and stable, without overshoots and oscillations.

The combination of quickness and stability in responding is desirable in many of the devices engineers deal with. It is also desirable in people. A person with very reactive emotions is prone to overreact and oscillate behaviorally. A person who is emotionally unreactive is slow to respond even to urgent events. A person whose reactions are between those extremes responds quickly but without behavioral overreaction and oscillation.

For biological entities, being able to respond quickly yet accurately confers a clear adaptive advantage. We believe this combination of quick and stable responding is a consequence of having both behavior-managing and affect-managing control systems. Affect causes people’s responses to be quicker (because this control system is time sensitive); as long as the affective system is not over-responsive, the responses are also stable.

Our focus here is on how affects influence behavior, emphasizing the extent to which they are interwoven. However, note that the behavioral responses that are linked to the affects also lead to *reduction of the affects*. We thus would suggest that the affect system is, in a very basic sense, self-regulating (cf. Campos et al., 2004). It is undeniable that people also engage in voluntary efforts to regulate their emotions (e.g., Gross, 2007; Ochsner and Gross, 2008), but the affect system does a good deal of that self-regulation on its own.

AFFECT ISSUES

There are at least two important ways in which this view of affect differs from other theories bearing on emotion. One difference

concerns the dimensional structure of affect (Carver, 2001).

Bipolarity

In some theories (though not all) affects are seen as having underlying dimensionality (e.g., Watson et al., 1999). Our view holds that affect generated through approach has the potential to be either positive or negative and that affect generated through avoidance also has the potential to be either positive or negative. Most dimensional models, however, ascribe to an approach system affects with positive valence and ascribe to an avoidance system affects with negative valence (e.g., Cacioppo et al., 1999; Lang et al., 1990; Watson et al., 1999).

There is at least some support for our view. There is evidence, albeit limited, that positive feelings of calmness and relief (as situationally relevant) relate to avoidance motivation (Carver, 2009; Higgins et al., 1997). There is far more evidence linking sadness to approach failure (for reviews see Carver, 2004; Higgins, 1996). There is also a good deal of evidence linking the approach system to the negative affect of anger (Carver and Harmon-Jones, 2009). Although it is clear that diverse negative feeling qualities coalesce with one another in moods (Watson, 2009), the evidence does not make that case with regard to situation-specific affective responses.

This issue is important, because it has implications for any attempt to identify a conceptual mechanism underlying creation of affect. Theories positing two unipolar dimensions assume that greater activation of a system translates to more affect of that valence (or more potential for affect of that valence). If the approach system relates both to positive and to negative feelings, however, this direct transformation of system activation to affect is not tenable. A conceptual mechanism is needed that naturally addresses both valences within the approach function (and, separately, the avoidance function). The mechanism described here does so.

Counterintuitive effect of positive affect

A second issue also differentiates this model from other views (Carver, 2003a). Recall our argument that affect reflects the error signal from a comparison in a feedback loop. If this is so, affect is a signal to adjust rate of progress. This would be true whether the rate is above the mark or below it – that is, whether affect is positive or negative. For negative feelings, this is intuitive. The first response to negative feelings is usually to try harder. If the person tries harder – and if more effort (or better effort) increases progress – the negative affect diminishes or ceases.

For positive feelings, prediction is counterintuitive. In this model, positive feelings arise when things are going better than they need to. But the feelings still reflect a discrepancy (albeit a positive one), and the function of a negative feedback loop is to keep discrepancies small. Such a system is organized in such a way that it “wants” to see neither negative nor positive affect. Either quality (deviation from the standard in either direction) would represent an “error” and lead to a change in output that would eventually reduce it. This view argues that people who exceed the criterion rate of progress (and who thus have positive feelings) will automatically tend to reduce subsequent effort in this domain. They will “coast” a little – ease back.

Expendig greater effort to catch up when behind, and coasting when ahead, are both presumed to be specific to the goal domain to which the affect is attached, usually the goal from which the affect arises in the first place. We are not arguing that positive affect creates a tendency to coast *in general*, but with respect to the activity producing the positive feelings. We should also be clear that we are talking about the current, ongoing episode of action. We are *not* arguing that positive affect makes people less likely to do the behavior later on.

Does positive affect lead to coasting? Evidence consistent with this idea has been

reported by Mizruchi (1991), by Louro et al. (2007), and by Fulford et al. (2010). Thus far, however, the issue is relatively uncontested. Some are skeptical about this idea, because that it is hard to see why a process would be built in that limits positive feelings – indeed, dampens them. We see at least two bases for it. The first lies in a basic biological principle: it is adaptive not to spend energy needlessly. Coasting prevents this. Brehm has similarly argued that people engage only as much effort as is needed to accomplish a given task, and no more (Brehm and Self, 1989).

The second basis for such a process stems from the fact that people have multiple simultaneous concerns. Given multiple concerns, people do not optimize their outcome on any one of them, but “satisfice” (Simon, 1953) – do a good enough job on each to deal with it satisfactorily. This permits them to handle the many concerns adequately, rather than just any one of them. Coasting facilitates satisficing. A tendency to coast virtually defines satisficing regarding that particular goal. A tendency to coast also fosters satisficing of a broader set of goals, by allowing easy shift to other domains at little or no cost (see Carver, 2003a, for detail).

Affects and priority management

This line of argument implicates positive feelings in a broad function that deserves further attention: the shifting from one goal to another as focal in behavior (Dreisbach and Goschke, 2004; Shallice, 1978). This basic and very important phenomenon is often overlooked. Many goals are typically under pursuit simultaneously, but only one has top priority at a given moment. People need to shield and maintain intentions that are being pursued (cf. Shah et al., 2002), but they also need to be able to shift flexibly among goals (Shin and Rosenbaum, 2002).

The issue of priority management was addressed very creatively many years ago by

Simon (1967). He proposed that emotions are calls for reprioritization. He suggested that emotion arising with respect to a goal that is out of awareness eventually induces people to interrupt their behavior and give that goal a higher priority than it had. The stronger the emotion, the stronger is the claim that the unattended goal should have higher priority than the goal that is presently focal.

Simon’s discussion focused on cases in which a nonfocal goal demands a higher priority and *intrudes* on awareness. By strong implication, his discussion dealt only with negative affect. However, there is another way for priority ordering to shift: the focal goal can *relinquish its place*. Perhaps positive feelings also pertain to reprioritization, but rather than a call for higher priority, they reflect *reduction* in priority. Positive affect regarding avoidance (relief or tranquility) indicates that a threat has dissipated, no longer requires so much attention, and can assume a lower priority. Positive feelings regarding approach (happiness, joy) indicate that an incentive is being attained and could temporarily be put on hold because you are doing so well; thus, this goal can assume a lower priority (see Carver, 2003a).

Priority management and feelings of depression

One more aspect of priority management must be addressed, concerning the idea that some goals are best abandoned. As noted earlier, we have long held that sufficient doubt about goal attainment yields a tendency to disengage from effort, and even to disengage from the goal itself. This is certainly a kind of priority shift, in that the abandoned goal now has an even lower priority than it had before. How does this case fit the ideas described thus far?

This case seems to contradict Simon’s (1967) view that negative affect is a call for higher priority. But there is an important difference between two classes of negative

affect associated with approach (Carver, 2003a, 2004; for this discussion we disregard avoidance). Some of these affects coalesce around frustration and anger. Others coalesce around sadness, depression, and dejection. The former relate to an increase in priority, the latter to a decrease.

In describing our view on affect, we said that approach-related affects fall on a dimension. However, the dimension is not a simple straight line. Progress below the criterion creates negative affect, as the incentive slips away. Inadequate movement gives rise to frustration, irritation, and anger, prompting more effort to overcome obstacles and reverse the inadequate current progress. But efforts sometimes do not change the situation. Indeed, losses preclude movement forward. Now the feelings are sadness, depression, despondency, and hopelessness. Behaviors also differ in this case. The person tends to disengage from – give up on – further effort.

In the first case, feelings of frustration and anger are a call for an upgrade in priority, an increase in effort, a struggle to gain the incentive despite setbacks. In the second case, feelings of sadness and depression accompany *reduction* of effort and a downgrade in priority. As described earlier, both the upgrade and the downgrade have adaptive functions in the appropriate situations.

CHANGES IN THE THEORETICAL LANDSCAPE: TWO MODES OF FUNCTIONING

During the last two decades, changes have occurred in how people view cognition and action. The implicit assumption that behavior is generally managed in a top-down, directive way has been challenged. Questions have been raised about the role of consciousness in many kinds of action. Interest has arisen in the idea that the mind has both explicit and implicit representations. These various ideas have also influenced

how we think about ideas we have been working with.

Two-mode models

Several literatures have developed around the possible existence of two modes of functioning (for reviews see Carver and Scheier, 2009a; Carver et al., 2008). In personality, Epstein (e.g., 1973, 1994) has long advocated such a view. He argues that people experience reality through two systems. What he calls the *rational* system operates mostly consciously, uses logical rules, is verbal and deliberative, and thus is fairly slow. The *experiential* system is intuitive and associative in nature. It provides a quick and dirty way of assessing and reacting to reality. It relies on salient information and uses shortcuts and heuristics. It functions automatically and quickly. It is considered to be emotional (or at least very responsive to emotions) and nonverbal.

The experiential system is presumably older and more primitive neurobiologically. It dominates when speed is needed (as when the situation is emotionally charged). The rational system evolved later, providing a more cautious, analytic, planful way of proceeding. Being able to operate in that way has important advantages, if there is sufficient time and freedom from pressure to think things through. Both systems are presumed to be always at work, jointly determining behavior, though the extent of each one's influence can vary by situation and disposition.

A model in many ways similar to this, but with different roots, was proposed by Metcalfe and Mischel (1999). Drawing on decades of work on delay of gratification, Metcalfe and Mischel (1999) proposed that two systems influence restraint. One they called a “hot” system: emotional, impulsive, and reflexive. The other they called a “cool” system: strategic, flexible, slower, and unemotional. How people respond to difficult situations depends on which system is in charge.

There are also several two-mode theories in social psychology (Chaiken and Trope, 1999). The essence of such a view has existed for a long time in the literature of persuasion. Strack and Deutsch (2004) have recently extended this reasoning more broadly into the range of behavioral phenomena of interest to social psychologists. They proposed a model in which overt social behavior is a joint output of two simultaneous modes of functioning, which they termed *reflective* and *impulsive*. Again, differences in the systems' operating characteristics lead to differences in behavior. The reflective system anticipates the future, makes decisions on the basis of those anticipations, and forms intentions. It is planful and wide-ranging in its search for relevant information. It is restrained and deliberative. The impulsive system acts spontaneously when its schemas or production systems are sufficiently activated. It acts without consideration for the future or for broader implications or consequences of the action. This depiction is very similar in some ways to the ideas of Epstein (1973, 1994) and Metcalfe and Mischel (1999).

The idea that there are two modes of functioning can also be linked to the burgeoning literature on explicit and implicit motives, knowledge structures, and attitudes. There often is little or no relation between explicit (self-report) measures and implicit measures of the same construct. Although it is fairly easy to see why that might be the case for variables such as prejudice (given the social desirability issues involved), it is less obvious why it would be so for such variables as the self-concept. Two-mode models suggest a possibility (Beevers, 2005; Fazio and Olson, 2003). Implicit measures assess only associative links between pairs of elements. Explicit measures are symbolic, products of deliberative processing. Implicit knowledge presumably accrues through associative learning; explicit knowledge presumably accrues through verbal, conceptual learning. Perhaps associative and conceptual sources of knowledge about the self (or the world) are more independent of one another

than is often assumed. Thus, the two sources may not agree well with each other over time, leading to different results from implicit and explicit measures.

Consistent with this line of thought, a number of studies have found that both implicit and explicit measures do predict aspects of behavior, but typically different aspects. Explicit measures predict deliberated decisions and intentions; implicit measures predict relatively automatic actions, nonverbal behaviors, and primed word completions (Dovidio et al., 1997; Neumann et al., 2004).

Two-mode thinking has also been very influential in developmental psychology. Rothbart and her colleagues have argued for the existence of three temperament systems: two for approach and avoidance, and a third termed *effortful control* (e.g., Derryberry and Rothbart, 1997; Rothbart and Posner, 1985; Rothbart et al., 2000; see also Nigg, 2000). Effortful control concerns (in part) the ability to suppress approach when it is situationally inappropriate. Effortful control is superordinate to approach and avoidance temperaments. The label *effortful* conveys the sense that this is an executive, planful activity, entailing the use of cognitive resources beyond those needed to react impulsively. This view of effortful control has substantial resemblance to depictions of the deliberative mode of the two-mode models outlined in previous sections.

Hierarchicality re-examined

Thus, several sources of theory propose that the mind functions in two modes. Indeed, the sources described above are far from an exhaustive list. All of them promote the inference that a deliberative mode of functioning uses symbolic and sequential processing and thus is relatively slow. They also suggest that a more impulsive or reactive mode of functioning uses associationist processing and is relatively fast. Many of the theories suggest that the two modes are

semiautonomous in their functioning, competing with each other to influence actions. Indeed, many also point to situational variables that influence which mode dominates at a given time.

These kinds of ideas have influenced how we construe the hierarchy of control that was proposed by Powers (1973). We said earlier that *programs* of action entail decisions. They seem to be managed top-down, using effortful processing. Planfulness, an element of programs, is also a common characterization of behavior managed by the reflective system. It seems reasonable to map program-level control onto the deliberative, reflective mode of functioning.

In contrast to this deliberative quality, well-learned *sequences* occur in a relatively automatic stream once they are triggered. Sequences (and all lower levels of control) are inevitably called up during the execution of programs. However, it may be that sequences can also be triggered more autonomously. Sequences may respond to cues that trigger them simply by virtue of associations in memory. In such cases, the operating characteristics would seem akin to those of the reactive mode of functioning.

We have often noted that the level of control that is functionally superordinate can vary by situations and persons (e.g., Carver and Scheier, 1998, 1999a). That is, it is possible to imagine cases in which a person is behaving intentionally according to a principle (e.g., a moral or ethical value), and it is possible to imagine cases in which the person is behaving according to a plan or program. It is also possible, however, to imagine cases in which the person is acting impulsively and spontaneously, without regard to either principle or plan.

In making this case in the past, our emphasis typically was simply on how sequences and programs differed. Now we are inclined to wonder if this particular differentiation is not perhaps more important than we had previously realized. Perhaps we have underappreciated the extent to which lower levels of self-regulatory structures can be triggered

autonomously and their outputs enter the stream of ongoing action without oversight from higher levels (Carver and Scheier, 2002), and potentially even in conflict with values at higher levels. This seems an important question for further exploration.

Self-control: impulse and restraint

The idea that there are conflicts between longer-term and shorter-term goals is also part of a literature on self-control and self-control failure (e.g., Baumeister et al., 1994). This literature focuses on cases in which a person is both motivated to act and motivated to restrain that action. This is essentially the case that is the focus of work on children's effortful control, and it also resembles the logical structure of the delay of gratification paradigm. A difference is that in the self-control literature the intent often is to delay indefinitely rather than temporarily.

Although the self-control situation is often portrayed as pitting longer- and shorter-term goals against each other, the preceding discussion suggests a somewhat different view. The self-control situation may be pitting the two modes of processing against each other. This reframing would be consistent with the literature on self-control failure, which tends to portray such failures as involving a relatively automatic tendency to act in one way, opposed by a planful effort to restrain that act. The action that is being inhibited is often characterized as an impulse, a desire that would automatically be translated into action unless it is controlled (perhaps because this action is habitual, perhaps because it is more primal). The restraint is presumed to be effortful, and to depend on limited resources. If the planful part of the mind is fully enough able to attend to the conflict, the person can resist the impulse. If not, the impulse is more likely to be expressed. This portrayal seems quite consonant with the two-mode models of functioning.

APPLICABILITY TO SOCIAL ISSUES

A broader purpose of this volume is to suggest how theories may bear on applied topics. Authors were asked to evaluate the applicability of the theory they described to understanding and solving social issues and problems. This is a particularly difficult task for us, partly because our minds seem not to work that way, and partly because of the very nature of the ideas we have been writing about. As personality psychologists, most of what comes to us when we think about these ideas is what is going on in the person's mind. If we construe the idea of "social issues" quite broadly, however, a couple of applications do come to mind.

The affect-related part of the model provides what we think are useful ideas for understanding the nature of human distress: the idea that distress follows from the perception of not reaching desired goals (or not avoiding threats) combined with a continuing commitment to those goals. Put more simply, the model points to the bind that is inherent in a commitment to the unattainable. Reducing distress sometimes requires finding better ways to move forward, but it sometimes requires the person to abandon goals and values and adopt new ones. Unfortunately, with few exceptions the model does not provide clear guidance about which of these options will be more profitable in a given case. An exception is that pursuit of some goals pulls and tears at the fabric of the self, enlarging discrepancies at a higher level even while closing them at a lower level. Pursuit of those goals will inevitably be problematic. We believe these ideas are useful for conceptualizing (and potentially treating) debilitating distress.

Another place where the ideas under discussion have had applied implications (if not exactly implications for social problems) concerns the portion of the theory bearing on expectations for future outcomes. Although we did not say so earlier in the chapter, we have used that theoretical principle to conceptualize and measure the individual

difference dimension of optimism versus pessimism. This variable has been studied in a good deal of research in health psychology and related areas (Carver and Scheier, 2009b). It turns out to have important implications for how people respond to adversity, both psychologically and physically (Rasmussen et al., 2009; Solberg Nes and Segerstrom, 2006). Confidence about one's future keeps one in the struggle to adapt and thrive, and leads to better outcomes. This application indicates how the ideas embedded in this model are relevant to the broad sweep of human wellbeing.

ACKNOWLEDGMENTS

Preparation of this chapter was supported by grants from the National Cancer Institute (CA64710), the National Science Foundation (BCS0544617), and the National Heart, Lung, and Blood Institute (HL65111, HL65112, HL076852, and HL076858).

REFERENCES

Ashby, W.R. (1940) Adaptiveness and equilibrium. *Journal of Mental Science*, 86, 478–483.

Atkinson, J.W. (1964) *An Introduction to Motivation*. Princeton, NJ: Van Nostrand.

Baltes, P.B., Cornelius, S.W. and Nesselroade, J.R. (1979) Cohort effects in developmental psychology. In J.R. Nesselroade and P.B. Baltes (eds), *Longitudinal Research in the Study of Behavior and Development*, pp. 61–87. New York: Academic Press.

Bandura, A. (1986) *Social Foundations of Thought and Action: A Social Cognitive Theory*. Englewood Cliffs, NJ: Prentice-Hall.

Baumeister, R.F., Heatherton, T.F. and Tice, D.M. (1994) *Losing Control: Why People Fail at Self-regulation*. San Diego: Academic Press.

Baumeister, R.F. and Vohs, K.D. (eds) (2004) *Handbook of Self-regulation: Research, Theory, and Applications*. New York: Guilford Press.

Beevers, C.G. (2005) Cognitive vulnerability to depression: A dual-process model. *Clinical Psychology Review*, 25, 975–1002.

Brehm, J.W. (1966) *A Theory of Psychological Reactance*. New York: Academic Press.

Brehm, J.W. and Self, E.A. (1989) The intensity of motivation. *Annual Review of Psychology*, 40, 109–131.

Brendl, C.M. and Higgins, E.T. (1996) Principles of judging valence: What makes events positive or negative? *Advances in Experimental Social Psychology*, 28, 95–160.

Cacioppo, J.T., Gardner, W.L. and Berntson, G.G. (1999) The affect system has parallel and integrative processing components: Form follows function. *Journal of Personality and Social Psychology*, 76, 839–855.

Campos, J.J., Frankel, C.B. and Camras, L. (2004) On the nature of emotion regulation. *Child Development*, 75, 377–394.

Cannon, W.B. (1932) *The Wisdom of the Body*. New York: Norton.

Carver, C.S. (1974) Facilitation of physical aggression through objective self-awareness. *Journal of Experimental Social Psychology*, 10, 365–370.

Carver, C.S. (2001) Affect and the functional bases of behavior: On the dimensional structure of affective experience. *Personality and Social Psychology Review*, 5, 345–356.

Carver, C.S. (2003a) Pleasure as a sign you can attend to something else: Placing positive feelings within a general model of affect. *Cognition and Emotion*, 17, 241–261.

Carver, C.S. (2003b) Self-awareness. In M.R. Leary and J. Tangney (eds), *Handbook of Self and Identity*, pp. 179–196. New York: Guilford Press.

Carver, C.S. (2004) Negative affects deriving from the behavioral approach system. *Emotion*, 4, 3–22.

Carver, C.S. (2009) Threat sensitivity, incentive sensitivity, and the experience of relief. *Journal of Personality*, 77, 125–138.

Carver, C.S. and Harmon-Jones, E. (2009) Anger is an approach-related affect: Evidence and implications. *Psychological Bulletin*, 135, 183–204.

Carver, C.S. and Humphries, C. (1981) Havana daydreaming: A study of self-consciousness and the negative reference group among Cuban Americans. *Journal of Personality and Social Psychology*, 40, 545–552.

Carver, C.S., Johnson, S.L. and Joormann, J. (2008) Serotonergic function, two-mode models of self-regulation, and vulnerability to depression: What depression has in common with impulsive aggression. *Psychological Bulletin*, 134, 912–943.

Carver, C.S., Lawrence, J.W. and Scheier, M.F. (1999) Self-discrepancies and affect: Incorporating the role of feared selves. *Personality and Social Psychology Bulletin*, 25, 783–792.

Carver, C.S. and Scheier, M.F. (1981) *Attention and Self-regulation: A Control-theory Approach to Human Behavior*. New York: Springer Verlag.

Carver, C.S. and Scheier, M.F. (1990) Origins and functions of positive and negative affect: A control-process view. *Psychological Review*, 97, 19–35.

Carver, C.S. and Scheier, M.F. (1998) *On the Self-regulation of Behavior*. New York: Cambridge University Press.

Carver, C.S. and Scheier, M.F. (1999a) Themes and issues in the self-regulation of behavior. In R.S. Wyer, Jr. (ed.), *Advances in Social Cognition*, 12, 1–105. Mahwah, NJ: Erlbaum.

Carver, C.S. and Scheier, M.F. (1999b) Several more themes, a lot more issues: Commentary on the commentaries. In R.S. Wyer, Jr. (ed.), *Advances in Social Cognition*, 12, 261–302. Mahwah, NJ: Erlbaum.

Carver, C.S. and Scheier, M.F. (2000) Scaling back goals and recalibration of the affect system are processes in normal adaptive self-regulation: Understanding 'response shift' phenomena. *Social Science and Medicine*, 50, 1715–1722.

Carver, C.S. and Scheier, M.F. (2002) Control processes and self-organization as complementary principles underlying behavior. *Personality and Social Psychology Review*, 6, 304–315.

Carver, C.S. and Scheier, M.F. (2009a) Action, affect, multi-tasking, and layers of control. In J.P. Forgas, R.F. Baumeister, and D. Tice (eds), *The Psychology of Self-regulation*, pp. 109–126. New York: Psychology Press.

Carver, C.S. and Scheier, M.F. (2009b) Optimism. In M.R. Leary and R.H. Hoyle (eds), *Handbook of Individual Differences in Social Behavior*, pp. 330–342. New York: Guilford Press.

Carver, C.S. and White, T.L. (1994) Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of Personality and Social Psychology*, 67, 319–333.

Chaiken, S.L. and Trope, Y. (eds) (1999) *Dual-process Theories in Social Psychology*. New York: Guilford Press.

Clark, R.N. (1996) *Control System Dynamics*. New York: Cambridge University Press.

Cleiren, M. (1993) *Bereavement and Adaptation: A Comparative Study of the Aftermath of Death*. Washington, DC: Hemisphere.

Clore, G.C. (1994) Why emotions are felt. In P. Ekman and R.J. Davidson (eds), *The Nature of Emotion: Fundamental Questions*, pp. 103–111. New York: Oxford University Press.

Davidson, R.J. (1992) Anterior cerebral asymmetry and the nature of emotion. *Brain and Cognition*, 20, 125–151.

Davidson, R.J. (1998) Affective style and affective disorders: Perspectives from affective neuroscience. *Cognition and Emotion*, 12, 307–330.

DeAngelis, D.L., Post, W.M. and Travis, C.C. (1986) *Positive Feedback in Natural Systems (Biomathematics, Vol. 15)*. Berlin: Springer-Verlag.

Derryberry, D. and Rothbart, M.K. (1997) Reactive and effortful processes in the organization of temperament. *Development and Psychopathology*, 9, 633–652.

Dovidio, J.F., Kawakami, K., Johnson, C., Johnson, B. and Howard, A. (1997) On the nature of prejudice: Automatic and controlled processes. *Journal of Experimental Social Psychology*, 33, 510–540.

Dreisbach, G. and Goschke, T. (2004) How positive affect modulates cognitive control: Reduced perseveration at the cost of increased distractibility. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30, 343–353.

Duval, S. and Wicklund, R.A. (1972) *A Theory of Objective Self-awareness*. New York: Academic Press.

Elliot, A. (ed.) (2008) *Handbook of Approach and Avoidance Motivation*. Mahwah, NJ: Erlbaum.

Epstein, S. (1973) The self-concept revisited: Or a theory of a theory. *American Psychologist*, 28, 404–416.

Epstein, S. (1994) Integration of the cognitive and the psychodynamic unconscious. *American Psychologist*, 49, 709–724.

Fazio, R.H. and Olson, M.A. (2003) Implicit measures in social cognition research: Their meaning and use. *Annual Review of Psychology*, 54, 297–327.

Feather, N.T. (1982) *Expectations and Actions: Expectancy-value Models in Psychology*. Hillsdale: Erlbaum.

Festinger, L. (1957) *A Theory of Cognitive Dissonance*. Evanston, IL: Row, Peterson.

Ford, D.H. (1987) *Humans as Self-constructing Living Systems: A Developmental Perspective on Behavior and Personality*. Hillsdale, NJ: Erlbaum.

Frijda, N.H. (1986) *The Emotions*. Cambridge: Cambridge University Press.

Frijda, N.H. (1988) The laws of emotion. *American Psychologist*, 43, 349–358.

Fulford, D., Johnson, S.L., Llabre, M.M. and Carver, C.S. (2010) Pushing and coasting in dynamic goal pursuit: Coasting is attenuated in bipolar disorder. *Psychological Science*, 21, 1021–1027.

Gross, J.J. (ed.) (2007) *Handbook of Emotion Regulation*. New York: Guilford Press.

Heckhausen, J. and Schulz, R. (1995) A life-span theory of control. *Psychological Review*, 102, 284–304.

Higgins, E.T. (1987) Self-discrepancy: A theory relating self and affect. *Psychological Review*, 94, 319–340.

Higgins, E.T. (1996) Ideals, oughts, and regulatory focus: Affect and motivation from distinct pains and pleasures. In P.M. Gollwitzer and J.A. Bargh (eds), *The Psychology of Action: Linking Cognition and Motivation to Behavior*, pp. 91–114. New York: Guilford Press.

Higgins, E.T., Shah, J. and Friedman, R. (1997) Emotional responses to goal attainment: Strength of regulatory focus as moderator. *Journal of Personality and Social Psychology*, 72, 515–525.

Klinger, E. (1975) Consequences of commitment to and disengagement from incentives. *Psychological Review*, 82, 1–25.

Lang, P.J., Bradley, M.M. and Cuthbert, B.N. (1990). Emotion, attention, and the startle reflex. *Psychological Review*, 97, 377–395.

Louro, M.J., Pieters, R. and Zeelenberg, M. (2007) Dynamics of multiple-goal pursuit. *Journal of Personality and Social Psychology*, 93, 174–193.

MacKay, D.M. (1956) Toward an information-flow model of human behavior. *British Journal of Psychology*, 47, 30–43.

MacKay, D.M. (1966) Cerebral organization and the conscious control of action. In J.C. Eccles (ed.), *Brain and Conscious Experience*, pp. 422–445. Berlin: Springer-Verlag.

Markus, H. and Nurius, P. (1986) Possible selves. *American Psychologist*, 41, 954–969.

Maruyama, M. (1963) The second cybernetics: Deviation-amplifying mutual causal processes. *American Scientist*, 51, 164–179.

McFarland, D.J. (1971) *Feedback Mechanisms in Animal Behavior*. New York: Academic Press.

Metcalfe, J. and Mischel, W. (1999) A hot/cool-system analysis of delay of gratification: Dynamics of willpower. *Psychological Review*, 106, 3–19.

Miller, G.A., Galanter, E. and Pribram, K.H. (1960) *Plans and the Structure of Behavior*. New York: Holt, Rinehart, and Winston.

Miller, G.E. and Wrosch, C. (2007) You've gotta know when to fold 'em: Goal disengagement and

systemic inflammation in adolescence. *Psychological Science*, 18, 773–777.

Miller, N.E. (1944) Experimental studies of conflict. In J. McV. Hunt (ed.), *Personality and the Behavior Disorders*, 1, 431–465. New York: Ronald Press.

Miller, N.E. and Dollard, J. (1941) *Social Learning and Imitation*. New Haven: Yale University Press.

Mizruchi, M.S. (1991) Urgency, motivation, and group performance: The effect of prior success on current success among professional basketball teams. *Social Psychology Quarterly*, 54, 181–189.

Nesse, R.M. (2000) Is depression an adaptation? *Archives of General Psychiatry*, 57, 14–20.

Neumann, R., Hülsenbeck, K. and Seibt, B. (2004) Attitudes towards people with AIDS and avoidance behavior: Automatic and reflective bases of behavior. *Journal of Experimental Social Psychology*, 40, 543–550.

Nigg, J.T. (2000) On inhibition/disinhibition in developmental psychopathology: Views from cognitive and personality psychology as a working inhibition taxonomy. *Psychological Bulletin*, 126, 220–246.

Ochsner, K.N. and Gross, J.J. (2008) Cognitive emotion regulation: Insights from social cognitive and affective neuroscience. *Current Directions in Psychological Science*, 17, 153–158.

Ogilvie, D.M. (1987) The undesired self: A neglected variable in personality research. *Journal of Personality and Social Psychology*, 52, 379–385.

Ortony, A., Clore, G.L. and Collins, A. (1988) *The Cognitive Structure of Emotions*. New York: Cambridge University Press.

Powers, W.T. (1973) *Behavior: The Control of Perception*. Chicago: Aldine.

Pyszcynski, T. and Greenberg, J. (1992). *Hanging On and Letting Go: Understanding the Onset, Progression, and Remission of Depression*. New York: Springer-Verlag.

Rasmussen, H.N., Scheier, M.F. and Greenhouse, J.B. (2009) Optimism and physical health: A meta-analytic review. *Annals of Behavioral Medicine*, 37, 239–256.

Rosenbaum, D.A., Carlson, R.A. and Gilmore, R.O. (2001) Acquisition of intellectual and perceptual-motor skills. *Annual Review of Psychology*, 52, 453–470.

Rosenblueth, A., Wiener, N. and Bigelow, J. (1943) Behavior, purpose, and teleology. *Philosophy of Science*, 10, 18–24.

Rothbart, M.K., Ahadi, S.A. and Evans, D.E. (2000). Temperament and personality: Origins and outcomes. *Journal of Personality and Social Psychology*, 78, 122–135.

Rothbart, M.K. and Posner, M. (1985) Temperament and the development of self-regulation. In L.C. Hartlage and C.F. Telzrow (eds), *The Neuropsychology of Individual Differences: A Developmental Perspective*, pp. 93–123. New York: Plenum.

Rotter, J.B. (1954) *Social Learning and Clinical Psychology*. New York: Prentice-Hall.

Ryan, R.M. and Deci, E.L. (2000) Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55, 68–78.

Scheier, M.F., Fenigstein, A. and Buss, A.H. (1974) Self-awareness and physical aggression. *Journal of Experimental Social Psychology*, 10, 264–273.

Schwartz, S.H. and Bilsky, W. (1990) Toward a theory of the universal content and structure of values: Extensions and cross-cultural replications. *Journal of Personality and Social Psychology*, 58, 878–891.

Schwartz, S.H. and Rubel, T. (2005) Sex differences in value priorities: Cross-cultural and multimethod studies. *Journal of Personality and Social Psychology*, 89, 1010–1028.

Shah, J.Y., Friedman, R. and Kruglanski, A.W. (2002) Forgetting all else: On the antecedents and consequences of goal shielding. *Journal of Personality and Social Psychology*, 83, 1261–1280.

Shallice, T. (1978) The dominant action system: An information-processing approach to consciousness. In K.S. Pope and J.L. Singer (eds), *The Stream of Consciousness: Scientific Investigations into the Flow of Human Experience*, pp. 117–157. New York: Wiley.

Shin, J.C. and Rosenbaum, D.A. (2002) Reaching while calculating: Scheduling of cognitive and perceptual-motor processes. *Journal of Experimental Psychology: General*, 131, 206–219.

Simon, H.A. (1953) *Models of Man*. New York: Wiley.

Simon, H.A. (1967) Motivational and emotional controls of cognition. *Psychology Review*, 74, 29–39.

Solberg Nes, L. and Segerstrom, S.C. (2006) Dispositional optimism and coping: A meta-analytic review. *Personality and Social Psychology Review*, 10, 235–251.

Steenbarger, B.N. and Aderman, D. (1979) Objective self-awareness as a nonaversive state: Effect of anticipating discrepancy reduction. *Journal of Personality*, 47, 330–339.

Strack, F. and Deutsch, R. (2004) Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, 8, 220–247.

Tolman, E.C. (1938) The determiners of behavior at a choice point. *Psychological Review*, 45, 1–41.

Vallacher, R.R. and Wegner, D.M. (1987) What do people think they're doing? Action identification and human behavior. *Psychological Review*, 94, 3–15.

von Bertalanffy, L. (1968) *General Systems Theory*. New York: Braziller.

Vroom, V.H. (1964) *Work and Motivation*. New York: Wiley.

Watson, D. (2009) Locating anger in the hierarchical structure of affect: Comment on Carver and Harmon-Jones. *Psychological Bulletin*, 135, 205–208.

Watson, D., Wiese, D., Vaidya, J. and Tellegen, A. (1999) The two general activation systems of affect: Structural findings, evolutionary considerations, and psychobiological evidence. *Journal of Personality and Social Psychology*, 76, 820–838.

Weiss, R.S. (1988) Loss and recovery. *Journal of Social Issues*, 44, 37–52.

Wiener, N. (1948) *Cybernetics: Control and Communication in the Animal and the Machine*. Cambridge, MA: MIT Press.

Wortman, C.B. and Brehm, J.W. (1975) Responses to uncontrollable outcomes: An integration of reactance theory and the learned helplessness model. In L. Berkowitz (ed.), *Advances in Experimental Social Psychology*, 8, 277–336. New York: Academic Press.

Wrosch, C., Scheier, M.F., Carver, C.S. and Schulz, R. (2003a) The importance of goal disengagement in adaptive self-regulation: When giving up is beneficial. *Self and Identity*, 2, 1–20.

Wrosch, C., Scheier, M.F., Miller, G.E., Schulz, R. and Carver, C.S. (2003b) Adaptive self-regulation of unattainable goals: Goal disengagement, goal re-engagement, and subjective well-being. *Personality and Social Psychology Bulletin*, 29, 1494–1508.

Mindset Theory of Action Phases

Peter M. Gollwitzer

ABSTRACT

Mindset theory of action phases is based on the distinction between motivation and volition as proposed by the Rubicon model which claims that prior to crossing the Rubicon (i.e., making a goal decision) motivational principles apply whereas thereafter volitional principles set in. The latter are concerned with goal implementation, whereas the former relate to the choosing of goals. Mindset theory of action phases proposes that different cognitive procedures are activated when people tackle the task of choosing goals versus implementing them. The respective task demands determine the features that characterize the deliberative versus implemental mindset. These pertain to what type of information is preferably processed and how it is analyzed. Mindset research has produced findings that not only support the motivation versus volition distinction but also enlighten various debates and theories in social psychology (e.g., optimism versus realism debate, dual process theories, goal theory). Mindset theory of action phases has also spurred research on effective planning by pointing to implementation intentions (i.e., if-then plans). This research has had much applied impact. When it was linked up to research on strategies of motivationally smart goal setting (i.e., mental contrasting), it initiated the development of a time- and cost-effective behavior change intervention.

INTRODUCTION

During my graduate education in the late 1970s at the University of Texas at Austin, my mentor, Robert Wicklund, and I started to conceive of people's selves or identities as goals. We thought that people can very well set themselves goals to become a good parent, a brilliant scientist, or a great athlete. If one takes this perspective, the self of a person is no longer just something to understand (self-concept) and like (self-esteem) but something to be achieved (identity goal). We turned to the writings of Kurt Lewin (1926) and his students, whose tension system theory of goal pursuit, with its notion of substitution, was very helpful to developing our theory of symbolic self-completion (Gollwitzer and Kirchhof, 1998; Wicklund and Gollwitzer, 1982). The main proposition of self-completion theory is that once people have set themselves certain identity or self-defining goals, they respond to failure experiences, shortcomings, or barriers not with retreat but instead with intensified efforts to reach the goal. These efforts, however, do not

have to alleviate the problem at hand but may involve resorting to any substitute that indicates goal attainment (e.g., showing off relevant status symbols, engaging in identity-relevant activities, describing oneself as having the required personal attributes; for recent research on self-completion theory see Gollwitzer et al., 2009; Harmon-Jones et al., 2009; Ledgerwood et al., 2007).

In the early 1980s, Heinz Heckhausen invited me to join the newly founded Max Planck Institute for Psychological Research at Munich to start a research unit called Motivation and Action. We quickly realized that we had a directly opposed conceptual view of motivation. Whereas Heinz Heckhausen's motivation was that of an expectancy-value theorist in the tradition of Atkinson (1957) and Heckhausen (1977) and was thus fueled by the perceived feasibility and desirability of a given action, my motivation was that of Lewin's (1926) tension system and was resting in the determination or commitment a person holds with respect to the action goal at hand. Apparently, in the research on self-completion (Wicklund and Gollwitzer, 1982) I had been studying issues of goal striving (i.e., thoughts and behavior directed toward existing goals), whereas Heinz Heckhausen in his work on achievement motivation (Heckhausen, 1977) had focused on issues of goal setting (i.e., what goals people find attractive and feasible, and thus choose for themselves).

THE RUBICON MODEL OF ACTION PHASES

To highlight this insight, we suggested making a distinction between motivation and volition. Following the conceptual terms used by Lewin (1926) and Narziß Ach (1935), we dubbed the goal-striving-related motivation with the term *volition*, and kept the term *motivation* for the goal-setting-related motivation. More importantly, in an attempt to integrate these two kinds of phenomena

(i.e., motivation and volition) we developed the Rubicon model of action phases (Heckhausen, 1987; Heckhausen and Gollwitzer, 1987). This model suggests that the course of action can be segmented into four different, consecutive phases that differ in terms of the tasks that are to be solved by the individual given that s/he wants to execute a given course of action successfully. The first phase (predecision phase) is said to pose the task of setting preferences among wishes and desires by deliberating their desirability and feasibility. As people's motives and needs produce more wishes and desires than can possibly be realized, the individual is forced to choose among these desires and by doing so turn them into goals. Once goals are set (i.e., the Rubicon has been crossed), the individual faces the second task (preaction phase), which is getting started with goal-directed behaviors. This may be simple if the necessary goal-directed actions are well practiced and routine but complex if the individual is still undecided about where, when, and how to act. In such complex cases, the execution of goal-directed action has to be planned by deciding on when, where, and how to act. The third task (action phase) is bringing the initiated goal-directed action to a successful ending, and this is best achieved by determined and persistent pursuit of goal completion. Finally, in the fourth task (postaction phase), the individual needs to decide whether the desired goal has indeed been achieved or whether further striving is needed.

The Rubicon model of action phases postulates that a person's psychological functioning in each of these phases is governed by different principles. Classic theories of motivation (adhering to the restricted definition of motivation as determined by feasibility and desirability; Atkinson, 1957; Feather and Newton, 1982; Heckhausen, 1977) are said to be well suited to explicate the psychological processes associated with the predecision and postaction phases, whereas theories of volition (i.e., theories on the self-regulation of goal attainment; Lewin, 1926;

Mischel, 1974; Mischel and Patterson, 1978) are most appropriate to explaining the psychological processes that characterize the preaction and action phase. In other words, the predecision and postaction phases are expected to encompass motivational phenomena and processes in the classic sense of the term, whereas in the phases in between volitional phenomena and processes are thought to occur.

This radical statement needed empirical support, and therefore Heckhausen and I conducted an early experiment aimed at demonstrating that individuals placed in the predecision phase evidence different cognitive functioning than do individuals in the preaction phase (Heckhausen and Gollwitzer, 1987, Study 2). Assuming that deliberation of the desirability and feasibility of wishes and desires (the task of the predecision phase) is cognitively more demanding than committing to a plan that specifies, when, where, and how one wants to perform goal-directed actions (the task of the preaction phase), we expected that deliberating individuals experienced a higher cognitive load than planning (i.e., preaction) individuals. We therefore interrupted experimental participants who were either in the middle of deliberating a choice between two different tests that presumably measured their creative potential or in the middle of planning how to perform the test they had just chosen and then asked them to take a short-term memory test (i.e., a noun span test that presented nouns irrelevant to the creativity tests at hand). We expected that deliberating participants, because of heightened cognitive load, would evidence a reduced noun span, compared with their span as measured at the beginning of the experiment. We also expected that deliberating participants would evidence a comparatively more reduced noun span than planning participants because laying down a plan on how to act was expected to take up less cognitive resources than deliberating the pros and cons of a goal decision.

To our surprise, the results were just opposite to what we had predicted (Heckhausen

and Gollwitzer, 1987, Study 2). The deliberating participants showed an increase in their short-term memory capacity, compared with both their own prior span and the span of the planning participants. In an effort to reduce our confusion about these unexpected findings, I turned to Gerhard Strube, at the time a cognitive psychologist at the Max-Planck-Institute for Psychological Research, and he pointed me to the classic concept of mindset as originally advanced at the turn of the century by the German psychologists Külpe (1904), Marbe (1915), Orth (1903), and Watt (1905), all members of the Würzburg school. These early cognitive psychologists had discovered that becoming intensively involved with performing a given task activates exactly those cognitive procedures that help task completion. The created mindset (i.e., the sum total of the activated cognitive procedures) is the cognitive orientation most conducive to successful task performance.

The mindset notion allows interpreting the observed noun span data as follows: deliberating between potential action goals activates cognitive procedures (*the deliberative mindset*) that facilitate the task of the predecision phase, which is to set preferences. As undecided individuals do not know yet in which direction their decisions will finally take them, a heightened receptiveness to all kinds of information (open-mindedness) seems appropriate and functional to task solution. Similarly, planning out the implementation of a chosen goal should activate cognitive procedures (*the implemental mindset*) that facilitate the task of the preaction phase (i.e., getting started on the chosen goal). As this requires a more focused and selective orientation to processing information, closed-mindedness rather than open-mindedness with respect to available information seems called for. This postulated difference in receptiveness between deliberating and planning individuals is expressed in the fact that the experimental participants in the Heckhausen and Gollwitzer (1987, Study 2) noun span study processed the presented information in the noun span task

faster than planning participants (i.e., the deliberating participants demonstrated a broader noun span than the planning participants).

MINDSET THEORY OF ACTION PHASES

However, isn't all of this post hoc? This is exactly the type of worry that made me use the mindset notion as a hypothesis-generating device for subsequent research and thus for developing a comprehensive mindset theory of action phases in my *Habilitationsschrift* (i.e., a second, more extensive doctoral thesis that in Germany is a prerequisite for attaining a tenured professorship; Gollwitzer, 1987). A summary of this thesis (Gollwitzer, 1990) can be found in a chapter in *Motivation and Cognition* edited by E. Tory Higgins and Richard M. Sorrentino (1990), and a more extensive version in a German book, *Abwägen und Planen* [Deliberating and Planning] (Gollwitzer, 1991). If one analyzes the unique demands of the task of choosing between wishes and desires in the predecision phase versus the typical demands of the task of getting started on a chosen goal in the preaction phase, it becomes possible to detect further cognitive features of the deliberative as compared to the implemental mindset that can then be tested in new experiments. The task of deliberating in the predecisional phase is to choose, from among various wishes and desires, those few that one wants to realize (Gollwitzer, 1990). The criteria for selection should be the feasibility and desirability of the wishes and desires at issue. The systematic analysis of the chances of realization as well as the desirability of realization requires that relevant information be preferentially encoded and retrieved. But such cognitive tuning to this information should not suffice, as feasibility-related information needs to be analyzed objectively (and not in a self-serving manner), and desirability-related information needs to be analyzed in an

impartial manner (and not in a biased manner). Only if feasibility-related information is analyzed realistically, and the pros and cons are weighed impartially, can the individual turn those desires into binding goals that can potentially be realized and possess a genuine attractiveness. Moreover, deliberating requires a general open-mindedness (as was demonstrated in the Heckhausen and Gollwitzer [1987] study described above) with respect to any available information, as undecided individuals do not know yet in which direction their decision will finally take them.

Once a goal decision has been made, the task of planning is to promote the initiation of goal-directed behavior. This requires committing oneself to when, where, and how to get started. Accordingly, one needs to discover good opportunities and link them to appropriate goal-directed actions, thus creating plans for action. For this purpose, cognitive tuning toward implementation-related issues should be beneficial. Feasibility-related and desirability-related issues should no longer matter, and, if forced on the individual, they are avoided by distorting the relevant information in support of the goal decision made: the person sees the feasibility of the chosen goal in an overly optimistic way, and views the desirability of the chosen goal in a partial manner (i.e., pros exceed cons). Finally, processing all of the available information in an open-minded manner should be dysfunctional, as it might derail the individual from the chosen course of action. Accordingly, a reduced open-mindedness (closed-mindedness) favouring the selective processing of information in support of the chosen goal is to be expected.

Given these different features of the deliberative and implemental mindsets, one should not forget that the two different mindsets also possess many similar attributes. For instance, the mindset theory of action phases assumes that both deliberative and implemental mindsets become more pronounced as a person gets more involved with deliberating between potential goals and with planning chosen

goals, respectively. Moreover, neither mindset should immediately vanish when the task activity that produced it is ended; instead, the mindset should show a moment of inertia. This implies that the cognitive orientations associated with the deliberative and implemental mindsets can be detected in their effects on performing temporally subsequent tasks of a different nature. These ideas regarding the similarity between the deliberative and implemental mindset have been used to develop a research program aimed at testing the proposed different cognitive features of the deliberative and implemental mindsets.

In this research, the following method of inducing the deliberative and implemental mindsets turned out to be most effective: experimental participants are asked either to extensively deliberate an unresolved personal problem to be named by the participants (who indicate problems such as, "Should I move to another city or not?," "Should I change my major?," "Should I buy a new car?," or "Should I get involved with somebody?") or to plan the implementation of a chosen goal indicated by the participants (projects such as, "I will move to another city," "I will change my major," etc., are named). These requests create a deliberative and an implemental mindset, respectively. To intensify these mindsets, deliberating participants are asked to list the short-term and long-term pros and cons of making and not making a decision, in order to get heavily involved with deliberating. Planning participants, on the other hand, are asked to list the five most important steps of implementing the chosen goal, and then to specify when, where, and how they intend to execute each step, all of which serves the purpose of creating an intensive involvement with planning. Thereafter, both the deliberating and the planning participants are asked to perform presumably unrelated tasks (usually presented by a different experimenter in a different situational context), which are designed to measure the very cognitive features hypothesized to differ between the deliberative and implemental mindsets.

This procedure of inducing the deliberative and implemental mindsets in one situational context and assessing their cognitive and behavioral consequences in a different setting, has been referred to as procedural priming or mindset priming (Bargh and Chartrand, 2000), as research participants commonly stay unaware of the mindset effects they evidence.

Deliberative versus implemental mindsets and cognitive tuning

The hypothesis that the deliberative mindset creates cognitive tuning toward information relevant to making goal decisions (information on feasibility and desirability), whereas the implemental mindset tunes a person's cognitions to implementation-related information (information on where, when, and how to act), was tested most critically by Gollwitzer et al. (1990). Participants were placed into either a deliberative or an implemental mindset by having them deliberate on unresolved personal problems or plan chosen goal projects, respectively (the standard procedure described above was used). In a presumably unrelated second part of the experiment, participants were presented with the first few lines of a number of novel fairy tales and were instructed to complete each tale. Even though participants were allowed to continue the stories in any way they liked, deliberating participants had the protagonists of the tales reflect on reasons for choosing or not choosing certain action goals to a greater degree than planning participants did. Thoughts about how to accomplish a chosen goal, however, were more frequently attributed to the protagonists by planning participants than by deliberating participants.

Focusing on the processing of mindset-congruent information, Gollwitzer et al. (1990, Study 2) conducted an experiment in which participants had to recall the presented deliberative and implemental thoughts of others. Participants were placed into either a deliberative or an implemental mindset by

having them reflect the choice of one of two tests (i.e., decide between two different creativity tests) or plan to perform a chosen test. While participants were involved in deliberating or planning, slides were presented that depicted different persons mulling over personal decisions. For example, a depicted elderly lady was thinking of the pros (i.e., "It would be good because ...") and cons (i.e., "It would be bad because ...") of having her grandchildren spend their summer vacation at her home. For each of these slides, next to the pros and cons of making a decision, potential plans of implementation were also presented. These specified how the person would get started with the particular goal-directed actions (i.e., "If I decide to do it, then I will first ... and then ...!"); "If I decide to do it, then I won't ... before ...!"). A cued-recall test of this information was given following a distractor task; it provided participants with the pictures of the persons they had viewed and the stems of the sentences (as above) describing their thoughts. The deliberating participants, who had to view the slides and to recall the information depicted on the slides prior to making a decision about the two types of creativity tests, recalled pros and cons better than they recalled information on the when, where, and how of implementation. The recall performance of the planning participants, who had received and recalled the information after a decision on the creativity tests had been made, showed the reverse pattern.

All of these findings corroborate the cognitive-tuning hypothesis. But how do these differential recall performances observed in the last study (Gollwitzer et al., 1990, Study 2) come about? If one assumes that individuals' retrieval attempts necessitate constructing descriptions of what they are trying to retrieve (Norman and Bobrow, 1979), it seems possible that mindsets provide perspectives (Bobrow and Winograd, 1977) that allow the easy construction of specific descriptions. The deliberative mindset for instance should favour descriptions

phrased in terms of pros and cons, benefits and costs, and so forth. In other words, the deliberative mindset supports the ready construction of descriptions that specify desirability-related information, whereas the implemental mindset supports the construction of descriptions that specify implementation-related information. As Norman and Bobrow (1979) point out, quick construction of specific descriptions at the time of retrieval facilitate further successful retrieval. Norman and Bobrow also assume that whenever the description of the information sought matches the elaboration of the information at the time of encoding, recall performance is particularly enhanced. It seems possible, then, that deliberative and implemental mindsets favor congruent recall through both congruent elaboration at the time of encoding and ready construction of congruent descriptions at the time of retrieval.

Deliberative versus implemental mindsets and biased inferences

Deliberative and implemental mindsets are also postulated to differentially affect the way in which feasibility-related and desirability-related information is handled. In a deliberative mindset, information related to desirability should be analyzed impartially; in an implemental mindset, an analysis partial to the chosen goal is expected. Also, feasibility-related information is expected to be analyzed rather accurately in a deliberative mindset, whereas optimistic inferences that overestimate the actual feasibility of the chosen goal are expected in an implemental mindset.

Desirability-related information

With respect to testing the postulated impartial versus partial analysis of desirability-related information, a first study (reported by Taylor and Gollwitzer, 1995, Study 3) was conducted by asking participants to name either potential goals or chosen goals and

subsequently attempt to achieve clarity on the question of whether they should make an affirmative decision or had made the correct decision, respectively. Whereas the predecisional participants reported on positive and negative consequences with the same frequency, postdecisional participants failed to do so. The latter reported about five times more thoughts about pros than about cons, indicating a strong partiality in favor of the chosen goal in postdecisional participants.

Evidence for differences between the deliberative and implemental mindset in processing pros and cons is also provided by Harmon-Jones and Harmon-Jones (2002, Study 2). They tested the effects of mindsets on the postdecisional spreading of alternatives, a classic cognitive dissonance paradigm (Brehm and Cohen, 1962). Using this paradigm, dissonance researchers have found that after making a choice between two options, the chosen option becomes evaluated more positively whereas the nonchosen option becomes evaluated more negatively. Harmon-Jones and Harmon-Jones found that the implemental mindset increased postdecisional spreading of alternatives, whereas the deliberative mindset reduced it.

There is an important set of studies by Gagné and Lydon (2001a) suggesting that deliberation only then leads to an impartial analysis of pros and cons when deliberation is linked to the predecisional action phase. Deliberation over goal decisions that have already been made can initiate defensive processing of information that leads to even greater biasing. In one study, they asked participants involved in romantic relationships to deliberate a relationship or a nonrelationship goal decision. They found that when asked to rate how their partner compared with the average, those individuals asked to deliberate over a relationship goal decision gave much higher ratings than those who were asked to deliberate over a nonrelationship goal decision. Of interest, these ratings were also higher than those of implemental participants who had been planning the implementation of a relationship goal.

Gagné and Lydon (2001a) argue that the deliberation of a relationship goal may have been perceived as threatening, resulting in greater enhancement of the partner's attributes. In a second study, they measured the commitment participants had to their relationship and found that high-commitment but not low-commitment participants defended against the threat of a deliberative mindset by increasing their positive views of their partner. This pattern of findings indeed supports the assumption that deliberation may have threatened the participants' perceived ability to attain the goal of maintaining the relationship. In response, these individuals reasserted their commitment to the relationship by boosting the ratings of their partner.

Feasibility-related information

The hypothesized accurate analysis of feasibility-related information in the deliberative mindset, and the expected overly optimistic assessment in the implemental mindset, were observed in experiments by Gollwitzer and Kinney (1989) using the contingency-learning task designed by Alloy and Abramson (1979). In this task, participants are asked to determine to what degree they can influence the onset of a target light (outcome) by choosing to press or not to press a button (alternative actions). Participants commonly go through a series of trials (at least 40); the start of each trial is indicated by a warning light. By observing whether or not the target light comes on after they have pressed or not pressed the button, participants estimate how much control they have over the target light onset. The experimenter varies the actual control by manipulating the frequency of the target light onset associated with each of the two action alternatives (pressing or not pressing). The smaller the difference between these two frequencies, the less objective control participants have over the target light onset.

Nondepressed individuals commonly claim to possess control over target light onset that is noncontingent on their actions,

whenever the target light onset occurs frequently (e.g., in the “75/75” problem, where the target light comes on in 75 percent of pressing and 75 percent of nonpressing responses; see Alloy and Abramson, 1979). Gollwitzer and Kinney (1989, Study 2) asked deliberating, planning, and control participants to work on a contingency problem that presented frequent and noncontingent target light onset (i.e., the 75/75 problem). Participants were given the instruction to discover how to produce the target light onset. A set of 40 trials was offered, and participants were then asked to judge how much control they could exert over the target light onset.

Deliberating participants showed the most accurate judgment of control; their judgments of control were lower than those of either the control group or the planning group. The planning participants' judgments of control tended to be even higher than those of the control participants. The mindsets were created via the standard procedure described above. A mindset interpretation of these findings is supported by the additional observation that deliberating participants' judgments of control correlated negatively with the personal importance of the unresolved personal problems these participants were mulling over. Apparently, the more involved participants were in deliberating, the more realistic their subsequent judgments of control. A parallel finding was observed for planning participants, whose judgments of control were positively related to the participants' anticipated frustration in case they should fail to implement their chosen goals.

When Shelly Taylor asked me to collaborate on studies testing whether deliberative and implemental mindsets might even differentially affect the perceived controllability of events in everyday life, I was happy to assent. We observed (Taylor and Gollwitzer, 1995, Study 1) that the deliberative and implemental mindset indeed manage to affect people's judgments of the controllability of everyday risks (e.g., the risks involved being

in an automobile accident, becoming divorced, becoming depressed, developing a drinking problem, and being mugged). Participants were college students who had to judge these risks for themselves and for the average college student. Mindsets were induced via the standard procedure just before participants had to judge the named risks. Even though all participants perceived themselves as less vulnerable to these risks than the average college student, deliberating participants did this to a lesser degree than planning participants. This more pronounced illusion of invulnerability in the implemental mindset than in the deliberative mindset held no matter whether the critical events to be considered were of a more or less controllable (e.g., developing an addiction to prescription drugs, having a drinking problem) versus uncontrollable kind (e.g., developing diabetes, losing a partner to an early death). For both types of events, planning participants reported a higher invulnerability as compared to the average college student than deliberating participants did. The fact that deliberative and implemental mindsets even managed to modify the perceived vulnerability of rather uncontrollable events attests again to their enormous influence on the analysis of feasibility-related information.

We reasoned that assessing the feasibility of potential goals in the predecisional action phase not only requires that people accurately assess whether their actions could effectively control desired outcomes, they also need to know whether they are in the position to perform these instrumental actions. To answer this question, they have to assess correctly whether they possess the relevant aptitudes and skills. This implies that people in a deliberative mindset should show a relatively accurate evaluation of their personal attributes. Accordingly, we (Taylor and Gollwitzer, 1995, Study 2) also asked deliberating and planning participants to rate themselves on 21 qualities and skills (e.g., cheerfulness, athletic ability, writing ability, popularity, artistic ability) in comparison with the average college student of the same

age and gender. Even though all participants perceived themselves as more capable than the average college student, planning participants did so to a higher degree than deliberating participants. More recent research by Bayer and Gollwitzer (2005) suggests, however, that deliberative and implemental mindset effects on perceived ability may be moderated by people's relevant original low versus high self-views. The deliberative mindset helped in particular people with originally high self-views to arrive at modest ability appraisals (that foster the setting of realistic goals); and it was again in particular people with originally high self-views that were helped by the implemental mindset to arrive at optimistic ability appraisal (that foster the attainment of chosen goals).

A study by Puca (2001) also speaks to the biased analysis of feasibility-related information in the implemental as compared to the deliberative mindset. She studied realism versus optimism in terms of choosing test materials of different difficulty (Study 1) and predicting their own future task performance (Study 2). Deliberative participants preferred tasks of medium difficulty, whereas implemental participants opted for the too difficult tasks; also, implemental mindset participants overestimated their probability of success more than deliberative participants. Moreover, deliberative participants referred more than implemental participants to their past performance when selecting levels of difficulty or predicting future performance. Finally, when Gagné and Lydon (2001b) moved this biased inferences research to the real world by studying the issue of relationship predictions they found that individuals with a deliberative mindset were more accurate in their forecasts of survival of their romantic relationships than individuals with an implemental mindset. This effect was even more pronounced for long-term than for short-term relationship survival. Of most interest, participants with a deliberative mindset did not achieve this heightened accuracy by simply taking a pessimistic attitude.

Deliberative versus implemental mindsets and open-mindedness

Beyond differences in cognitive tuning and biased inferences, deliberative and implemental mindsets should also differ in openness to information. Task analysis of the demands of making a goal decision suggests that deliberative mindsets should be associated with enhanced receptivity to all sources and types of information. To make good decisions, one should be open to any available information that might potentially inform one's decision-making. One should be careful not to dismiss information prematurely as it may ultimately be useful or helpful in making good goal decisions. Implemental mindsets, in contrast, should be associated with more selective information processing. Once a goal is set, successful goal implementation requires more particular filtering of information, selectively processing goal-relevant stimuli while ignoring goal-irrelevant stimuli (e.g., Gollwitzer, 1990; Kuhl, 1984). For these reasons, the deliberative mindset should be associated with greater openness to information incidental to one's goals.

The early study by Heckhausen and Gollwitzer (1987, Study 2) reported above bears some relevance to this hypothesis. Participants were interrupted either while they were deliberating a choice between two different creativity tests (deliberative mindset), or just after having chosen one of them (implemental mindset), and verbally presented with lists of five to seven one-syllable nouns (e.g., house, art, and tree). Immediately after each list had been presented, participants had to recall the words in order. Participants' performance in this task was used to compute their working memory span (i.e., noun span), and results indicated that deliberative mindset participants evidenced a broader span (about half a word more) than implemental mindset participants.

The superior noun span by deliberative as compared with implemental mindset participants, however, only suggests that

deliberative mindset participants are more capable of storing information (i.e., they have a broader working memory span). Although broadened working memory suggests an enhanced capacity to process information, it does not directly address the hypothesis that deliberative mindsets, as compared with implemental mindsets, are associated with heightened processing of information that is incidental to one's goals. The information in the word lists used in the Heckhausen and Gollwitzer (1987) study cannot be considered incidental. Participants were explicitly asked to correctly reproduce as many words as possible of each presented word list. Moreover, broader working memory span by itself does not necessarily lead to more or less selective processing of incidental information.

Accordingly, Fujita et al. (2007) attempted a more critical test of the hypothesis that there are differences between the deliberative and implemental mindset in the selective processing of incidental information. In three experiments, participants in deliberative and implemental mindsets performed a primary task (*d2-concentration test*; Brickenkamp, 1981) while randomly presented incidental, unavoidable words. A subsequent unexpected recognition memory test assessed selective processing of these incidental words. In Study 1, they observed that participants in a deliberative mindset took less time than those in an implemental mindset to recognize whether or not they had previously been exposed to incidental words presented in a concentration task. In Studies 2 and 3, participants in a deliberative mindset had higher recognition accuracy of these words as compared with those in an implemental mindset. The results from all three studies indicate that deliberative individuals more easily accessed memory traces of information incidental to the ongoing task than implemental individuals. This occurred even when the mindsets induced were unrelated to the performance task that measured the cognitive differences (Studies 2 and 3). This "carryover" effect of mindsets suggests that

whereas the implemental mindset is more selective, the deliberative mindset is more open-minded to incidental information available in one's immediate environment. Study 3 also allowed clarifying whether the effect of mindset on selective information processing was due to enhanced open-mindedness in the deliberative mindset, enhanced closed-mindedness in the implemental mindset, or both. Results from this study suggested that the change in selective processing as a function of mindset is attributable to less selective filtering of incidental information in the deliberative mindset, as opposed to greater selective filtering in the implemental mindset.

Finally, the three studies taken together indicate that changes in selective processing as a function of mindset occur preconsciously. Preconscious cognitive processes are those that are initiated and do operate outside of conscious intent (Bargh, 1994). Researchers have argued that reactions to stimuli that require a response within 300 ms are not consciously controlled (e.g., Bargh and Chartrand, 2000; Greenwald and Banaji, 1995). Participants in all three studies were presented with incidental stimuli for only 300 ms prior to the primary performance task materials. Apparently, the act of deliberating a goal decision produces dramatic changes in the cognitive processing of information even when individuals do not intend such changes.

Deliberative versus implemental mindsets and behavior

Gollwitzer and Bayer (1999) pointed out that mindsets have been analyzed primarily in terms of their cognitive features, whereby the effects of these features on the control of behavior were ignored. As an exception, they reported a study by Pösl (1994), who found that participants in the implemental mindset were faster to initiate goal-directed behavior than those in the deliberative mindset. This was particularly true when participants experienced a behavioral conflict (i.e., whether

they had a choice to perform Behavior A or B, or only one of these), suggesting that the implemental mindset's closed mindedness allows participants to stay on track even in the case of behavioral conflict.

There is also evidence that the implemental mindset generates greater persistence in goal-directed behavior. Brandstätter and Frank (2002) found that participants in the implemental mindset persisted longer at an unsolvable puzzle task (Study 1) and at a self-paced computer task (Study 2). Similar to the findings of Pösl (1994), the impact of the implemental mindset on persistence was present in particular in situations of behavioral conflict. When both the perceived feasibility and desirability of the tasks were either uniformly high or low, persistence on the persistence tasks did not differ by mindset. However, when the perceived feasibility and desirability of the tasks were in opposite directions (i.e., one was high whereas the other was low) the implemental mindset participants persisted longer than did the deliberative mindset participants. Interestingly, the persistence in the implemental mindset was not executed insensitively or in a blind fashion. Brandstätter and Frank (2002, Study 3) obtained evidence that when a task is perceived as impossible or when persistence is not beneficial, individuals in the implemental mindset disengage much more quickly than individuals in the deliberative mindset.

Finally, Armor and Taylor (2003) report that an implemental mindset facilitates task performance (i.e., a scavenger hunt to be performed on campus) as compared with a deliberative mindset, and that this effect is mediated by the cognitive features of the implemental mindset (i.e., enhanced self-efficacy, optimistic outcome expectations, perceiving the task as easy). Such a study had been missing so far. Note that the Gollwitzer and Kinney (1989) study predicted cognitive changes (i.e., strong illusions of control) as a consequence of the implemental as compared with the deliberative mindset, because such illusions should benefit acting on the goal.

However, this inference had not been tested within one and the same study as was done by Armor and Taylor.

Recent research by Henderson et al. (2008) suggests that the beneficial effects of the implemental mindset on task performance might also be mediated by changes in respective attitude strength. The unambiguous, evaluative polarized or one-sided assessment of information in support of the chosen goal in the implemental mindset should foster the strength of the attitude toward the goal. Given the carryover properties of mindsets, Henderson et al. hypothesized that as people adopt an implemental mindset, they should experience an increase in attitude strength toward objects even if these are unrelated to their current goal pursuit. In a series of experiments they found support for this hypothesis. They observed that implemental mindset participants more than deliberative mindset participants adopt an extreme position towards an issue that is irrelevant to their goal concern. Moreover, implemental mindset participants evidenced lower levels of ambivalence toward a variety of unrelated objects than deliberative and neutral mindset participants. Implemental mindset participants were also characterized by more accessible evaluations of unrelated objects than deliberative and neutral mindset participants. And, importantly, implemental mindset participants showed a greater correspondence between their attitude and behavior than neutral mindset participants.

Finally, in order to investigate the process that is hypothesized to underlie the effects of an implemental mindset on attitude strength (i.e., one-sided focus on pros), Henderson et al. had implemental mindset participants either focus on the pros only or on both the pros and cons behind their decision. Critically, it was only the evaluative one-sided analysis of information that fostered attitude strength. As the authors had taken two groups who made a decision on how to act and only varied their evaluative focus (i.e., one group engaged in a one-sided evaluation of their

decision, whereas the other group engaged in a two-sided evaluation of their decision), the act of deciding itself does not seem to be sufficient to increase attitude strength. Otherwise, those individuals who made a decision on how to act and who analyzed both sides of their decision would have evidenced the same level of attitude strength as those who analyzed only one side of their decision.

Summary

Under the assumption that the course of goal pursuit presents itself to the individual as a series of consecutive tasks that need to be solved in order to promote goal attainment, the concept of mindset has been introduced. Mindset theory of action phases argues that becoming involved in these tasks leads to characteristic cognitive orientations (mindsets) that are beneficial for solving these tasks effectively, and the features of the cognitive orientations associated with the tasks of choosing between potential action goals (the deliberative mindset) and preparing the implementation of chosen goals (the implemental mindset) are spelled out. Various experiments tested the postulated characteristics of the deliberative and implemental mindsets. This research shows that the deliberative mindset is characterized by cognitive tuning toward desirability-related and feasibility-related thoughts and information, by an accurate analysis of feasibility-related information and an impartial analysis of desirability-related information, and, finally, by a heightened general receptivity to available information. The implemental mindset, on the other hand, is characterized by cognitive tuning toward implemental thoughts and information, by an overly optimistic analysis of feasibility-related information and a partial analysis of desirability-related information, and, finally, by a comparatively reduced receptivity (closed-mindedness) to available information.

IMPLICATIONS FOR OTHER THEORIES AND CONCEPTUAL DEBATES IN SOCIAL PSYCHOLOGY

Mindset theory of action phases has received much attention in other theories and conceptual debates in social psychology. These pertain to the optimism versus realism debate and to dual process theories. How the mindset theory of action phases can help to clarify issues in these areas of research will be discussed next.

Optimism versus realism

The results of mindset research on the processing of feasibility-related information pertain to the illusionary optimism versus realism controversy triggered by Taylor and Brown's (1988) article on positive illusions. Taylor and Brown proposed that mentally healthy people are not characterized by accurate assessments of their personal qualities, realistic estimates of personal control, and a realistic outlook on the future; instead, they maintain overly positive, self-aggrandizing perspectives of the self, the world, and the future. More specifically, mentally healthy people are said to be characterized by unrealistically positive self-perceptions, an illusion of a high degree of personal control, and unrealistic optimism about the future. Instead of being maladaptive, these positively distorted perceptions foster the criteria normally associated with mental health: positive regard, the ability to care for and about other people, and the ability to manage stress effectively (Taylor and Brown, 1988). Despite empirical support for the model (Taylor and Armor, 1996; Taylor et al., 2000), this portrait of the healthy person raises a disturbing question: if healthy people's perceptions are marked by positive bias, how do they effectively identify and make use of negative feedback they may encounter? If people are capable of explaining away, compartmentalizing, or otherwise dismissing or minimizing negative feedback, as Taylor and Brown

(1988) suggested, these self-serving illusions that bolster self-esteem and produce a positive mood in the short run may ultimately set people up for long-term disappointment and failure as they fail to incorporate negative feedback into their goal setting and planning (Colvin and Block, 1994; Weinstein, 1984).

The mindset research on illusion of control offers the following insights to the debate about positive illusions versus realism. First, neither realism nor positive illusions seem adaptive in general to a person's psychological functioning. Realistic thinking seems functional when it comes to making goal decisions, whereas positive illusions seem functional when the chosen goals are to be implemented. Second, people can easily open the window to realism provided by the deliberative mindset. People do not have to go through the effortful mental exercises we have induced in our experiments to create a deliberative mindset; simply trying to achieve clarity in regard to an unresolved personal problem will trigger an intensive deliberation of pros and cons (Taylor and Gollwitzer, 1995, Study 3). Third, postdecisional individuals who plan the implementation of a chosen goal seem to be protected from an accurate analysis of feasibility-related information and thus can benefit from illusionary optimism that makes them strive harder to reach their goals, especially in the face of hindrances and barriers. It appears, then, that the individual's cognitive apparatus readily adjusts to the various demands of the control of action: choosing between action goals leads to realism, and implementing chosen goals leads to positive illusions.

Dual process theories

The ideas and research originating within the framework of the notion of deliberative and implemental mindsets seem to constitute a dual process theory in the realm of goal pursuit. The approach taken is to juxtapose a cognitive orientation that is functional to

choosing goals with a cognitive orientation that is functional to the implementation of chosen goals. In other words, the ideal information-processing styles for solving two different tasks that serve one end (i.e., the effective control of action) are analyzed in contrast to each other. This is different from those dual process models that compare two different styles of information processing in the service of one and the same task, such as perceiving another person (Bargh, 1984; Brewer, 1988; Fiske and Neuberg, 1990), making attributions (Gilbert, 1989; Gilbert and Malone, 1995), or forming attitudes (Chaiken et al., 1989; Fazio, 1990). The approach taken in those models is to analyze how the two forms of information processing delineated differ in meeting the task at hand as is also true for those dual process lines of research that explicitly adopt the mindset notion for conceptual and methodological reasons, such as research on the counterfactual mindset (e.g., Galinsky and Moskowitz, 2000; Wong et al., 2009) and near versus distal mental construal (e.g., Liberman and Trope, 2008; Freitas et al., 2004). Counterfactual mindset research induces a counterfactual mindset by having participants read one and the same behavioral episode experienced by another person that is known to trigger counterfactual musings ("If only he had done ...," or "What if he had done ..."), and then checks whether reading as compared to not reading this episode improves performance in subsequent classic problem solving, creativity, or negotiation tasks. Mental construal research, on the other hand, studies how inducing psychologically distant versus near construals of certain events affects the perception, categorization, judgmental inferences, evaluations, and behaviors with respect to these events as well as related events. Note that the attempted gain in knowledge of these two approaches refers to how a given cognitive orientation (counterfactual, near versus distal mental construal) affects respective cognitive, affective, and behavioral performances. With respect to the mindset theory of action phases,

however, the attempted gain of knowledge refers to the typical characteristics of deliberating versus planning in terms of the underlying cognitive procedures.

However, there are some similarities between the deliberative versus implemental mindset distinction and those dual process notions that are construed as stage theories (e.g., Gilbert and Malone, 1995). People in everyday life should experience deliberative mindsets prior to implemental mindsets as people mostly prefer to make plans on how to achieve a goal only after they have made a binding goal choice. In this temporal sense therefore the deliberative versus implemental mindset model qualifies as a stage model. This is not true, however, with respect to the quality of cognitive processes associated with the two mindsets. For instance, in the two-step model of the attribution process – a stage model suggested by Gilbert (1989; Gilbert and Malone, 1995) – the first step is simple and automatic (i.e., a quick personal attribution), whereas the second step requires attention, thought, and effort (i.e., adjusting that inference to account for situational influences). The notion of deliberative versus implemental mindsets, on the other hand, does not assume that the deliberative mindset is associated with more rudimentary cognitive processes than the implemental mindset (or vice versa). In both the deliberative mindset and the implemental mindset, highly complex cognitive procedures are activated that determine the individual's cognitive and behavioral functioning. Moreover, in both deliberative and implemental mindsets these procedures can, but do not need to, reach consciousness to unfold their effects, and their effects can, but do not have to, be detected by the individual (e.g., the illusion of control in the implemental mindset).

Also, mindset theory sees the deliberative and implemental mindsets as distinct and independent of each other. Whereas it is assumed in some dual process theories (e.g., Chaiken et al., 1989) that the postulated modes of information processing can operate at the same time, the deliberative and

implemental mindsets are assumed to preclude each other. This is because the strength with which the cognitive procedures associated with the deliberative mindset are activated is positively related to the degree of involvement with the task of choosing between potential goals, whereas the strength with which the cognitive procedures associated with the implemental mindset are activated is positively related to the degree of involvement with the task of planning the implementation of a chosen goal. Because a person cannot become intensely involved in both of these tasks at one and the same time but only successively, pronounced deliberative and implemental mindsets cannot coexist. They also do not affect each other in the sense that a preceding strong deliberative mindset makes for a strong succeeding implemental mindset; it all depends on how intensely people become involved with solving the task of choosing between potential goals and with planning the implementation of a chosen goal, respectively.

Summary

The deliberative versus implemental mindset distinction had an impact on various theoretical discussions in social psychology. First, with respect to the realism versus optimism discussion, it made clear that people are capable of flexibly adopting the type of orientation that is demanded by the task at hand. When decisions between goals are to be made (e.g., whether to go to college at home or abroad), getting involved with reflections on which goal to choose spurs realism that in turn allows making the more appropriate (feasible) choice. And if a chosen goal (e.g., studying abroad) is to be implemented, getting involved with planning out the course of goal realization spurs optimism that promotes the necessary persistence for goal attainment. Second, with respect to classic dual process theories, the mindset theory of action phases points out that modes of information processing may not only be studied in

terms of what determines taking one or the other mode (e.g., heuristic versus systematic information processing is found to be affected by time pressure, likeability of the source, strong versus weak arguments), and what effects taking one or the other mode has on grasping the information at hand (e.g., leads to more or less respective attitude change). Mindset theory of action phases suggests that the mere involvement with one or the other type of reasoning task (deliberating the choice of potential goals versus planning the implementation of a chosen goal) already activates different task-facilitating cognitive procedures that affect the processing of both task-relevant and task-irrelevant information in a unique way.

THE APPLIED IMPACT OF THE MINDSET THEORY OF ACTION PHASES

Mindset theory of action phases had set aside the question of what makes for good planning of goal implementation. So when I was invited to create a research unit on “Intention and Action” at the Max Planck Institute for Psychological Research at Munich in 1989, this question became our central concern and has stayed with me ever since. As we had induced the implemental mindset by asking research participants to list a series of steps toward goal attainment and then specify for each individual step exactly when, where, and how one wants to realize it, we wondered whether people could facilitate goal striving by planning out goal-directed action in this fashion. We referred to this type of planning as forming *implementation intentions* (Gollwitzer, 1993, 1999). Whereas goals (or goal intentions) merely specify desired end states (“I want to achieve goal X!”), implementation intentions in the format of “If situation Y arises, then I will initiate behavior Z!” additionally specify when, where, and how a person intends to strive for the goal. Implementation intentions thus

delegate control over the initiation of the intended goal-directed behavior to a specified situational cue by creating a strong mental link between this cue and a goal-directed response. For example, a person who has chosen the goal to eat more healthily can form the implementation intention, “When I’m at my favorite restaurant and the waiter asks me for my order, then I’ll request a vegetarian meal!” The mental links created by implementation intentions were assumed (Gollwitzer, 1999) to facilitate goal attainment on the basis of psychological processes that relate to both the anticipated situation (i.e., enhanced activation of the mental representation of the situation specified in the if-part of the plan) and the intended behavior (i.e., automatic initiation of the response specified in the then-part of the plan once the critical situation is encountered).

I was fortunate to get to know Paschal Sheeran in the mid 1990s. His students and colleagues in England as well as mine in Germany set out to test the claimed positive effects of implementation intentions on goal attainment as well as the assumed underlying processes. Because forming an implementation intention implies the selection of a critical future situation, the mental representation of this situation becomes highly activated and hence more accessible. This heightened accessibility of the “if” part of the plan has been observed in numerous studies; it helps people to easily recall the specified situation and it leads to swift allocation of attention when the situation arises. Various studies also observed that the initiation of the goal-directed response specified in the then-component of an implementation intention exhibits features of automaticity; if-then planners are found to act quickly in the face of the critical situation, they deal effectively with cognitive demands, and do not need to consciously intend to act in the critical moment.

A meta-analysis involving over 8,000 participants in 94 independent studies revealed a medium-to-large effect size of implementation intentions on goal achievement (Gollwitzer and Sheeran, 2006).

Importantly, implementation intentions were found to benefit goal-directed responses no matter whether these were cognitive, emotional, or behavioral in nature, and this held true in all kinds of goal domains (e.g., consumer, environmental, antiracist, prosocial, academic, and health). Apparently, implementation intentions help people to better cope with the major problems of goal striving: getting started, staying on track, calling a halt to a futile goal striving, and not overextending oneself. Importantly, implementation intentions still show their beneficial effects when the going gets tough (Gollwitzer et al., 2010; Gollwitzer & Oettingen, 2011); that is, when goal striving is limited by conditions that are very resistant to change by self-regulatory strategies (e.g., a low level of competence, a fierce competitor, or strong competing habitual responses).

A recent fMRI study conducted by Gilbert et al. (2009) provides an answer to the puzzling power of implementation intentions. Brain activity in the lateral area 10 was observed to move toward the medial area 10 when participants switched from performing an executive-function task by the guidance of a goal intention to performing the very same type of task by the guidance of an implementation intention. On the basis of an extensive meta-analysis on various executive-function tasks it is known that lateral and medial area 10 are implicated in top-down and bottom-up action control, respectively (Burgess et al., 2005). Apparently, implementation intentions induce a switch in action control from top-down to bottom-up control of action. This explains why even habitual responses can be broken by implementation intentions, and why special populations that are known to suffer from ineffective conscious control of their thoughts, feelings, and actions are also found to benefit from forming implementation intentions (e.g., heroin addicts during withdrawal, schizophrenic patients, children with attention deficit hyperactivity disorder [ADHD]).

But will the discovery of an effective self-regulation strategy of goal striving and the

in-depth analysis of how it works make people use it to solve personal, interpersonal, and even societal problems? Simply disseminating the good news of the existence of a powerful strategy for goal implementation may not suffice. It needs the second step of developing an intervention that facilitates the acquisition of this strategy so that people can use it on their own in everyday life. In other words, one needs to develop an intervention that effectively teaches the forming of implementation intentions as a meta-cognitive strategy.

To come up with such an effective intervention, Gabriele Oettingen and I first considered all of the studies that had analyzed potential moderators of implementation effects (see Gollwitzer and Sheeran, 2006). Importantly, implementation intentions only then unfold their beneficial effects if strong goal commitments are in place. Moreover, implementation intention effects seem to be stronger when people put exactly those cues in the if-part of an implementation intention which they personally consider to be most critical (Adriaanse et al., 2009). Accordingly, we looked for a procedure that established these prerequisites, and we found it in *mental contrasting* as this self-regulation strategy is also known to motivate if-then planning.

What is mental contrasting and how does it work? When people mentally contrast (Oettingen et al., 2001), they first imagine a desired future (e.g., to improve one's health behavior), and then reflect on the present reality that stands in the way of reaching this desired future (e.g., feeling the urge to give in to a temptation). Thereby, mental contrasting turns desired futures that are perceived as feasible into strong goal commitments. The beneficial effects of mental contrasting evince in various domains (achievement, interpersonal, and health), for cognitive (e.g., making plans), affective (e.g., feelings of anticipated disappointment in case of failure), motivational (e.g., feelings of energization, systolic blood pressure), and behavioral indicators of goal commitment (e.g., invested effort and

actual achievement such as obtained course grades), directly after or weeks later (summary by Oettingen and Stephens, 2009). Mental contrasting facilitates strong goal commitments by changing implicit cognition. The increased strength of association between future and reality induced by mental contrasting has been found to mediate the cognitive, emotional, and behavioral indicators of strong goal commitment, measured by self-report and other-rated performance (Kappes and Oettingen, 2011).

We therefore integrated the forming of implementation intentions with mental contrasting into one self-regulation strategy called Mental Contrasting with Implementation Intentions (MCII; Oettingen and Gollwitzer, 2010). When taught as a meta-cognitive strategy that is then applied by participants in everyday life, it supports behavior change more than mental contrasting and implementation intentions used in isolation; moreover, MCII effects hold for various life domains (academic achievement, interpersonal relations, romantic satisfaction, and health, eating and regular exercise), and over time periods of up to two years (Adriaanse et al., 2010; Christiansen et al., 2010; Stadler et al., 2009, 2010). When engaging in MCII, participants first go through a mental contrasting exercise to create strong goal commitment and to identify the obstacles that truly stand in the way of goal attainment. Implementation intentions (if-then plans) are then formed to help translate the goal commitment into instrumental behavior by putting the obstacles identified in mental contrasting as critical cues in the if-part of the plan and linking it to an instrumental coping response specified in the then-part.

CONCLUSION

There are times when people need to make decisions, and there are times when the decisions made have to be implemented.

From the perspective of effective action control in everyday life, then, it seems helpful to activate the respective cognitive procedures that facilitate goal setting and goal implementation when making decisions versus acting on them is at issue. In other words, people should allow for and become involved in deliberative or implemental mindsets, depending on whether a goal decision or the implementation of a chosen goal is called for.

Moreover, there exist powerful strategies of pre- and postdecisional reasoning that are more effective than others. Accordingly, interventions geared at helping people to maximize their goal setting and goal striving should not merely confront people with the tasks of committing to goals and implementing the chosen goals. Rather, they should go one step further and equip people with those goal setting (e.g., mental contrasting) and goal implementation strategies (e.g., forming implementation intentions) that are known to be most effective in promoting appropriate goal commitments and successful goal attainment, respectively. Future research might explore how such strategies are taught best in time- and cost-effective behavior change interventions.

REFERENCES

Ach, N. (1935) Analyse des Willens (The analysis of willing). In E. Abderhalden (ed.), *Handbuch der biologischen Arbeitsmethoden*, Vol. 6. Berlin: Urban u. Schwarzenberg.

Adriaanse, M.A., de Ridder, D.T.D. and de Wit, J.B.F. (2009) Finding the critical cue: Implementation intentions to change one's diet work best when tailored to personally relevant reasons for unhealthy eating. *Personality and Social Psychology Bulletin*, 35, 60–71.

Adriaanse, M.A., Oettingen, G., Gollwitzer, P.M., Hennes, E.P., de Ridder, D.T.D. and de Wit, J.B.F. (2010) When planning is not enough: Fighting unhealthy snacking habits by Mental Contrasting with Implementation Intentions (MCII). *European Journal of Social Psychology*, 40, 1277–1293.

Alloy, L.B. and Abramson, L.Y. (1979) Judgement of contingency in depressed and nondepressed students: Sadder but wiser? *Journal of Experimental Psychology: General*, 108, 449–485.

Armor, D.A. and Taylor, S.E. (2003) The effects of mindset on behavior: Self-regulation in deliberative and implemental frames of mind. *Personality and Social Psychology Bulletin*, 29, 86–95.

Atkinson, J.W. (1957) Motivational determinants of risk-taking behavior. *Psychological Review*, 64, 359–372.

Bargh, J.A. (1984) Automatic and conscious processing of social information. In R.S. Wyer, Jr. and T.K. Srull (eds), *Handbook of Social Cognition*, 3, 1–43. Hillsdale, NJ: Erlbaum.

Bargh, J.A. (1994) The Four Horsemen of automaticity: Awareness, efficiency, intention, and control in social cognition. In R.S. Wyer Jr. and T.K. Srull (eds), *Handbook of Social Cognition*, 2nd Edition, pp. 1–40. Hillsdale, NJ: Erlbaum.

Bargh, J.A. and Chartrand, T.L. (2000) A practical guide to priming and automaticity research. In H. Reis and C. Judd (eds), *Handbook of Research Methods in Social Psychology*, pp. 253–285. New York: Cambridge University Press.

Bayer, U.C. and Gollwitzer, P.M. (2005) Mindset effects on information search in self-evaluation. *European Journal of Social Psychology*, 35, 313–327.

Bobrow, D.G. and Winograd, T. (1977) An overview of KRL: A knowledge representation language. *Cognitive Science*, 1, 3–46.

Brandstätter, V. and Frank, E. (2002) Effects of deliberative and implemental mindsets on persistence in goal-directed behavior. *Personality and Social Psychology Bulletin*, 28, 1366–1378.

Brehm, J.W. and Cohen, A.R. (1962) *Explorations in Cognitive Dissonance*. New York: Wiley.

Brewer, M.B. (1988) A dual process model of impression formation. In T.K. Srull and R.S. Wyer (eds), *Advances in Social Cognition*, 1, 1–36. Hillsdale, NJ: Erlbaum.

Brickenkamp, R. (1981) *Test d2*, 7th Edition. Goettingen: Hogrefe.

Burgess, P.W., Simons, J.S., Dumontheil, I. and Gilbert, S.J. (2005) The gateway hypothesis of rostral PFC function. In J. Duncan, L. Phillips, and P. McLeod (eds) *Measuring the Mind: Speed Control and Age*, pp. 215–246. Oxford: Oxford University Press.

Chaiken, S., Liberman, A. and Eagly, A.H. (1989) Heuristic and systematic information processing within and beyond the persuasion context. In J.S. Uleman and J.A. Bargh (eds), *Unintended Thought*, pp. 212–252. New York: Guilford Press.

Christiansen, S., Oettingen, G., Dahme, B. and Klinger, R. (2010) A short goal-pursuit intervention to improve physical capacity: A randomized clinical trial in chronic back pain patients. *Pain*, 149, 444–452.

Colvin, C.R. and Block, J. (1994) Do positive illusions foster mental health?: An examination of the Taylor and Brown formulation. *Psychological Bulletin*, 116, 3–20.

Fazio, R.H. (1990) Multiple processes by which attitudes guide behavior: The MODE model as an integrative framework. In M.P. Zanna (ed.), *Advances in Experimental Social Psychology*, 23, 75–109. San Diego: Academic Press.

Feather, N.T. and Newton, J.W. (1982) Values, expectations, and the prediction of social action: An expectancy-valence analysis. *Motivation and Emotion*, 6, 217–244.

Fiske, S.T. and Neuberg, S.L. (1990) A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. *Advances in Experimental Social Psychology*, 23, 1–73.

Freitas, A.L., Gollwitzer, P.M. and Trope, Y. (2004) The influence of abstract and concrete mindsets on anticipating and guiding others' self-regulatory efforts. *Journal of Experimental Social Psychology*, 40, 739–752.

Fujita, K., Gollwitzer, P.M. and Oettingen, G. (2007) Mindsets and preconscious open-mindedness to incidental information. *Journal of Experimental Social Psychology*, 43, 48–61.

Gagné, F.M. and Lydon, J.E. (2001a) Mindset and relationship illusions: The moderating effects of domain specificity and relationship commitment. *Personality and Social Psychology Bulletin*, 27, 1144–1155.

Gagné, F.M. and Lydon, J.E. (2001b) Mindset and close relationships: When bias leads to (in)accurate predictions. *Journal of Personality and Social Psychology*, 81, 85–96.

Galinsky, A.D. and Moskowitz, G.B. (2000) Counterfactuals as behavioral primes: Priming the simulation heuristic and consideration of alternatives. *Journal of Experimental Social Psychology*, 36, 384–409.

Gilbert, D.T. (1989) Thinking lightly about others: Automatic components of the social inference process. In J.S. Uleman and J.A. Bargh (eds), *Unintended Thought*, pp. 189–211. New York: Guilford Press.

Gilbert, D.T. and Malone, P.S. (1995) The correspondence bias. *Psychological Bulletin*, 117, 21–38.

Gilbert, S.J., Gollwitzer, P.M., Cohen, A-L., Oettingen, G. and Burgess, P.W. (2009) Separable brain systems supporting cued versus self-initiated realization of delayed intentions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35, 905–915.

Gollwitzer, P.M. (1987) *Motivationale vs. volitionale Bewußtseinslage*. (Motivational vs. volitional mindset.) *Habilitationsschrift*, Ludwig-Maximilians-Universität München.

Gollwitzer, P.M. (1990) Action phases and mindsets. In E.T. Higgins and R.M. Sorrentino (eds), *Handbook of Motivation and Cognition: Foundations of Social Behavior*, 2, 53–92. New York: Guilford Press.

Gollwitzer, P.M. (1991) Abwägen und Planen: Bewußtseinslagen in verschiedenen Handlungsphasen (Deliberating and planning: Mindsets in different action phases). Goettingen: Hogrefe.

Gollwitzer, P.M. (1993) Goal achievement: The role of intentions. *European Review of Social Psychology*, 4, 141–185.

Gollwitzer, P.M. (1999) Implementation intentions: Strong effects of simple plans. *American Psychologist*, 54, 493–503.

Gollwitzer, P.M. and Bayer, U.C. (1999) Deliberative and implemental mindsets in the control of action. In S. Chaiken and Y. Trope (eds), *Dual-process Theories in Social Psychology*, pp. 403–422. New York: Guilford Press.

Gollwitzer, P.M., Gawrilow, C. and Oettingen, G. (2010) The power of planning: Effective self-regulation of goal striving. In R. Hassin, K. Ochsner, and Y. Trope (eds), *Self-control in Society, Mind, and Brain*, pp. 279–296. New York: Oxford University Press.

Gollwitzer, P.M., Heckhausen, H. and Steller, B. (1990) Deliberative and implemental mindsets: Cognitive tuning toward congruous thoughts and information. *Journal of Personality and Social Psychology*, 59, 1119–1127.

Gollwitzer, P.M. and Kinney, R.F. (1989) Effects of deliberative and implemental mindsets on illusion of control. *Journal of Personality and Social Psychology*, 56, 531–542.

Gollwitzer, P.M. and Kirchhof, O. (1998) The willful pursuit of identity. In J. Heckhausen and C.S. Dweck (eds), *Life-span Perspectives on Motivation and Control*, pp. 389–423. New York: Cambridge University Press.

Gollwitzer, P.M. and Oettingen, G. (2011) Planning promotes goal striving. In K.D. Vohs and R.F. Baumeister (eds), *Handbook of Self-regulation: Research, Theory, and Applications*, 2nd Edition, pp. 162–185. New York: Guilford Press.

Gollwitzer, P.M. and Sheeran, P. (2006) Implementation intentions and goal achievement: A meta-analysis of effects and processes. *Advances in Experimental Social Psychology*, 38, 69–119.

Gollwitzer, P.M., Sheeran, P., Michalski, V. and Seifert, A.E. (2009) When intentions go public: Does social reality widen the intention-behavior gap? *Psychological Science*, 20, 612–618.

Greenwald, A.G. and Banaji, M.R. (1995) Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, 102, 4–27.

Harmon-Jones, C., Schmeichel, B.J. and Harmon-Jones, E. (2009) Symbolic self-completion in academia: Evidence from department web pages and email signature files. *European Journal of Experimental Social Psychology*, 39, 311–316.

Harmon-Jones, E. and Harmon-Jones, C. (2002) Testing the action-based model of cognitive dissonance: The effect of action orientation on postdecisional attitudes. *Personality and Social Psychology Bulletin*, 28, 711–723.

Heckhausen, H. (1977) Achievement motivation and its constructs: A cognitive model. *Motivation and Emotion*, 4, 283–329.

Heckhausen, H. (1987) Wünschen – Wählen – Wollen. In H. Heckhausen, P.M. Gollwitzer, and F.E. Weinert (eds), *Jenseits des Rubikon. Der Wille in den Humanwissenschaften* (Crossing the Rubicon: The concept of will in the life sciences), pp. 3–9. Berlin: Springer Verlag.

Heckhausen, H. and Gollwitzer, P.M. (1987) Thought contents and cognitive functioning in motivational versus volitional states of mind. *Motivation and Emotion*, 11, 101–120.

Henderson, M.D., de Liver, Y. and Gollwitzer, P.M. (2008) The effects of implemental mindset on attitude strength. *Journal of Personality and Social Psychology*, 94, 394–411.

Higgins, E.T. and Sorrentino, R.M. (1990) (eds) *The Handbook of Motivation and Cognition: Foundations of Social Behavior*, Vol. 2. New York: Guilford Press.

Kappes, A. and Oettingen, G. (2011) *From wishes to goals: Mental contrasting connects future and reality*. Submitted for publication.

Külpe, O. (1904) Versuche über Abstraktion (Experiments on abstraction). Bericht über den 1. Kongreß für experimentelle Psychologie, pp. 56–68. Leipzig: Barth.

Kuhl, J. (1984) Volitional aspects of achievement motivation and learned helplessness: Toward a

comprehensive theory of action control. In B.A. Maher (ed.), *Progress in Experimental Personality Research*, 13, 99–171. New York: Academic Press.

Ledgerwood, A., Liviatan, I. and Carnevale, P.J. (2007) Group-identity completion and the symbolic value of property. *Psychological Science*, 18, 873–878.

Lewin, K. (1926) Vorsatz, Wille und Bedürfnis (Intention, will, and need). *Psychologische Forschung*, 7, 330–385.

Liberman, N. and Trope, Y. (2008) The psychology of transcending the here and now. *Science*, 322, 1201–1205.

Marbe, K. (1915) Zur Psychologie des Denkens (The psychology of reasoning). *Fortschritte der Psychologie und ihre Anwendungen*, 3, 1–42.

Mischel, W. (1974) Processes in delay of gratification. *Advances in Experimental Social Psychology*, 7, 249–292.

Mischel, W. and Patterson, C.J. (1978) Effective plans for self-control in children. *Minnesota Symposia on Child Psychology*, 11, 199–230.

Norman, D.A. and Bobrow, D.G. (1979) Descriptions: An intermediate stage in memory retrieval. *Cognitive Psychology*, 11, 107–123.

Oettingen, G., Pak, H. and Schnetter, K. (2001) Self-regulation of goal-setting: Turning free fantasies about the future into binding goals. *Journal of Personality and Social Psychology*, 80, 736.

Oettingen, G. and Gollwitzer, P.M. (2010) Strategies of setting and implementing goals: Mental contrasting and implementation intentions. In J.E. Maddux and J.P. Tangney (eds), *Social Psychological Foundations of Clinical Psychology*, pp. 114–135. New York: Guilford Press.

Oettingen, G. and Stephens, E.J. (2009) Fantasies and motivationally intelligent goal setting. In G.B. Moskowitz and H. Grant (eds), *The Psychology of Goals*, pp. 153–178. New York: Guilford Press.

Orth, J. (1903) Gefühl und Bewußtseinslage: Eine kritisch-experimentelle Studie (Feelings and mindsets: An experimental study). In T. Ziegler and T. Ziehen (eds), *Sammlung von Abhandlungen aus dem Gebiet der Pädagogischen Psychologie und Physiologie*, 4, 225–353. Berlin: Verlag von Reuter and Reichard.

Pösl, I. (1994) Wiederaufnahme unterbrochener Handlungen: Effekte der Bewusstseinslagen des Abwägens und Planens (Mindset effects on the resumption of disrupted activities). Unpublished master's thesis, University of Munich, Germany.

Puca, R.M. (2001) Preferred difficulty and subjective probability in different action phases. *Motivation and Emotion*, 25, 307–326.

Stadler, G., Oettingen, G. and Gollwitzer, P.M. (2009) Physical activity in women. Effects of a self-regulation intervention. *American Journal of Preventive Medicine*, 36, 29–34.

Stadler, G., Oettingen, G. and Gollwitzer, P.M. (2010) Intervention effects of information and self-regulation on eating fruits and vegetables over two years. *Health Psychology*, 29, 274–283.

Taylor, S.E. and Armor, D.A. (1996) Positive illusions and coping with adversity. *Journal of Personality*, 64, 873–898.

Taylor, S.E. and Brown, J.D. (1988) Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, 103, 193–210.

Taylor, S.E. and Gollwitzer, P.M. (1995) The effects of mindsets on positive illusions. *Journal of Personality and Social Psychology*, 69, 213–226.

Taylor, S.E., Kemeny, M.E., Reed, G.M., Bower, J.E. and Gruenewald, T.L. (2000) Psychological resources, positive illusions, and health. *American Psychologist*, 55, 99–109.

Watt, H.J. (1905) Experimentelle Beiträge zu einer Theorie des Denkens (Experimental analyses of a theory of reasoning). *Archiv für die gesamte Psychologie*, 4, 289–436.

Weinstein, N.D. (1984) Why it won't happen to me: Perception of risk factors and susceptibility. *Health Psychology*, 3, 431–457.

Wicklund, R.A. and Gollwitzer, P.M. (1982) *Symbolic Self-completion*. Hillsdale, NJ: Erlbaum.

Wong, E.M., Galinsky, A.D. and Kray, L.J. (2009) The counterfactual mindset: A decade of research. In K.D. Markman, W.M.P. Klein, and J.A. Suhr (eds), *Handbook of Imagination and Mental Simulation*, pp. 161–174. New York: Psychology Press.

Name Index

AABT *see* Associates for the Advancement and Behavior Therapy

Aaker, J.L. 500

Aarts, H. 90, 452-3

Abelson, R.P. 182, 184-5, 279, 302, 317

Abramson, L.Y. 217, 532-3

Ackerman, J.M. 23-4

Adams, N.E. 447

Adaval, R. 165, 304

Addis, M.E. 122

Aderman, D. 189, 510

Adriaanse, M.A. 541-2

Agnoli, F. 316

Ahrensdorf, P.J. 401

Ainslie, G. 99, 119, 122

Ajzen, I. 93, 107, 439-40, 444-6, 449-52, 454-5

Albaracín, D. 451

Alcock, J. 12

Alexander, M.J. 498

Alfieri, T. 206-7

Alicke, M.D. 186-7

Allan, J.L. 109

Allen, K. 40, 193, 197, 461, 478

Allen, V.L. 193, 461

Allen-Arave, W. 15

Allik, J. 186-7

Alloy, L.B. 532-3

Allport, G.W. 1, 224, 247, 257, 268, 284

Altemus, M.P. 40

Amarel, D. 44

American Psychological Association (APA) 356, 497

Anckarsäter, H. 42

Andersen, S.M. 93, 214, 272

Anderson, J.R. 164, 173, 209-10, 252

Anderson, N.H. 157-8, 179, 269-71

Anderson, R.C. 252

Antaki, C. 152

APA *see* American Psychological Association

Arcuri, L. 313

Aristotle 185, 197

Armitage, C.J. 441, 444, 448-9, 451

Armor, D.A. 536-7

Arndt, J. 404, 406-11

Aronson, E. 180, 185, 194, 379, 382, 384-6, 392

Arrowood, A.J. 464, 467

Asch, S.E. 118, 191, 193-4, 271-2, 279, 311

Asendorpf, J.B. 99

Ashby, W.R. 506

Associates for the Advancement and Behavior Therapy (AABT) 356

Assor, A. 422

Atkinson, J.W. 135-6, 138-9, 143-4, 489, 492, 510, 527

Au, T.K. 317

Axsom, D. 255, 392

Baard, P.P. 421, 424, 432

Baddeley, A.D. 106

Baldwin, J. 401

Baltes, P.B. 511

Bamber, S. 454

Banaji, M.R. 102-3, 170, 283, 316, 453, 535

Bandura, A. 6, 107, 349-62, 364-5, 367-70, 439, 446-7, 453, 506, 510

Bao, X. 425

Bar, R. 212, 217

Bar-Anan, Y. 124-5

Bar-Tal, D. 216

Barden, J. 234

Bargh, J.A. 14, 83-5, 88-90, 92, 100, 102-4, 170-2, 252, 440, 452-4, 487, 494, 530, 535, 538

Barndollar, K. 452

Baron, R. 229

Barrett, L.F. 27, 296

Barrowclough, C. 152

Barsalou, L.W. 298

Bartlett, F.C. 76, 210

Bassili, J.N. 491

Baumeister, R.F. 14, 64, 103, 108, 186, 190, 192, 196-7, 257, 335, 337, 339, 399, 421, 427, 506, 520

Baxter, S.D. 298

Bayer, U.C. 534-5

Beale, D.A. 454

Beardslee, D. 454

Beauvois, J. 388

Becerra, A. 237

Beck, A.T. 218, 340

Becker, D.V. 23, 401, 408, 410

Becker, E. 399, 401, 408, 410

Beer, J. 190

Beavers, C.G. 519

Bem, D.J. 104, 232, 239, 250, 418

Benesh-Weiner, M. 184

Bentham, J. 489

Benware, C. 431

Bergsieker, H.B. 281

Berkman, L.F. 41

Berkowitz, L. 354

Bernberg, R.E. 444

Bernton, G.G. 51, 55, 57, 59-60, 62, 65-7

Berscheid, E. 180

Betley, G. 417-18

Bianco, A.T. 497

Biernat, M. 473

Bilsky, W. 509

Bin Kader, A.H. 217

Birch, D. 136

Blankenship, K.L. 236

Blanton, H. 102, 476

Bless, H. 100, 111, 235, 292-3, 297-8, 300, 302-4

Block, J. 98, 538

Blumer, H. 444

Bobak, M. 43

Bobrow, D.G. 531

Bodenhausen, G.V. 209, 236, 254, 295, 304, 472

Boffey, P. 355

Bogart, L. 354

Bohner, G. 169

Bohns, V.K. 498

Bolger, N. 44

Bollnow, O.F. 290, 293-4

Bond, R.N. 193

Borgida, E. 282-4

Boudreau, L.A. 215

Bourgeois, M.J. 26

Bower, G.H. 290-1

Bowlby, J. 37

Bradley, E. 153

Bradley, M.M. 55

Braithwaite, J. 351

Brandstätter, V. 453, 536

Branscombe, N.R. 209

Brehm, J. 57, 194, 231, 379, 382, 399, 507, 510, 517, 532

Brekke, N. 295

Brendl, C.M. 104

Brendl, M. 514

Brennan, P.L. 69

Brewer, M. 494-5

Brewer, N. 259

Brewer, W.F. 252, 439

Brewin, C.R. 152

Brickenkamp, R. 535

Brückman, P. 466-7

Brñol, P. 224, 231, 233, 235, 237-41

Britton, P.C. 431

Brock, T.C. 184, 194, 225

Brown, N. 401

Brown, R. 76, 312

Brubaker, R.G. 454

Brunel, F.F. 170

Bruner, G. 185-6

Bruner, J.S. 76-7, 88, 210, 252, 471

Budd, R.J. 444

Budesheim, T.L. 168

Buehler, R. 119

Bulman, R.J. 466-7

Burgess, P.W. 541

Bush, G.W. 163, 409

Buss, D.M. 19, 21, 26-7

Buunk, B.P. 4-5, 468-9, 471, 474, 476

Byrne, D. 180-1, 187, 195, 290

Cacioppo, J.T. 32, 41, 51, 54-5, 57, 59-65, 68, 101, 104, 108, 170, 205, 211-12, 225-6, 228-33, 235, 239-41, 253, 277, 295, 302-3, 337, 384, 451, 516

Calder, B.J. 419

Cameron, C. 19, 419

Cameron, J. 419

Campbell, D. 201, 466, 495

Campbell, J. 257

Campos, J.J. 515

Cannon, W.B. 33, 506

Cantor, N. 270

Cantril, H. 249

Capraeal, L.R. 14

Caprara, G. 363

Caprarillo, P.A. 282

Caramazza, A. 317

Carlsmith, J.M. 197, 378-80, 384, 386
 Carlson, M. 37, 441
 Carlson, D.E. 63, 158-9, 164, 166, 290-2
 Carmichael, L. 78
 Carnegie, D. 180
 Carruthers, P. 342
 Carter, C.S. 36, 39
 Cartwright, D. 178-80, 182
 Carver, C.S. 57, 88, 98-9, 328, 331, 336, 361, 451, 489, 492-3, 496, 505, 507, 509-18, 520-1
 Cascio, W.F. 417
 Caspi, A. 39
 Catania, C.A. 356, 361
 Cater, D. 355
 Catheart, T. 460
 Cattell, R.B. 213
 Cervone, D. 446
 Cesario, J. 89, 91, 104, 453-4, 500
 Chaiken, S. 100-1, 169-70, 205, 211-12, 225-6, 232, 235, 246-62, 278, 451, 454, 519, 538-9
 Chaires, W.M. 80-1, 88
 Champagne, F.A. 39
 Chan, D.K. 451
 Chartrand, T.L. 170, 172, 440, 453, 530, 535
 Chassein, J. 311
 Chatzisarantis, N.L. 432
 Chave, E.L. 56
 Chen, M. 102, 257
 Chen, S. 93, 246-7, 249-53, 257, 261-2, 272, 278
 Cheung, S. 451
 Chiodera, P. 40
 Chirkov, V.I. 425-6, 431
 Chong, J. 391
 Christiansen, S. 542
 Chun, W.Y. 207
 Cialdini, R.B. 14, 183, 443
 Cikara, M. 282
 Clark, H. 77
 Clark, M.S. 272
 Clark, R.N. 515
 Clausell, E. 282
 Claypool, H.M. 472
 Cleiren, M. 511
 Cloninger, C.R. 42
 Clore, G.L. 289-90, 293, 297-8, 513
 Cohen, A.R. 532
 Cohen, C.E. 158-9
 Cohen, F. 408-9
 Cohen, G.L. 262
 Cohen, M.R. 185
 Cohen, S. 41, 45
 Colcombe, S.J. 165
 Cole, S.W. 214
 Coleman, P.T. 344
 Collins, B. 201
 Collins, J.E. 160, 172
 Collins, N.L. 40
 Collins, R.L. 38, 471, 475
 Colvin, C.R. 538
 Connell, J.P. 422, 431
 Conner, M. 444, 448-9, 451
 Coombs, C.H. 454
 Cooper, E.T. 152
 Cooper, J. 290, 377, 380-1, 383-8, 391-2
 Corneille, O. 110
 Corrigan, P.W. 152
 Cosmides, L. 23, 416
 Costa, P.T. 99, 358
 Costanzo, P.R. 1-3, 6
 Cottrell, C.A. 27
 Courneya, K.S. 445
 Crandall, C.S. 152
 Crawford, C.B. 12, 26-7
 Crawford, L.E. 55, 64
 Crick, N. 153
 Crites, S.L., Jr. 55
 Crocker, J. 252, 468
 Croyle, R. 384
 Csikszentmihalyi, M. 337
 Cuddy, A.J.C. 279-82
 Custers, R. 90
 Dabholkar, P.A. 442
 Daly, M. 21, 23, 26-7
 Damasio, A.R. 296
 D'Andrade, R.G. 311
 Danto, A. 329
 Darke, P.R. 247, 256
 Darley, J. 399, 466-7, 469-71
 Darwin, C. 11, 18, 23, 367
 Daubman, K.A. 322
 Davidson, R.J. 513
 Davis, F.D. 439
 Davis, K.E. 250, 380
 De Busk, Robert 362
 de Charms
 R. 137, 417
 De Dreu, C.K.W. 129, 216, 259-60
 De Fleur, M.L. 444
 de Gelder, B. 61
 De Grada, E. 216
 De Houwer, J. 102, 104
 De Poot, C.J. 320
 DeAngelis, D.L. 512
 Deci, E.L. 14, 417-22, 425-32, 453, 512
 DeCoster, J. 254, 472
 DeGrada, E. 216
 Delillo, D. 401
 Dépret, E.F. 276-7
 Derryberry, D. 519
 Detillion, C.E. 41
 Deutsch, M. 1, 191, 257, 462
 Deutsch, R. 100, 103, 105, 109, 111, 169, 209, 226, 519
 Deutscher, I. 444
 Devine, P.G. 92, 101, 254, 384, 454
 Dewey, J. 212
 Dickerson, C.A. 385
 Diehl, M. 250
 Dijksterhuis, A. 89, 170, 453
 Dimberg, U. 102
 Ditto, P.H. 205, 268
 DiVesta, F. 183
 Dodge, K.A. 153
 Dollard, J. 350, 513
 Douglas, K.M. 314
 Dovidio, J.F. 519
 Drebien, E.K. 269, 274
 Dreisbach, G. 517
 Dulany, D.E. 439-41, 443
 Dunbar, R.I.M. 27
 Duncker, K. 80-1
 Dunton, B.C. 101
 Durante, F. 280-1
 Duval, S. 506-8, 510, 512
 Eagly, A.H. 20, 225, 248-51, 253, 257
 Earls, F. 37
 Ebbesen, E. 158-9
 Echterhoff, G. 93
 Edwards, J.D. 57
 Edwards, W. 454, 489
 Efron, E. 353
 Ehrlich, K. 317
 Eibl-Eibesfeldt, I. 15
 Ein-Dor, T. 409
 Eisenberg, L. 354
 Eitam, B. 90-1
 Elliott, A. 384, 513
 Ellis, A. 218
 Ellis, S. 213, 217
 Ellsworth, P.C. 293, 299
 Elster, J. 103, 119
 Emerson, R. W. 183, 401
 Emery, L. 284
 Epley, N. 342-3
 Epstein, A.R. 213
 Epstein, S. 103, 518-19
 Erb, H.P. 206, 257
 Erber, R. 274
 Erlbaum, L. 466
 Ermer, E. 27
 Eyal, T. 131
 Eysenck, H.J. 355
 Faber, R.J. 109
 Farah, M.J. 100, 111
 Faulkner, J. 27, 100
 Fazendeiro, T. 302
 Fazio, R.H. 91, 93, 101-2, 212-13, 234, 254, 377, 381-2, 385, 388, 491, 519, 538
 Feather, N.T. 442, 506, 510, 527
 Feeney, B.C. 38
 Feinberg, O. 431
 Fekadu, Z. 455
 Feldman, D.B. 152
 Felson, R.B. 339
 Ferguson, M.J. 90, 103
 Fernandez, N. 393
 Fernandez, S. 412
 Festinger, L. 191, 197, 237, 239, 301, 377-84, 386-8, 394, 444, 460-4, 466-7, 471, 478, 489, 507
 Feyerabend, P. 219
 Fiedler, K. 3-5, 102, 309, 311-13, 317, 319, 322
 Fischer, E.F. 21
 Fishbein, M. 157, 183, 439-42, 444-5, 450-1, 454
 Fisher, J.D. & W.A. 439
 Fiske, S.T. 2-3, 14, 252, 254, 257, 268-84, 343, 538
 Fleming, M.A. 262
 Fleming, R. 40
 Ford, D.H. 506
 Ford, T.E. 90
 Forehand, M. 260-1
 Forgas, J.P. 235, 296
 Förster, J. 90-1, 111, 123-4, 169, 293, 295, 453, 495-7, 499
 Fortier, M.S. 430
 Fowler, C. 454
 Fraley, R.C. 38
 Francis, D. 37
 Franco, F.M. 313
 Frank, E. 283, 285, 536
 Frederick, C.M. 427
 Frederick, S. 99
 Fredrickson, B.L. 343
 Freedman, J.L. 379
 Freitas, A.L. 127, 538
 Freud, S. 18, 78, 98, 100, 214, 399, 401, 489
 Freund, T. 219
 Frey, D. 337
 Fried, C.B. 385
 Friedman, M. 404
 Friedman, R.S. 293, 295
 Fries, M. 102, 108-9
 Frijsda, N.H. 293, 513
 Frith, C.D. 342
 Frith, U. 342
 Fujita, K. 110, 126-7, 535
 Fulford, D. 517
 Funder, D.C. 98
 Gagné, C. 448
 Gagné, F.M. 532, 534
 Galinsky, A.D. 278, 538
 Gambone, M.A. 431
 Gangestad, S.W. 21-2
 Garcia, J. 16
 Gardner, H. 17
 Gardner, W.L. 17, 55, 60, 64, 388
 Garnham, A. 317
 Garvey, C. 317
 Gastorf, J.W. 467
 Gaunt, R. 207
 Gawronski, B. 102, 109, 185, 209, 254
 Gayer, C. 218
 Geary, D.C. 20
 Gelfand, M.J. 216-17
 Gendolla, G.H.E. 141
 Geraerts, E. 451
 Gerard, H.B. 1, 191, 257
 Gergen, K.J. 330, 465
 Gerrard, M. 468

Gibbons, F.X. 468, 476
 Gibbons, R. 468
 Gigerenzer, G. 206-8
 Gilbert, D.T. 109, 122, 130, 254, 538-9, 541
 Gilovich, T. 119, 186
 Gilson, C. 317
 Giner-Sorolla, R. 253, 258
 Giuliano, T. 336
 Glick, P. 281-2
 Glucksberg, S. 81
 Godin, G. 448
 Goethals, G.R. 382-3, 466-7, 469-71
 Goffman, E. 401
 Goldenberg, J.L. 406-8
 Goldman, A.I. 328-9, 331
 Goldsen, R. 354
 Goldstein, K. 428
 Gollan, J.K. 55
 Gollwitzer, P.M. 106, 494, 526-36, 538, 540-2
 Goodwin, S.A. 276-8
 Gordon, B.F. 464
 Gorn, G.J. 298
 Gortner, E.T. 69
 Goschke, T. 517
 Grabenhorst, F. 57
 Graham, S. 151, 153
 Gray, H.M. 343
 Gray, J.A. 57, 489, 492
 Gray, K. 343
 Greenberg, J. 399, 404, 406, 410-12, 511
 Greene, S. 317
 Greenslade, L. 311, 313
 Greenwald, A.G. 102-3, 108, 170, 184, 225, 241, 453, 535
 Gregory, R.L. 210
 Greifeneder, R. 100, 297, 300
 Grice, H.P. 165, 319
 Griffin, D. 120
 Griffitt, W. 180, 290
 Grippo, A.J. 36
 Griskevicius, V. 14, 25
 Grolnick, W.S. 422, 431-2
 Gruenfeld, D.H. 160
 Guastello, S. 343
 Guinote, A. 276, 278
 Gunnar, M.R. 38
 Guthrie, E.R. 269
 Haack, S. 196
 Haddock, G. 300
 Hager, J.L. 16
 Hagger, M.S. 432, 451
 Hakmiller, K.L. 464-5, 467-8
 Hall, C.C. 260
 Halvari, A.M. 430
 Halvari, H. 430
 Hamilton, D. 158-9, 168, 174, 271
 Hamilton, W.D. 43
 Hammen, C. 69
 Hampshire, S. 328
 Han, H.A. 110
 Haney, C. 363
 Hannover, B. 295
 Hansen, J. 296
 Hansen, N. 499
 Harary, F. 178-9, 182
 Harary, R. 180, 185
 Harborne, A. 152
 Hardin, C. 93, 205, 211, 216, 219, 316
 Harinck, F. 129
 Harlow, H.F. 428
 Harmon-Jones, E. 383-4, 388, 404, 410, 515-16, 527, 532
 Harré, R. 365
 Harris, L.T. 278, 282, 343
 Harris, P.R. 259
 Hart, A.J. 278
 Hart, J. 38
 Hartmann, H. 428
 Harton, H.C. 26
 Hartwick, J. 157, 167
 Harvey, J. 225
 Haselton, M.G. 22-3, 26
 Haslam, N. 343
 Hass, R.G. 60, 281
 Hassin, R.R. 452-3
 Hastie, R. 158-9, 168, 259, 269, 272
 Hatfield, E. 55
 Hau, K. 475
 Haugvedt, C. 234, 240
 Hausenblas, H.A. 451
 Hawkins, D.L. 392
 Hawley, L.C. 32, 41
 Hayes, J. 409
 Haynes, M.C. 284
 Heatherton, T.F. 108, 196, 335, 337
 Heckhausen, H. 527-9, 534-5
 Heckhausen, J. 511
 Hedberg, P.H. 91
 Heider, F. 135, 138, 178-84, 186-8, 231, 257, 279, 378, 417, 428, 489
 Heiman, M.E. 284
 Heine, S.J. 391
 Heinrichs, M. 40, 42
 Helgeson, V.S. 477
 Helmholtz, H. 210
 Henderson, M.D. 124, 126, 129, 131, 536
 Henle, M. 185-7, 189
 Hennigan, K. 477
 Herr, P.M. 471
 Higgins, E.T. 2, 4-6, 14, 25, 77-8, 81-93, 106-7, 159, 169, 205, 211, 216, 219, 249, 252-3, 295, 323, 331, 386, 394, 452, 472, 484-9, 491-5, 497-8, 500-1, 512, 514-16, 529
 Himelstein, P. 444
 Hirschberger, G. 406, 409
 Hirshleifer, D. 298, 301-2
 Hirst, W. 106
 Hodgins, H.S. 428
 Hoefling, A. 109
 Hoffman, C. 316, 322
 Hoffmann, W. 102-3, 108-10
 Hofstede, G. 193
 Hogg, A.M. 388-90
 Holahan, C.J. & C.K. 40
 Holbrook, A.L. 55
 Holland, P.C. 209
 Holyoak, K.J. 209-10
 Hooley, J.M. 152
 Hoshino-Browne, E. 391-2
 Houston, D.A. 498
 Hovland, C.I. 212, 224, 229, 255, 387
 Howard, M.V.A. 259
 Hrubes, D. 449-50
 Huang, J.Y. 14
 Hudley, C. 153
 Huguet, P. 476
 Hull, C.L. 135, 379, 383, 478
 Humphries, C. 512
 Hung, I.W.P. 172
 Hunter, R. 157
 Hupfeld, R.M. 259
 Husaini, B.A. 152
 Hutchins, E. 318
 Huttenlocher, J. 75-7, 80
 Idson, L.C. 342, 490-1, 496-7
 Imada, T. 391
 Imber, L.G. 335
 Insel, T.R. 39
 Insko, C.A. 179-81, 183-4, 187-92, 195-7, 468
 Isbell, L.M. 297-8
 Isen, A.M. 290, 302, 322
 Ito, T.A. 55-6, 65
 Jaccard, J. 102
 Jackson, J.H. 65
 Jacobson, N.S. 69
 Jaffe, Y. 217-18
 Jahoda, M. 193
 Jalowicz, J.E. 36
 James, W. 88, 104, 186, 210, 298, 331, 359, 401
 Jamner, L.D. 37
 Jang, H. 425
 Jankowiak, W.R. 21
 Janoff-Bulman, R. 342
 Jepson, C. 259
 Jiang, Y. 165, 172
 Johnson, B.T. 249
 Johnson, E. 301-2
 Johnson, M.K. 106
 Johnson-Laird, P.N. 312
 Jonas, E. 406
 Jones, C. 78
 Jones, E.E. 1, 4-5, 250, 379-80
 Jones, S. 463, 470-1
 Jordan, J.V. 426
 Jordan, N. 180, 188-9
 Jost, J.T. 217-18, 409
 Joule, R.V. 388
 Joyce, J. 401
 Kage, M. 425
 Kahneman, D. 57, 61, 101, 119-20, 207, 209, 489
 Kandel, E.R. 84
 Kang, J. 283
 Kanouse, D.E. 317
 Kaplan, H.S. 20
 Kappes, A. 542
 Kashima 322
 Kasmier, J.A. 240
 Kasoff, J. 317
 Kasser, T. 424
 Katkin, E.S. 384
 Katz, D. 257
 Katz, I. 281
 Kaufman, J.C. 39
 Kay, P. 320, 322
 Keefe, R.C. 19-20
 Keller, J. 486, 499
 Keller, M.C. 27
 Keller, P.A. 500
 Kelley, H.H. 2, 78, 141, 201-3, 211, 218, 250, 291, 465-7
 Kelly, G.A. 82
 Kelman, H. 251
 Keltner, D. 278, 294, 299
 Kendrick, K.M. 36
 Kennedy, K. 268, 284
 Kenrick, D.T. 12, 19-27
 Kervyn, N. 280
 Kessen, W. 2
 Kessler, R.C. 44
 Keverne, E.B. 36
 Kierkegaard, S. 399, 401
 Kiesler, C.A. 225, 249
 Kim, S. 152
 Kim, U. 367
 Kim, Y. 500
 Kimble, G.A. 335, 337
 Kimchi, R. 322
 King, G. 82-3
 King, K.B. 41
 Kinney, R.F. 532-3, 536
 Kirby, K.N. 119
 Kirchhof, O. 526
 Kitayama, S. 193, 322-3, 390-2
 Klapper, J. 355
 Klayman, J. 107, 472
 Klein, D. 460
 Klinger, E. 510-11
 Kluckhohn, C. 368
 Knafo, A. 42
 Knee, C.R. 428
 Knowles, E.S. 341
 Koelling, R.A. 16
 Koestner, R. 420, 467
 Kohler, W. 252
 Koller, M. 141
 Konrad L. 15
 Koole, S.L. 216, 406
 Koomen, W. 473
 Koriati, A. 101
 Kosfeld, M. 37, 40
 Kosic, A. 217
 Kosloff, S. 409
 Kosslyn, S.M. 100

Kozak, M.N. 342-3
 Kraft, P. 455
 Krahé, B. 311
 Krauss, D.H. 152
 Krauss, R. 1, 75, 462
 Krieger, D.L. 27
 Krieger, L.H. 283
 Krieger, L.S. 424
 Kriegelmeyer, R. 110
 Krosnick, J.A. 55, 234
 Kruglanski, A.W. 2, 5-6, 90, 92-3, 101, 170-1,
 202, 204-8, 211-19, 232-3, 237, 259, 314,
 320, 328, 341, 452-3, 471, 494
 Kuhl, J. 534
 Kuhn, T.S. 213-14
 Kukla, A. 138, 144
 Kulik, J.A. 41, 476
 Kumar, P.A. 158-9, 168
 Kunda, Z. 236, 277, 394, 469, 472
 Kundera, M. 401
 Kurzban, R. 26
 La Guardia, J.G. 426-7
 La Rouche, L. 356
 Lakatos, I. 219
 Lam, S. 425
 Lancaster, J.B. 12-13
 Landau, M.J. 404, 406, 409, 412
 Lang, P.J. 55, 57, 516
 Langer, E.J. 335, 401
 Langer, S. 401
 LaPiere, R.T. 119
 Larsen, J.T. 55, 60, 65
 Larsen, O. 354
 Latané, B. 25, 464
 Latham, G.P. 143, 439, 453
 Latimer, A.E. 450, 500
 Lau, R.R. 91, 260-1
 Le Bon, G. 100
 Leary, M.R. 14, 186, 190, 257, 388, 421
 Ledgerwood, A. 131, 250, 259-60, 527
 LeDoux, J.E. 51
 Lee, A.Y. 500
 Lee, J.Y. 317
 Legros, J.J. 40
 Lehman, D.R. 391
 Lenneberg, E.H. 316
 Lepper, M.R. 418
 Lerner, J.S. 257, 260, 299, 302, 343
 Levesque, C. 428
 Levine, J.M. 5, 499
 Levine, S. 38
 Levinson, S. 128
 Levy-Sadot, R. 101
 Lewin, K. 89-90, 118, 120, 123, 135-6, 143,
 149-50, 214, 330, 341, 377-8, 496, 526-7
 Lewinsohn, P.M. 69
 Li, N.P. 22, 27
 Liberman, A. 258-9
 Liberman, N. 90-1, 121-2, 124, 126-7,
 169, 538
 Lichtman, R. 468
 Lieberman, M.D. 278
 Lieberman, N. 333
 Lim, M.M. 42
 Lin, M.H. 40, 282
 Linde, D.E. 380
 Lingle, J.H. 158, 164
 Linn, J.A. 341
 Linn, L.S. 444
 Linville, P.W. 252
 Lippman, W. 313
 Litt, M.D. 447
 Liu, D. 37
 Liu, T.J. 194
 Livi, S. 217
 Liviatan, I. 124, 126-8
 Lobel, K.L. 147
 Lobel, M. 469
 Locke, E.A. 143, 357, 362, 439, 446, 453
 Lockwood, P. 469, 472
 Loewinger, J. 429
 Loewenstein, G.F. 99, 103, 119-20, 127
 Logan, G.D. 210
 Lombardi, W.J. 85-6, 472
 Lopes, L.L. 492
 Lopez, D.F. 205, 258
 Lord, C.G. 258
 Lord, R.G. 361
 Lorig, K. 362
 Losch, M.E. 384
 Lotze, H. 104
 Louro, M.J. 517
 Lovaldo, D. 119-20
 Lummaa, V. 15
 Lydon, J.E. 532, 534
 Lykins, E.L.B. 411
 Lynam, D.R. 99
 Lynch, M.F. 426
 Maass, A. 313-14, 323, 345
 MacDorman, K.F. 406
 McAlister, A. 363
 Mackie, D.M. 250, 388-9
 McCann, C.D. 219, 249
 McCarthy, M.M. 39-40
 McClure, S.M. 99
 McCrae, R.R. 99, 358
 McEwen, B.S. 33
 McFarland, D.J. 512
 McGregor, H. 404
 McGuire, W.J. 157, 159, 166, 174, 182, 187,
 194, 224, 250-1, 253, 257, 312
 MacKay, D.M. 508
 McKay, G. 45
 McKoon, G. 317
 McMillan, D. 274
 McPherson-Franz, C. 342
 Maheswaran, D. 255
 Maher, H.I.M. 41, 476
 Majid, A. 322
 Malone, P.S. 538-9
 Mandler, G. 2
 Maner, J.K. 24-5
 Manian, N. 486
 Manis, M. 471
 Manstead, A.S.R. 454
 Mantella, R.C. 40
 Marbe, K. 528
 Marilyn, G.E. 260
 Markel, P. 110
 Markman, A.B. 104, 474
 Markus, H. 82, 193, 270, 390-1, 425-6, 512
 Marmot, M. 43
 Marsh, A.A. 342
 Marsh, H.W. 475
 Marteau, T.M. 41
 Martel, F.L. 36
 Martin, L. 86, 168, 235, 292, 297, 299
 Martin, M.A. 241
 Martin, R. 470
 Martin Ginis, K.A. 450
 Maruyama, M. 512
 Marx, K. 399
 Maslow, A.H. 429
 Mather, M. 278
 Matlin, M.W. 292
 Mayr, E. 16
 Mayseless, O. 471
 Mead, G.H. 339
 Meaney, M.J. 37, 39
 Medvec, V.H. 119
 Medvene, L.J. 152
 Meichenbaum, D. 340
 Menasco, M.B. 392
 Mendelson, T. 368
 Merton, R.K. 2
 Metcalfe, J. 103, 120, 518-19
 Mikulincer, M. 14, 406, 410
 Milgram, S. 118-19, 353
 Mill, J. S. 185, 187
 Miller, G.A. 312, 331, 507-9, 512
 Miller, J.G. 390
 Miller, N.E. 56, 179, 350, 496, 513
 Miller, N.R. 120, 179, 312, 331, 350, 466, 468,
 496, 507-9, 512-13
 Miller, R.L. 465-6, 468
 Mills, J. 379, 382, 384, 392
 Mineka, S. 14, 16
 Mischel, W. 82, 103, 119-20, 270, 342, 528
 Mita, T.H. 179
 Mitchell, C.J. 104, 110, 209
 Miyake, K. 468-9
 Mizruchi, M.S. 517
 Modahl, C. 40
 Molden, D.C. 499
 Moles, A. 35
 Moller, A.C. 427
 Mondak, J.J. 260-1
 Monin, B. 390, 393
 Monteith, M.J. 92, 454
 Montoya, M. 180
 Moore, J. 444
 Moors, A. 102
 Moreland, R.L. 5, 499
 Moretti, M.M. 486, 488
 Morris, W.N. 294
 Morse, S. 465
 Moskowitz, G.B. 92, 252, 538
 Motyl, M. 411
 Movshon, T. 83
 Mowrer, O.H. 489, 492
 Muñoz, R. 368
 Munson, P.A. 225
 Muraven, M. 427
 Murray, D.R. 14, 368
 Murstein, B.I. 17, 21
 Mussweiler, T. 80, 86, 107, 236, 472-5
 Nachmias, M. 38
 Nagel, E. 360
 Namiki, H. 425
 Napper, L. 259
 Neal, D.T. 104, 452
 Neely, J.H. 471
 Neff, J.A. 152
 Nesse, R.M. 27, 511
 Nettle, D. 23
 Neuberger, S.L. 24, 27, 271, 275, 277, 538
 Neumann, R. 102, 104, 519
 Newcomb, T.M. 185, 378
 Newman, B. 210, 260
 Newman, L.S. 445
 Newsom, J.T. 277
 Newton, I. 136
 Newton, J.W. 527
 Newton, N. 40
 Newtonson, D. 328
 Niemiec, C.P. 424
 Nigg, J.T. 519
 Niko T. 15
 Nisan, M. 119
 Nisbett, R.E. 159, 290
 Nix, G.A. 427
 Norman, D.A. 112, 531
 Norris, C. 55, 65-6
 Norton, M.I. 383, 389, 391, 393
 Notani, A.S. 449, 451
 Novemsky, N. 301-2
 Nowak, A. 26, 336, 344
 Nurius, P. 512
 Nussbaum, S. 126
 Oakhill, J. 317
 Ochsmann, R. 405
 Ochsner, K.N. 515
 Oettingen, G. 541-2
 Ogilvie, D.M. 512
 Ohman, A. 14, 16, 56
 Oldmeadow, J. 280
 Olson, J.M. 499
 Olson, M.A. 519
 Operario, D. 277
 Oppenheimer, D.M. 236, 301-2
 Orehek, E. 218
 Orth, J. 528
 Orton, R. 359
 Ortony, A. 293-4, 299, 513
 Osgood, C.E. 182

Ostrom, T.M. 57, 158-9, 164, 174, 271
 Ottati, V. 293, 297-8, 303
 Overbeck, J.R. 278
 Özgen, E. 316, 322

Paisley, M. 354-5
 Paivio, A. 111
 Pallak, M.S. 255, 384
 Palmer, S.E. 322
 Palmonari, A. 494
 Panksepp, J. 35-6, 42
 Parducci, A. 471
 Park, B. 269, 278
 Pascal H. 476
 Paschal S. 540
 Pastorelli, T. 363
 Patrick, C. 56
 Patrick, H. 426-7
 Patterson, C.J. 528
 Pavelchak, M.A. 272-3
 Peake, P.K. 446
 Pechmann, C. 500
 Peizer, D.B. 184
 Pelletier, L.G. 422, 428, 431-2
 Penn, D.J. 26
 Penninx, B.W.J.H. 40
 Penton-Voak, I.S. 22
 Perlmutter, L.C. 335, 337
 Perloff, R.M. 260
 Perry, R.P. 150
 Peters, R. 153
 Petersson, M. 39
 Pettigrew, T. 465
 Petty, R.E. 54-5, 101, 169-70, 205, 211-12,
 225-6, 228-41, 253, 255, 262, 303, 451
 Pfander, A. 428
 Pham, M.T. 292, 296, 298
 Phelps, E.A. 278, 284
 Phillips, D.C. 359
 Piaget, J. 429
 Pichert, J.W. 252
 Pierro, A. 206, 208, 214, 216
 Pietromonaco, P. 83, 171
 Pillsworth, E.G. 22
 Pinker, S. 15
 Pitschel-Walz, G. 152
 Pittman, T.S. 14, 384
 Pizlo, Z. 210
 Plaks, J.E. 89
 Popper, K.R. 92, 201-4, 213-14, 218-19
 Port, R.F. 344
 Pösl, I. 535-6
 Posner, M.I. 82-3, 519
 Potepan, P.A. 137
 Powers, W.T. 361-2, 508-9, 512, 520
 Pratto, F. 27
 Prelec, D. 119, 127
 Pritchard, R.D. 418
 Przybylski, A.K. 432
 Puca, R.M. 534
 Pyszczynski, T. 398, 405, 409, 412, 511

Quinn, K.A. 499

Rachlin, H. 119
 Radloff, R. 464-5, 475
 Radvansky, G.A. 160, 165, 172
 Raghunathan, R. 296
 Randall, D.M. 449
 Rank, O. 399, 401-2, 410
 Rasmussen, H.N. 521
 Ratcliff, R. 166
 Ratneshwar, S. 253
 Raviv, A. 212
 Reber, R. 301-2
 Redlawsk, D.P. 260-1
 Redlmeier, D.A. 298
 Reeve, J. 431
 Regan, D. 463, 470-1
 Reis, H.T. 424
 Reisenzein, R. 296-8
 Reitsma-van Rooijen, M. 314
 Repetti, R.L. 32, 38

Rescorla, R.A. 209-10
 Rhamey, R. 180
 Rholes, W.S. 78, 404
 Richter, L. 215
 Ricoeur, P. 315
 Rimai, R.N. 455
 Rivis, A. 451
 Rizzolatti, G. 104
 Roberts, J.M. 316
 Roberts, T.A. 343
 Robinson, M.E. 283-4
 Rock, I. 210
 Rodin, G.C. 430
 Rodrigues, A. 147
 Roese, N. 23
 Roets, A. 214, 219
 Rogers, C.R. 212, 429
 Rogers, E.M. 369-70
 Roney, C.J.R. 490, 494
 Rook, K.S. 45
 Rosen, N.A. 167
 Rosenbaum, D.A. 509, 517
 Rosenberg, M.J. 182-4, 194, 379-80, 441-2
 Rosenberg, S. 279
 Rosenblatt, A. 400, 405
 Rosenblueth, A. 506
 Rosenstock, I.M. 439
 Rosenthal, T.L. 351
 Ross, L. 159
 Ross, M. 119-20
 Rostral 53
 Roth, G. 423
 Roth, P. 401
 Rothbart, M.K. 519
 Rothermund, K. 102
 Rothman, A.J. 297, 300
 Rothschild, Z.K. 410
 Rotter, J.B. 135, 137-8, 140, 350, 420, 506, 510
 Routledge, C. 409, 411-12
 Rowe, C.J. 97
 Rowe, D.C. 16
 Rozanski, A. 427
 Rubel, T. 509
 Rubini, M. 215, 314, 320
 Rucker, D.D. 240-1
 Ruder, M. 300
 Rudman, L.A. 283
 Rudolph, U. 146-7
 Runkel, P.J. 184
 Ruscher, J.B. 275-6, 280
 Russell, A.M. 282
 Russell, B. 76
 Russell, J.A. 51
 Russell Sage Foundation 284
 Rutledge, T. 41
 Ryan, R.M. 14, 417-32, 453, 512
 Ryle, G. 329

Sabido, M. 369
 Sabini, J. 193
 Sachs, P. 471
 Sadalla, E.K. 13
 Saegert, S. 179
 Sagristano, M. 126
 Salovey, P. 296
 Salzman, M.B. 406
 Sampson, E.E. 367
 Sande, G.N. 388
 Sanna, L. 295, 300
 Santelli, A.G. 498
 Sarason, I.G. 40, 337
 Sarkisian, C.A. 151
 Sassenberg, K. 499
 Savitsky, K. 298
 Schachter, S. 75, 193, 290, 463-4, 466
 Schacter, D.L. 24, 122
 Schaller, M. 14, 25, 27
 Schank, R. 302
 Scheier, M.F. 328, 331, 336, 361, 451, 489,
 492-3, 496, 505, 507, 509-15, 518, 520-1
 Schelling, T. 119, 122
 Scher, S. 383
 Scherer, K.R. 293, 299

Schimel, J. 404, 406
 Schlenker, B.R. 194
 Schmidt, G. 146
 Schmidt, P. 454
 Schmitt, D.P. 21, 186-7
 Schmitt, D.R. 137, 144
 Schneider, W. 83, 100
 Scholer, A.A. 500
 Schooler, C. 368
 Schulz, R. 511
 Schulze, R. 451
 Schumway 301
 Schupp, H.T. 56
 Schwartz, A.J. 419
 Schwartz, S.H. 509
 Schwarz, N. 101-2, 111, 169, 235-6, 289,
 291-304
 Scott, C.G. 259
 Scott, W.E. 419
 Sears, D.O. 63
 Secord, P. 381
 Sedikides, C. 14, 188, 193, 297-8
 Sedlak, A. 279
 Seeman, T.E. 40
 Segerstrale, U. 13
 Seibt, B. 105, 499
 Seligman, M.E.P. 16
 Selz, K. 341
 Semin, G.R. 297, 309, 311-14, 317-22
 Sexton, J. 93, 449, 451
 Shafer, J.B.P. 212
 Shah, J.Y. 88, 171, 216, 490-1, 495, 499, 517
 Shallie, T. 112, 517
 Shanahan, J. 152
 Shaver, P.R. 1, 14
 Shaw, M.E. 1-3, 6
 Sheeran, P. 449-52, 540-1
 Sheets, V.L. 26
 Sheldon, K.M. 424
 Sheldon, T. 399
 Shen, H. 173, 296
 Sheppard, B.H. 444, 449
 Sherif, M. 249, 255
 Sherman, J. 23, 312
 Sherman, P.W. 17
 Sherman-Williams, B. 236
 Sherry, D.F. 24
 Shiffrin, R.M. 83, 100
 Shin, J.C. 517
 Shiota, M.N. 24
 Shotland, R.L. 353
 Shultz, T.R. 194
 Shumway, T. 298, 302
 Shweder, R. 311
 Shytenberg, G. 215
 Sidanius, J. 27
 Siemer, M. 296-8
 Silberman, I. 486
 Simeon, D. 40
 Simon, H.A. 247, 336, 517
 Sinclair, R.C. 292, 298, 303
 Singer, J. 290, 464
 Skinner, B.F. 16, 150, 356
 Skinner, M.L. 379, 383
 Skowronski, J.J. 14, 63, 188
 Sloman, S.A. 100-1, 204, 209
 Slovic, P. 235, 296
 Smith, D.M. 55, 209
 Smith, E.R. 83, 164, 209-10, 254, 297, 317-18
 Smith, K. 55
 Smith, M.B. 249, 257
 Smith, N.K. 55
 Smith, P.K. 123-4, 128, 342
 Smith, P.M. 193
 Smith, R.A. 468
 Smith, W. 471
 Snyder, M. 274
 Solarz, A.K. 102
 Solberg N. 521
 Soldat, A.S. 303
 Solomon, S. 398, 403
 Song, H. 293, 301-3
 Sorabji, R. 97

Sorkin, D. 40
 Sorrentino, R.M. 81, 529
 Spelke, E. 61
 Spence, K.W. 135, 379, 383, 478
 Spiegel, D. 477, 500
 Spitz, R.A. 37
 Srivastava, S. 190
 Srull, T.K. 80, 83-4, 158-60, 164, 167-8, 274, 452, 472
 Staw, B.M. 392, 419
 Stearns, S.C. 20
 Steele, C.M. 194, 231, 388, 499
 Steenbarger, B.N. 510
 Stephan, E. 128
 Stephens, E.J. 542
 Stepper, S. 292, 295
 Stevens, L.E. 276-7
 Stone, A.A. 40, 43
 Stone, J. 377, 385-8
 Strack, F. 97, 100-3, 105, 107, 109-11, 169, 209, 226, 236, 292, 295, 472, 519
 Strahan, E.J. 90
 Strauman, T. 485-6, 489, 498
 Strickland, S. 355
 Stroebe, W.H.M. 181, 250
 Stroop, J.R. 125
 Sullivan, D. 406
 Suls, J. 466-8, 470-1, 473-6
 Suomi, S.J. 37-8
 Sutton, R.M. 314, 452
 Swann, W.B.Jnr. 196, 274, 339-40
 Swanson, J. 453
 Syme, S.L. 41
 Szfy, M. 39

Tajfel, H. 389
 Talaska, C.A. 283
 Tamres, L. 35
 Tangney, J.P. 186, 388
 Tannenbaum, P.H. 182, 354
 Tashakkori, A. 195-6
 Taylor, S.E. 14, 23, 32-6, 39, 41-2, 44-5, 186, 252, 268-70, 272, 274, 278-9, 451, 469, 531, 533, 536-8
 Tedeschi, J.T. 231
 Tellegen, A. 51
 Tennen, H. 468
 Tesch, F. 468
 Tesser, A. 257, 334, 339, 394, 468, 474-5
 Testa, M. 468
 Tetlock, P.E. 257, 259-60
 Thagard, P. 277
 Thaler, R.H. 99, 120, 496
 Thayer, R.E. 427
 Thibaut, J.W. 211
 Thibodeau, R. 384-5
 Thistlethwaite, D.L. 195
 Thoits, P.A. 45
 Thomas, A.W. 466
 Thompson, E.P. 101, 206, 212, 232-3
 Thompson, L.L. 259
 Thorndike, E.L. 136, 139, 489
 Thornhill, R. 22
 Thurstone, L.L. 56
 Todorov, A. 126, 260, 279
 Tolman, E.C. 135, 269, 379, 510
 Tolstoy, L. 54, 401
 Tooby, J. 23, 416
 Topolinski, S. 107
 Tormala, Z.L. 238, 241
 Towles-Schwen, T. 254
 Treyens, J.C. 252
 Triandis, H.C. 453
 Trivers, R.L. 18, 43

Trope, Y. 100, 121-4, 126-9, 169, 206-7, 226, 250, 252, 254, 278, 333, 451, 454, 519, 538
 Turner, J.C. 389, 461
 Turner, R.A. 36, 461
 Turner-Cobb, J.M. 40
 Tversky, A. 27, 57, 100-1, 107, 119-20, 207, 236, 252, 300-2, 446, 471, 489
 Tybur, J.M. 12
 Tykocinski, O. 486

Uhlmann, E. 453
 Uleman, J.S. 210
 Upshaw, H.S. 157
 Uvnäs-Moberg, K. 39-40

Vallacher, R.R. 26, 298, 328, 331, 333-8, 341-4, 509
 Vallerand, R.J. 422
 van den Berg, A.E. 406
 van den Bos, K. 404
 Van Lange, P.A.M. 1-2, 4, 14
 Van Overwalle, F. 150
 Van Ryn, M. 454
 Van Vugt, M. 5, 14
 VanderPlatte, C. 40
 Vansteenkiste, M. 424, 431
 Vaughan, P.W. 370
 Veitch, R. 290
 Vershure, B. 13, 19
 Vess, M. 407, 412
 Vieth, A.Z. 498
 Vlachopoulos, S.P. 432
 Vohs, K.D. 109, 506
 von Bertalanffy, L. 506
 Von Frisch, K. 15
 Von Hendy, H.M. 275
 Von Hippel, W. 313
 von Humboldt, W. 319
 Vonnegut, K. 401
 Vroom, V.H. 444, 510
 Vroomen, J. 61

Wagner, A.R. 209-10
 Waksak, C.J. 123-4, 131
 Walsh, J. 355
 Walster, P. 180
 Walters, R.H. 355
 Walther, E. 185, 209
 Wänke, M. 300
 Warren, R.E. 83
 Waterman, C.K. 384
 Watson, D. 51, 69, 516
 Watt, H.J. 528
 Waytz, A. 342-3
 Weaver, K.D. 301, 387
 Webb, T.L. 450
 Webster, D.M. 214-16, 259, 341
 Webster-Nelson, D. 215
 Weeden, J. 27
 Wegener, D.T. 226, 229, 231, 234-7, 240, 255, 303
 Wegner, D.M. 298, 328, 333, 336, 339, 342-3, 440, 453-4
 Weigel, R.H. 445
 Weinberg, L.E. 447
 Weiner, B. 137-47, 202
 Weinstein, N.D. 432, 538
 Weisberg, R.W. 81
 Weise, D.R. 410
 Weishaar, M. 340
 Weiss, R.S. 511
 Wellens, A.R. 195
 Wells, G.L. 237
 Wentura, D. 102

Werkman, W.M. 314
 Werner, C.M. 259
 Wertheimer, M. 179
 West, S.G. 1, 6
 Westie, F.R. 444
 Wetzel, C. 181, 187
 Wheatley, T. 440, 453
 Wheeler, L. 464-5, 467-71, 474-6
 Wheeler, M.E. 278
 Whiteside, S.P. 99
 Whorf, B.L. 310, 316, 321
 Wichman, A.L. 242
 Wicker, A.W. 225, 444
 Wicklund, R.A. 1, 6, 337, 506-8, 510, 512, 526-7
 Wiener, N. 506
 Wiener, S.G. 38
 Wiers, R.W. 109
 Wiest, W.M. 180, 186, 195
 Wigboldus, D.H. 314
 Wiklund, I. 41
 Wilcoxon, H.C. 16
 Wilhelm, H. 109
 Williams, G.C. 421-2, 424, 430
 Willis, J. 279
 Wills, T.A. 40, 464, 467-8, 471
 Wilson, E.O. 13, 15, 17, 25, 366
 Wilson, M.I. 21, 23, 26-7, 179
 Wilson, T.D. 102-3, 122, 130, 151, 290, 295
 Wine, J.D. 337
 Winkelmann, P. 108, 294-5, 301-2, 337
 Winograd, T. 531
 Wisman, A. 406
 Wittenbrink, B. 102
 Wittmann, W.W. 451
 Wojciszke, B. 279
 Wolff, J.A. 449
 Wolff, K.M. 37
 Wong, E.M. 538
 Wood, J.V. 255, 460, 467-8, 476, 478
 Wood, W. 20, 104, 253, 255, 452
 Woodruff, H. 183
 Worchsel, P. 180
 Worchsel, S. 381, 383
 Wortman, C.B. 510
 Wrosch, C. 511
 Wyer, R.S. 156-60, 165-70, 172-3, 194, 290-2

Xu, A.J. 173
 Xu, J. 173, 300

Yalom, I. 401
 Yang-Wallentin, F. 449
 Yates, S.M. 252
 Ybema, J.F. 474
 Yeager, R.J. 218
 Yeh, W. 298

Zajonc, R.B. 160, 179, 188-9, 196, 290-1, 301
 Zak, P.J. 37, 40
 Zakary, T. 238
 Zanna, M.P. 5, 212-13, 271, 290, 384, 388, 467
 Zeldman, A. 431
 Zhao, G. 500
 Ziegler, R. 257
 Zilboorg, G. 401
 Zillman, D. 290
 Zimbardo, P.G. 118
 Zimmerman, B.J. 351
 Zsolni, L. 363
 Zuckerman, A. 259
 Zuckerman, M. 99, 417-18, 428
 Zuroff, D.C. 431

Subject Index

ability appraisals 462, 534
absolute judgments 107, 473–4
abstract language 215, 314, 323
abstract predicates 313, 322
abstract thinking 110, 124, 131
abstraction 2–4, 123, 129, 320, 332, 508–9
 levels of 124, 320, 365, 506, 508–9
academic self-concepts 475–6
accessibility 75–8, 80–3, 85–8, 90–3, 404–5,
 487–8
 beliefs 448–9
 category 78
 chronic 75, 81–3, 85–8, 90, 487, 491
 death-thought 404–5
 decay 83, 91
 effects 75, 77–8, 80–1, 83, 92
 experiences 294, 299–300
 of goal-related constructs 90–1
 higher 76–7, 298, 491
 knowledge 159, 169, 460, 473
 relative 76, 170
 selective 107, 473
 sources of 85, 92–3
 temporary 77–8, 81–2, 487
 theory 75–93, 484
accessible beliefs 442, 447–9
accessible constructs 82, 91
accessible knowledge 87, 472
accountability 229, 260, 363, 431
accuracy 90, 92, 130, 246, 251, 275–6
accuracy motivation 249, 251, 255–6,
 258, 260
accurate judgments 130, 533
achievement behaviors 89, 143, 202
achievement failure 148, 150
achievement motivation 143–4, 489, 527
achievement strivings 142–4, 150, 362
act identities 330–2, 334, 337–8
action, reasoned 438–45, 447, 451–4
action approach, reasoned 438–9, 451
action contexts 333, 339
action control 335, 337, 541
action identification theory 327–46, 509
action loop 513–15
action models, reasoned 438, 449–50, 453–4
action perspectives, reasoned 452–3
action phases, mindset theory of 526–42
action verbs 313, 319–20, 322–3
actions
 goal-directed 89, 162, 171, 173, 527–9,
 535–6
 negative 153, 343
activation 52–3, 56, 58–9, 64, 88–92, 106–7
 automatic 83, 453
 evaluative 53, 58–60, 63
 functions 50, 52–3, 55–6, 58, 61–4
 knowledge 79, 82, 85–8, 92–3
 reciprocal 50, 52, 55–6, 58, 60
actual behavior 108–9, 444, 450
actual control 446–8, 532
actual-ideal discrepancies 485, 490–1, 498,
 500–1

actual-ought discrepancies 485, 487,
 490–1, 498
actual self 187, 485, 488
adaptation 64, 105, 318, 349, 351, 356
 dual 318, 321
adaptive functions 11, 17–18, 518
adaptive responses 51–2, 54, 66–7
addiction 109–10, 533
additivity 85–6, 157
addressees 128, 318
adjectives 77–8, 311–13, 322–3
adjustments 430, 468, 472, 476, 506, 512
 psychological 40, 426
adversity 505, 521
advocacy 241, 384, 409
 counterattitudinal 380, 383–4, 388, 390–1
 proattitudinal 383, 385
affect 50–2, 55–6, 146–8, 228–31, 234–5,
 513–16
 loop 513, 515
 prior models of 52, 58
 systems 51–2, 54–9, 63–5, 67–9, 515
affection 36, 422, 427
affective consequences 312, 440, 513
affective disorders 55, 69
affective forecasting 130
affective information 64, 296
affective inputs 57, 62
affective processes 69, 362, 475
affective reactions/responses 58–9, 108, 130,
 146, 295–6, 301
affective states 65, 67, 290, 292, 337
affective stimuli 53, 64–5
affective valence 209, 514
affiliations 24, 32, 34–6, 45, 463, 477
affiliative behaviors 32, 34, 37
age 20, 26, 35, 121, 151, 267–8
 preferences 20
agency 4, 317, 319, 341, 358, 367
 human 349, 358, 364–5, 367
 moral 362–4, 370
 personal 338–9, 349, 357
agenetic capabilities 356, 361, 366
agenetic theory 357, 361–2, 364
agents 293, 319, 349, 356, 358, 364
 high-level 3–9
 influence 340–1
 low-level 338–9
 opioid-blocking 36–7, 40
aggression 12, 38, 147–50, 152–3, 352,
 399–400
 modeled 352, 354
aggressive behavior 89, 353–4
agitation-related emotions 485, 490
agreement-contact semicycles 189–90
agreement effects 188–91, 196
agreements, integrative 129
alcohol 33, 102, 108–9, 441
alternative inputs 295, 300, 302
altruism 12, 14, 43, 146, 406, 477
altruistic behavior 42–3
ambiguous behaviors 206, 473

ambivalence 52–3, 59–62, 280, 283, 392, 536
ambivalent sexism theory (AST) 281
American culture 18–21, 403
American Psychologist 399, 497
amotivation 420–1, 423
ancestors 11, 14–15, 26, 402
anger 14, 23, 145–8, 152–3, 293–4, 518
antagonistic effects 56–8
antecedents 109, 140, 142, 144, 211, 280
 cognitive 141, 153
 unique 57–8
anthropology 11–13, 17–18, 401, 403
anxiety 33, 36–7, 40, 67, 69, 404
 disorders 357, 485, 498
appeals, persuasive 172, 215, 241, 250,
 254–5, 257
appetitive predispositions 51–2
applicability 2, 4–6, 87, 91–2, 236, 521
approach-avoidance models 475, 492
approach motivation 63, 105, 136
approach system 489, 492–3, 516
approach/withdrawal 56–8
aptitude 137, 139–41, 144–5, 148, 150–1, 533
argument quality 227, 229–30, 232, 237, 255
argument strength 101, 247, 303
arguments 187–9, 228–30, 232–5, 254–6, 258,
 310–11
 message 206, 208, 225, 229, 235
 strong 93, 225, 229–30, 238–9, 256, 303
 weak 230, 233, 238–9, 256, 303, 540
arousal 33, 65, 290, 336–7, 384, 391
 dissonance 380–3, 385, 391
 sexual 14, 24, 512
aspirations 119, 136, 424, 483–5, 487, 494
 intrinsic 424
 level of aspiration 119, 466, 512
assimilation 86, 206, 403, 429, 471–5
 downward 475–7
 effect 86–7, 476
 upward 475
associative learning 112, 519
associative links 104, 519
AST *see* ambivalent sexism theory
asymmetries 62, 483, 496–7
attachments 38, 44, 406, 410, 426
 security of 38, 426
attainment 76, 130–1, 166, 171, 216, 424
 goal 90, 130, 517, 527, 539–40, 542
attention 268–72, 274–7, 295–7, 318–19,
 321–2, 506–9
 direct 186, 292, 315, 321–2
 focal 405–7
 focus of 271, 294, 319
 listeners 318–19
 self-focused 336–7, 509
attention-driving devices 319, 322
attentional resources 210–11, 271
attitude change 54, 224–5, 239–41, 254–5,
 380–1, 383–5
attitude discrepant behavior 380, 383
attitude formation 169, 234, 239, 442
attitude functions 257, 438, 447

attitude objects 183, 212, 228, 232–4, 256, 442
 attitude strength 55, 233–4, 491, 536–7
 attitudes 224–9, 231–9, 253–8, 390–3, 441–5, 447–51
 expectancy-value model of 441–3, 447, 450
 explicit 102, 170
 implicit 102, 170, 453–4
 initial 227
 negative 441, 443
 personal 443, 445
 positive 256, 340
 social 441, 454
 unfavorable 442, 445, 448
 verbal 444–5
 attraction 22, 63, 128, 179–81, 187, 195–6
 effects 189–91
 attractive sources 228, 233–4
 attractive women 23–5
 attractiveness 101, 121–2, 126, 226, 251, 501
 physical 22
 attribute-based processes 271–4, 277, 284
 attribute evaluations 442, 474
 attributes 163, 228, 272–4, 442, 455, 466–70
 personal 108, 343, 358, 388, 533
 uninformative 272–3
 attribution change programs 143, 150–1
 attribution manipulations 292, 298, 301
 attribution models 209, 250
 attribution processes 141, 203–5, 539
 attribution theory 202–3, 328, 428, 465–6
 of motivation 135–53
 attributional analyses 203, 293, 296
 attributional categories 202–3
 attributional reformulation 460, 467, 469
 attributions 143–4, 146–7, 150–1, 201–3, 299–300, 342
 causal 141, 150, 203, 219
 dispositional 206–7, 275, 280
 personal 390, 539
 of success 139, 141, 153
 augmentation effects 291–2, 295–6
 automatic activation 83, 453
 automatic processing 83, 160, 471
 automatic social behavior 103
 automaticity 209–10, 283, 438, 451–2, 454, 540
 autonomous extrinsic motivation 416, 420
 autonomous motivation 420–2, 427, 430–1
 autonomous regulation 427–8, 432
 autonomy 417–30, 432
 support 418–20, 422–7, 430–2
 aversive consequences 381, 383, 385
 aversive events 357, 381–3, 385–6, 391
 aversive stimuli 55, 60, 65, 69
 aversiveness 506–7
 avoidance 51, 104, 489, 496, 514–16, 518–19
 motivation 516
 systems 16, 492–3, 497, 516
 balance interpretations 187, 194
 balance-logic theory 178–97
 balance theory 178, 184–6, 196, 227
 balanced semicycles 188–9
 befriend theory 32–7, 39, 41–5
 behavior change 150, 234, 379, 385, 431, 542
 Behavior Identification Form (BIF) 338, 342
 behavior(s)
 achievement 89, 143, 202
 activated 88–9
 actual 108–9, 444, 450
 affiliative 32, 34, 37
 aggressive 89, 353–4
 altruistic 42–3
 ambiguous 206, 473
 bad 98, 343
 change 150, 234, 355–6, 379, 385, 431
 control 112, 535
 cooperative 89, 340
 defensive/protective 65, 67
 determinants of 136, 138
 everyday 339, 451
 exploratory 37, 53, 63, 428
 flows 188, 510
 goal-directed 171, 173, 453, 508, 527, 535–6
 health 93, 392, 541
 impulsive 98–9, 103, 109–10
 inconsistent 167–8, 383
 individual 167, 270, 445
 ingroup 314, 345
 interpersonal 144, 148, 332
 modification *see* behavior, change
 motivated 135–6, 138, 140, 143–4, 148–9, 420
 negative 153, 309, 314, 345
 negotiation 215–17
 nurturant 39, 43
 observable 165, 270
 planned 93, 438–55
 positive 314, 345
 outgroup 313–14
 prediction 445, 447–50, 452
 prosocial 367, 152, 412, 416, 429–30, 432–3
 sexual 146–7
 social 11–19, 21, 25–7, 400–1, 438–40, 451–4
 verbal 312
 voting 260–1
 behavioral beliefs 440–1, 443, 448
 behavioral change *see* behavior(s), change
 behavioral choices 93, 430
 behavioral conflict 108, 535–6
 behavioral control 356, 445, 448
 perceived 438, 446–51
 behavioral decisions 105–7, 156, 170, 451
 behavioral determinants 135–6, 138, 146
 behavioral economists 99, 120
 behavioral effects 60, 90–1
 behavioral hypothesis (BH) 440–1, 443
 behavioral indicators 541–2
 behavioral intentions 93, 440, 447–8, 452
 behavioral manifestations 56–7, 102, 110
 behavioral performance 112, 438, 446–8, 453, 538
 behavioral predispositions 50, 52, 58
 behavioral responses 39, 51, 68, 515
 behavioral self-regulation 505–21
 behaviorism 349–50, 355
 beliefs 166–7, 217–19, 357–8, 441–3, 446–51, 467
 accessibility 448–9
 accessible 442, 447–9
 behavioral 440–1, 443, 448
 causal 135–6, 140–1, 143, 149–50
 cherished 258, 382, 412
 control 145, 438, 448, 451
 efficacy 357, 368
 formation 166, 204
 normative 441, 443–4, 448, 455
 benign situations 293–4, 302
 ‘better than average’ effect 186–7, 193
 BFLPE *see* big-fish-little-pond effect
 BH *see* behavioral hypothesis
 bias 11, 22–4, 63, 86, 231, 295
 hypothesis 256–7
 positivity 63, 276
 biased processing 231, 236, 256
 BIF *see* Behavior Identification Form
 big-fish-little-pond effect (BFLPE) 475–6
 big picture 123, 465
 biological mechanisms 38, 43–5
 biological signaling system 34–5
 biological stress 32–3, 44
 responses 32–4
 biology 6, 12–13, 17, 25, 365–6
 evolutionary 11–12, 18, 21, 401
 blood pressure 33, 39, 41–2
 bodily sensations/experiences 289, 292–3, 295, 302–3
 brands 304, 340
 broadcast industry 352–5
 cancer 33, 40, 393, 430
 patients 445, 468–9, 477
 capabilities 106, 122, 309, 318, 337, 446
 agentic 356, 361, 366
 caregivers 33, 35, 37–8, 43
 categorical knowledge 105, 111
 categories
 attributional 202–3
 lexical 313, 319, 321
 linguistic 322–3
 social 68, 267–8, 270, 304
 category-based processes/processing 267, 271–2, 278
 causal attributions 141, 150, 203, 219
 causal beliefs 135–6, 140–1, 143, 149–50
 causal control 140, 142, 144–5, 148
 causal inferences 149, 203–4
 causal origin 206, 320
 causal stability 139–40, 142
 causal structures 359–60
 causality 139, 141, 149, 203, 218–19, 312
 internal perceived locus of 418, 420
 uncontrollable 141–2
 causality orientations theory (COT) 420–1, 423
 central nervous system 65, 67
 central route 226, 232, 237, 255
 CET *see* cognitive evaluation theory
 change
 attitude 54, 224–5, 239–41, 254–5, 380–1, 383–5
 behavior 150, 234, 379, 385, 431, 542
 self-concept 339–40, 344
 volitional 137, 144–5
 characteristics
 personal 108, 388
 source 212, 250
 cherished-dejected dimension 490–1
 cherished beliefs 258, 382, 412
 children 15, 37–8, 120–1, 352–4, 422–4, 486
 choices, behavioral 93, 430
 chronic accessibility 75, 81–3, 85–8, 90, 487, 491
 chronic self-discrepancies 484, 487, 490, 492
 chronic strength of individuals 490–1
 closeness, psychological 472–3
 closure
 cognitive 201, 204–5, 214, 217, 259, 341
 need for 204–5, 213–19, 259, 320
 nonspecific 204–5, 320
 CLT *see* construal level theory
 coactivation of positivity 52–3, 61
 coasting 505, 516–17
 coffee drinking 167, 333, 452
 cognition 24, 50, 315–19, 321–2, 377–8, 388
 death-related 398, 406–7
 human 293, 297, 304
 social *see* social cognition
 cognitive antecedents 141, 153
 cognitive capacity 100–1, 103, 108–9, 206–7, 219, 248–9
 cognitive closure 201, 204–5, 214, 217, 259, 341
 cognitive dissonance 240, 465
 theory 377–94, 463, 507
 cognitive effort 247, 253, 393, 451
 cognitive evaluation theory (CET) 418–19, 423
 cognitive foundations 442–3, 447–8
 cognitive inconsistency 377, 387
 cognitive load 207–8, 236, 317
 cognitive modeling 277, 352, 475
 cognitive operations 156, 169, 294, 297, 299–300, 352
 cognitive processes 24, 51, 316–17, 350, 475, 539
 cognitive properties of language 309, 312
 cognitive psychology 25, 76, 78, 87–8, 100, 269
 cognitive representations 290, 378
 cognitive resources 24, 108, 170, 207, 211, 232
 cognitive scales 190–1
 cognitive structures 157–8
 cognitive theory, social 6, 201, 349–70, 439, 446
 cognitive tuning 297, 529–30, 534, 537
 commitments, goal 541–2
 communications
 health 235, 240

persuasive 109, 169, 227
 communities, linguistic 310, 321–3
 comparison
 downward 460, 464, 467–9, 475–6
 social *see* social comparison
 upward 464, 468–9, 471, 474, 476, 478
 competing predictions 84–5
 conceptual learning 422, 519
 conceptual replications 85, 275, 278, 298
 conceptual structures 109, 153
 conceptualization 5, 43, 122, 157, 159–60,
 167–72
 concrete predicates 314, 322
 conditioning, evaluative 110, 209
 conditions, high relevance 232–3, 256–7
 confidence
 actual 248–9, 253
 desired 248–50, 253, 258, 260
 gap 248–9, 258
 judgmental 247–9, 253, 258, 261
 conflicts 67–8, 97–100, 103, 111, 344–5, 520
 behavioral 108, 535–6
 intergroup 26, 218, 398, 403, 408–9
 social 127, 129, 344
 conformity 25, 178, 187, 191–3, 399, 444
 conscious deliberation 164, 452–3
 consensus, social 237, 301
 consensus information 256–8, 260
 conservatism, political 216, 218
 constitutional 118, 121–5, 127–32, 472, 538
 high-level 120–2, 124–32, 509
 low-level 120–2, 124, 126, 128, 130–2
 mental 120, 123, 128, 283, 538
 temporal 119–21, 333
 construal level theory (CLT) 118–32
 construction
 knowledge 204, 218–19
 theory 4–5, 202, 328–30, 488, 490
 contacts, social 33–6, 40–2, 44–5
 content
 domains 136, 148, 208
 message 254–6
 stereotype 267–85
 contexts
 action 333, 339
 autonomy-supportive 418–20, 431
 controlling 419–20
 interpersonal 144, 314, 317, 319, 321, 432
 persuasion 230, 232, 235–6, 252–3, 257
 situated 317, 320, 322
 situational 166, 168, 530
 social 69, 99, 345, 416, 420–1, 423–5
 contingencies, self-esteem 407, 421
 continuum
 elaboration 226, 228, 232, 236, 238–9
 model 267–85
 thinking 228, 234
 valence 56–7, 59, 68
 contrastive self-evaluations 473–4
 control
 action 335, 337, 541
 actual 446–8, 532
 behavior 112, 356, 445, 448, 535
 beliefs 145, 438, 448, 451
 causal 140, 142, 144–5, 148
 condition 295, 354, 409, 411, 454, 469
 degree of 445–6, 449
 effortful 519–20
 ego 98–9
 external 137, 421–2
 factors 447–8
 groups 474–5, 533
 no-comparison 468, 475
 illusion of 364, 538–9
 locus of 137–9, 358
 mental 328, 334, 336, 342
 perceived 93, 276, 438, 446–51, 472
 personal 145–6, 152, 537
 propositional 438, 440
 reflective 100, 109
 second-order 364–5
 stimulus 360–1
 theory 361–2

controllability 142, 146–7, 149, 533
 causal 145, 148
 controllable causes 141, 144, 148
 controlled motivation 420–3, 428
 controlled processes 100–1, 453–4
 controlled regulation 427–8
 cooperation 12, 90, 284
 cooperative behavior 89, 340
 coordination 335–7
 corrections 86–7, 101, 111
 costs 25, 43, 64, 66, 166, 187–8
 COT *see* causality orientations theory
 counterattitudinal advocacy 380, 383–4, 388,
 390–1
 creativity 293, 352, 411, 538
 tests 528, 531, 534
 credibility, source 54, 211–12, 238–40, 256
 critical cues 23, 542
 crowds 23–4, 99, 360, 424, 475
 cues 125, 206–8, 232–4, 239–40, 268–70,
 339–40
 critical 23, 542
 environmental 103, 302–3
 heuristic 205, 253–6, 259, 296
 source 250–1, 254, 256, 260
 cultural norms 19–21, 26, 281
 cultural values 411, 425
 cultural worldviews 402–3, 412
 cultures 11–12, 18–19, 366–8, 390–1, 409–10,
 424–6
 American 18–21, 403
 collectivist 391
 individualistic 391–2
 Western 391, 401, 425, 462
 cybernetics 506–7

DAVs *see* Descriptive Action Verbs
 death 20, 32, 36, 42–3, 398–402, 404–12
 reminders of 385, 398, 404–10
 unconscious thoughts of 405, 407
 death-related cognition 398, 406–7
 death-thought accessibility 404–5
 decision freedom 380, 382
 decisions
 behavioral 105–7, 156, 170, 451
 goal 526, 528–9, 532, 534–5, 542
 health 406–7
 defense motivation 249, 257–8
 defensive/protective behaviors 65, 67
 dehumanization 282, 343, 364
 dejection-related emotions 485, 490
 delay 84, 88, 99, 127, 405–7, 520
 of gratification 119–20, 127, 518
 long 84, 92
 short 85–6, 92
 deliberating participants 528–31, 533–4
 deliberation 451, 528, 532
 conscious 164, 452–3
 deliberative mindset 528, 530–9
 participants 534–6
 deliberative mode 251, 519
 deliberative processing 251, 471, 519
 dependency, outcome 275–6
 depletion 108, 416, 427
 depression 37, 39–40, 69, 357, 484–5, 517–18
 Descriptive Action Verbs (DAVs) 312–13
 design 5, 14, 23, 64, 84, 274
 processes 313–15, 318, 321
 desirability 107, 121, 126, 194, 527–31, 536
 desirability-related information 529, 531, 537
 desired confidence 248–50, 253, 258, 260
 desired end-states 77, 90, 489, 492–6, 500
 determinants 103, 107, 139–41, 143, 350–1,
 438–9
 behavioral 135–6, 138, 146
 development
 evolutionary 65, 188
 personal 123, 349, 366
 developmental psychology 15, 76, 416, 519
 devices
 attention-driving 319, 322
 linguistic 312, 318–19, 321–2
 diets

healthy 393, 422, 444, 449, 454
 low-fat 443, 448
 differences
 cultural 322, 390
 sex 18–21, 24, 42
 strategic 483, 493, 495–6
 differential parental investment 18–19, 21
 difficult tasks 136, 139, 143, 337, 521, 534
 difficulty
 personal 332, 336–7, 342
 task 138–40, 202, 206, 211
 diffusion of responsibility 363–4
 direct attention 186, 292, 315, 321–2
 direct experience 122, 212, 350
 disagreement 190, 195–6, 216, 249, 385, 467
 discomfort 231, 384, 389, 408
 discount rate 99, 120
 discrepancies
 actual-ideal 485, 490–1, 498, 500–1
 actual-ought 485, 487, 490–1, 498
 discrepancy reduction 512–13
 discriminations 24, 27, 51, 110, 283, 499
 disease model 349, 362
 disengagement 364, 509–12, 517–18
 limited 511–12
 moral 363–4, 370
 disgust 280, 282, 284
 dislikes 82, 178, 180–1, 183, 188–90, 195–6
 dispositional attributions 206–7, 275, 280
 dissimilarity 122, 124, 127–8, 180, 215, 471–3
 dissonance 178, 194, 377–94
 arousal 380–3, 385, 391
 cognitive *see* cognitive dissonance
 New Look model 380–4
 processes 382–3, 385–6, 388, 391, 393–4
 reduction 194, 383–4, 389
 theory 12, 229, 239, 378–81, 384, 393
 vicarious 377, 388, 390–1, 393–4
 distance 64, 104, 118, 123–6, 128–9, 513
 dimensions 118, 124–5, 131–2
 psychological 118, 122, 124–6, 131
 social 122–6, 128
 spatial 122, 125–6, 128
 temporal 118–20, 122–3, 125–6, 131,
 297, 300
 distraction 227, 229–31, 234, 237, 335
 distress 36, 42, 283, 511, 521
 psychological 39–40
 separation 35–6
 distribution of reinforcement 439–40
 diversity 6, 93, 216, 235, 322, 367–8
 doctors 152, 183, 273, 443
 dominance 13, 19, 98
 hierarchies 14, 39
 downward assimilation 475–7
 downward comparison 460, 464, 467–9, 475–6
 drive concept 379, 383–4
 dual processes 103, 205, 232–3, 254–5,
 405, 538
 dual systems model 209, 211
 duality 97–8, 100–2, 111–12
 duties 483, 485, 487–8, 494
 dynamic range 52–3, 60, 64
 dynamical systems 25–6, 328, 342–4
 dynamics 98, 103, 147, 310, 327, 417
 group 460–1, 464–5
 psychological 103, 283

early family environment 38–9
 early mortality 32, 42
 economics 6–7, 22, 27, 76, 99, 119–20
 education 212, 241, 354, 357, 416, 419
 efficacy 213, 217, 357, 368
 beliefs 357, 368
 personal 357–8, 368
 effortful control 519–20
 effortful processing 254, 261, 520
 ego control 98–9
 elaboration continuum 226, 228, 232, 236,
 238–9
 elaboration likelihood model (ELM) 54,
 224–42, 255, 400

'elderly' study 88–9, 104–5
 elections 131, 174, 258, 261, 409, 444
 ELM *see* elaboration likelihood model
 emergence
 avenue of 333–4, 339
 process 333, 340, 344
 scenario 339–41, 344
 emotion research 51, 293
 emotional contagion 50, 55
 emotional experiences 144, 489, 491
 emotional expressions 11, 23
 emotional life 141, 143
 emotional reactions/responses 27, 57, 145, 282, 463, 485
 emotional states 24, 202, 322
 emotional support 40, 43
 emotional vulnerabilities 485–6
 emotions 51–2, 54–8, 144–9, 235, 297–9, 336
 agitation-related 485, 490
 dejection-related 485, 490
 negative 35, 51–2, 108, 142, 336–7
 positive 35, 51–2, 54, 228, 231, 337
 empathy 42, 215, 217, 341, 412
 encoding 23, 104, 159, 269, 283–4, 531
 end-states
 desired 77, 90, 489, 492–6, 500
 undesired 492–3, 496
 endorsements 260–2, 354
 energy 19–20, 33, 83, 103, 108, 427
 engagement 33, 422, 431–2, 497, 499, 501
 engineers 207, 272, 515
 entities 203, 315, 408, 467, 509
 environmental cues 103, 302–3
 environments 14–15, 39, 76–7, 105–6, 358–9, 444–5
 social 15, 43, 69, 217, 317–18, 416–18
 envy 280–1, 284
 epistemic authority 201, 205, 211–13, 217, 219
 self-ascribed 212–13
 epistemic questions 289, 293–4, 297
 epistemic theory, lay 201–19
 equal weights model 195–6
 equivalence, functional 329–30
 errors 92–3, 185, 189, 323, 507, 514
 ESM *see* evaluative space model
 evaluations
 attribute 442, 474
 explicit 108–9
 negative 98, 180, 292, 404
 evaluative activation 53, 58–60, 63
 postulate 56, 58
 evaluative conditioning 110, 209
 evaluative judgment 228, 289, 292
 evaluative processes 50–1, 53, 59, 102, 364
 evaluative space model (ESM) 50–69
 events
 aversive 357, 381–3, 385–6, 391
 external 417, 420
 interpersonal 309, 316
 priming 75, 86–7
 social 312, 318–19, 322
 unwanted 381–3
 everyday life 4, 185, 284, 327, 451, 541–2
 evolutionary psychology 26–7
 evolutionary social psychology 11–12, 14–15, 17–18, 25
 evolutionary theorists 12, 14, 21
 evolutionary theory 6, 11–27, 34
 excitement 378, 515
 exogenous administration of oxytocin 36, 39–40, 44
 expectancies 76, 135–43, 149–51, 189–90, 274–5, 510–12
 expectancy shifts 139, 142
 expectancy-value models
 of attitudes 441–3, 447, 450
 of motivation 506, 510
 expectancy/value theory 136, 139, 143
 experience 104–6, 212–13, 289–92, 299–301, 334–7, 485–6
 direct 122, 212, 350
 hedonic 489, 513
 human 328, 345, 405, 428, 432
 metacognitive 289, 292–5, 297–300, 302
 perceptual 100–1, 111
 personal 5, 120, 213, 225, 290
 subjective 33, 101, 289, 291, 402
 experiential information 296–7
 experimental psychology 112, 428
 experimental social psychology 12, 118, 274, 428
 experimenters 78, 81, 192, 215–16, 389–90, 440–1
 experiments 137–8, 190–2, 272–4, 319–20, 322, 463–8
 laboratory 17, 19, 21, 60, 416, 419
 explicit attitudes 102, 170
 explicit evaluations 108–9
 explicit instructions 80, 303
 exploratory behavior 37, 53, 63, 428
 expressions
 behavioral 50, 68
 emotional 11, 23
 external causes 138–41
 external control 137, 421–2
 external events 417, 420
 external rewards 202, 422
 extrinsic motivation 418, 421–3, 429, 431
 autonomous 416, 420
 internalization of 421–3, 431
 extrinsic rewards 416–17, 419
 faces 23–4, 440
 failure, achievement 148, 150
 failure feedback 490
 faith 403, 412
 fame 404, 424
 familiarity 180, 268, 294, 301
 families 11, 14, 16, 38, 58, 500–1
 family members 69, 152, 491
 fans 282, 379, 399
 favoritism 216, 499
 feasibility 107, 121, 126, 527–30, 533
 perceived 527, 536
 feasibility-related information 529, 531–4, 537–8
 feedback 105, 334, 339–40, 418–19, 494, 508
 loops 343, 507–8, 511–13, 516
 negative 361–2, 516
 positive 418, 537–8
 positive 417–20, 430, 512
 processes 505, 508, 512–13
 social 339–40, 344
 feelings
 incidental 294–6, 302
 mixed 60–1
 negative 60, 291, 505, 511, 513–14, 516
 feelings-as-information theory 289–304
 fellow group members 377–8, 389–90, 393–4
 females 13, 15–16, 18–23, 25, 27, 486
 'fight or flight' 32–3, 42, 404
 fixed-pie assumptions 259–60
 flexible thinking 411–12
 flexors 65, 67–8
 fluency 101, 107, 289, 301–2
 processing 294, 300–3
 speech 335–7
 focal attention 405–7
 focus 269–72, 294–6, 298–9, 310–11, 428–9, 472–3
 prevention 91, 490–3, 495–500
 promotion 490–3, 495–7, 499
 regulatory 25, 88, 483–501
 foods 16, 18, 20, 64, 88, 105
 poisonous 16, 24
 formation
 attitude 169, 234, 239, 442
 belief 166, 204
 impression 63, 157–8, 166–7, 269–70, 272–3, 278–9
 knowledge 201, 205–6, 211, 213, 218–19
 freedom 344, 356, 380, 412, 507, 518
 decision 380, 382
 freezing 214, 218
 friendship 12, 21, 427, 493 *see also* befriend theory
 functional approach 121–2, 315, 323
 functional architecture 51, 67, 69
 functional equivalence 329–30
 functions
 activation 50, 52–3, 55–6, 58, 61–4
 adaptive 11, 17–18, 518
 attitude 257, 438, 447
 fundamental dimensions 267, 281, 284
 fundamentalists, religious 27, 407, 410–11
 GCT *see* goal content theory
 gender 202, 268, 270, 278, 284, 322
 generality, level of 123–4, 129
 genes 22, 25–6, 38–9, 43–4, 188
 goal attainment 90, 130, 517, 527, 539–40, 542
 goal commitments 541–2
 goal content theory (GCT) 424
 goal decisions 526, 528–9, 532, 534–5, 542
 goal-directed actions/activity/behaviors 89, 162, 171, 173, 527–9, 535–6
 goal-directed processing 161–4, 166, 171
 goal-directed responses 540–1
 goal implementation 526, 534, 540–2
 goal intentions 540–1
 goal pursuit 91, 483, 490, 493, 497, 537–8
 goal-related constructs, accessibility of 90–1
 goal schemas 160–4, 166–71
 Goal Specification Box 160–2
 goals
 individual 298, 303
 ingroup 279–80
 low-level 131, 511
 personal 90, 339
 grades 143, 476, 511, 542
 grandchildren 20, 500–1, 531
 grandparents 500–1
 gratification, delay of 119–20, 127, 518
 group centrism 216
 group dynamics 460–1, 464–5
 group endorsements 261–2
 group members 101, 193, 388–9, 393–4
 fellow 377–8, 389–90, 393–4
 group membership 247, 260, 388
 group minds 100
 groups 279–82, 310, 388–91, 393–4, 441–4, 477
 control 474–5, 533
 social 27, 32, 35–6, 282, 388–9, 392
 societal 279–80
 guilt 98, 141–4, 341, 411, 421, 423
 habits 104, 135, 379, 383, 452
 health 33, 41, 362
 habituation 63, 440, 451–2
 happiness 23, 60, 101, 141, 235, 389
 happy individuals 302–4
 happy moods 292–3, 300, 302, 304
 happy participants 292, 297, 299, 302–3
 harmony 189–90, 193, 391
 head nodding 237–8
 health 4, 32–3, 41–4, 362, 406–7, 541–2
 behavior 93, 392, 541
 change 416, 422, 433
 benefits 35, 40–1
 communications 235, 240
 decisions 406–7
 habits 33, 41, 362
 maintenance 150–1
 mental 70, 135, 406, 537
 messages 500
 model 349, 362
 physical 45, 406, 427, 430, 441
 professionals 151–2, 441
 promotion 357, 362, 432
 psychological 424, 426
 health psychology 109, 468, 521
 healthcare 42, 382
 healthy diets 393, 422, 444, 449, 454
 heart rate 33, 365
 hedonic experiences 489, 513
 hedonism 178, 187–8

here-and-now 93, 118–19, 122, 125, 127
 heuristic, affect 235, 240, 296
 heuristic cues 205, 253–6, 259, 296
 heuristic information processing 246–62
 heuristic processing 246–9, 253, 255–6, 260, 262, 300
 heuristic-systematic model (HSM) 226, 232, 246–7, 249–50, 252–5, 257–6
 multiple-motive 249, 251
 heuristics 100–1, 205–6, 246–7, 249, 251–3, 255–61
 judgmental 100–1
 hierarchies 26, 171, 217, 331, 429, 508
 dominance 14, 39
 high accessibility 76, 84
 high chronic accessibility 88, 90, 92
 high effort 138–9, 151
 high-level agents 338–9
 high-level condition 339–40
 high-level construals 120–2, 124–32, 509
 high-level identification 340–4, 342–3
 high-level identities 333–5, 338, 340–1, 343–4
 high processing motivation 207–8
 high processing resources 206, 208
 high self-esteem 186–7, 387, 467–8
 high thinking conditions 233, 235, 238
 higher accessibility 76–7, 298, 491
 higher joint outcomes 259–60
 higher-level act identities 331–2, 334
 higher-level construals 121, 128–9, 509
 higher-level identities 330–4, 338, 341, 343–4
 hobbies 15, 338–9
 holistic impression 472–3
 homogeneity, perceived 179
 hopelessness 139, 141–2, 518
 hostility 452, 464, 468, 472
 HSM *see* heuristic-systematic model
 human agency 349, 358, 364–5, 367
 human behavior *see* behavior
 human condition 328, 357, 412
 human culture 12, 401
 human experience 328, 345, 405, 428, 432
 human judgment 205, 289–90
 human knowledge 213–14
 human language 15, 17
 human motivation *see* motivation
 human nature 12, 23, 26, 182, 365–6, 416–17
 hypothesis generation 204–5, 214
 hypothetical situations 146, 189–90, 196
 hypotheticality 118, 122, 125, 131

IAT *see* Implicit Association Test
 IAVs *see* Interpretative Action Verbs
 ideal self 187, 509
 ideal self-guides 485–6
 ideals 88, 485, 487–8, 491–2, 495–6
 personal 90, 491, 493
 identification, high-level 340, 342–3
 identification level 330, 334–6, 338, 342
 optimal 334, 336–8
 identification theory, action 327–46, 509
 identifications 4, 207, 334–8, 350, 421, 455
 identities 24, 231, 269, 329–34, 389–90, 526
 high-level 333–5, 338, 340–1, 343–4
 higher-level 330–4, 338, 341, 343–4
 lower-level 330, 332, 334–5, 339, 342, 344
 multiple 329–31
 social 249, 388
 ideo-motor action 88–9
 ideologies 217, 261, 282, 408
 ill-being 423–4
 illusions 249, 341, 358, 408, 536–7
 of control 364, 538–9
 positive 537–8
 images 104, 165, 179, 282, 321, 424
 visual 111, 163, 165, 172–3
 imbalance 181–2, 194, 196
 imitation 350–1
 immediate antecedent of behavior 438, 448
 immediate situations 13, 118–19
 immortality 400–1
 implemental mindset participants 534, 536
 implemental mindsets 526, 528–40, 542
 implemental participants 532, 534
 implementation 4, 216, 318, 369, 530–2, 537–40
 goal 526, 534, 540–2
 implementation intentions 526, 540–2
 implication, salient 465, 469
 implications 11–12, 23–7, 156–60, 163–9, 171–2, 319–20
 applied 110, 303, 521
 behavioral 187, 297
 important 26, 91, 425, 521
 psychological 315, 320
 real-world 217, 298
 self-evaluative 334, 337–8
 Implicit Association Test (IAT) 102, 108–9, 124, 283
 implicit attitudes 102, 170, 453–4
 implicit measures 108, 124, 239, 491, 519
 implicit social comparisons 465, 474
 impression formation 63, 157–8, 166–7, 269–70, 272–3, 278–9
 impression motivations 249, 257–8
 impressions 63, 78–9, 82, 86, 167–8, 268–70
 holistic 472–3
 initial 236, 472–3
 person 157–8
 impulse and reflection theory 97–112
 impulses 97–100, 103, 109, 361, 506, 520
 impulsive behaviors 98–9, 103, 109–10
 impulsive mechanisms 97–8
 impulsive system 103–7, 109, 169
 impulsivity 100, 103, 109, 338
 incentives 135–6, 139, 150, 378–9, 513–14, 517–18
 large 239, 378–9
 magnitude 380–1, 384
 incidental feelings 294–6, 302
 incidental information 535
 incidental source 294–5, 299, 301–2
 inconsistency 167, 182, 194, 196–7, 378, 381–4
 bias 274
 cognitive 377, 387
 inconsistent behaviors 167–8, 383
 independence 69, 269, 276, 425–6
 individual differences, measurable 53, 65
 individual variation 64, 338, 345
 individualism 193, 367–8, 390–2, 425
 individualistic cultures 391–2
 individuals, promotion-focused 91, 493, 495, 497
 individuating information 284, 304
 individuating processes 267, 271, 275–7
 infants 12, 15, 18, 33, 35–8, 402
 Inference Maker 160–1, 166–7
 inference rules 204–5, 210–11, 217, 300, 302
 inferences 105–6, 156, 158–60, 299–301, 316–17, 342
 causal 149, 203–4
 influence
 normative 191
 social 250, 327–8, 336, 340–1, 344–5, 350–1
 influence agents 340–1
 influence judgments 76, 247, 295
 information
 affective 64, 296
 consensus 256–8, 260
 desirability-related 529, 531, 537
 experiential 296–7
 feasibility-related 529, 531–4, 537–8
 feelings-as-information theory 289–304
 incidental 535
 individuating 284, 304
 input 79, 87, 160–1, 164
 representativeness 207–9
 self-relevant 344, 473
 information processing 11, 160, 165–6, 231, 241–2, 538–9
 heuristic 246–62
 selective 249, 529, 534–5
 social 156–74
 systematic 246–62, 540
 informational social influence 191
 informational value 289, 292, 294–6, 298–302
 perceived 293–6
 ingroup, active 382, 393
 ingroup goals 279–80
 ingroup members 182, 345, 389, 393, 499
 ingroups 131, 216–18, 262, 281, 314, 389–90
 initial impressions 236, 472–3
 initial judgments 158–9
 innate mechanisms 15–16
 innovation 351–2, 369
 input information 79, 87, 160–1, 164
 inputs 15, 66–7, 76, 87, 106, 513
 affective 57, 62
 alternative 295, 300, 302
 perceptual 103, 105
 stimulus 76, 360
 vague 86–7, 92
 instructions 81, 161–2, 166–7, 189, 213, 439–40
 explicit 80, 303
 insufficient effort 138, 141, 150–1
 integration 12, 159–60, 239–40, 416–17, 421–2, 429
 integrative agreements 129
 integrative processes 417, 429
 intended behavior 447, 540
 intentionality 98, 262, 342, 358
 intentions 83, 341–2, 407, 438, 440, 445–53
 behavioral 93, 440, 447–8, 452
 goal 540–1
 implementation 526, 540–2
 prediction of 450–1
 interactionism 359
 interactions 14–15, 88–9, 106–7, 110–11, 192, 359
 social 14, 17, 241, 315
 interdependence 271, 275–6, 282, 367, 425
 interference 67, 81, 91, 336
 intergroup biases 313, 315, 345
 intergroup conflict 26, 218, 398, 403, 408–9
 intergroup relations 498–9
 internal locus of control 339, 420
 internal perceived locus of causality 418, 420
 internalization of extrinsic motivation 416, 421–3, 429–31
 interpersonal behavior 144, 148, 332
 interpersonal contexts 144, 314, 317, 319, 321, 410, 439
 interpersonal theory 135, 146–7, 149
 interpersonal verbs 317, 319–20
 interpretation-comparison model 460, 473
 interpretations, balance 187, 194
 Interpretative Action Verbs (IAVs) 312–13
 intersubjectivity 389
 intrapersonal phenomena 214–15
 intrapersonal theory 144–6, 149
 intrinsic motivation 416–23, 425, 428–30
 investment, parental 20–1, 24
 individual comparisons 476–7
 isolation, social 32, 36, 41
 joint outcomes 129, 259–60
 judgment tasks 107, 297
 judgmental confidence 247–9, 253, 258, 261
 judgmental heuristics 100
 judgmental outcomes 224, 239
 judgments
 absolute 107, 473–4
 accurate 130, 533
 correct 192, 212, 260
 evaluative 228, 289, 292
 first 107, 158, 164
 initial 158–9
 life-satisfaction 291, 295
 perceptual 210, 393
 personal standards of judgment 386–7
 positive 292, 298
 social 78, 112, 205, 210, 214–15, 344–5
 target of 76, 289, 295, 299
 juries 241, 259, 283

knowledge
 accessibility 159, 169, 460, 473
 accessible 87, 472
 activation 79, 82, 85–8, 92–3
 categorical 105, 111
 construction 204, 218–19
 formation 201, 205–6, 211, 213, 218–19
 procedural 164, 173
 representations 159, 162, 164, 173
 scientific 4, 202, 204, 214
 stored 87, 93, 253
 structures 247, 253, 293, 452–3, 519
 pre-existing 252, 298, 304

ability, response 52, 60
 laboratory experiments 17, 19, 21, 60, 416, 419
 language 15, 17, 75–6, 128–9, 309–19, 321–3
 abstract 215, 314, 323
 cognitive properties 309, 312
 common 487–8
 interpersonal 309, 311–12, 322
 language-cognition interface 309, 314–18, 321, 323
 lay epistemic theory 201–19
 LCM *see* linguistic category model
 leaders 408–9
 learning 12, 15–16, 350–1, 387, 416–17, 430–1
 associative 112, 519
 conceptual 422, 519
 observational 350, 352, 354

life
 emotional 141, 143
 everyday 4, 185, 284, 327, 451, 541–2
 expectancy 42–3
 history 20–1
 social 4, 26, 122, 367

life-satisfaction judgments 291, 295

lifespan 32, 38–9, 410, 422, 429

likability 270, 273–4, 284

likelihood
 of goal attainment 90
 of thinking 228, 235–6

liking 82, 178–81, 188, 190, 301–2, 342–3

limited disengagement 511–12

linguistic category model (LCM) 309–23

linguistic communities 310, 321–3

linguistic devices 312, 318–19, 321–2

linguistic forms 77, 311, 351

linguistic intergroup bias 313, 315

locus of control 137–9, 358
 internal 339, 420

logic 2, 4, 24, 184–5, 187, 194
 deductive 184
 fuzzy 196
 two-valued 178, 184

logical thinking 100, 185, 187

long-term relationships 21–2, 24

longevity 41–2

loops
 action 513–15
 affect 513, 515
 feedback 343, 507–8, 511–13, 516

low ability 137–9, 150–1

low effort 138, 148, 236

low-fat diets 443, 448

low level 62, 131, 151, 248, 330, 336
 agents 338–9
 condition 339–41
 construals 120–2, 124, 126, 128, 130–2
 goals 131, 511

low motivation 208, 233

lower-level elements 344, 509

lower-level identities 330, 332, 334–5, 339, 342, 344

males 13, 19–21, 23, 25, 27, 42–3

malleability 317, 339

management, terror 398–412

managing stress 38, 44

manipulations
 attribution 292, 298, 301
 experimental 137, 146, 488

misattribution 290–1

marriages 20–1, 26, 484

MAS *see* Mind Attribution Scale

maternal nurturance 37–8, 44

mates 15, 19–20, 22, 25, 42

mating goals 24–5

members
 family 69, 152, 491
 group 101, 193, 388–9, 393–4
 ingroup 182, 345, 389, 393, 499
 outgroup 24, 345, 389–90, 393, 444, 499

membership 196, 388–9
 group 247, 260, 388

memory 23–4, 106, 156–60, 166–71, 269, 274
 capacity 108–9
 persons 168, 210–11
 span 534–5

men 21–2, 42–3 *see also* males

mental construal 120, 123, 128, 283, 538

mental control 328, 334, 336, 342

mental health 70, 135, 406, 537

mental processes 91, 294, 297, 299–300, 342, 453

mental representations 45, 52, 57, 388, 473, 540

mental systems 103, 333, 344

message arguments 206, 208, 225, 229, 235

message content 254–6

message source 224, 232

messages, persuasive 101, 157, 169, 184, 235, 255–6

meta-analysis 4, 146–7, 193, 419, 448–52, 540–1

metacognitive experiences 289, 292–5, 297–300, 302

metacognitive strategy 541–2

metaphors 26, 83, 108, 159, 166, 379

metatheory 239, 328, 345, 429, 432, 505

Mind Attribution Scale (MAS) 342

mind perception 341–3

mindsets 173, 528, 535–7, 539
 deliberative 528, 530–9
 implemental 526, 528–40, 542
 mindset theory of action phases 526–42

mini-theories 13, 400, 416, 420, 423–4, 432

misattribution manipulations 290–1

mismatches 45, 77, 130, 272, 492–3
 avoidance 492–5

mixed feelings 60–1

modeling
 cognitive 277, 352, 475
 social 350–1, 353, 367–8

momentary situations 485, 490

monetary rewards 417–18

money 76, 143–4, 146, 172, 329, 382

moods 232, 240, 289–94, 296–9, 302–3, 428
 happy 292–3, 300, 302, 304
 negative 79, 299
 sad 289, 291–3, 295, 298, 304

moral agency 362–4, 370

moral disengagement 363–4, 370

moral self-sanctions 362–3

mortality 40–3, 398, 401–2, 404–12 *see also* death
 early 32, 42
 reminders of 404–5, 408

mothers 12, 15, 18, 35–7, 39, 43

motivated behavior 135–6, 138, 140, 143–4, 148–9, 420

motivation 251–4, 274–8, 416–22, 428–32, 443–4, 526–7
 accuracy 249, 251, 255–6, 258, 260
 achievement 143–4, 489, 527
 approach 63, 105, 136
 attribution theory of 135–53
 autonomous 420–2, 427, 430–1
 controlled 420–3, 428
 defense 249, 257–8
 expectancy-value models of 506, 510
 extrinsic 418, 421–3, 429, 431
 impression 249, 257–8
 intrinsic 416–23, 425, 428–30

low 208, 233

processing 207–8, 294, 304

sequences 136, 146–7

social 144, 146–7

motivational force, strength of 496–7

motivational orientations 104–5, 109, 416

motivational states 211, 421

motivational substrates 56–7

multiple identities 329–31

multiple-motive heuristic-systematic model 249, 251

multiplicative rule 178, 182, 184–5, 188, 194, 196

multistability 344

music 300, 333, 335, 401

natural selection 11, 17, 20, 54, 64, 366

necessity 22, 43, 92, 333–4, 385, 424

need for closure 204–5, 213–19, 259, 320

need for nonspecific closure 204–5, 320

negative actions 153, 343

negative activation function 53, 63

negative attitudes 441, 443

negative behaviors 153, 309, 314, 345

negative contrast effect 475–6

negative emotions 35, 51–2, 108, 142, 336–7

negative evaluations 98, 180, 292, 404

negative feedback 418, 537–8
 loops 361–2, 516

negative feelings 60, 291, 505, 511, 513–14, 516

negative outcomes 63, 120, 148, 280–1, 486, 494–6

negative reactions 404, 408, 469

negative stereotypes 272, 499

negative stimuli 53, 61–2, 64–5, 102

negative thoughts 231, 238, 404

negative unit relations 179–80, 183

negative valence 514, 516

negativity 50, 52–4, 58–61, 63–5, 409
 bias 53, 62–6, 69

negotiation 215–17, 259, 538

negotiators 129, 259–60

nervous system 65–7, 508
 sympathetic 33, 41

networks, social 40, 69, 351, 391

neuraxial organization 53, 65, 67

neurobehavioral organizations 53, 64

neutral stimuli 57, 65, 179, 301

New Look 6, 77, 252, 323, 377, 381–6
 model of dissonance 380–4

no-comparison control groups 468, 475

nodding 237–8

nonspecific closure, need for 204–5, 320

normative beliefs 441, 443–4, 448, 455

normative influence 191

normative standards 386–7

norms 26, 93, 275, 329, 368, 423
 cultural 19–21, 26, 281
 social 98, 142, 144
 subjective 438, 443–5, 447–51

nurturance 25, 37–9, 43–4, 51

obesity 41, 146, 392, 430

objective self-awareness 12, 506

obligations 40, 483, 485, 487, 490, 494

observable behavior 165, 270

observational learning 350, 352, 354

observers 145, 148, 296, 329, 332, 352

offspring 16–22, 25, 32–5, 37–9, 42–4, 488
see also children

opioid-blocking agents 36–7, 40

opioid mechanisms 36–7

opioid system 32, 34–5

opioids 35–6, 41, 44

opponents 139, 275, 499

optimal level of identification 334, 336–8

optimality 334, 336–7

optimism 119, 407, 521, 526, 534, 537

organization
 bipolar 50, 57
 neuraxial 53, 65, 67

neurobehavioral 53, 64

oscillations 53, 61, 334, 515
 oughts 88, 90, 485, 487–8, 491–2, 495–6
 outcome dependency 275–6
 outcomes 126–7, 137–8, 275–6, 420–3, 440–2, 473–4
 behavioral 139, 474
 given 98–9, 364
 judgmental 224, 239
 negative 63, 120, 148, 280–1, 486, 494–6
 persuasion 233, 239
 positive 120, 280–1, 293, 423, 485–6, 494–6
 outgroups 24, 122, 216, 218–19, 280–1, 313–14
 members 24, 345, 389–90, 393, 444, 499
 overcorrection 86, 295
 oversimplification 180–1, 189
 overt behaviors 104, 107, 334, 444
 overt responses 156, 160–1, 171
 oxytocin 32, 34, 36–7, 39–44
 levels 36, 40
 plasma 32, 36, 42, 44

pain 16, 52, 65, 489, 491–2, 499
 physical 35–6
 social 36

parental investment 20–1, 24
 differential 18–19, 21

parenting 432, 486

parents 14–15, 212–14, 402–3, 422–4, 486–7, 500–1

participants 77–82, 123–8, 188–97, 253–62, 387–93, 530–6
 control 453, 533
 deliberating 528–31, 533–4
 experimental 86, 528, 530
 happy 292, 297, 299, 302–3
 implemental 532, 534
 motivated 256–7
 outcome-dependent 275–6
 planning 528–31, 533–4
 prevention-focused 493, 495
 promotion-focused 493, 495
 sad 292, 297, 299, 303

partners 19–22, 258, 275–6, 278, 426–7, 532–3

partnerships 369–70

patients 41, 393, 430, 445, 476–7, 485
 cancer 445, 468–9, 477

perceived control 93, 276, 438, 446–51, 472

perceived feasibility 527, 536

perceived homogeneity 179

perceived informational value 293–6

perceived relevance 210, 294, 296, 298, 302

perceived self-efficacy 357, 446, 450

perceived social pressure 443–4, 448

perceivers 78, 124–5, 271, 304, 378, 498–9

perceptions 79, 123–4, 150–2, 212–14, 252, 379–81
 of control 447, 449–51
 mind perception 341–3
 person perception 63, 76, 102–3, 120, 208, 271

perceptual experience 100–1, 111

perceptual input 103, 105

perceptual judgments 210, 393

performance 142–3, 151, 304, 337, 445–9, 466
 behavioral 112, 438, 446–8, 453, 538
 school 143, 150

performance-contingent rewards 418–19

peripheral routes 54–5, 226, 231–2, 255

Permanent Storage 161–6

perseverance 99, 446–8

persistence 140, 225, 234, 284, 417, 536

person impressions 157–8

person perception 63, 76, 102–3, 120, 208, 271

person representation 167–8

personal agency 338–9, 349, 357

personal attitudes 443, 445

personal attributes/characteristics 108, 343, 358, 388, 533

personal attributions 390, 539

personal control 145–6, 152, 537

personal development 123, 349, 366

personal difficulty 332, 336–7, 342

personal efficacy 357–8, 368

personal experience 5, 120, 213, 225, 290

personal goals 90, 339

personal history 110, 141–2, 157, 174, 246, 250

personal responsibility 145–6, 152, 343, 363, 382, 393

personal standards 327, 358, 361–2, 386–7, 487

personal values 284, 385

personality 97–8, 311–12, 335–8, 357–8, 489–90, 508–9
 psychology 81, 358, 506, 512, 521
 theory 157, 358–9, 488
 traits 157, 167, 270, 327, 338–9, 358

persuasion 101, 224–6, 228–35, 237–41, 254–5, 303–4
 context 230, 232, 235–6, 252–3, 257
 effective 217, 241
 research 205, 208, 215, 304
 theory 225, 235, 253
 variables 228, 234, 250

persuasion paradigm 236, 253

persuasive appeals 172, 215, 241, 250, 254–5, 257

persuasive communications 109, 169, 227

persuasive messages 101, 157, 169, 184, 235, 255–6

pets 36, 339

physical activities 41, 151, 428, 430, 432, 449–51

physical health 45, 406, 427, 430, 441
 outcomes 38, 41

physical pain 35–6

pity 148, 280–1

planned behavior 93, 438–55

planning 91, 119, 327, 334, 528–33, 538–40
 fallacy 119, 121
 task 529, 539

plasma oxytocin 32, 36, 42, 44

pleasantness 37, 184, 188–90, 290

poisonous foods 16, 24

politeness 127–9, 453

political conservatism 216, 218

political science 7, 12, 76, 119

politics 241, 355, 403, 408, 444

polls 257–8, 260

positive activation function 53, 63

positive attitudes 256, 340

positive behaviors 314, 345

positive emotions 35, 51–2, 54, 228, 231, 337

positive feedback 417–20, 430, 512

positive illusions 537–8

positive judgments 292, 298

positive outcomes 120, 280–1, 293, 423, 485–6, 494–6

positive outgroup behaviors 313–14

positive relations 36, 182, 216, 353–4

positive self 187–8, 190–1, 194

positive self-evaluation 187, 190

positive sentiment relations 178, 182, 186

positive unit relations 178–80, 186, 189, 194

positive valence 494, 514, 516

positivity 50, 52–4, 58–61, 63–5, 67, 128
 bias 63, 276
 coactivation of 52–3, 61
 offset 53, 62–6, 69

postaction phase 527–8

poverty 88, 147, 368, 370

power
 social 127–8, 147–9
 theory of 277–8

power-as-control theory 278, 283

powerholders 277–8

pre-existing knowledge structures 252, 298, 304

preaction phase 527–9

predators 60, 64, 188

predecision phase 527–9

predicates
 abstract 313, 322
 concrete 314, 322

predicted attitudes 183, 242

predictions 68–9, 84–5, 125–7, 129–30, 291–2, 450
 of behavior 445, 447–50, 452
 competing 84–5
 of intentions 450–1

predictive power 349, 355, 358

predispositions
 appetitive 51–2
 behavioral 50, 52, 58
 psychological 11, 14, 26

preferences 20, 22–3, 36, 179, 332–4, 340

pregnancy 18, 39–40

prejudice 11–12, 26–7, 101, 241, 283–4, 398–400

presidential candidates 218, 260–1, 279

prevention 25, 408, 483, 492–7, 499–500
 focus 91, 490–3, 495–500
 states 490–2
 strength 88, 491–2
 system 88, 492–4, 497

pride 136, 139–42, 267, 280, 399

priming 75–8, 80–1, 83–92, 453–4, 471–2, 478–9
 effects 80, 83, 88
 events 75, 86–7
 recent 84–7, 91–2
 situational 78, 89, 92, 487
 subliminal 89–90
 verbal 78–80

proattitudinal advocacy 383, 385

probabilities, subjective 440–1, 447–8, 455

problem solving 269, 352, 488

procedural knowledge 164, 173

processing, top-down 293–4, 302–4

processing fluency 294, 300–3

processing motivation 207–8, 294, 304

processing objectives 162–3, 166, 173

processing resources 205–8, 219

processing strategies 293–4, 298, 303–4

processing style 293–5, 297–8, 302–4

processing units 159–62, 164
 special-purpose 161–2, 169, 171

promotion 25, 88, 270, 283, 483, 489–501
 focus 490–3, 495–7, 499
 health 357, 362, 432

promotion-focused individuals 91, 493, 495, 497

propositional categorization 105–6

propositional control 438, 440

prosocial behavior 36–7, 152, 412, 416, 429–30, 432–3

protection 20, 35, 39, 281, 398, 402

prototypes 417, 425

proxies 368, 382, 447–8, 470–1, 476

proximal perspective 129–31

proximity, psychological 37, 124

proxy model 460, 467

psychodynamic theory 98, 355

psycholinguistics 75, 87

psychological adjustment 40, 426

psychological closeness 472–3

psychological distance 118–19, 121–8, 131
 dimensions 118, 122, 124–6, 131

psychological distress 39–40

psychological dynamics 103, 283

psychological predispositions 11, 14, 26

psychological processes 45, 101, 311–12, 314–16, 321, 527–8

psychological proximity 37, 124

psychological reactance 231, 379, 507

psychological reality 318, 322

psychological security 402–3, 408, 410

psychological well-being 417, 425–6, 430, 432

psychotherapy 241, 349, 355, 392–3, 416, 431
 punishment 98, 136, 144, 148, 150, 202

question-answer paradigm 319–20
 quiescence 496–7
 quiescent-agitated dimension 490–1

rats 16, 36, 39, 350, 353, 379
 reactance 231, 379, 507, 510
 reactions 14, 102, 130, 144–6, 181–2, 406
 affective 58–9, 108, 130, 146, 295–6, 301
 behavioral 67, 130, 148
 emotional 27, 57, 145, 282, 463, 485
 negative 404, 408, 469
 readiness 15, 76, 217–18, 338, 430
 realism 537–9
 reality
 psychological 318, 322
 shared 93, 131, 205, 216, 330, 488
 social 191, 350–1
 reasoned action 438–45, 447, 451–4
 models 438, 449–50, 453–4
 recalibrations 53, 64, 514
 recall 82, 167–8, 174, 291–2, 294, 299–300
 recency 90, 104, 106, 163, 169, 300–1
 recent activation 78, 84
 recent evidence 323, 486
 recent priming 84–7, 91–2
 recipients 45, 172–3, 205, 216–17, 240, 303
 reciprocal activation 50, 52, 55–6, 58, 60
 recognition 136, 165, 332, 401, 463, 468
 referents 163, 165–7, 169–70, 441, 443, 448
 social 391, 443–4, 455
 reflection, impulse and reflection theory 97–112
 reflective control 100, 109
 Reflective-Impulsive Model (RIM) 97, 103–11
 reflective operations 106–9
 reflective processes 98, 107
 reflective system 103, 105–8, 169, 519–20
 reformulation, attributional 460, 467, 469
 refreezing 214, 218
 regulation
 autonomous 427–8, 432
 behavioral 416, 421
 cognitive 357, 364
 of social behavior 314–15
 regulatory fit theory 483, 493, 495, 497, 500
 regulatory focus 483–501
 regulatory systems 32–3, 44
 reinforcement 69, 136, 157, 301, 350, 440
 distribution of 439–40
 relatedness 417, 421, 423–7, 429, 432
 relations
 intergroup 498–9
 interpersonal 11, 93, 178, 321, 410, 439
 positive 36, 182, 216, 353–4
 sentiment 178–9, 182, 186, 190
 social 32–4, 36, 41, 127, 129, 320
 unit 178–80, 182
 relationships
 interpersonal 11, 321, 410
 long-term 21–2, 24
 personal 43, 218, 426
 romantic 21, 426–7, 532, 534
 social 32–4, 36, 41, 367, 390
 relative accessibility 76, 170
 release 16, 33, 354
 relevance
 conditions 232–3, 256–7
 perceived 210, 294, 296, 298, 302
 personal 227, 229–33, 253, 255–6, 260
 subjective 208, 210–11
 Relevance of a Representation (ROAR) 91
 reliance 185, 206, 232, 259, 296, 303–4
 relief 515–17
 religion 241, 406, 432, 445
 religious fundamentalists 27, 407, 410–11
 reminders of death/mortality 385, 398, 404–10
 replications, conceptual 85, 275, 278, 298
 representations 91, 111, 139–40, 159–60,
 162–4, 167–8
 cognitive 290, 378
 dual 100, 102, 111

knowledge 159, 162, 164, 173
 mental 45, 52, 57, 388, 473, 540
 person 167–8
 representativeness information 207–9
 residential colleges 389–90
 resources 14, 19–22, 25, 32, 209, 211
 attentional 210–11, 271
 cognitive 24, 108, 170, 207, 211, 232
 limited 108, 520
 processing 205–8, 219
 response lability 52, 60
 response selectors 161, 171
 responses
 adaptive 51–2, 54, 66–7
 affective 58, 272, 295–6, 469, 516
 behavioral 39, 51, 68, 515
 emotional 57, 463, 487
 flight 33, 42
 goal-directed 540–1
 overt 156, 160–1, 171
 social 32, 34–5, 41, 44–5
 stress 32–4, 37–8, 41–2
 responsibility 98, 145–6, 148–9, 152–3, 341–2,
 381–2
 diffusion of 363–4
 personal 145–6, 152, 343, 363, 382, 393
 retrieval 23, 68, 156, 163–4, 166, 531
 processes 159, 162–3, 168, 290
 rewards 103, 148–9, 187–8, 290, 379–80,
 417–20
 external 202, 422
 extrinsic 416–17, 419
 monetary 417–18
 performance-contingent 418–19
 task-contingent 418–19

RIM *see* Reflective-Impulsive Model
 risk 25, 33, 39, 42, 301, 533
 ROAR *see* Relevance of a Representation
 romantic relationships 21, 426–7, 532, 534
 routes 54, 101, 156, 194, 225, 255
 central 226, 232, 237, 255
 peripheral 54–5, 226, 231–2, 255
 routinization 209–10, 219

sad moods 289, 291–3, 295, 298, 304
 sad participants 292, 297, 299, 303
 sadness 60, 235, 294, 389, 515, 518
 SAEA *see* self-ascribed epistemic authority
 salience 191, 207, 272, 424
 SAM *see* selective accessibility model
 SAVs *see* State Action Verbs
 schoolwork 268, 422–3
 scientific knowledge 4, 202, 204, 214
 SCM *see* stereotype content model
 scripts 164, 302–3
 SDT *see* self-determination theory
 second-order control 364–5
 secondary tasks 108, 237, 303
 security 25, 91, 403, 426, 494, 499
 of attachment 38, 426
 psychological 402–3, 408, 410
 selection
 natural 11, 17, 20, 54, 64, 366
 sexual 13, 18–19, 24
 selective accessibility 107, 473
 model (SAM) 460, 472–5
 selective information processing 249, 529,
 534–5

self
 actual 187, 485, 488
 ideal 187, 509
 physical 407
 positive 187–8, 190–1, 194
 self-actualization 391, 420, 429
 self-ascribed epistemic authority (SAEA)
 212–13
 self-assessment 339–40, 460
 self-awareness 247, 262, 402,
 506–11, 513
 objective 12, 506
 theory 506–9
 self-completion theory 526–7
 self-concept change 339–40, 344

self-concepts 186, 275, 327–8, 339, 344–5, 473
 academic 475–6
 self-control 103, 108, 127, 131, 427, 520
 self-definition 339, 473
 self-descriptions 340, 391
 self-determination theory (SDT) 416–33
 self-development 351, 356, 368
 self-discrepancies 394, 484–5, 487–8, 490
 chronic 484, 487, 490, 492
 self-discrepancy theory 88, 483–9, 491–2, 494,
 498, 500
 testing 487–8
 self-doubt 174, 368
 self-efficacy 357–8, 368, 446–7, 450, 477
 perceived 357, 446, 450
 self-enhancement 460, 467, 475, 478
 self-esteem 141–2, 387, 398–9, 403–4, 407–8,
 446–5
 contingencies 407, 421
 high 186–7, 387, 467–8
 low 196, 387, 467
 self-evaluations 335, 460, 466–7, 469,
 472–5, 478
 contrastive 473–4
 negative 196, 506
 positive 187, 190
 self-evaluative implications 334, 337–8
 self-focus 507, 509–12
 self-focused attention 336–7, 509
 self-guides 88, 483–90, 494
 ideal 485–6
 self-handicapping 399, 428
 self-improvement 127, 478
 self-interest 361, 369
 self-organization 26, 344
 self-perception 104, 150, 232, 239–40, 250–1
 theory 105, 239, 250, 418
 self-protection 24–5, 466
 self-regulation 88, 360–2, 427, 483–6, 488–9,
 492–5
 autonomous 427, 432
 behavioral 505–21
 systems 489–90, 492, 494
 self-regulatory skills 360, 362
 self-relevant information 344, 473
 self-relevant semicycle 190–1, 196–7
 self-reports 193, 282, 322, 519, 542
 self-sanctions, moral 362–3
 self-standard model 377, 386–7, 394
 self-system therapy 489, 498
 self-validation hypothesis 237–8
 semantic concepts 161, 163, 165
 semantics 163, 165, 301, 310–13
 semicycles 182, 185–6, 188–92, 195–6
 agreement-contact 189–90
 balanced 188–9
 manipulated 189–90
 self-relevant 190–1, 196–7
 sensations, bodily 289, 292–3, 295, 302–3
 sensitivity 39–40, 53, 64, 337
 sentences 316, 319, 363, 404, 439–40, 453
 sentiment relations 178–9, 182, 186, 190
 separate situations 78, 80
 separation distress 35–6
 sex differences 18–21, 24, 42
 sexual arousal 14, 24, 512
 sexual behavior 146–7
 sexual selection 13, 18–19, 24
 shame 141–4, 423
 shared reality 93, 131, 205, 216, 330, 488
 shift attributions 150, 152
 shock(s) 42, 209, 399, 463
 signals 32, 42, 293–4, 302, 304, 344
 situated contexts 317, 320, 322
 situated features 311, 313, 322
 situation models 165, 172–3
 situational contexts 166, 168, 530
 situational priming 78, 89, 92, 487
 skills 5, 137–9, 334–5, 362, 446, 533
 self-regulatory 360, 362
 socioemotional 38–9, 44
 SN *see* subjective norms
 SNS *see* sympathetic nervous system

sociability 270, 280, 338, 343
 social accommodation 130–1
 social attitudes 441, 454
 social behavior 11–19, 21, 25–7, 400–1, 438–40, 451–4
 regulation of 314–15
 social categories 68, 267–8, 270, 304
 social cognition 11, 87, 251–2, 267–9, 315–16, 471
 research 290, 465
 social cognitive theory 6, 201, 349–70, 439, 446
 social comparison 144
 implicit 465, 474
 processes 460–1, 466
 theory 191–2, 460–79
 social conflicts 127, 129, 344
 social consensus 237, 301
 social contacts 33–6, 40–2, 44–5
 social contexts 69, 99, 345, 416, 420–1, 423–5
 social discrimination 92, 499
 social distance 122–9, 314
 social environments 15, 43, 69, 217, 317–18, 416–18
 social events 312, 318–19, 322
 social feedback 339–40, 344
 social groups 27, 32, 35–6, 282, 388–9, 392
 social identity 249, 388
 theory 388–9
 social influence 250, 327–8, 336, 340–1, 344–5, 350–1
 informational 191
 social information processing 156–74
 social interactions 14, 17, 241, 315
 social isolation 32, 36, 41
 social judgment 78, 112, 205, 210, 214–15, 344–5
 social life 4, 26, 122, 367
 social modeling 350–1, 353, 367–8
 social motivation 144, 146–7
 social networks 40, 69, 351, 391
 social norms 98, 142, 144
 social power 127–8, 147–9
 social pressure, perceived 443–4, 448
 social psychology *see Introductory Note*
 social reality 191, 350–1
 social referents 391, 443–4, 455
 social relations/relationships 32–4, 36, 41, 127, 129, 320
 social responses 32, 34–5, 41, 44–5
 social structure 280–2, 349, 367
 social support 32, 35, 40–1, 43–5, 69, 135
 social systems 349–50, 363–5
 social threats 32, 34–5
 societal groups 279–80
 socioemotional skills 38–9, 44
 solace 34–5, 38
 source characteristics 212, 250
 source credibility 54, 211–12, 238–40, 256
 source cues 250–1, 254, 256, 260
 sources 85, 92, 210–12, 226–8, 232–3, 289–96
 of accessibility 85, 92–3
 attractive 228, 233–4
 incidental 294–5, 299, 301–2
 message 224, 232
 unrelated 290, 292, 303
 space 51, 58, 124, 163, 172, 322
 evaluative 51, 58–9, 67
 spatial distance 122, 125–6, 128
 spatial learning 65
 special-purpose processing units 161–2, 169, 171
 speech 336–7, 385, 389–90, 393, 404, 497
 counterattitudinal 390–1
 fluency 335–7
 spinal cord 53, 65, 67, 450
 reflexes 50–1, 65, 67
 spirit 97–8, 101, 173, 330, 477
 sports 121, 268, 416, 432, 469
 stability 140, 142, 149, 189–90, 339, 345
 causal 139–40, 142
 standards 2, 4, 386, 486–8, 506–7, 512
 normative 386–7
 personal 327, 358, 361–2, 386–7, 487
 State Action Verbs (SAVs) 312
 State Verbs (SVs) 312, 317, 319–20
 stereotype content model (SCM) 267–85
 stereotype threat 400, 499
 stereotypes 101, 104, 236, 271–4, 282–3, 313
 cultural 23
 default 281–2
 elderly 104–5
 ethnic 452
 negative 272, 499
 stereotyping 12, 92, 101–2, 236, 277, 303–4
 stigmas 146–7, 150, 152
 stimuli 51–4, 56–8, 60–1, 63–5, 102, 121
 affective 53, 64–5
 aversive 55, 60, 65, 69
 hospitable 50–1, 54, 67
 negative 53, 61–2, 64–5, 102
 neutral 57, 65, 179, 301
 target 52, 57, 59, 102, 472
 stimulus control 360–1
 stimulus inputs 76, 360
 stimulus sets 273, 311
 stock market 298, 301–2, 477
 storage 83–5, 162–4, 290
 memory 156, 159, 166
 permanent 161–6
 storage bin model 82–3
 stored knowledge 87, 93, 253
 strangers 15, 17, 24, 40, 122, 124
 strategic differences 483, 493, 495–6
 strength
 argument 101, 247, 303
 attitude 55, 233–4, 491, 536–7
 of motivational force 496–7
 prevention 88, 491–2
 stress 32–5, 37–45, 52, 55, 102, 464
 biological 32–3, 44
 responses 32–4, 37–8, 41–2
 stressors 34, 41
 subjective experiences 33, 101, 289, 291, 402
 subjective norms (SN) 438, 443–5, 447–51
 subjective probabilities 440–1, 447–8, 455
 subjective relevance 208, 210–11
 subjective values 440–1, 455
 subjective wellbeing 460, 467, 477
 subliminal priming 89–90
 substance abuse 42, 357, 410
 substitutable tasks 90–1
 substrates, motivational 56–7
 success 136–44, 469–70, 483, 490–2, 496–7, 510–11
 attributions of 139, 141, 153
 expectancy of 139, 151
 sufficiency threshold 212, 248, 253
 sunscreen 393, 500
 superstars 469, 476
 support
 autonomy 418–20, 422–7, 430–2
 emotional 40, 43
 empirical 142, 148–9, 355, 400, 438, 467
 social 32, 35, 40–1, 43–5, 69, 135
 surprise 77, 141, 157, 292, 294, 312
 survival 14, 18, 20, 32, 41, 188
 SVs *see* State Verbs
 syllogisms 184–6, 189, 194, 197
 syllogistic model of belief organization 157, 166
 sympathetic nervous system (SNS) 33, 35, 41
 sympathy 145–8, 151–2
 sympathy model 82–5
 systematic information processing 246–62, 540
 target light onset 532–3
 target person information 78, 82
 target stimuli 52, 57, 59, 102, 472
 targets of judgment 76, 289, 295, 299
 task-contingent rewards 418–19
 task demands 208, 302–4
 task difficulty 138–40, 202, 206, 211
 tasks 137–9, 202, 335–7, 498–9, 526–30, 534–9
 cognitive 292, 303
 difficult 136, 139, 143, 337, 521, 534
 easy 136, 139, 151, 337
 inference 207, 211
 judgment 107, 297
 negotiation 259, 538
 secondary 108, 237, 303
 substitutable 90–1
 TAT *see* Thematic Apperception Test
 teachers 139–40, 144, 151, 212, 360, 431
 technological evolution 366–7
 television 218, 251, 352–5, 477–8
 temporal construal 119–21, 333
 temporal distance 118–20, 122–3, 125–6, 131, 297, 300
 temporary accessibility 77–8, 81–2, 487
 tend and befriend theory 32–45
 terror management
 health model (TMHM) 406–7
 theory (TMT) 398–412
 tetrahedron model 178, 194–6
 Thematic Apperception Test (TAT) 136–7
 theory construction 4–5, 202, 328–30, 488, 490
 thoughts
 death-related 404–5
 negative 231, 238, 404
 threats 32–5, 37–8, 45, 404, 407–10, 464–6
 of death 406–7
 social 32, 34–5
 stereotype 400, 499
 time pressure 214, 216, 219, 253, 259–60, 296
 TMHM *see* terror management, health model
 TMT *see* terror management, theory
 tolerance 410–12
 top-down processing 293–4, 302–4
 TPB *see* planned behavior
 TRA *see* reasoned action
 traditional women 281–2
 traditions 7, 180, 216, 428–9, 510, 527
 trait-behavior clusters 167–8
 trait concepts 162–3, 167, 452
 trait judgments 270
 traits 82, 167–8, 179, 186–7, 193, 338–9
 personality 157, 167, 270, 327, 338–9, 358
 transference 214–15
 triads 179–86, 190
 trials 16, 77, 179, 191, 439, 532–3
 truth 2–4, 92–3, 141, 165, 196, 301–2
 tuning, cognitive 297, 529–30, 534, 537
 two-mode models 277, 518–20
 unconscious thoughts of death 405, 407
 uncontrollability 146, 152
 uncontrollable causality 141–2
 undecided individuals 528–9
 undesired end-states 492–3, 496
 unequal weights model 195–7
 unfavorable attitudes 442, 445, 448
 unfreezing 214, 218
 unimodel 170, 201, 205, 213
 uninformative attributes 272–3
 unique antecedents 57–8
 unit relations 178–80, 182
 negative 179–80, 183
 positive 178–80, 186, 189, 194
 unresolved personal problems 530, 533, 538
 unwanted events 381–3
 upward comparison 464, 468–9, 471, 474, 476, 478
 utility 99, 168, 173, 183, 259, 454–5
 utterances 314–15, 319, 321, 339, 351
 vague inputs 86–7, 92
 valence 89, 102, 104, 128, 514, 516
 affective 209, 514
 continuum 56–7, 59, 68
 dimension 54, 56, 58–9
 bipolar 51, 58
 negative 514, 516
 positive 494, 514, 516
 valent activation functions 56, 58
 validity 93, 111, 165, 167–8, 249, 255–7
 value theory 135–6, 140, 150

values
 cultural 411, 425
 personal 284, 385
 subjective 440–1, 455

variance 143, 279, 359, 426, 446, 444, 449–50

vasopressin 42, 44

verbal attitudes 444–5

verbal behavior 312

verbal priming 78–80

verbal reasoning 75, 77

verbs 160, 162, 197, 311–12, 319–20, 322–3
 interpersonal 317, 319–20

vicarious dissonance 377, 388, 390–1, 393–4

videogames 334, 432

vigilance 304, 496–7

vigilant strategies 495, 497, 499–500

violence, televised 352–5

visual images 111, 163, 165, 172–3

vitality 416, 424, 427

volition 421, 426–7, 432, 440, 526–7

volitional change 137, 144–5

voting behavior 260–1

vulnerability 402, 407
 emotional 485–6

warmth 40, 152, 181, 280–1, 284

weak arguments 230, 233, 238–9, 256, 303, 540

wealth 22, 119, 282, 321, 345, 424

well-being 43, 45, 377, 416, 422–7, 430–2
 psychological 417, 425–6, 430, 432
 subjective 460, 467, 477

wellness 416, 422, 424, 428

Western cultures 391, 401, 425, 462

'which-to-buy' mindset 173

willpower 103, 145

women 19–26, 32, 35–7, 42–4, 281–3, 393–4
 attractive 23–5
 traditional 281–2
 young 20, 506
 younger 19, 26

Work Space 160–4, 171, 174

workplace 283, 394, 424, 432

worldviews 398, 402–5, 408–12
 cultural 402–3, 412

wound healing 41

youth 338, 353, 355, 368, 454