
MANUSCRIPT SUBMISSION:

Until December 31, 2001, five copies of the manuscript should be submitted to:

Jerry Suls, Editor, **PSPB**
Department of Psychology
11 Seashore Hall E
University of Iowa
Iowa City, Iowa, 52242-1407

As of January 1, 2002, five copies of the manuscript and one electronic copy on disk should be submitted to:

Frederick Rhodewalt
Editor-Elect, **PSPB**
Department of Psychology
University of Utah
390 S. 1530 East, Rm 502
Salt Lake City, UT 84112-0250

Editor:

Jerry Suls, University of Iowa

Assistant to the Editor:

Patricia Noeth, University of Iowa

Associate Editors:

Nyla R. Branscombe, University of Kansas
Brad J. Bushman, Iowa State University
Victoria M. Esses, The University of Western Ontario
Vicki Helgeson, Carnegie-Mellon University
Alan Lambert, Washington University
Carol T. Miller, University of Vermont
Delroy L. Paulhus, The University of British Columbia
Paula Pietromonaco, University of Massachusetts-Amherst
Duane Wegener, Purdue University
Vincent Yzerbyt, Catholic University of Louvain

Consulting Editors:

Mark Alicke, Ohio University
Lisa Feldman Barrett, Boston College
Anne Bettencourt, University of Missouri-Columbia
Monica Biernat, University of Kansas
Irene Blair, University of Colorado
Galen Bodenhausen, Northwestern University
Charles Carver, University of Miami
William Chaplin, University of Alabama
Margaret Clark, Carnegie-Mellon University
Chris Crandall, University of Kansas
Kay Deaux, Graduate Center of the City University of New York
Kenneth Dion, University of Toronto
John Dovidio, Colgate University
Naomi Ellemers, Leiden University
Lee Fabrigar, Queen's University (Ontario)
James Gross, Stanford University
Cindy Hazan, Cornell University
Verlin Hinsz, North Dakota State University
Michael Hogg, University of Queensland
Linda Jackson, Michigan State University
Lee Jussim, Rutgers University
Dacher Keltner, University of California, Berkeley
Michael Kernis, University of Georgia
Joachim Krueger, Brown University
James Kulik, University of California, San Diego
Ziva Kunda, University of Waterloo
James Larson, University of Illinois at Chicago
Michael Leippe, St. Louis University
Greg Maio, University of Cardiff
René Martin, University of Iowa
Margo Montieth, University of Kentucky
Richard Moreland, University of Pittsburgh
Gordon Moskowitz, Princeton University
Amelie Mummendy, Universität of Jena
Steven Neuberg, Arizona State University
John B. Nezlek, College of William & Mary
Susan Nolen-Hoeksema, University of Michigan
Craig D. Parks, Washington State University
Louis Penner, University of South Florida
Joseph Priester, University of Michigan Business School
Tracey Revenson, Graduate Center
of the City University of New York
Brent Roberts, University of Illinois at Urbana-Champaign
Richard W. Robins, University of California, Davis
Neal Roesse, Northwestern University
Janet Ruscher, Tulane University
Daniel Russell, Iowa State University

Carey S. Ryan, University of Pittsburgh
Lawrence J. Sanna, University of North Carolina at Chapel Hill
Gerard Saucier, University of Oregon
Barry Schlenker, University of Florida
James Shepperd, University of Florida
Jeffrey Sherman, Northwestern University
Heather J. Smith, Sonoma State University
Stephen M. Smith, North Georgia College and State University
Russell Spears, University of Amsterdam
Susan Sprecher, Illinois State University
Charles Stangor, University of Maryland
Janet Swim, Pennsylvania State University
Howard Tennen, University of Connecticut Medical Center
R. Scott Tindale, Loyola University
David Watson, University of Iowa
Ladd Wheeler, Macquarie University
Thomas Wills, Albert Einstein College of Medicine
Dawn Wilson, Medical College of Virginia
Bernd Wittenbrink, University of Chicago
Barbara Woike, Barnard College
Joanne Wood, University of Waterloo
Michael Zarate, University of Texas at El Paso

SUBMISSION GUIDELINES:

The word count (see below) including the abstract, references and notes should appear on the title page. Authors should also submit two copies of their submission letter that includes a statement of compliance with APA ethical standards in the treatment of their sample, human or animal. The submission letter should also contain a statement that the manuscript or data have not been published previously and that they are not under consideration for publication elsewhere.

Two copies of a separate page should also be submitted that includes:

- (1) title of the manuscript,
- (2) names of authors, and
- (3) mailing address, phone number, and e-mail address of the corresponding author.

Receipt of manuscript will be acknowledged by e-mail.

PSPB aspires to publish papers on a wide range of topics in personality and social psychology in each issue. To accomplish this goal, submitted manuscripts must not exceed 10,000 words in length. Manuscripts exceeding the 10,000 word limit will be returned to authors without review. All manuscripts should be typed in accordance with the guidelines set forth in the 1994 (i.e., 4th edition) *Publication Manual of the American Psychological Association*. Manuscripts should include an abstract of no more than 150 words typed on a separate sheet of paper. Journal policy prohibits and author from submitting the same manuscript for consideration by another journal and does not allow duplicate publication (i.e., publication of a manuscript that has been published in whole or part in another journal). All submitted copies should be clear and legible. Authors should keep a copy of the manuscript to guard against loss. Masked reviews are optional for **PSPB**. For masked review, authors' names, institutions, and other identifying information should appear on only two of the five submitted copies. Three copies should have the authors' names and affiliations removed from the title page and author notes and/or footnotes that contain identifying information.

For Sage Publications:

Jason Dean, Angela Emery, Rose Tylak, and Elena Nikitina

PERSONALITY and SOCIAL PSYCHOLOGY bulletin

Volume 27, Number 12
December 2001

- 1567 The Association of Intimacy Goals and Marital Satisfaction: A Test of Four Mediation Hypotheses**
Catherine A. Sanderson and Nancy Cantor
- 1578 The Effect of Fleeting Attraction on Compliance to Requests**
Jerry M. Burger, Shelley Soroka, Katrina Gonzago, Emily Murphy, and Emily Somervell
- 1587 Environmental and Genetic Influences on Sex-Typed Behaviors and Attitudes of Male and Female Adolescents**
H. Harrington Cleveland, J. Richard Udry, and Kim Chantala
- 1599 On-Line Versus Memory-Based Processing: The Role of "Need to Evaluate" in Person Perception**
Zakary L. Tormala and Richard E. Petty
- 1613 How One Cause Discounts or Augments Another: A Connectionist Account of Causal Competition**
Frank Van Overwalle and Dirk Van Rooy
- 1627 Pessimistic Bias in Comparative Evaluations: A Case of Perceived Vulnerability to the Effects of Negative Life Events**
Hart Blanton, Danny Axsom, Kimberly P. McClive, and Simani Price
- 1637 Social Anxiety and the Cued Activation of Relational Knowledge**
Mark W. Baldwin and Kelley J. Main
- 1648 Lessons From the Past: Do People Learn From Experience That Emotional Reactions Are Short-Lived?**
Timothy D. Wilson, Jay Meyers, and Daniel T. Gilbert
- 1662 Incremental Commitment and Reciprocity in a Real-Time Public Goods Game**
Robert Kurzban, Kevin McCabe, Vernon L. Smith, and Bart J. Wilson
- 1674 Goals, Culture, and Subjective Well-Being**
Shigehiro Oishi and Ed Diener
- 1683 Assessing Generalization in Perceived Self-Efficacy: Multidomain and Global Assessments of the Effects of Self-Defense Training for Women**
Julie C. Weitlauf, Daniel Cervone, Ronald E. Smith, and Paul M. Wright
- 1692 Depression as a Moderator of Relationships Between Positive Daily Events and Day-to-Day Psychological Adjustment**
John B. Nezlek and Shelly L. Gable
- 1705 Self-Enhancement: Is It Restricted to Individualistic Cultures?**
Jenny Kurman
- 1717 Reviewers for 2000**
- 1721 Index**

The **PERSONALITY AND SOCIAL PSYCHOLOGY BULLETIN** is an official journal of the Society for Personality and Social Psychology. The journal is an international outlet for original empirical papers in all areas of personality and social psychology. Occasionally, **PSPB** publishes other pieces of particular interest to members of the Society, such as special topical issues, selected symposia, and invited addresses. All papers are reviewed with respect to their scholarly merit.

MANUSCRIPT SUBMISSION—Masked reviews are optional for **PSPB**. For masked reviews, authors' names, institutions, and other identifying information should appear on only two of the five submitted copies. Three copies should have the authors' names and affiliations removed from the title page and author notes and/or footnotes that contain identifying information deleted. All manuscripts should be typed in accordance with the guidelines set forth in the 4th edition of the *Publication Manual of the American Psychological Association*. Manuscripts should include an abstract of no more than 150 words typed on a separate sheet of paper. Journal policy prohibits an author from submitting the same manuscript for concurrent consideration by another journal and does not allow duplicate publication (i.e., publication of a manuscript that has been published in whole or in substantial part in another journal). All submitted copies should be clear and legible. Authors should keep a copy of their manuscript to guard against loss.

The **PERSONALITY AND SOCIAL PSYCHOLOGY BULLETIN** (ISSN 0146-1672) is published monthly by Sage Publications, 2455 Teller Road, Thousand Oaks, CA 91320; telephone: (800) 818-SAGE (7243) and (805) 499-9774; fax: (805) 375-1700; e-mail: order@sagepub.com; <http://www.sagepub.com>. Copyright © 2001 by the Society for Personality and Social Psychology, Inc. All rights reserved. No portion of the contents may be reproduced in any form without written permission from the publisher (address below) and author(s).

Subscriptions: Address all nonmember subscription correspondence to the publisher, Sage Publications, 2455 Teller Road, Thousand Oaks, CA 91320; telephone: (800) 818-SAGE (7243) and (805) 499-9774; fax/order line: (805) 375-1700. Regular institutional rate \$795.00; single issue, \$77.00. Those individuals who are not members or affiliates of the Society may subscribe at a one-year rate of \$155.00; single issue, \$22.00. Add \$48.00 for subscriptions outside the United States. Noninstitutional orders **must** be paid by personal check, VISA, MasterCard, or Discover. Periodicals postage paid at Thousand Oaks, California, and at additional mailing offices.

This journal is abstracted in **ASSIA (Applied Social Science Index & Abstracts)**, **Automatic Subject Citation Alert**, **Communication Abstracts**, **Current Contents/Social & Behavioral Sciences**, **Family Resources Database**, **Health Instrument File**, **Human Resources Abstracts**, **Linguistics & Language Behavior Abstracts**, **Psychological Abstracts**, **Psyc-INFO Database**, **Research Alert**, **Social Planning Policy & Development Abstracts**, **Social Sciences Citation Index**, **Social SciSearch**, **Sociological Abstracts**, and **Violence and Abuse Abstracts**; indexed in **Academic Search**, **Applied Social Sciences Index & Abstracts (ASSIA)**, **Biological Sciences Abstracts**, **Corporate ResourceNET**, **Current Citations Express**, **Current Contents/Social & Behavioral Sciences**, **Expanded Academic Index**, **Health & Safety Sciences Abstracts**, **IBZ (International Bibliography of Periodical Literature on the Humanities and Social Sciences)**, **Indian Psychological Abstracts and Reviews**

(IPAR), **ISI Basic Social Sciences Index**, **MasterFILE FullTEXT**, **Periodical Abstracts**, **PsycINFO**, **PsycLIT**, **Risk Abstracts**, **Social Science Source**, **Social Sciences Citation Index**, **Sociological Abstracts**, **Standard Periodical Directory (SPD)**, **TOPICsearch**, **Wilson OmniFile V**, and **Wilson Social Sciences Index/Abstracts** and is available on microfilm from University Microfilms, Ann Arbor, Michigan.

Back Issues: Information about availability and prices of back issues may be obtained from the publisher's order department (address above). Single-issue orders for five or more copies will receive a special adoption discount. Contact the order department for details. Write to London office for sterling prices.

Membership and Member Subscription Information: All paid-up members and affiliates of the Society for Personality and Social Psychology, Inc., including members of Division 8 of the American Psychological Association, receive the **PERSONALITY AND SOCIAL PSYCHOLOGY BULLETIN**. APA and non-APA members who are interested in joining the Society should write for information and application materials from Harry Reis, Executive Officer, SPSP, Department of Psychology, University of Rochester, Rochester, NY 14627; e-mail spsp@scp.rochester.edu. Special student memberships are available. Individuals who want to subscribe to **PSPB** without joining the Society should write to Sage Publications (address above; see nonmember subscription information).

Inquiries: All subscription inquiries, orders, and renewals with ship-to addresses in North America, South America, Australia, China, Indonesia, Japan, Korea, New Zealand, and the Philippines must be addressed to Sage Publications, 2455 Teller Road, Thousand Oaks, CA 91320, U.S.A.; telephone: (800) 818-SAGE (7243) and (805) 499-9774; fax: (805) 375-1700; e-mail: order@sagepub.com; <http://www.sagepub.com>. All subscription inquiries, orders, and renewals with ship-to addresses in the U.K., Europe, the Middle East, and Africa must be addressed to Sage Publications, Ltd., 6 Bonhill Street, London EC2A 4PU, England; telephone: +44 (0)20 7374 0645; fax: +44 (0)20 7374 8741. All subscription inquiries, orders, and renewals with ship-to addresses in India and South Asia must be addressed to Sage Publications Private Ltd., P.O. Box 4215, New Delhi 110 048, India; telephone: (91-11) 641-9884; fax: (91-11) 647-2426. Address all permissions requests to the Thousand Oaks office.

Authorization to photocopy items for internal or personal use, or the internal or personal use of specific clients, is granted by Sage Publications for libraries and other users registered with the Copyright Clearance Center (CCC) Transactional Reporting Service, provided that the base fee of 50¢ per copy, plus 10¢ per copy page, is paid directly to CCC, 21 Congress St., Salem, MA 01970. 0146-1672/2001 \$.50 + .10.

Advertising: Current rates and specifications may be obtained by writing to the Advertising Manager at the Thousand Oaks office (address above).

Claims: Claims for undelivered copies must be made no later than six months following month of publication. The publisher will supply missing copies when losses have been sustained in transit and when the reserve stock will permit.

Change of Address, nonmembers: Six weeks' advance notice must be given when notifying of change of address. Please send old address label along with new address to ensure proper identification. Please specify name of journal. POSTMASTER: Send address changes to Personality and Social Psychology Bulletin, c/o 2455 Teller Road, Thousand Oaks, CA 91320.

Members: All members of the Society should send their address changes to Harry Reis at the address above. Allow six weeks for change of address.

The Association of Intimacy Goals and Marital Satisfaction: A Test of Four Mediational Hypotheses

Catherine A. Sanderson

Amherst College

Nancy Cantor

University of Michigan

This research examines the association of intimacy goals and marital satisfaction and tests four potential mediators of the goals-satisfaction link. Forty-four married couples completed measures of their own intimacy goals, their perceptions of their spouse's goals, patterning of marital interactions (e.g., social support, time spent together, number of activities engaged in together, mutual influence), and marital satisfaction. As predicted, both individuals' own and their spouses' pursuit of intimacy goals were associated with marital satisfaction. However, these associations between goals and satisfaction were eliminated when individuals' perceptions of their spouses' goals were included in the analysis, indicating that the link between intimacy goals and marital satisfaction was mediated by individuals' perceptions of their spouses' goals. Discussion focuses on the theoretical and applied implications of these findings.

Several prominent models of relationship functioning have described the potential impact of various types of individual difference factors on marital satisfaction (Bradbury & Fincham, 1988; Reis & Shaver, 1988). For example, Bradbury and Fincham's (1988) contextual model of marriage describes the role of both the distal context (e.g., attachment styles, traits, goals) and the proximal context (e.g., attributions, behaviors, cognitions) in influencing marital satisfaction. Similarly, the model of the intimacy process proposed by Reis and Shaver (1988) describes how individuals' own particular needs, motives, and goals influence how they act toward their partner, which in turn is interpreted and responded to by his or her partner based on their own unique tendencies. These models therefore emphasize the role of individual differences in influencing interac-

tion and interpretation within close relationships, and thereby influencing the quality of such relationships.

Although these models have described the potential impact of a variety of factors on marital satisfaction (e.g., needs, fears, cognitions, traits, goals, etc.), research has focused almost entirely on the impact of only a few factors, namely, attributions (Bradbury & Fincham, 1992; Karney & Bradbury, 2000), relationship beliefs and thinking (Acitelli, 1992; Cate, Koval, Lloyd, & Wilson, 1995), and attachment styles (Davila, Bradbury, & Fincham, 1998; Feeney, Noller, & Callan, 1994). For example, in a series of studies, Fincham and Bradbury (1989) have shown that a spouse's maladaptive attributions are related to relationship dysfunction, with dissatisfied spouses offering attributions that accentuate the impact of negative events and diminish the impact of positive events. Research also has shown that individuals' unrealistic expectancies about close relationships (e.g., that partners should be able to sense what each other is feeling) are associated with both causal attributions and satisfaction (Fincham & Bradbury, 1989) and that individuals' attachment styles influence patterns of marital interaction (e.g., negative affectivity), which in turn affects satisfaction (Davila et al., 1998). Thus, both theory and research in close relationships indicate that par-

Authors' Note: We wish to acknowledge the assistance of Allison Baer and Darren Yopyk in conducting this research. Nancy Cantor is now at the University of Illinois. Correspondence should be addressed to Catherine A. Sanderson, Department of Psychology, Box 2236, Amherst College, Amherst, MA 01002-5000; e-mail: casanderson@amherst.edu.

PSPB, Vol. 27 No. 12, December 2001 1567-1577

© 2001 by the Society for Personality and Social Psychology, Inc.

ticular individual difference factors are associated directly and indirectly with marital satisfaction.

Despite the prevailing emphasis in empirical work on only these few factors, some researchers have specifically recognized the value of incorporating the study of individuals' goals and needs into marital research (Cantor & Malley, 1991; Fincham & Bradbury, 1990; Read & Miller, 1989). For example, McAdams (1988) describes the importance of identifying personal goals, needs, and motives to understand individual differences in behavior, and Cantor and Malley (1991) note that behavior in close relationships will depend on the goals and needs of both partners in the relationship. However, virtually all research examining the impact of goals on relationship functioning has been based entirely on college student samples in relatively short dating relationships (Sanderson & Cantor, 1997; Sanderson & Evans, 2001; Sanderson & Karetsky, in press). This work has shown that individuals with intimacy goals in dating experience greater relationship satisfaction and are more likely to maintain their relationships over time (Sanderson & Cantor, 1997). Moreover, a recent study by Sanderson and Evans (2001) indicates that women with intimacy goals perceive their partners as sharing these goals, which in turn is associated with greater relationship satisfaction.

Although this prior work certainly suggests that intimacy goals are associated with satisfaction both directly and indirectly (e.g., through their impact on individuals' perceptions of their partner's goals), marital relationships clearly differ in many ways from dating relationships in college (e.g., length of relationship, age of individuals, opportunities to cooperate on various household tasks and decisions, commitment, etc.); hence, it is also important to examine both the direct and indirect impact of intimacy goals on marital satisfaction. The focus of the present research is therefore on examining four distinct processes that may mediate the link between a novel distal factor (i.e., intimacy goals) and marital satisfaction: having spouses with intimacy goals, structuring marital interactions in intimacy-relevant ways, perceiving spouses as having intimacy goals, and seeing spouses' intimacy goals accurately.

The Presence of Intimacy-Focused Spouses

As described in Gollwitzer's (1993) analysis of goal implementation, individuals have considerable choice in terms of the contexts in which they spend time and, hence, they are more likely to seek out and select particular interaction partners, namely, those who will assist them in fulfilling their own needs and goals (Buss, 1987; Cantor, 1994; Mischel, Cantor, & Feldman, 1996; Snyder & Ickes, 1985; Snyder & Simpson, 1984). In fact, considerable research by Niedenthal and colleagues (Niedenthal, Cantor, & Kihlstrom, 1985; Niedenthal &

Mordkoff, 1991) has shown that people prefer to spend time with others who are similar to themselves on various dimensions. For example, college students prefer campus housing in which the prototypic resident (e.g., fraternity type, co-op type) is similar to themselves (Niedenthal et al., 1985), and individuals prefer to see therapists whose patients are more similar to themselves in terms of personality traits (Niedenthal & Mordkoff, 1991). Within the domain of close relationships, Snyder's work on self-monitoring indicates that individuals with a relatively public orientation may attempt to date "high-status individuals," whereas those with a more private self-focus may attempt to date partners who share their own values and attitudes (Snyder & Simpson, 1984). This preference for particular types of relationship partners makes sense because the goals and plans of each partner in a relationship are likely to affect the fulfillment of the other person's needs (Miller & Read, 1991).

Because creating an intimacy-focused marriage necessarily requires the cooperation of both partners, individuals with strong intimacy goals may have similarly oriented spouses (i.e., engaging in open self-disclosure and mutual dependence requires the cooperation of both partners) (cf. Miller, 1990). An individual who is trying to fulfill intimacy goals, for example, may be frustrated when her or his spouse is primarily focused on independence and self-reliance and lacks the ability or comfort to engage in such self-disclosure (Miller, 1990; Miller & Read, 1991). As Miller (1990) describes, the amount of self-disclosure in a relationship is influenced by both participants. Because individuals with intimacy goals should find it easiest to fulfill their goals with a partner who is receptive to such behavior, they should be motivated to have intimacy-focused spouses.

The Patterning of Marital Interaction

Just as people select partners for daily life situations that should facilitate their goal pursuit, they also may engage in particular behaviors that facilitate such pursuit (Buss, 1987; Snyder & Simpson, 1984); for example, a beeper study by McAdams and Constantian (1983) demonstrated that individuals who are high in intimacy and affiliation motives engage in more interpersonal conversations and letter-writing than others. Similarly, research has shown that the traits of extraversion, affiliation, and sociability are all positively correlated with the amount of time individuals choose to spend in social situations (Emmons, Diener, & Larsen, 1986). Thus, both research and theory indicate that individuals intentionally structure their lives to spend time in goal-fulfilling situations (Gollwitzer, 1993; Mischel et al., 1996).

In turn, one possible explanation for why individuals with intimacy goals may experience greater marital satis-

faction is that they actively engage in a variety of patterns of marital interaction that facilitate intimacy with their spouse (e.g., Buss, 1987; Emmons et al., 1986; Snyder, 1981). First, they may choose to spend considerable time with their spouse and to engage in many activities together as a way of fostering interdependence; for example, intimacy-focused individuals may prefer to work together with their spouse on various household tasks (e.g., balancing the checkbook, washing dishes, planning vacations) as opposed to dividing up such responsibilities. Second, individuals with intimacy goals may give substantial social support to their spouses as a way of strengthening the marital bond and enhancing feelings of responsibility for and commitment to the relationship (Sarason, Shearlin, Pierce, & Sarason, 1987). Third, given their greater focus on creating interdependence, their spouse's feelings and needs may have more impact on their own thoughts, feelings, behaviors, and plans (Fincham & Bradbury, 1990). In sum, individuals with intimacy goals may create opportunities within their marital relationships to engage in particular types of intimacy-enhancing interactions, which in turn may be associated with greater satisfaction.

The Perception of Spouses As Intimacy Focused

Considerable research in close relationships has shown that individuals develop detailed models of their partner's goals and beliefs (Miller & Read, 1991) and these beliefs (regardless of their accuracy) may in turn influence relationship satisfaction (Murray, Holmes, & Griffin, 1996a, 1996b; Ptacek & Dodge, 1998). Murray and colleagues, for example, have found that both dating and married couples experience more satisfying and longer-lasting relationships when individuals hold idealized views about their partners (Murray et al., 1996a, 1996b). Moreover, the social support literature indicates that individuals' perceptions of the amount of social support they receive can be a stronger predictor of well-being than the actual amount of support received (Dunkel-Schetter & Bennett, 1990), and a recent study by Ptacek and Dodge (1998) found that satisfaction in both dating and married relationships was predicted by the perceived similarity of coping styles between partners. This research all demonstrates that the mere perception of one's partner is associated with relationship satisfaction, regardless of its accuracy.

Because prior research indicates that individuals may project their own traits onto their partners (Murray et al., 1996a), individuals with intimacy goals may perceive their spouses as sharing their intense focus on communion. In fact, research by Ruvalo and Fabin (1999) indicates that such projection increases as intimacy increases in a relationship, suggesting that individuals with intimacy goals may be especially likely to engage in

such biased perception. This perception (regardless of its accuracy) may lead those with intimacy goals to simply perceive their spouse's behavior accordingly (e.g., one could interpret the behavior of a hard-working spouse who is rarely home as intensely focused on providing for his or her family and thereby quite concerned with communion) as well as enable them to act in goal-affording ways (e.g., because they believe their spouse would be receptive to such behavior). Thus, individuals with intimacy goals may experience marital satisfaction because they merely believe their spouse shares these goals.

The Accurate Perception of a Spouse's Intimacy Goals

Several lines of research have investigated aspects of interpersonal accuracy and its association with relationship satisfaction (Ickes, 1993). This work indicates that individuals whose own self-perceptions are in line with the way they are perceived by their partners have greater satisfaction (e.g., Sillars, Pike, Jones, & Murphy, 1984; Sillars & Scott, 1983; Swann & Predmore, 1985). For example, Swann and colleagues (Swann, De La Ronde, & Hixon, 1994) found that individuals in marital relationships who saw their partner as they saw themselves had greater intimacy. Similarly, a study by Tuckers and Anders (1999) on attachment style and interpersonal perception accuracy revealed that anxiously attached men consistently showed lower accuracy in perceiving their partner's feelings of love, which in turn partially accounted for their decreased relationship satisfaction. The ability to accurately understand and interpret one's partner's feelings and behavior therefore seems to play an important role in predicting relationship satisfaction.

Because individuals with intimacy goals are particularly focused on fostering open communication in their relationship (e.g., engaging in self-disclosure, eliciting self-disclosure from their partner) (Sanderson & Evans, 2001), it is conceivable that they would have greater awareness, and hence greater accuracy, of their spouse's intimacy goals. In turn, individuals who have more accurate perceptions of their spouse's goals should be able to both engage in intimacy-relevant behaviors (i.e., in cases in which the spouse is also strongly focused on intimacy) and concentrate on compensating for their spouse's relative lack of skill in this domain (i.e., in cases in which the spouse is not particularly focused on intimacy). In fact, prior research has shown that those with strong intimacy goals are most satisfied when their perceptions of their partner's goals are accurate, regardless of what those goals are (Sanderson & Cantor, 1997). These findings suggest that the association between intimacy goals and satisfaction may be mediated by the greater accuracy in perceiving their spouse's intimacy goals shown by individuals who themselves have intimacy goals.

Overview

This study examines four distinct mediational hypotheses that may explain the goals-satisfaction link. First, do individuals with intimacy goals have intimacy-focused spouses, which in turn is associated with greater marital satisfaction? Second, do individuals with intimacy goals engage in particular patterns of marital interaction (e.g., spend more time together, participate in more activities together, give more social support, have greater influence on each other's thoughts and behaviors, have greater influence on each other's future goals and plans), which in turn is associated with greater marital satisfaction? Third, do individuals with intimacy goals perceive their spouses as having intimacy goals, which in turn is associated with greater marital satisfaction? Finally, do individuals with intimacy goals have more accurate perceptions of their spouse's intimacy goals, which in turn is associated with greater marital satisfaction?

METHOD

Participants

Forty-four married couples served as participants in this research (husbands: $M = 42.20$, $SD = 9.46$; wives: $M = 41.64$, $SD = 9.69$). Of these couples, 80% had at least one child, with the majority of these reporting two children ($M = 2.17$, $SD = 0.75$, range of 0 to 4). The couples had been married an average of 16.35 years ($SD = 8.94$). The majority of the couples were White ($N = 39$, 89%), with 2 African American couples (4%) and 3 Asian couples (7%). About half of the couples ($N = 24$, 59%) had a mean income level of more than \$90,000, although there was a range from less than \$30,000 ($N = 2$, 5%). The majority of husbands (91%, $N = 40$) worked full-time outside the home, with the remainder working part-time (9%, $N = 4$). Of the husbands, 89% ($N = 39$) had graduated from college, with 27% ($N = 12$) having completed an advanced degree. A little more than half of the wives worked full-time outside the home (57%, $N = 25$), with 16% ($N = 7$) working part-time and 27% ($N = 12$) not working outside the home. Of the wives, 89% ($N = 39$) had completed college, with 27% ($N = 12$) having completed some graduate work.

Procedure

Participants were recruited as part of a fund-raising drive at a local elementary school. Parents of elementary schoolchildren were contacted by letter and told that a \$20 contribution would be made to a school trip fund for each pair of parents who completed a 20-min survey. Interested parents called to request materials, which were then sent to them along with self-addressed

TABLE 1: Means and Standard Deviations by Gender of All Variables

Variable	Wives		Husbands	
	M	SD	M	SD
Intimacy goals ^a	3.66	0.56	3.91	0.44
Marital interaction patterns				
Time spent ^b	5.16	4.88	5.57	6.00
Number of activities engaged in together ^c	10.47	5.00	9.77	4.33
Influence on thoughts, mood, behaviors ^d	4.26	0.71	4.84	0.74
Influence on future plans and goals ^d	4.93	1.53	5.47	1.06
Social support provided ^d	5.43	1.22	5.32	1.01
Perceived spouse's intimacy goals ^a	3.55	0.51	3.81	0.58
Marital satisfaction ^e	105.77	15.34	107.67	12.28

NOTE: $N = 88$ (44 couples).

a. Measured on a 1 to 5 scale.

b. Measured in hours.

c. Measured yes or no on a checklist of 38 items.

d. Measured on a 1 to 7 scale.

e. Measured on a 32-item scale, with most items measured on a 1 to 6 scale.

stamped envelopes. They were asked to complete the questionnaire without discussing the questions or their answers with their spouse; separate envelopes were provided for each spouse.

Measures

Means and standard deviations for all measures by gender are provided in Table 1.

Intimacy Goals Scale. To examine participants' orientation toward intimacy in their marital relationships, we used a revised version of the Social Dating Goals Scale to focus on marital as opposed to dating relationships (Sanderson & Cantor, 1995). This scale is scored on a 1 to 5 scale (1 = *strongly disagree* to 5 = *strongly agree*) and consists of items assessing individuals' concerns with self-disclosure and dependence (e.g., "In my relationship, I try to share my most intimate thoughts and feelings with my partner," "In my relationship, I try to depend on my partner," "In my relationship, I try to take care of my partner"). This scale meets the standard criteria for determining unidimensionality of a scale (Briggs & Cheek, 1986) and has been used in a series of studies with high school and college students (e.g., Sanderson & Cantor, 1995, 1997; L. T. Volenski, personal communication, March 10, 1995). The scale had acceptable internal coherence ($\alpha = .73$ for wives, .68 for husbands).

Perceived spouse's intimacy goals. Participants also were asked to complete the Intimacy Goals Scale based on how they thought their spouse would respond to this scale (e.g., their perceptions of their spouse's intimacy

goals). The internal coherence of these two measures was also acceptable (wives' $\alpha = .61$, husbands' $\alpha = .80$).

Relationship closeness inventory. This measure assesses the closeness of a particular relationship and in particular its strength, diversity, and frequency of interaction (Berscheid, Snyder, & Omoto, 1989). First, participants reported the time they spent with their spouse during the last week in the morning, the afternoon, and the evening (responses to these three questions were summed to get an overall mean of time spent per week). Second, participants noted whether they had engaged in each of 38 activities with their spouse in the past week (e.g., did laundry, watched TV, engaged in sexual relations, ate a meal, discussed things of a personal nature). Third, participants rated the extent to which their spouse affected 27 different types of thoughts, feelings, and behaviors (e.g., my moods, the basic values I hold, how I dress, the type of career I have) using a 1 to 7 scale (1 = *disagree strongly* to 7 = *agree strongly*) (wives' $\alpha = .71$, husbands' $\alpha = .86$). Finally, participants rated the extent to which their spouse affects their future plans and goals (e.g., having children, vacationing, achieving a given financial standard of living), again using a 1 to 7 scale (1 = *not at all* to 7 = *a great extent*) (wives' $\alpha = .83$, husbands' $\alpha = .71$).

Social support provided. Participants rated the extent to which they provided their spouse with each of four types of social support that are widely used in the close relationships literature (House & Kahn, 1985): listening, instrumental, informational, emotional ($\alpha = .88$ for wives, .86 for husbands). Ratings were based on a 1 to 5 scale (1 = *not at all* to 5 = *a lot*).

Dyadic Adjustment Scale. The Spanier Dyadic Adjustment Scale (DAS) (Spanier, 1976) is a commonly used 32-item measure of marital satisfaction. The scale is composed of items rated on both 6- and 7-point scales and items that are answered yes or no. When summed, the items indicate satisfaction with the relationship, with higher scores indicating greater marital satisfaction (range of 0 to 151). Sample items include, "How often do you discuss or consider divorce or terminating your relationship?" "How often do you kiss your mate?" and "Indicate the approximate extent of agreement or disagreement between you and your partner on handling family finances." Although many studies of married couples use participants who are likely to be above average in marital satisfaction (e.g., those who respond to advertisements seeking married couples are presumably relatively high on satisfaction), the couples in this study represent a relatively broad range of marital quality. For example, 20% of the wives ($N=9$) and 14% of the husbands ($N=6$) reported that they had received marital or family counseling. Moreover, the mean marital satisfaction scores of

participants in this study were somewhat lower than that reported by many other studies.¹

Demographic information. Participants provided a variety of demographic information, including age, gender, race, income, education, number of children, and length of marriage.

RESULTS

To examine various processes that may lead those with a strong focus on intimacy goals in their relationships to experience greater marital satisfaction, we used structural equation modeling (SEM) to estimate simultaneous regression equations (Edwards, 1994). This approach allows us to simultaneously examine the association of each person's intimacy goals with their own and their spouse's satisfaction as well as the association of various potential mediators (e.g., time spent, perceived goals, etc.). In each case, we used the SEM program within the AMOS procedure of SPSS for Windows. The SEM approach also had the advantage of allowing us to test for gender differences in our path models by constraining certain paths to be equal and then comparing the fit of models that estimate common path coefficients for men and women versus those that allow these paths to vary (Kenny, 1996). In all cases, we found no significant differences by gender in the size of the paths or fit of the models; hence, in describing our results, we present standardized path coefficients that are pooled across gender. In each case, we specified that husbands' and wives' goals could be correlated and that all other variables were associated in the particular causal direction we hypothesized (e.g., that individuals with strong intimacy goals would spend more time with their spouse, which in turn would lead to marital satisfaction). Although in constructing these structural models one must specify causal links among the variables in the model, these causal arrows are only hypothetical and hence cannot directly test any causal assumptions we made in constructing each model.

Table 2 presents the zero-order correlations between all variables. As predicted, there were strong correlations between individuals' intimacy goals and their marital satisfaction for both husbands ($r = .50$, $p < .0001$) and wives ($r = .70$, $p < .0001$). Thus, in line with prior research, these findings indicate that there is indeed a link between intimacy goals and marital satisfaction.

Hypothesis 1: Do Spouses' Intimacy Goals Mediate the Goals-Satisfaction Link?

The first model examined the association between each spouse's pursuit of intimacy goals and marital satisfaction to examine whether individuals' focus on intimacy goals was associated with marital satisfaction for

TABLE 2: Zero-Order Correlations Among Intimacy Goals, Interaction Patterns, Perceived Spouse's Intimacy Goals, and Marital Satisfaction

	1	2	3	4	5	6	7	8	9
1. Intimacy goals	—	.53	.03	.39	.45	.39	.68	.70	.70
2. Spouse's intimacy goals	.53	—	.09	.31	.32	.35	.46	.57	.50
3. Time spent together	-.11	-.25	—	.19	.05	.14	.05	.17	-.06
4. Number of activities together	.35	.26	.33	—	.11	.12	.11	.27	.36
5. Influence on thoughts, feelings, behavior	.58	.30	-.05	.34	—	.56	.25	.21	.37
6. Influence on future plans and goals	.33	.43	-.33	.11	.45	—	.17	.35	.37
7. Giving social support	.38	.50	.11	.36	.17	.22	—	.67	.55
8. Perceived spouse's intimacy goals	.66	.64	-.06	.37	.40	.23	.61	—	.63
9. Marital satisfaction	.50	.44	-.08	.23	.26	.15	.53	.59	—

NOTE: $N = 88$ (44 couples). Correlations above the diagonal are for wives; correlations below the diagonal are for husbands.

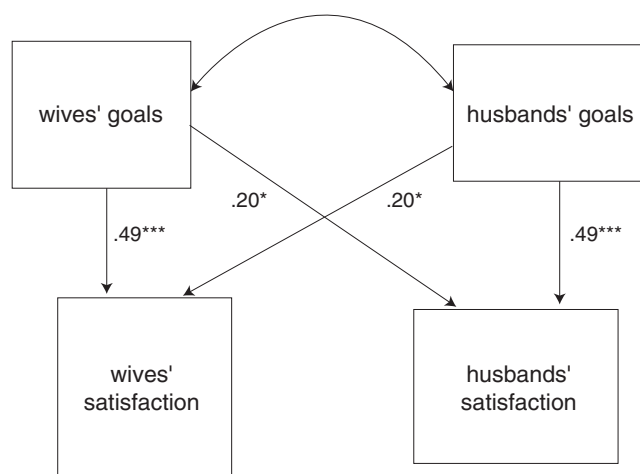


Figure 1 Marital satisfaction as a function of individuals' own intimacy goals and their partners' intimacy goals.

NOTE: $\chi^2(3, N = 44) = 20.05, p < .0001$ (Goodness of Fit Index [GFI] = .84, Comparative Fit Index [CFI] = .71).

* $p < .05$. *** $p < .001$.

themselves as well as for their spouses. As shown in Figure 1, these analyses revealed that one's own pursuit of intimacy goals was a strong predictor of marital satisfaction for both husbands and wives. Moreover, individuals with a strong focus on intimacy goals also had spouses with greater marital satisfaction. These analyses indicate that both individuals' own pursuit of intimacy goals as well as their spouses' pursuit were associated with marital satisfaction but that individuals with strong intimacy goals do not depend on having spouses with intimacy-focused goals to experience satisfaction (i.e., the link between own goals and own satisfaction was significant even when the model included the link between spouse's goals and own satisfaction).

Hypothesis 2: Do Patterns of Marital Interaction Mediate the Goals-Satisfaction Link?

The next model tested examined the association of individuals' intimacy goals with patterns of marital interaction (e.g., strength of influence, social support given, time spent, etc.) as well as with satisfaction to assess whether individuals with a strong focus on intimacy experience greater satisfaction because they structure and interact in their marital relationships in distinct (i.e., intimacy-relevant) ways. As shown in Table 3, the pursuit of intimacy goals was associated with engaging in more activities with one's spouse, giving more social support, and having greater influence on spouse's thoughts and feelings as well as on future plans. Although there were associations between social support given and number of activities participated in together and satisfaction, one's own pursuit of intimacy goals remained a significant predictor of marital satisfaction in all five cases. These findings therefore indicate that although individuals with intimacy goals may act in certain intimacy-conducive ways, these patterns of interaction do not mediate the link between intimacy goals and marital satisfaction.

Hypothesis 3: Do Perceptions of Spouse's Intimacy Goals Mediate the Goals-Satisfaction Link?

Our third model examined the association between the perception of one's spouse's intimacy goals and marital satisfaction and, specifically, whether individuals with strong intimacy goals perceive their spouses as also pursuing such goals, which in turn leads to greater marital satisfaction. First, this analysis demonstrated that individuals' own pursuit of intimacy goals was strongly associated with their perception of their spouse's intimacy goals as well as with their spouse's perception of their goals (see Figure 2). However, individuals' pursuit of intimacy goals had no impact on either their own or their spouse's marital satisfaction once perception was

TABLE 3: Intimacy Goals, Interaction Patterns, and Marital Satisfaction

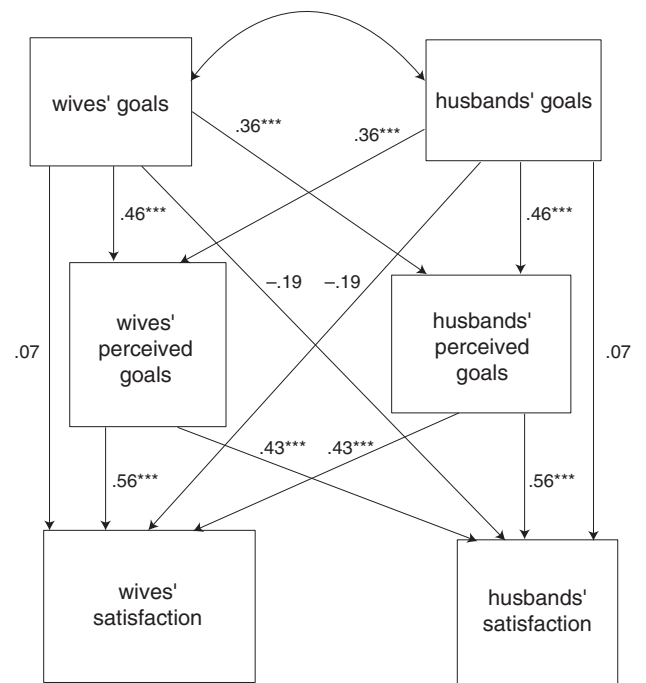
Path	Interaction Pattern Variable				
	Time Spent	Activities	Social Support	Strength of Influence	Future Influence
Own goals to own interaction	.03	.30**	.43***	.50***	.22*
Spouse's goals to own interaction	-.07	.13	.24*	.07	.28**
Own interaction to own satisfaction	-.03	.18*	.31**	-.04	.03
Own interaction to spouse's satisfaction	-.23**	-.16	.09	.20*	.06
Own goals to own satisfaction	.46***	.46***	.32**	.49***	.46***
Spouse's goals to own satisfaction	.20*	.22*	.09	.10	.18
Measures of fit					
Chi-square (8, N = 44)	24.98**	39.01***	31.55***	27.92***	23.42**
Goodness-of-fit index (GFI)	.86	.80	.85	.85	.87
Comparative fit index (CFI)	.73	.64	.77	.78	.79

* $p < .05$. ** $p < .01$. *** $p < .001$.

entered in the analysis. Both individuals' own perception of their spouse's goals and their spouse's perception of their own goals were associated with satisfaction, indicating that perception does indeed mediate the link between intimacy goals and marital satisfaction.

Hypothesis 4: Does Accuracy of Spouse's Intimacy Goals Mediate the Goals-Satisfaction Link?

To examine our fourth model, namely, whether individuals with greater intimacy goals see their spouse's goals more accurately and hence experience greater marital satisfaction, we used hierarchical linear regression (following Murray et al., 1996a). The path analyses we presented thus far examine only main effects and hence could not determine whether individuals with a stronger focus on intimacy goals perceive their spouse's goals more accurately as well as whether such accuracy is a more important predictor of marital satisfaction for those with a strong versus a weak focus on intimacy.² As a first step, we examined whether individuals with a strong focus on intimacy goals see their spouse's goals more accurately than do those with a weak focus on intimacy. To examine this question, we conducted regression analyses predicting perception of one's spouse's goals as a function of own intimacy goals and spouse's actual goals (entered on Step 1) and the interaction representing the match between each spouse's actual goals (entered on Step 2).³ In the case of predicting wives' perceptions

**Figure 2** Marital satisfaction as a function of both spouses' intimacy goals and their perceptions of their spouses' intimacy goals.

NOTE: $\chi^2(8, N = 44) = 31.81, p < .0001$ (Goodness of Fit Index [GFI] = .83, Comparative Fit Index [CFI] = .83).

*** $p < .001$.

of their husbands' goals, there was a main effect only of wives' own goals, indicating that wives with a strong focus on intimacy see their husbands as sharing this focus. In the case of predicting husbands' perceptions of their wives' goals, there were main effects of both husbands' goals (with husbands with strong intimacy goals seeing their wives as sharing this focus) and wives' goals (with wives with strong intimacy goals being seen by their husbands as more intimacy focused). Contrary to our predictions, the Own Goals \times Spouse's Goals interaction did not enter the equation in either analysis, indicating that individuals with a strong focus on intimacy in their marital relationships are not more accurate in assessing their spouse's goals.

DISCUSSION

This study examined the association of intimacy goals with marital satisfaction and specifically tested four mediational hypotheses that may account for the goals-satisfaction link. First, although our data demonstrated that those with a strong focus on intimacy goals do engage in particular patterns of marital interaction that

are likely to facilitate communion, including engaging in more activities with their spouse, having greater interdependence and influence on each other, and providing more social support (cf. Buss, 1987; Snyder, 1981), there was little evidence that the presence of these intimacy-relevant behaviors mediated the link between intimacy goals and marital quality. Similarly, although there was some evidence that people experience greater marital satisfaction when they as well as their spouses have intimacy goals, the direct association between individuals' own goals and satisfaction (as well as between their spouse's goals and satisfaction) was eliminated when individuals' perceptions of their spouse's goals were included in the equation. Thus, individuals with a strong focus on intimacy goals in their marital relationships apparently experience greater satisfaction because they see their spouses as sharing this intense focus on intimacy. Finally, there was no evidence that individuals with intimacy goals are more accurate in their perceptions of their spouse's goals.

These results contribute to prior work in several ways. First, our finding that the pursuit of intimacy goals is associated with greater relationship satisfaction extends work from prior studies with college students in dating relationships (Sanderson & Cantor, 1997; Sanderson & Evans, 2001) by showing a similar goals-satisfaction link in adult couples in marital relationships. Second, this research extends prior work by Murray and colleagues (Murray et al., 1996a, 1996b) demonstrating that individuals in both dating and marital relationships benefit in terms of satisfaction from seeing their partners in idealized ways by showing that individuals may themselves differ in the extent to which they engage in such idealization. In fact, Ruvolo and Fabin (1999) found that the higher the degree of intimacy in the relationship, the more individuals projected their own views onto their partners, suggesting that individuals with intimacy goals may be particularly likely to hold such biased perceptions of their spouse's goals. Third, although our findings are in line with those from the close relationships literature demonstrating the importance of intimacy in determining the quality of marital relationships (Hendrick, 1981; Hendrick, Hendrick, & Adler, 1988), this research extends such work by examining various mediators of the goals-satisfaction link and specifically by demonstrating that this link is mediated by individuals' perceptions of their spouse's intimacy goals.

Limitations and Future Research

Although we believe this research makes some valuable contributions to prior work in both social-personality psychology and the close relationships literature, there are several limitations that should be addressed.

First, given the reliance in this study on cross-sectional data, the associations found in this research are clearly open to multiple alternative explanations. Although we have described this research as focusing on how intimacy goals are associated with marital satisfaction (e.g., via their effects on perceptions of spouse's goals), it is certainly possible that individuals who are in satisfying marital relationships develop intimacy goals over time. Similarly, research in social cognition has shown that individuals' cognitive structures and motivations can lead to biases in their recollections (Fiske & Taylor, 1991; Greenwald & Banaji, 1995; Nisbett & Ross, 1980; Ross, 1989); hence, our findings may be at least partially a result of the particular reconstructions and reports made by our participants. Specifically, individuals' reports of how much time they spent with their partner, social support provided, or influence of future plans and goals may be driven by their intimacy goals and/or their overall feelings of satisfaction in the marriage (e.g., "sentiment override"; see Weiss, 1990). For example, research by Jacobson and Moore (1981) demonstrates that satisfied spouses report more positive events and greater time spent, whereas dissatisfied spouses report more negative events and less time spent.

Although given the use of cross-sectional data we cannot determine the direction of the effects in this data, we believe that the particular directional pathway we hypothesize (e.g., goals lead to perception, which in turn leads to satisfaction) is most likely for several reasons. First, recent longitudinal work on the effects of personality on interpersonal relationships suggests that personality affects relationships, whereas relationships have little impact on personality (Asendorpf & Wilpers, 1998), and research on close relationships has shown that spouses' behavior and interpretations seem to contribute to marital satisfaction as opposed to the opposite (e.g., marital satisfaction influencing such perceptions) (Karney & Bradbury, 2000). Second, when we included marital satisfaction first in the equation as an exogenous variable predicting intimacy goals, we still found evidence for the association between intimacy goals and both patterns of marital interaction and perceived spouses' intimacy goals, indicating that individuals' focus on intimacy goals is not merely a reflection of their current feelings toward their spouse.⁴ However, further research tracking couples from courtship to later marriage is clearly needed to answer these questions about the distinct association between goals, interactions, perception, and satisfaction over time.

Second, although this work found little evidence that various types of intimacy-relevant marital interactions influenced satisfaction, other such variables may serve to mediate the intimacy goals-marital satisfaction link. Inti-

macy-focused individuals may structure their close relationships and interact with their partners in a variety of ways not addressed in this study, including resolving conflicts constructively, eliciting self-disclosure, making adaptive attributions, and/or engaging in "relationship talk" (e.g., Acitelli, 1992; Sanderson & Evans, 2001; Sanderson & Karetzky, *in press*), which in turn may be associated with marital satisfaction. For example, those who are pursuing intimacy goals may be more aware of and focused on their spouse's feelings regarding the marital relationship and may be more effective at resolving disagreements through open discussion. Relatedly, research by Franzoi and colleagues (Franzoi, Davis, & Young, 1985) has shown that perspective-taking scores are significantly related to relationship satisfaction in college student couples. Future research should examine these as well as other potential mediators of the intimacy goal-marital satisfaction link to provide insight into whether factors other than mere perception of spouse's goals are associated with marital quality.

Finally, although these findings are in line with those from prior research indicating the importance of perception of one's partner in creating satisfaction (Murray et al., 1996a, 1996b; Ptacek & Dodge, 1998), this research does not answer the question of exactly how perception serves to mediate the intimacy goals-marital satisfaction link. One possibility is that this perception leads people to act in ways that facilitate intimacy either directly (e.g., by actively attempting to elicit self-disclosure from their spouse) or indirectly (e.g., by spending considerable time alone with one's spouse as a way of providing opportunities for fulfilling intimacy goals) (Jussim, 1989; Snyder, Tanke, & Berscheid, 1977). For example, prior research with college students has shown that those individuals with intimacy goals report spending more time alone with their partners (although not more time with their partner and others) and both engaging in and eliciting more self-disclosure (Sanderson & Cantor, 1997; Sanderson & Evans, 2001). Correspondingly, recent longitudinal research by Downey and colleagues has shown that individuals who are especially sensitive to rejection in their dating relationships actually behave in ways that elicit rejection from their dating partners (Downey, Freitas, Michaelis, & Khouri, 1998). Alternatively, individuals with intimacy goals may feel content with their marital relationship merely through their belief that their spouse also is interested in engaging in interdependence and communion (e.g., the power of seeing one's spouse through "intimacy-colored glasses"). For example, individuals who believe (even erroneously) that their spouse shares their intense focus on intimacy may feel comfortable engaging in various personally relevant (e.g., intimacy-focused) behaviors

(e.g., sharing thoughts and feelings, eliciting self-disclosure, spending time alone with one's spouse); as a result, they may experience marital satisfaction. Future research is clearly needed to examine these and other potential explanations for how perception serves to mediate the intimacy goals-marital satisfaction link.

Conclusions

These findings have important implications for research on personality processes and close relationships and are particularly strong given their basis in adult couples in relatively long-term marriages with a range of marital satisfaction (e.g., as opposed to the typical college student dating couples and/or newlywed couples in relatively satisfying relationships samples). This research extends prior work focusing largely on attributions, relationship beliefs, and attachment styles by demonstrating the association of individuals' intimacy goals with marital satisfaction and suggests that future work needs to examine not only the characteristics (e.g., traits, goals, needs, styles) of one individual in the relationship but also the actual and perceived characteristics of her or his partners (Acitelli, Douvan, & Veroff, 1993; Murray et al., 1996a, 1996b; Ptacek & Dodge, 1998). Our findings suggest that the link between intimacy goals and marital satisfaction is largely (or even entirely) in the mind: Individuals with intimacy goals saw their spouses as sharing these goals, and this perception, in turn, was associated with satisfaction.

NOTES

1. Although one reviewer raised a question about the potential overlap in items between the Intimacy Goals Scale and the Dyadic Adjustment Scale (DAS), a review of these scales indicated only a single item assessing a similar construct ("Share my most intimate thoughts and feelings" in the Intimacy Goals Scale; "Do you confide in your mate?" in the DAS). We therefore reconducted all analyses with this item eliminated from the DAS. There were no differences in findings; hence, we present all analyses based on the full measures.

2. Although some prior research has examined issues of accuracy using difference or discrepancy scores, the use of such scores is problematic because the difference score is confounded with the two main effects (e.g., individuals' own and their spouse's goals) that go into the difference. See Edwards (1994) and Griffin, Murray, and Gonzalez (1999) for a thorough review of the problems associated with the use of discrepancy scores.

3. If, as individuals' intimacy goals increased, their accuracy of their spouse's goals increased, then we should find a significant effect of the interaction term. For example, if wives with strong intimacy goals are most accurate at perceiving their spouse's goals, then wives with high intimacy goals should have high perceptions for husbands with high intimacy goals and low perceptions for husbands with low goals. In turn, if wives with weak intimacy goals are less accurate, they should show a different pattern (e.g., perhaps by seeing their spouse's goals as neither particularly high nor low, or perhaps by seeing spouses with low goals having high goals, and vice versa). Although the specific form of the interaction may vary, if individuals with a strong focus on intimacy are more accurate in perceiving their spouse's goals, then the Own

Goals \times Spouse's Goals interaction should significantly predict the perception of one's spouse's goals.

4. When we included marital satisfaction as an exogenous variable in each of the analyses, the association between intimacy goals and each interaction variable as well as perception of spouse's goals remained significant in all cases. Thus, marital satisfaction does not appear to be functioning as a third variable, causing the observed link between goals and potential mediators.

REFERENCES

- Acitelli, L. K. (1992). Gender differences in relationship awareness and marital satisfaction among young married couples. *Personality and Social Psychology Bulletin*, 18, 102-110.
- Acitelli, L. K., Douvan, E., & Veroff, J. (1993). Perceptions of conflict in the first year of marriage: How important are similarity and understanding? *Journal of Social and Personal Relationships*, 10, 5-19.
- Asendorpf, J. B., & Wilpers, S. (1998). Personality effects on social relationships. *Journal of Personality and Social Psychology*, 74, 1531-1544.
- Berscheid, E., Snyder, M., & Omoto, A. M. (1989). The Relationship Closeness Inventory: Assessing the closeness of interpersonal relationships. *Journal of Personality and Social Psychology*, 57, 702-807.
- Bradbury, T. N., & Fincham, F. D. (1988). Individual difference variables in close relationships: A contextual model of marriage as an integrative framework. *Journal of Personality and Social Psychology*, 54, 713-721.
- Bradbury, T. N., & Fincham, F. D. (1992). Attributions and behavior in marital interaction. *Journal of Personality and Social Psychology*, 63, 613-628.
- Briggs, S. R., & Cheek, J. M. (1986). The role of factor analysis in the development and evaluation of personality scales. *Journal of Personality*, 54, 106-148.
- Buss, D. M. (1987). Selection, evocation, and manipulation. *Journal of Personality and Social Psychology*, 53, 1214-1221.
- Cantor, N. (1994). Life task problem-solving: Situational affordances and personal needs. *Personality and Social Psychology Bulletin*, 20, 235-243.
- Cantor, N., & Malley, J. (1991). Life tasks, personal needs, and close relationships. In G. J. O. Fletcher & F. D. Fincham (Eds.), *Cognition in close relationships* (pp. 101-125). Hillsdale, NJ: Lawrence Erlbaum.
- Cate, R. M., Koval, J. E., Lloyd, S. A., & Wilson, G. (1995). The assessment of relationship thinking in dating relationships. *Personal Relationships*, 2, 77-95.
- Davila, J., Bradbury, T. N., & Fincham, F. (1998). Negative affectivity as a mediator of the association between adult attachment and marital satisfaction. *Personal Relationships*, 5, 467-484.
- Downey, G., Freitas, A. L., Michaelis, B., & Khouri, H. (1998). The self-fulfilling prophecy in close relationships: Rejection sensitivity and rejection by romantic partners. *Journal of Personality and Social Psychology*, 75, 545-560.
- Dunkel-Schetter, C., & Bennett, T. L. (1990). Differentiating the cognitive and behavioral aspects of social support. In I. G. Sarason, B. R. Sarason, & G. R. Pierce (Eds.), *Social support: An interactional view* (pp. 267-296). New York: John Wiley.
- Edwards, J. R. (1994). The study of congruence in organizational behavior research: Critique and a proposed alternative. *Organizational Behavior and Human Decision Processes*, 58, 51-100.
- Emmons, R. A., Diener, E., & Larsen, R. J. (1986). Choice and avoidance of everyday situations and affect congruence: Two models of reciprocal interactionism. *Journal of Personality and Social Psychology*, 51, 815-826.
- Feeney, J. A., Noller, P., & Callan, V. J. (1994). Attachment style, communication and satisfaction in the early years of marriage. *Advances in Personal Relationships*, 5, 269-308.
- Fincham, F., & Bradbury, T. N. (1989). The impact of attributions in marriage: An individual difference analysis. *Journal of Social and Personal Relationships*, 6, 69-85.
- Fincham, F., & Bradbury, T. N. (1990). Social support in marriage: The role of social cognition. *Journal of Social and Clinical Psychology*, 9, 31-42.
- Fiske, S. T., & Taylor, S. E. (1991). *Social cognition* (2nd ed.). New York: Random House.
- Franzoi, S. L., Davis, M. H., & Young, R. D. (1985). The effects of private self-consciousness and perspective taking on satisfaction in close relationships. *Journal of Personality and Social Psychology*, 48, 1584-1594.
- Gollwitzer, P. (1993). Goal achievement: The role of intentions. In W. Stroebe & M. Hewstone (Eds.), *European review of social psychology* (Vol. 4, pp. 141-185). London: Wiley.
- Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, 102, 4-27.
- Griffin, D., Murray, S., & Gonzalez, R. (1999). Difference score correlations in relationship research: A conceptual primer. *Personal Relationships*, 6, 505-518.
- Hendrick, S. S. (1981). Self-disclosure and marital satisfaction. *Journal of Personality and Social Psychology*, 40, 1150-1159.
- Hendrick, S. S., Hendrick, C., & Adler, N. L. (1988). Romantic relationships: Love, satisfaction, and staying together. *Journal of Personality and Social Psychology*, 54, 930-988.
- House, J. S., & Kahn, R. L. (1985). Measures and concepts of social support. In S. Cohen & S. L. Syme (Eds.), *Social support and health* (pp. 83-108). Orlando, FL: Academic Press.
- Ickes, W. (1993). Empathetic accuracy. *Journal of Personality*, 61, 587-610.
- Jacobson, N. S., & Moore, D. (1981). Spouses as observers of the events in their relationship. *Journal of Consulting and Clinical Psychology*, 49, 269-277.
- Jussim, L. (1989). Teacher expectations: Self-fulfilling prophecies, perceptual biases, and accuracy. *Journal of Personality and Social Psychology*, 57, 469-480.
- Karney, B. R., & Bradbury, T. N. (2000). Attributions in marriage: State or trait? A growth curve analysis. *Journal of Personality and Social Psychology*, 78, 295-309.
- Kenny, D. A. (1996). Models of non-independence in dyadic research. *Journal of Social and Personal Relationships*, 13, 279-294.
- McAdams, D. P. (1988). Personal needs and personal relationships. In S. W. Duck (Ed.), *Handbook of personal relationships* (pp. 7-22). New York: John Wiley.
- McAdams, D. P., & Constantian, C. A. (1983). Intimacy and affiliation motives in daily living: An experience sampling analysis. *Journal of Personality and Social Psychology*, 45, 851-861.
- Miller, L. C. (1990). Intimacy and liking: Mutual influence and the role of unique relationships. *Journal of Personality and Social Psychology*, 59, 50-60.
- Miller, L. C., & Read, S. J. (1991). On the coherence of mental models of persons and relationships: A knowledge structure approach. In G. O. Fletcher & F. D. Fincham (Eds.), *Cognition in close relationships* (pp. 69-99). Hillsdale, NJ: Lawrence Erlbaum.
- Mischel, W., Cantor, N., & Feldman, S. (1996). Principles of self-regulation: The nature of willpower and self-control. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 329-360). New York: Guilford.
- Murray, S. L., Holmes, J. G., & Griffin, D. W. (1996a). The benefits of positive illusions: Idealization and the construction of satisfaction in close relationships. *Journal of Personality and Social Psychology*, 70, 79-98.
- Murray, S. L., Holmes, J. G., & Griffin, D. W. (1996b). The self-fulfilling nature of positive illusions in romantic relationships: Love is not blind but prescient. *Journal of Personality and Social Psychology*, 71, 1155-1180.
- Niedenthal, P. M., Cantor, N., & Kihlstrom, J. F. (1985). Prototype matching: A strategy for social decision-making. *Journal of Personality and Social Psychology*, 48, 575-584.
- Niedenthal, P. M., & Mordkoff, J. T. (1991). Prototype distancing: A strategy for choosing among threatening situations. *Personality and Social Psychology Bulletin*, 17, 483-493.

- Nisbett, R. E., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice Hall.
- Ptacek, J. T., & Dodge, K. L. (1998). Coping strategies and relationship satisfaction in couples. *Personality and Social Psychology Bulletin*, 21, 76-84.
- Read, S. J., & Miller, L. C. (1989). Inter-personalism: Towards a goal-based theory of persons in relationships. In L. Pervin (Ed.), *Goal concepts in personality and social psychology* (pp. 413-472). Hillsdale, NJ: Lawrence Erlbaum.
- Reis, H. T., & Shaver, P. (1988). Intimacy as an interpersonal process. In S. W. Duck (Ed.), *Handbook of personal relationships* (pp. 367-389). New York: John Wiley.
- Ross, M. (1989). The relation of implicit theories to the construction of personal histories. *Psychological Review*, 96, 341-357.
- Ruvolo, A. P., & Fabian, L. A. (1999). Two of a kind: Perceptions of own and partner's attachment characteristics. *Personal Relationships*, 6, 57-79.
- Sanderson, C. A., & Cantor, N. (1995). Social dating goals in late adolescence: Implications for safer sexual activity. *Journal of Personality and Social Psychology*, 68, 1121-1134.
- Sanderson, C. A., & Cantor, N. (1997). Creating satisfaction in steady dating relationships: The role of personal goals and situational affordances. *Journal of Personality and Social Psychology*, 73, 1424-1433.
- Sanderson, C. A., & Evans, S. M. (2001). Seeing one's partner through intimacy-colored glasses: The role of intimacy goals in predicting relationship satisfaction in college women. *Personality and Social Psychology Bulletin*, 27, 461-471.
- Sanderson, C. A., & Karetzky, K. H. (in press). Intimacy goals and strategies of conflict resolution in dating relationships: A mediational analysis. *Journal of Social and Personal Relationships*.
- Sarason, B. R., Shearlin, E. N., Pierce, G. R., & Sarason, I. G. (1987). Interrelationships between social support measures: Theoretical and practical implications. *Journal of Personality and Social Psychology*, 52, 813-832.
- Sillars, A. L., Pike, G. R., Jones, T. S., & Murphy, M. A. (1984). Communication and understanding in marriage. *Human Communication Research*, 10, 317-350.
- Sillars, A. L., & Scott, M. D. (1983). Interpersonal perception between intimates. *Human Communication Research*, 10, 153-176.
- Snyder, M. (1981). On the influence of individuals on situations. In N. Cantor & J. F. Kihlstrom (Eds.), *Personality, cognition, and social interaction* (pp. 309-329). Hillsdale, NJ: Lawrence Erlbaum.
- Snyder, M., & Ickes, W. (1985). Personality and social behavior. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology* (pp. 883-947). New York: Random House.
- Snyder, M., & Simpson, J. A. (1984). Self-monitoring and dating relationships. *Journal of Personality and Social Psychology*, 47, 1281-1291.
- Snyder, M., Tanke, E. D., & Berscheid, E. (1977). Social perception and interpersonal behavior: On the self-fulfilling nature of social stereotypes. *Journal of Personality and Social Psychology*, 35, 656-666.
- Spanier, G. B. (1976). Measuring dyadic adjustment: New scales for assessing the quality of marriage and similar dyads. *Journal of Marriage and the Family*, 38, 15-28.
- Swann, W. B., Jr., De La Ronde, C., & Hixon, J. G. (1994). Authenticity and positivity strivings in marriage and courtship. *Journal of Personality and Social Psychology*, 66, 857-869.
- Swann, W. B., Jr., & Predmore, S. C. (1985). Intimates as agents of social support: Sources of consolation or despair? *Journal of Personality and Social Psychology*, 49, 1609-1617.
- Tuckers, J. S., & Anders, S. L. (1999). Attachment style, interpersonal perception accuracy, and relationship satisfaction in dating couples. *Personality and Social Psychology Bulletin*, 25, 403-412.
- Weiss, R. S. (1990). Bringing work stress home. In J. Eckenrode & S. Gore (Eds.), *Stress between work and family: The Plenum series on stress and coping* (pp. 17-37). New York: Plenum.

Received May 10, 2000

Revision accepted December 6, 2000

The Effect of Fleeting Attraction on Compliance to Requests

Jerry M. Burger
Shelley Soroka
Katrina Gonzago
Emily Murphy
Emily Somervell
Santa Clara University

Three studies examined the effects of fleeting attraction on compliance to a request. Participants in Study 1 who either spoke with a confederate for a few minutes or sat quietly in a room with the confederate were more likely to agree to a request from the confederate than were participants not exposed to these manipulations. Findings from Study 2 replicated the mere exposure effect and argue against alternative interpretations based on priming and mood. Study 3 participants were more likely to agree with a request when led to believe the requester was similar to themselves. The findings support the notion of automatic responding to requests, with individuals reacting to fleeting feelings of attraction as if dealing with friends and long-term acquaintances.

Investigators find support for the commonsense notion that we are more likely to go along with requests from friends than with those from strangers (Cialdini & Trost, 1998; Clark, Ouellette, Powell, & Milberg, 1987; Frenzen & Davis, 1990; Williamson & Clark, 1992). Neighborhood children selling fund-raising candy, coworkers recruiting volunteers to organize an event, and relatives who invite us to Tupperware parties have an advantage over unfamiliar sellers and unacquainted solicitors making the same requests.

But why do we agree to requests from friends more often than when approached by a stranger? Several explanations for this effect can be advanced. First, it is pleasing to do nice things for those we care about. Studies find that agreeing to small requests from friends creates positive feelings (Williamson & Clark, 1992), whereas refusing these requests leads to negative affect and might even endanger the relationship (Williamson, Clark, Pegalis, & Behan, 1996). These reactions are espe-

cially likely in what Clark and Mills (1979) refer to as “communal” relationships, in which individuals are concerned about the needs of the other person. We are more likely to help people with whom we feel a communal relationship than those for whom we have a less intimate relationship (Clark et al., 1987). Another reason we comply with requests from friends is that we may anticipate future exchanges with these people (Frenzen & Davis, 1990). Most people in our society abide by the norm of reciprocity (Gouldner, 1960). This social rule maintains, among other things, that favors must be returned. Thus, we may comply with a request from a friend with the understanding that, when needed, we can count on that friend to help us. Finally, it also is possible that the tendency to help friends is influenced by our evolutionary heritage (Caporael, 1997). Evolutionary theorists argue that members of a society are more likely to survive when they help one another than when they act only in terms of self-interest (Wilson & Sober, 1994). Thus, from this perspective, we would expect a tendency to help those with whom we feel some sort of association. In sum, there are many reasons to believe that we are more likely to agree to a request from someone we know and like than from someone we do not know or do not like.

But what about more ephemeral experiences with liking? Are we more likely to buy a car from a friendly sales-

Authors' Note: We would like to thank Linda Chou, Danielle Cleveland, and Whitney Cole for their assistance with the data collection. Correspondence should be sent to Jerry M. Burger, Department of Psychology, Santa Clara University, Santa Clara, CA 95053; e-mail: jburger@scu.edu.

PSPB, Vol. 27 No. 12, December 2001 1578-1586
© 2001 by the Society for Personality and Social Psychology, Inc.

person who appears to be a great person or donate money to a pleasant solicitor who seems like the kind of individual we would like to know better? We propose that even short, fleeting feelings of attraction can lead to an increased likelihood of compliance to a request. These feelings of liking can be very brief and can occur in situations in which the association with the other person is obviously temporary. Nonetheless, we argue that fleeting increases in attraction can lead to significant increases in compliance.

But why might this be the case? Clearly, none of the reasons we comply with friends are present in these short-lived associations. The answer is that people typically do not invest a lot of thought into deciding their response when presented with a request. Investigators find that compliance to simple requests often reflects automatic or heuristic information processing rather than a thoughtful, reasoned consideration of relevant facts (Cialdini, 2001). Face-to-face requests in particular require a quick, somewhat mindless response (Langer, 1989). For example, we often respond to a free gift by donating money or buying a product from the person who gave us that gift (Burger, Horita, Kinoshita, Roberts, & Vera, 1997; Regan, 1971). This is true even if we did not ask for the gift and even if we do not like the gift giver. Rather than consider the obvious manipulation, people typically rely on the well-learned reciprocity principle that says favors must be returned in some form. On the other hand, in some cases, the practiced, automatic response to a request is refusal. Santos, Leve, and Pratkanis (1994) found mindless refusal to a panhandler's request for change was the typical response from passersby. The investigators succeeded in overcoming this mindless response, and thereby increasing the amount of money collected, only when the panhandler asked for an unusual amount (17 cents or 37 cents).

We propose that a similar mindless heuristic comes into play when presented with a request from someone for whom we feel a fleeting amount of attraction. Over the course of a lifetime, most of us come to understand that it is appropriate and rewarding to agree to requests from friends and others with whom we have close relationships. As a result, we rely on a heuristic that says we agree with requests from people we like. When the request comes from a friend, we respond appropriately with little or no thought. However, when the request comes from an individual for whom we have only a fleeting feeling of attraction, we may still apply the heuristic and thereby comply in an automatic fashion to the request. This is the experience we describe when we donate to a solicitor's cause "because he was so nice" or

purchase unnecessary cookies from the Girl Scouts "because they were so cute."

The kind of fleeting attraction that is likely to trigger this automatic response is, of course, different in many ways from the kind of emotional attachment we feel for long-term friends. However, studies demonstrate that people often experience a kind of liking for individuals with whom they share but a short conversation, a plane ride, or a common task. The history of social psychology provides ample examples of experimental manipulations designed to increase these temporary feelings of liking. For example, self-disclosure to and by a randomly assigned partner has been found to increase self-reported levels of attraction in structured laboratory conversations (Collins & Miller, 1994). Similarly, 10-min get-acquainted discussions have been found to produce higher levels of liking toward the discussion partner than toward someone not part of the conversation (Insko & Wilson, 1977). Researchers even find that repeated exposure to a student who sits in the same classroom without direct interaction increases self-reported liking for the unknown student (Moreland & Beach, 1992).

Although we know of no study that has demonstrated increased compliance with fleeting attraction, findings from several investigations are consistent with this notion. For example, we often are attracted to people who share similar values, interests, and personality characteristics (Byrne, 1997). This is true even when we have relatively little information or exposure to the person in question. Consistent with our reasoning, one study found that passersby were more willing to give money to a stranger for a phone call when that stranger was dressed in a manner similar to themselves (Emswiller, Deaux, & Willits, 1971). Other research finds that we tend to like physically attractive people more than less attractive people, at least during initial encounters (Hatfield & Sprecher, 1986). When individuals in one study were asked for donations to a charitable cause, they gave more money to physically attractive requesters than to less attractive requesters (Reingen & Kernan, 1993). Finally, waitresses in one study received higher tips when they used their first names with their customers, something we associate with friends (Garrity & Degelman, 1990).

We conducted three studies to test the hypothesis that small, ephemeral increases in liking toward a stranger will lead to an increased likelihood of complying to a request from that person. To test this prediction, we used procedures previously demonstrated to increase liking in laboratory studies. We then presented participants with a request from the newly attractive individual or

from a stranger. We predicted higher rates of compliance in the liking conditions.

STUDY 1

We manipulated liking in this study two ways. First, investigators find that simply talking to another person increases our attraction to him or her. Insko and Wilson (1977) demonstrated this phenomenon in a study in which three participants were seated in a small room. The three people—identified as Participants A, B, and C—were instructed to engage in get-acquainted conversations. First, A spoke with B while C listened. Then, B spoke with C while A listened. Although Participants A and C heard everything that was said in the room, these two never interacted directly. The researchers found that the participants reported greater liking for the individual they spoke to than for the one they did not. Insko and Wilson explained this increased attraction in terms of unit relationships (Heider, 1958); that is, the interacting participants perceived that they shared a special association that the noninteracting participants did not. Consistent with this explanation, Arkin and Burger (1980) found that strengthening or weakening the perceptual salience of the unit relationship in the three-person discussion situation (such as placing one of the participants out of view) affected the level of liking in the predicted direction.

Second, investigators also find increased liking for an individual with repeated exposure to that person. This “mere exposure” effect appears robust and is not limited to feelings about people (Bornstein, 1989; Harrison, 1977; Zajonc, 1968). Researchers find an increase in liking for people who are frequently seen but not spoken to (Moreland & Beach, 1992; Saegert, Swap, & Zajonc, 1973; Swap, 1977). For example, Saegert et al. (1973) had participants stand next to other participants either 0, 1, 2, 5, or 10 times as they moved about from station to station in what was described as a taste-test experiment. Participants were instructed to not talk or use nonverbal gestures among themselves. Each silent exposure lasted for no more than 40 seconds, yet this was sufficient to increase liking ratings for the individuals for whom the participants had been repeatedly exposed. Interestingly, researchers find that exposure effects work as well and sometimes even better when the individual is not aware that exposure has occurred (Bornstein, 1989).

We manipulated attraction to a stranger either by allowing some participants to talk with the stranger for a few minutes or by having the stranger sit silently in the same room with the participant. We predicted that participants in both of these conditions would be more likely to agree to a subsequent request from the stranger than participants not exposed to these liking manipulations.

Method

Participants. The study consisted of 114 female undergraduates who served as participants in exchange for class credit.

Procedure. Participants were randomly assigned to one of three conditions. After arriving to the study, participants in the interaction condition were seated at a table in the lab room. Approximately 30 seconds later, a female confederate posing as another participant entered the room and was seated across the table from the real participant. The two sat approximately 2 meters apart. The experimenter explained that the study concerned “manual dexterity and cognitive recognition skills.” She then gave each participant a sheet of paper that contained several rows of randomly typed letters. The experimenter explained that they would begin the study with a simple manual dexterity task. She explained that the participants would have 3 minutes to cross out as many of the letters *l*, *k*, and *s* as they could find on the sheet. The experimenter also said that she was running several participants simultaneously and that she would not be in the room during the task. She placed a timer in the middle of the table and instructed the participants to stop when the timer went off. The experimenter then started the participants on the task, started the timer, and left the room. The participant and the confederate worked on the task until the timer went off 3 minutes later, but the experimenter did not return for another 2 minutes. During that time, the confederate, who was blind to the hypotheses, initiated a conversation with the participant. Confederates were instructed to be appropriately friendly but to limit topics to school, classes, and professors. All participants joined in the conversations, which lasted approximately 2 minutes.

After returning, the experimenter explained that the participant and confederate had been part of a control group and that their role in the study was now over. She gave the two credit for their participation and dismissed them. After the participant and confederate left the room and were heading out of the laboratory area, the confederate presented the target request. She explained that her English instructor had required class members to get someone they did not know to critique an essay they had written. The confederate then pulled an essay from her backpack and said, “I wonder if you could read this eight-page essay for me and give me one page of written feedback on whether my arguments are persuasive and why?” The confederate added that she would need the written feedback by approximately this time the following day. After the participant agreed to or declined the request, the experimenter (who had been hiding out of sight) appeared and asked the participant and confederate to return to the lab room with her. The experi-

menter explained that she forgot to debrief them. She then probed for suspiciousness (none was found) and fully debriefed the participant.

Participants assigned to the mere exposure condition went through an identical procedure, except the experimenter instructed the participant and confederate that they were not to talk during the time between the end of the manual dexterity test and when she reentered the room. All participants obeyed these instructions. Thus, the confederate and the participant sat quietly throughout the study and exchanged no words until after they were out of the lab room. Confederates were instructed to spend most of their time looking over the completed test, making appropriate but limited eye contact with the participant. The first words spoken by the confederate to the participant were the target request.

Participants in the control condition took the manual dexterity test alone. As in the other conditions, they stopped when the timer went off and waited 2 minutes for the experimenter to return. As the participant was gathering up her things to leave, the confederate approached the lab room door (presumably from another lab room) and handed the experimenter her credit form to be signed. The experimenter signed the form in full view of the real participant, thanked both the confederate and the participant for their time, then quickly retreated out of view. At this point, the confederate presented the real participant with the target request.

Results and Discussion

We compared the number of participants who agreed to the request in each of the two experimental conditions against the number who complied in the control condition. As expected, participants in the interaction condition (19/39, 48.7%) were more likely to agree to the request than participants in the control condition (10/38, 26.3%), $\chi^2(1, N = 77) = 4.11, p < .05, \phi = .23$. Similarly, participants in the mere exposure condition (18/37, 48.6%) were more likely to comply with the request than participants in the control condition, $\chi^2(1, N = 75) = 4.00, p < .05, \phi = .23$. The interaction and mere exposure conditions did not differ significantly on this measure.

The findings thus support the notion that small, fleeting increases in liking between strangers can result in an increased tendency to comply with a request. In this study, 2 minutes of small talk with the confederate increased liking sufficiently to nearly double the rate of compliance to the confederate's request. Similarly, consistent with research on mere exposure, simply sitting across the table from the confederate momentarily increased liking enough to significantly increase compliance. We were a little surprised to find that the two liking manipulations were equally effective. That is, consistent

with the unit relationship notion and earlier research that finds increased attraction with interaction, we might have expected that talking with the confederate would lead to greater liking—and thus more compliance—than simply seeing the person across the table. It is possible that we hit an upper limit on the number of people who would go along with the request, regardless of the strength of the liking manipulation. However, we were not interested in the relative strength of the two liking manipulations but rather whether either or both of these would increase compliance relative to the control condition.

STUDY 2

The findings from Study 1 suggest that carrying on a short conversation or simply sitting in the same room with someone significantly increases compliance to a simple request from that other individual. We argue that this effect can be best explained in terms of fleeting feelings of attraction that trigger the use of heuristic information processing. However, it also is possible that other aspects of the situation used in the first study contributed to the increase in compliance. For example, engaging in conversation or being in the presence of others may have focused the participants' attention on social interactions. Another way to say this is that information about social interactions may have been primed by the experience, thus making thoughts about social interactions more accessible for these participants. This process is similar to that demonstrated in numerous investigations on automatic activation and nonconscious processing (Bargh, 1989; Bargh & Chartrand, 1999; Higgins, 1996). As a result of this increased accessibility of social information, it is possible that the participants in the interaction and mere exposure conditions more readily responded to the social request. This interpretation is not entirely inconsistent with our automatic processing explanation for the Study 1 findings. However, we argue that temporary feelings of attraction, rather than merely being in a social setting, trigger the automatic response.

Another possibility is that the conversation or mere exposure situation used in the first study altered the participants' moods. That is, because we typically find social interactions pleasant and solitude unpleasant (Larson, 1990), it is possible that participants in the two experimental conditions were in a more positive mood than the participants who sat in the room alone. Past studies have found a positive relation between good mood and helping behavior (Isen, Shalke, Clark, & Karp, 1978). Thus, at this point, we cannot rule out the possibility that the participants' mood was responsible for the higher rates of compliance in the two experimental conditions in Study 1.

Study 2 was designed to rule out these two alternative accounts for the effect uncovered in the first study. Specifically, we replicated the mere exposure condition and the control condition used in Study 1. We also included a condition similar to the mere exposure condition in which participants sat in a room without conversation. However, in this latter condition, the request was presented to participants by someone other than the individual to whom the participants were exposed. If fleeting feelings of attraction were responsible for the Study 1 results, then we should see increased compliance with mere exposure only when the request comes from the person who was actually in the room with the participant. If either of the alternative explanations are correct—priming social interaction information or mood—then the presence of the confederate should increase compliance regardless of who presents the request.

Method

Participants. The study consisted of 120 female undergraduates who served as participants in exchange for class credit.

Procedure. Female experimenters were used for all conditions. Participants were randomly assigned to three conditions. The mere exposure–same condition was identical to the mere exposure condition used in the first study. The control condition also was identical to the comparable condition in Study 1. The mere exposure–different condition was identical to the mere exposure–same condition with the following exceptions. As in the control condition, the experimenter gave the impression that she was conducting simultaneous experimental sessions. After the experimenter announced the study was over, the participant and the confederate who had been in the room with the participant were joined by another confederate who presumably had been part of the study but working in another room. The experimenter gave the three individuals experimental credit, always signing the credit form for the confederate who had been in the participant's room first. That confederate quickly left the setting while the experimenter signed the credit form for the confederate who presumably had been in another room. The real participant always had her credit form signed last so that the experimenter could leave the participant and the remaining confederate together. At that point, the confederate presented the request. As in Study 1, no participants expressed suspicion when asked and all were debriefed at the end of the study.

Results and Discussion

We examined the number of people who agreed to the request in each condition. Participants in the mere exposure–same condition (22/40, 55.0%) agreed to the

request significantly more often than did participants in the control condition (8/40, 20.0%), $\chi^2(1, N = 80) = 9.01, p < .004, \phi = .34$, thus replicating the findings from the first study. More important, the mere exposure–same participants complied with the request more often than did participants in the mere exposure–different condition (9/40, 22.5%), $\chi^2(1, N = 80) = 7.58, p < .007, \phi = .31$. The mere exposure–different and control conditions did not differ significantly.

The findings provide additional support for the notion that small, ephemeral increases in liking can lead to a significant increase in compliance. As in Study 1, participants who were simply exposed to another person for several minutes were more likely to comply with a request from that person than a request from someone they had not seen before. More important, the findings from the second study argue against alternative explanations for the effect uncovered in Study 1. Although it is possible that participants in the two-person conditions were better able to process social information or were in a different mood than participants who sat in the room alone, these reactions did not appear to affect compliance behavior in this study. Increased compliance was found only when the person to whom the participant was exposed asked the request. When the request came from someone for whom the participant had no exposure and thus no increased liking, she was no more likely to agree to the request than were participants in the control condition.

STUDY 3

The purpose of the third study was twofold. First, we wanted to replicate the liking–compliance effect produced in Studies 1 and 2 using another type of liking manipulation. In this way, we hoped to demonstrate a consistent pattern of results that could not be attributed to some unique feature of the procedures used in the first two studies. Second, we wanted to include a manipulation check measure to ensure that we were indeed manipulating liking. Directly assessing liking was not possible in the earlier studies without raising suspicions about the true nature of the investigation. Therefore, we developed a procedure that would provide us with a measure of liking without alerting participants to the hypothesis.

We manipulated liking in Study 3 by altering the degree of perceived similarity between the participant and the confederate who presents the request. Numerous investigations have found an increase in liking with perceived similarity (Byrne, 1971, 1997; Smeaton, Byrne, & Murnen, 1989). Although a variety of methods have been used to manipulate perceived similarity in attitudes, personality, values, and so forth, many researchers have produced increased liking simply by providing false

information about the other individual in a laboratory setting (Byrne, 1997). Thus, we also used a false information procedure to manipulate degree of liking. We predicted an increase in liking and subsequently greater compliance with an increase in perceived similarity of personalities.

Method

Participants. The study consisted of 90 female undergraduates who served as participants in exchange for class credit.

Procedure. Participants arrived at the experimental room alone. The experimenter explained that the study was concerned with first impressions. She explained that the participant and another female participant waiting in a nearby lab room would exchange information about themselves and form first impressions of one another. The two participants were being kept apart supposedly to control for the effects of physical appearance. The experimenter then handed the participant a questionnaire containing 50 adjectives. The participant was instructed to read over the adjectives and indicate the 20 that described her best. The experimenter explained that she would give the other participant the same questionnaire and that a little later the two participants would see each other's responses. The experimenter then left the room, presumably to administer the questionnaire to the other participant.

When completed, the experimenter collected the questionnaire, supposedly to give to the other participant. Once she left the room, the experimenter checked a list that randomly assigned participants to one of three conditions. She then quickly filled out a blank adjective checklist. If the participant was assigned to the similar condition, the responses on this new checklist indicated that 17 of the 20 adjectives selected by the "other" participant were identical to those the real participant used to describe herself. Questionnaire responses for participants in the dissimilar condition indicated that the other individual selected only 3 of 20 items that matched those of the real participant. Finally, if participants were assigned to the neutral condition, the experimenter selected items to indicate that the other participants' self-descriptive adjectives matched on 10 of 20 items.

The experimenter allowed the real participant about 2 minutes to study the "other participant's" questionnaire. She then administered a final questionnaire supposedly to assess first impressions. Participants answered several questions about the confederate on 7-point scales. The last three items on this questionnaire were designed to measure the participant's degree of liking for the confederate. These three questions were as follows: Do you think you would like this person if you got to know her better? Would you enjoy time spent with this

TABLE 1: Liking Scores and Compliance Rates—Study 3

	Liking	Compliance
Similar condition	17.93	23/30 (76.7%)
Neutral condition	16.63	18/30 (60.0%)
Dissimilar condition	15.23	13/30 (43.3%)

person? and Do you think you could be long-term friends with this person?

The experimenter collected the completed questionnaire and, as in the previous two studies, brought the real participant and a confederate playing the other participant together to give each credit. The experimenter then quickly excused herself, leaving the participant and the confederate alone to exit the laboratory setting. At this point, the confederate presented the same request used in the earlier studies. After responding to the request, the experimenter probed for suspicion (none was found) and debriefed the participants.

Results and Discussion

Responses to the three liking items on the second questionnaire were highly correlated (r s between .64 and .79). Thus, we summed the three response values to form an overall liking measure ($\alpha = .87$, $M = 16.6$, $SD = 2.96$). As shown in Table 1, the mean liking score differed significantly across the three conditions, $F(2, 87) = 7.12$, $p < .001$. Specific cell comparisons revealed a significant difference between the similar and dissimilar condition, $p < .001$, Tukey's honestly significant difference (HSD) test. The difference between the neutral condition and the similar and dissimilar conditions fell short of statistical significance, $p < .17$ and $p < .13$, respectively. Thus, the manipulation appears to have been successful.

Next, we compared the number of participants who agreed to the request in each condition. When each of the specific conditions was compared against the others, only the similar and dissimilar conditions were significantly different, $\chi^2(1, N = 60) = 6.94$, $p < .008$, $\phi = .34$.

Although the pattern in the data reported in Table 1 conforms nicely with our expectations, it is easier to interpret the results if we compare only the similar and dissimilar conditions for the moment. Consistent with past research findings, participants who thought the person in the other room was similar to themselves held greater liking for that person than when they thought they had little in common with that individual. More important, these participants also were more likely to agree to a request from that person than when they thought the person was dissimilar.

Finally, we used a series of regression analyses to look for evidence that the relationship between similarity and compliance is mediated by liking. As described by Baron

and Kenny (1986), mediation or partial mediation would be demonstrated in this study when each of the following relationships is found: (a) similarity (condition) is related to liking, (b) liking is related to compliance, and (c) a previously significant relation between similarity and compliance is no longer significant when the mediator variable (liking) is included in the analysis. Partial mediation is indicated if this relation is reduced but remains significant when the mediator variable is included. Because the strongest evidence for our hypothesis was found when comparing the similar and dissimilar conditions, we used only these two conditions in the mediation analyses.

The first necessary relationship—between similarity and liking—has already been presented. Participants in the similarity condition reported significantly higher liking than those in the dissimilar condition. Because the dependent variable in our study was a dichotomous variable, we used a series of logistic regressions to demonstrate the remaining relationships. The key statistic when examining mediation with logistic regression is the chi-square value. When we used liking to predict compliance—the second necessary relationship—we produced a significant chi-square value of 4.11 ($p < .05$). Finally, when predicting compliance from similarity without the liking score entered into the analysis, we obtain a chi-square of 7.11 ($p < .008$). However, when we entered the liking score into the regression equation first, the chi-square statistic for this relationship dropped to 5.28 ($p < .03$). The results of the logistic regressions thus suggest partial, but not full, mediation.

GENERAL DISCUSSION

In each of three studies, we found that small manipulations that intended to increase fleeting feelings of attraction resulted in increased levels of compliance with a request from the liked individual. Three procedures often used by social psychologists to increase liking—interaction, mere exposure, and perceived similarity—produced the effect. The findings are consistent with a general description of compliance behavior that emphasizes automatic processing over reasoned, thoughtful responding to requests (Cialdini, 2001). It appears that most people reply quickly and mindlessly to requests for small favors and purchases, relying on heuristics to determine if they should or should not go along with the request. Indeed, most of the explanations proposed for the effectiveness of the sequential-request compliance procedures (e.g., foot-in-the-door, door-in-the-face) assume that individuals are unaware of why they respond the way they do (Burger, 1999; Cialdini, 2001). In our studies, participants appeared to rely on a heuristic that says, “I go along with requests from friends and people I like.” We argue that the heuristic was acti-

vated by the presence of fleeting feelings of liking toward the confederates. A thoughtful analysis of the request and situation probably would have led participants to realize that their feelings of attraction were fleeting and that the reasons they typically help friends were not present. However, because they relied on the simple heuristic, participants in the liking conditions complied with the request almost as if they had been asked by a friend.

We also need to acknowledge some of the limitations and unanswered questions about the research reported here. To avoid practical complications arising from women asking men for favors, we used only female participants and female requesters in our investigations. Although we have no reason to expect the effect would not also be found with men, the question of gender differences remains open. Another concern has to do with the possibility of experimenter bias; that is, although confederates were kept blind to hypotheses in Studies 1 and 2, we cannot rule out that they nonetheless suspected the hypotheses and inadvertently altered the way they delivered the request in some conditions. Arguing against this possibility is the finding that the interaction and mere exposure conditions did not differ in their rates of compliance. If our confederates suspected any hypothesis, most likely they would have guessed that interaction would lead to greater compliance than noninteraction. Finally, Study 3 confederates did not know which condition the participant was in and thus could not have altered their behavior according to condition. Nonetheless, the relation between liking and compliance was demonstrated in this study.

Another set of questions concerns unanticipated participant reactions to the liking manipulation in the first two studies. For example, it is possible that the brief encounter with participants in Studies 1 and 2 did more than create a fleeting feeling of attraction. The short encounter with the confederate also might have generated an increased sense of trust. If that were the case, participants in these conditions may have agreed to the request more often because they trusted the confederate and felt more assured that the request was a valid one. Arguing against this interpretation is the fact that none of the participants voiced any suspicion about the request and its validity during the debriefing.

The manipulation of liking in Study 3 also raises some questions. First, it is possible that participants described themselves with flattering adjectives on the checklist. If that were the case, then the adjectives used to describe the confederate would have been more desirable in the similar condition than in the other two conditions. Therefore, one might argue that participants liked the confederate because she possessed desirable characteristics, not because she was similar. Although future studies are needed to tease out the answer to this question,

we should note that we are concerned here with the effects of fleeting feelings of liking regardless of their source. Another concern about the Study 3 procedures has to do with the possibility that we primed cognitions related to attraction when we asked participants how much they liked the confederate and that such priming could have had an effect on compliance rates. Although this is a possibility, because participants answered these questions in each condition, we cannot attribute the differences among conditions to possible priming.

Finally, there is some disagreement about whether participants in similarity-attraction studies are attracted to similar people or reject dissimilar people (Byrne, Clore, & Smeaton, 1986; Rosenbaum, 1986); that is, although we argue that similarity leads to attraction, we cannot rule out that our effects are due to an increased *dislike* of the confederate in the dissimilar condition in Study 3. Of course, this concern does not apply to the findings from the first two studies, in which there is no dislike condition and no reason to suspect that control condition participants disliked the confederate. Nonetheless, if future studies demonstrate that dissimilarity leads to a decrease in liking and thus a decrease in compliance, we might have to expand our notion of fleeting attraction to include fleeting repulsion.

We also can identify some issues that might be explored in future investigations. One of these issues concerns the size of the request. It is reasonable to speculate that people are more likely to respond in a mindless fashion to a small request than a large request; that is, at some point the request is so costly that the individual will be forced to consider it carefully before responding. One can only determine where mindlessness ends and mindfulness begins empirically. However, it is interesting that the request used in our studies—writing a page of criticism—was not trivial yet apparently was not sufficient to shake our participants out of heuristic processing. Another issue worth exploring is the relationship between similarity, familiarity, and what we call fleeting attraction. Investigators find that these concepts are highly related and may combine to form a psychological connection called affinity (Moreland & Beach, 1992; Moreland & Zajonc, 1982). We have been operating on the assumption that familiarity (Studies 1 and 2) and similarity (Study 3) lead to feelings of liking. However, further examination into how these concepts affect one another might prove useful.

Our final observations have to do with some of the broader implications of the research. Our description of the mindless manner in which individuals respond to requests is consistent with other social psychological models that assume people often rely on relatively effortless shortcuts when processing information (Chaiken &

Trope, 1999; Smith & DeCoster, 2000). In most cases, relying on heuristics allows us to proceed through life smoothly and efficiently. Automatically agreeing to requests from people we like probably works well in most cases. Unfortunately, this type of mindless responding also makes us vulnerable to those who understand how to exploit these tendencies (Cialdini, 2001). This observation leads to our final point, that our results have obvious practical applications. To those interested in sales, recruiting, and the like, the findings suggest yet another way to increase agreement to requests. Any action that makes the requester appear to be a likable person may trigger the liking-compliance heuristic described here. Thus, in the hands of a clever salesperson, a hearty smile, a friendly conversation, or some similarity in appearance or background can become a very valuable tool.

REFERENCES

- Arkin, R. M., & Burger, J. M. (1980). Effects of unit relation tendencies on interpersonal attraction. *Social Psychology Quarterly*, 43, 380-391.
- Bargh, J. A. (1989). Conditional automaticity: Varieties of automatic influence in social perception and cognition. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 3-51). New York: Guilford.
- Bargh, J. A., & Chartrand, T. L. (1999). The unbearable automaticity of being. *American Psychologist*, 54, 462-479.
- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51, 1173-1182.
- Bornstein, R. F. (1989). Exposure and affect: Overview and meta-analysis of research, 1968-1987. *Psychological Bulletin*, 106, 265-289.
- Burger, J. M. (1999). The foot-in-the-door compliance procedure: A multiple-process analysis and review. *Personality and Social Psychology Review*, 3, 303-325.
- Burger, J. M., Horita, M., Kinoshita, L., Roberts, K., & Vera, C. (1997). The effects of time on the norm of reciprocity. *Basic and Applied Social Psychology*, 19, 91-100.
- Byrne, D. (1971). *The attraction paradigm*. New York: Academic Press.
- Byrne, D. (1997). An overview (and underview) of research and theory within the attraction paradigm. *Journal of Social and Personal Relationships*, 14, 417-431.
- Byrne, D., Clore, G. L., & Smeaton, G. (1986). The attraction hypothesis: Do similar attitudes affect anything? *Journal of Personality and Social Psychology*, 51, 1167-1170.
- Caporael, L. R. (1997). The evolution of truly social cognition: The core configurations model. *Personality and Social Psychology Review*, 1, 276-298.
- Chaiken, S., & Trope, Y. (Eds.). (1999). *Dual-process theories in social psychology*. New York: Guilford.
- Cialdini, R. B., & Trost, M. R. (1998). Social influence: Social norms, conformity, and compliance. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology*, Vol. 2 (4th ed., pp. 151-192). Boston: McGraw-Hill.
- Cialdini, R. B. (2001). *Influence: Science and practice* (4th ed.). Boston: Allyn & Bacon.
- Clark, M. S., & Mills, J. (1979). Interpersonal attraction in exchange and communal relationships. *Journal of Personality and Social Psychology*, 37, 12-24.
- Clark, M. S., Ouellette, R., Powell, M. C., & Milberg, S. (1987). Recipient's mood, relationship type, and helping. *Journal of Personality and Social Psychology*, 53, 94-103.

- Collins, N. L., & Miller, L. C. (1994). Self-disclosure and liking: A meta-analytic review. *Psychological Bulletin*, 116, 457-475.
- Emswiller, T., Deaux, K., & Willits, J. E. (1971). Similarity, sex, and requests for small favors. *Journal of Applied Social Psychology*, 1, 284-291.
- Frenzen, J. K., & Davis, H. L. (1990). Purchasing behavior in embedded markets. *Journal of Consumer Research*, 17, 1-12.
- Garritty, K., & Degelman, D. (1990). Effect of server introduction on restaurant tipping. *Journal of Applied Social Psychology*, 20, 168-172.
- Gouldner, A. W. (1960). The norm of reciprocity: A preliminary statement. *American Sociological Review*, 25, 161-178.
- Harrison, A. A. (1977). Mere exposure. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 10, pp. 39-83). New York: Academic Press.
- Hatfield, E., & Sprecher, S. (1986). *Mirror, mirror . . . : The importance of looks in everyday life*. Albany: State University of New York Press.
- Heider, F. (1958). *The psychology of interpersonal relations*. New York: John Wiley.
- Higgins, E. T. (1996). Knowledge activation: Accessibility, applicability, and salience. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 133-168). New York: Guilford.
- Insko, C. A., & Wilson, M. (1977). Interpersonal attraction as a function of social interaction. *Journal of Personality and Social Psychology*, 35, 903-911.
- Isen, A. M., Shalke, T. E., Clark, M., & Karp, L. (1978). Affect, accessibility of material in memory, and behavior: A cognitive loop? *Journal of Personality and Social Psychology*, 36, 1-12.
- Langer, E. J. (1989). *Mindfulness*. Reading, MA: Addison-Wesley.
- Larson, R. W. (1990). The solitary side of life: An examination of the time people spend alone from childhood to old age. *Developmental Review*, 10, 155-183.
- Moreland, R. L., & Beach, S. R. (1992). Exposure effects in the classroom: The development of affinity among students. *Journal of Experimental Social Psychology*, 28, 255-276.
- Moreland, R. L., & Zajonc, R. B. (1982). Exposure effects in person perception. *Journal of Experimental Social Psychology*, 18, 395-415.
- Regan, D. T. (1971). Effects of a favor and liking on compliance. *Journal of Experimental Social Psychology*, 7, 627-639.
- Reingen, P. H., & Kernan, J. B. (1993). Social perception and interpersonal influence: Some consequences of the physical attractiveness stereotype in a personal selling setting. *Journal of Consumer Psychology*, 2, 25-38.
- Rosenbaum, M. E. (1986). The repulsion hypothesis: On the nondevelopment of relationships. *Journal of Personality and Social Psychology*, 51, 1156-1166.
- Saegert, S., Swap, S., & Zajonc, R. B. (1973). Exposure, context, and interpersonal attraction. *Journal of Personality and Social Psychology*, 25, 234-242.
- Santos, M. D., Leve, C., & Pratkanis, A. R. (1994). Hey buddy, can you spare seventeen cents? Mindful persuasion and the pique technique. *Journal of Applied Social Psychology*, 24, 755-764.
- Smeaton, G., Byrne, D., & Murnen, S. K. (1989). The repulsion hypothesis revisited: Similarity irrelevance or dissimilarity bias? *Journal of Personality and Social Psychology*, 56, 54-59.
- Smith, E. R., & 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.
- Swap, W. C. (1977). Interpersonal attraction and repeated exposure to rewarders and punishers. *Personality and Social Psychology Bulletin*, 3, 248-251.
- Williamson, G. M., & Clark, M. S. (1992). Impact of desired relationship type on affective reactions to choosing and being required to help. *Personality and Social Psychology Bulletin*, 18, 10-18.
- Williamson, G. M., Clark, M. S., Pegalis, L. J., & Behan, A. (1996). Affective consequences of refusing to help in communal and exchange relationships. *Personality and Social Psychology Bulletin*, 22, 34-47.
- Wilson, D. S., & Sober, E. (1994). Reintroducing group selection to the human behavioral sciences. *Behavioral and Brain Sciences*, 17, 585-654.
- Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, 9, 1-27.

June 2, 2000

November 15, 2000

Environmental and Genetic Influences on Sex-Typed Behaviors and Attitudes of Male and Female Adolescents

H. Harrington Cleveland

University of Nevada, Las Vegas

J. Richard Udry

Kim Chantala

Carolina Population Center; University of North Carolina, Chapel Hill

This study uses behavioral genetics to examine the contributions of environmental and genetic influences to the sex-typed behaviors and attitudes of adolescent males and females. Data were drawn from 1,301 sibling pairs ranging from monozygotic twins to half-sibling pairs from the National Longitudinal Study of Adolescent Health. The measure used was a probability score of being a male based on responses to 16 behavioral and attitudinal items. Using this scale, separate covariance matrices were computed across different levels of zygosity and fit to behavioral genetic models. It was estimated that for males, 25% of the variance was accounted for by genetic influences and 75% by nonshared environmental influences and measurement error. For females, 38% of the variance was accounted for by genetic influences and the remaining 62% by nonshared environmental influences and measurement error. No shared environmental influences were found for either sex's sex-typed behaviors and attitudes.

Parents are often concerned with the sex typicality of their children. This concern may stem from beliefs that sex-atypical behaviors make it difficult for their children to fit into school, strictly held norms about how males and females should act, or even assumptions about associations between being a sissy or a tomboy and later sexual orientation. The goal of this study is to evaluate potential influences on variation in sex-typical behaviors and attitudes among adolescents. Using behavioral genetic methods, this study estimates shared environmental, nonshared environmental, and additive genetic influences on sex-typical behaviors of male and female adolescents. These components of variance roughly correspond to three perspectives on how sex-typical behav-

iors and attitudes are acquired: vertical transmission from parents to children (Bem, 1981), experiences of children with peers and nonfamily adults outside the home (Bem, 1979; Maccoby, 1998; Mitchell, Baker, & Jacklin, 1989), and biological influences (Collaer & Hines, 1995; Udry, 2000).

At the outset, it is important to emphasize that the construct targeted by our study is sex typicality of behaviors and attitudes. Our measure of this construct derives directly from differences in male and female behaviors

Authors' Note: This research is based on data from the National Longitudinal Study of Adolescent Health (Add Health). The Add Health is a program project designed by J. Richard Udry (PI) and Peter Bearman and is funded by Grant P01-HD31921 from the National Institute of Child Health and Human Development to the Carolina Population Center, University of North Carolina at Chapel Hill, with cooperative funding participation by the National Cancer Institute; the National Institute of Alcohol Abuse and Alcoholism; the National Institute on Deafness and Other Communication Disorders; the National Institute on Drug Abuse; the National Institute of General Medical Sciences; the National Institute of Mental Health; the National Institute of Nursing Research; the Office of AIDS Research, National Institute of Health (NIH); the Office of Behavior and Social Science Research, NIH; the Office of the Director, NIH; the Office of Research on Women's Health, NIH; the Office of Population Affairs, Department of Health and Human Services (DHHS); the National Center for Health Statistics, Centers for Disease Control and Prevention, DHHS; the Office of Minority Health, Centers for Disease Control and Prevention, DHHS; the Office of Minority Health, Office of Public Health and Science, DHHS; the Office of the Assistant Secretary for Planning and Evaluation, DHHS; and the National Science Foundation. Correspondence concerning this article should be addressed to H. Harrington Cleveland, Department of Criminal Justice, University of Nevada, Las Vegas, 4505 Maryland Ave., Las Vegas, NV 89154-5009; e-mail: bo.cleveland@cmail.nevada.edu.

PSPB, Vol. 27 No. 12, December 2001 1587-1598

© 2001 by the Society for Personality and Social Psychology, Inc.

and attitudes observed in a nationally representative data set of adolescents. Prior to the 1970s, such measures of sex-typical behaviors and attitudes were common (Terman & Miles, 1936). This construct differs from, but is intricately related to, clusters of traits that are viewed as desirable for one sex or the other, which since the 1970s have become associated with the term *gender* (Bem, 1979; Spence & Helmreich, 1979). Whether one of these constructs is the appropriate definition of gender is not important for this study. What is important is to recognize the theoretical link between these two concepts. Theories on the acquisition of the latter construct, the most well known of which is Bem's (1981) Gender Schema Theory (GST), provide that underlying differences in average behaviors of males and females help create the context for children's acquisition of gender schema and that compliance with these resulting schema support the expression of sex-typical behaviors. This theoretical connection makes the sex-desirable traits literature relevant for understanding potential influences on the expression of sex-typical behaviors and attitudes. Accordingly, our review of the literature, especially in the areas of parent-to-child transmission and experiences outside the home that may influence the acquisition of sex-typical behaviors and attitudes, draws on theories and studies whose focus is the gender construct that corresponds to traits that are thought to be desirable or appropriate for one sex or the other.

Direct Transmission Via Family Socialization

There has been considerable research examining environmental influences on children's acquisition of sex-appropriate behaviors and attitudes (Bem, 1981; Spence & Helmreich, 1979). Following the perspective that gender-role development is primarily determined by sociocultural factors (Block, 1973; Maccoby, 1988), much of this research has focused on the role of parents as the primary agents of socialization (Huston, 1983). This research correctly points out that beginning with infants' very first days at home, parents immerse their children in gender-related expectations (Condry & Condry, 1976). Parents react positively to sex-appropriate play and negatively to sex-inappropriate play (Caldera, Huston, & O'Brien, 1989; see Lynton & Romney, 1991, for meta-analysis) and furnish their children's rooms with sex-appropriate images and objects (Polmearau, Bolduc, Malcuit, & Cossette, 1990).

The research that emphasizes the role of family socialization on the acquisition of gender-typed behaviors has been strongly influenced by Bem's GST. GST provides that gender-role identity originates within a developing network of gender-related cognitive associations (Bem, 1981). As a child perceives gender-related information, he or she selects and organizes this information into a

cognitive schema that then regulates behaviors in accordance with internalized definitions of femininity and masculinity (Bem, 1981). That actual differences in average male and female behaviors and attitudes exist is not only recognized by GST but is required for the intergenerational continuation of gender dichotomous cognitive schemas.

Because children begin to actively construct their gender schemata at a very early age, parents' behaviors—both purposeful instruction about the gender appropriateness of behaviors and attitudes and modeling of acceptance for cross-sexed behaviors—are important, if not primary, influences on children's development of gender schema. Addressing the importance of family environment in explaining variation in gender socialization, Bem (1981) explained that "not everyone becomes equally sex-typed, of course, and individual differences presumably derive from the extent to which one's particular socialization history has stressed the functional importance of the gender dichotomy" (p. 362; see also Maccoby, 1988).

The importance of differences between families in conveying information about the gender dichotomy and its legitimacy for creating differences in the acquisition of gender schema and resulting sex-typed behaviors is supported by the association between parents' general attitudes toward gender appropriate behaviors and children's attitudes and behaviors. For example, in their influential study of parents' role in the sex-typing of children, Fagot and Leinbach (1989) found that infants who mastered gender labeling at 27 months (early-labelers) were more likely to have been previously provided positive and negative feedback from their parents for sex-appropriate and sex-inappropriate play than toddlers who had not mastered at the same age (late-labelers). Although the two groups had not differed in sex-typed toy choices when observed 10 months earlier, the early-labeling toddlers were more sexed-typed in their toy choices at 27 months than were late-labelers. Fagot and Leinbach emphasized that parents in their study showed individual differences in their tendencies to use young children's biological sex for interpreting and responding to their behavior. Some parents paid little attention to the child's sex in interpreting behavior, whereas others employed the child's sex to interpret the child's behavior and emit affective responses accordingly (Fagot & Leinbach, 1989).

The impact of differences in family socialization for gender roles and sex-appropriate behaviors appears to continue beyond early childhood. For example, Bronstein, Briones, Brooks, and Cowen (1996) found that children whose parents were more accepting of wide ranges of emotions when they were in the fifth grade were more likely during adolescence to report

showing emotional behaviors not traditionally associated with their gender roles—such as female anger or male crying. Growing up in different family structures may provide adolescents with further variance in sex-role socialization. Adolescents from mother-led households report more egalitarian attitudes than adolescents from intact households (Barber & Eccles, 1992). This suggests that more relative exposure to mothers, rather than to fathers, who tend to be more traditional, leads to the adoption of less traditional beliefs about gender roles. It also may be that witnessing their mothers fill multiple gender roles leads to these attitudes. Being raised inside single-parent homes also reduces children's exposure to fathers' rough-and-tumble play style, which has been found to be associated with later rough-and-tumble play (McBride-Chang & Jacklin, 1993). The findings of Fagot and Leinbach (1989), Bronstein et al. (1996), Barber and Eccles (1992), and McBride-Chang and Jacklin (1993) underscore the existence of between-family variation in gender-socializing environments.

Nonshared Environmental Influences

Although associations between parenting practices and child and adolescent behaviors have been reliably demonstrated, the magnitude of these associations makes it unlikely that they can explain a large proportion of the variance in adolescents' sex-typed behaviors and attitudes. For example, when squared, even the relatively large correlations between fifth-grade family emotionality and parent-child expressiveness and the later child outcomes of male crying and female anger of .42 and .54 (Bronstein et al., 1996) explain only a quarter of the variance in these sex-typical behaviors. Three quarters of the variance remains, suggesting that any complete explanation of the acquisition of sex-typical behaviors and attitudes requires more than simple models of parent-to-child transmission that equally affect all children in the same household. A first place to look for these influences is within a household where the children's experiences are beyond those that are reliably shared by siblings. These influences, which in the parlance of behavioral genetics are referred to as nonshared environmental influences, vary for siblings within the same family and potentially contribute to differences in behaviors and attitudes among siblings within the same family. Examinations of family environments suggest that siblings, although they share the same household, can live very different lives. The nonshared aspects of environmental influences are not limited to family environments. Experiences outside the family—as siblings make their own way in the world—are even more likely candidates for nonshared influences. Correlations across siblings for experiences across domains such as peers, social support, and life events range from .10 to

.40, suggesting that experiences outside the home are important candidates for nonshared environmental influences.

Research has shown that nonshared influences, in contrast to shared environmental influences, make consistent and substantial contributions to variance in personality traits, often contributing the largest proportion of variance (Loehlin, 1985). Major sources of differential influences on siblings include differential treatment by parents (Baker & Daniels, 1990) and from each other (Reiss et al., 1994). It is important to realize that one sibling's environment differs from his or her sibling's by virtue of being exposed to that other sibling (Daniels, 1987). Outside the household is where siblings' lives can most dramatically diverge. Perhaps the greatest potential source of differential influences is nonshared exposure to peers (Baker & Daniels, 1990; Daniels, 1987; Rowe, Woulbroun, & Gully, 1994).

Differential peer experiences may be particularly important for the acquisition of sex-typed attitudes and behaviors. A difficulty in assessing specific nonshared environments is that it is always possible that differences in experiences across siblings are due to genetic differences between them. Even if siblings' different experiences are due to genetically guided self-selection, however, once selected, these differing environments may increase differences between siblings. According to Bem (1979), differing experiences with same-sex peers are major influences on differences in the acquisition of gender schema and therefore may contribute to the expression of sex-typical behaviors and attitudes. If one sibling in a household spent more time with same-sex peers than another, this could explain his or her greater compliance with sex-typical behaviors and attitudes than his or her sibling. The importance of nonshared environmental influences on Bem-type measures of masculinity and femininity has been demonstrated. A twin study by Mitchell et al. (1989) revealed that both differential experiences and genetic influences played substantial roles in variation on child versions of Spence and Helmreich's (1979) masculinity and femininity measures. The researchers emphasized the roles of teachers and other adult role models as important influences on masculinity and femininity.

Biological Influences

Influence on adolescents' sex-typed behaviors and attitudes may not be due solely to either direct transmission via parenting or idiosyncratic differences in adolescents' life experiences. Sex typicality, similar to most personality traits, is open to the influence of genetics. A substantial body of work suggests that inherited differences between individuals make important contributions to variance in sex-typed behaviors and attitudes.

Forty years of research on primates and other animals shows that genetically controlled biological mechanisms influence differences in sex-dimorphic behaviors (Ehrhardt & Meyer-Bahlburg, 1981; Goy, 1970). Across species, both within-sex and across-sex differences are controlled by differential exposure to testosterone during a critical developmental period. In primates, this critical period is midgestation. Absence of testosterone exposure produces offspring that behave in female-typical ways and have female genitalia. Increases in testosterone produce offspring that are increasingly masculinized. Testosterone permanently masculinizes the structure of the fetal brain. Male fetuses' testes produce large amounts of testosterone in midgestation, masculinizing both the fetal brain and the genitalia. Female fetuses have no internal source of appreciable testosterone but receive a small variable exposure from their mothers. Males are born with masculinized brains and predispositions to species-typical male behaviors, some to a greater extent than others. The small amounts and small differences in testosterone in the fetal experience of females do not affect the genitalia of normal females at birth but differentially masculinize subsequent sex-dimorphic behavior. Studies of human males and females with genetic clinical conditions that cause excess testosterone exposure in fetal life confirm that this process works in the same way for humans as for other primates (Collaer & Hines, 1995). In humans, we will use the term "sex-typed behaviors and attitudes" for sex-dimorphic behavior.

If variation in sex-typed behaviors and attitudes is heritable, then without behavior-genetic models we will be unable to determine whether the correlation between parental behavior and the gender attributes of children is better explained by social or genetic transmission. One implication of possible heritability of sex-typed behaviors and attitudes is that associations between parental socialization practices and later offspring behaviors and attitudes could be due the shared genetic inheritance of parents and children, a process that is known as the passive gene-environment correlation (Plomin, 1994). However, other nonheritable biological routes of influence are possible, for example, maternal drug treatments during pregnancy (Reinisch, 1977). In behavior-genetic models, these nonheritable biological influences, as well as any event that influences the sex-typed behaviors and attitudes of one sibling but not the other, are potential contributors to nonshared environmental variance (Plomin & Daniels, 1987).

Measuring Sex-Typed Behavior and Attitudes

As mentioned above, masculinity/femininity was historically theorized as a bipolar concept. Its psychometric measurement was dominated by a bipolar Male-Female

(M-F) Scale constructed by Terman and Miles (1936). This scale consisted of multiple items that had been shown to distinguish the responses of males and females. In the 1970s, orthogonal scales of masculinity and femininity became favored as researchers, perhaps uncomfortable with the conceptualization of the sexes as "opposite," examined the co-occurrence of stereotypically male traits and stereotypically female traits. This line of research did not intend to measure male-typical or female-typical behaviors. Instead, the goal was to examine the psychological implications of individuals' compliance with culturally proscribed definitions of socially desirable traits for males and females. The two dimensions of traits measured by Bem's Sex Role Inventory and other measures of this type, such as Spence and Helmreich's Personal Attributes Questionnaire (PAQ), clustered around instrumentality, which was given the label masculine, and expressiveness, which was labeled femininity (Spence & Helmreich, 1979). Although the emphasis of this program of research was understanding the implications of the co-occurrence of these two largely orthogonal dimensions of attitudes, the theorized mechanisms behind their acquisition presupposed the existence of average differences in behaviors and attitudes between the sexes (Bem, 1981). These average differences in behavior are the empirical basis for the construction of the sex-typicality scale used by this study.

To evaluate the sources of the influences in sex-typical behaviors and attitudes, we have constructed a bipolar measure of sex-typed behaviors and attitudes. A bipolar structure is consistent with both the genetic/hormonal biological perspective and the average sex differences that Bem and others invoke as the backdrop to the acquisition of gender schema (Bem, 1979). Among the sources that Bem suggests influence the acquisition of gender schema are differential exposures to gender role socialization in different households as well as experiences outside the home, such as differential exposures to same-sex peer groups. Taken together with biological influences, these influences cover the spectrum of influences estimated by behavioral genetic methodology: additive genetic, shared environmental, and nonshared environmental influences.

This study is not the only study that makes use of the bipolar measurement of sex-typed behavior and attitudes. Lippa and colleagues have used similar approaches to explore the relationship between sex-typed behavior and attitudes and other patterns of behavior (Lippa, 1995). The bipolar strategy used by Lippa builds scales from questionnaire items in surveys that show sex differences in response. Referred to by Lippa as gender diagnosticity, this approach allows the construction of scales that are sensitive to culture, con-

text, and age cohorts. Using behavior genetic methods, Lippa and Hershberger (1999) examined the genetic and environmental influences on variation of gender diagnosticity (GD) for 839 same-sex pairs of twins. The measure of gender diagnosticity used was a factor score contributed to by gender diagnosticity scales constructed from 839 twin pairs' degree of participation in 324 everyday activities, preferences for 160 occupations, responses to the California Psychological Inventory (CPI), and scores on the Femininity subscale of the CPI. Produced using a varimax rotation, the extracted GD factor score was orthogonal to both Feminine Expressiveness and Masculine Instrumentality factors. Results revealed that variance in the GD factor was contributed to by both genetic and environment influences, but similar to Mitchell et al.'s (1989) examination of Bem-type M and F scales, Lippa and Hershberger (1999) found that the preponderance of the environmental variance was attributed not to the shared environment but to the nonshared environmental influences.

Study Approach and Hypotheses

As the reviewed literature makes clear, there are reasons to predict substantial contributions from shared environmental, nonshared environmental, and genetic influences on sex-typed behaviors and attitudes. The goal of this study is to examine the relative contributions to the variance of sex-typed behaviors and attitudes of both male and female adolescents for these influences. To do so, we have approached the measurement of sex-typed behaviors and attitudes with a method similar to that used by Lippa and colleagues to construct their gender diagnosticity measure (Lippa, 1991; Lippa & Connelly, 1990). Our approach, which is explained in more detail in the Method section, uses a one-dimensional, bipolar scale of sex-typed behaviors and attitudes to compute covariance matrices for Sex-Type \times Zygosity sibling pair groups. Fitting structural equations models to these covariance matrices allows the estimation of three components of variance: additive genetic (A), shared environmental (C), and nonshared environmental influences (E). Together these components of variance make up what is known as the ACE model by behavioral geneticists. Genetic influences also may include nonadditive genetic influences. Parameters estimating these influences will be added to the models if the more parsimonious ACE fails to adequately fit the observed data. Specific hypotheses for the ACE model are examined by performing nested model comparisons between the relative fit of models that differ in their specification of additive genetic (A), shared environmental (C), and nonshared environmental (E) parameters. If the likelihood ratio χ^2 (the significance of the difference in χ^2 according to the difference in *df* between models) for

the comparison is significant, then the model with the lower χ^2 is accepted. If the likelihood ratio χ^2 is not significant, the more parsimonious model (i.e., the model with fewer parameters) is accepted.

This nested model comparison approach will be used to examine hypotheses addressing the significance and equivalence of influences on the variance in sex-typed behaviors and attitudes. Two hypothesis-testing steps are required. The first examines the equivalence of parameters across the sexes and the second examines the significance of parameters. The first set of hypotheses will examine the cross-sex equivalence of nonshared environmental, shared environmental, and genetic influences on sex-typed behaviors and attitudes by comparing pairs of models that differ in whether a given parameter is freely estimated across the sexes or set to equivalence across sexes. For example, by comparing the fit of two models that are equivalent to each other except that one requires shared environmental influences for males and females to be equivalent and the other allows shared environmental influence to be estimated freely across the sexes, we examine the equivalence of shared environmental influences for males and females. After determining if the nonshared environmental, shared environmental, and additive genetic influence parameters are statistically equivalent or different for males and females, the second hypothesis-testing step will examine the effect of dropping particular parameters on model fit, thereby determining the significance of different parameters. For example, by comparing the fit of two models that are equivalent to each other except that one estimates shared environmental influences and the other does not, we examine the significance of shared environmental influences. Using this nested model approach allows us to address the following hypotheses: (a) additive genetic, shared environmental, and nonshared environmental influences are equivalent for males and females; (b) there are nonshared environmental influences on the variance in adolescents' sex-typed behaviors and attitudes; (c) there are shared environmental influences on the variance in adolescents' sex-typed behaviors and attitudes; and (d) there are additive genetic influences on the variance in adolescents' sex-typed behaviors and attitudes.

METHOD

The Add Health Project

The data used for this study are drawn from Wave II of the National Longitudinal Study of Adolescent Health (Add Health). The Add Health project consists of multiple data sets organized around a school sample that is representative of U.S. schools with respect to region of country, urbanicity, school type, ethnicity, and school

size. In the In-School sample, questionnaires were collected from more than 90,000 adolescents. All students who completed an In-School questionnaire, or who were listed on a school roster, were eligible for inclusion in the Wave I In-Home sample ($N = 20,745$), which was collected between April and December 1995. Between April and August 1996, approximately 1 year after the collection of the Wave I In-Home data set, more than 14,000 participants were assessed for a second time as part of the Wave II In-Home sample ($N = 14,738$). Also included in the project are school administrator, parent and household, contextual, and sibling data sets. More specific information about the Add Health research design, sampling, and data instruments is available in Bearman, Jones, and Udry (1998).

Genetically Informative Add Health Data

The specific data analyzed here were drawn from pairs of monozygotic twins, dizygotic twins, full-siblings, and half-siblings from the Add Health's sibling sample. Respondents with twins, full-siblings, and half-siblings, as well as those who lived with "biologically unrelated siblings" and cousins, were identified using information from the In-School questionnaire and school rosters. All respondents with twins, half-siblings, and "unrelated siblings" and cousins—along with a probability sample of full-sibling pairs—were selected for the Wave I In-Home interview. In-Home data were obtained for both the target respondents and their siblings. The resulting sibling data set, including unrelated siblings raised together, twin pairs of undetermined zygosity, and cousins, consisted of 3,139 sibling pairs. The analyses reported here used only data from twins and full- and half-siblings that had nonmissing data for the sex-typed behaviors and attitudes scale. Zygosity of the majority of the same-sex twins pairs were classified on the basis of their self-reports of confusability of appearance. Confusability of appearance scales has been found to have greater than a 90% agreement with zygosity determined by DNA. In the 89 cases of twins for whom the confusability of appearance measure provided uncertain zygosity, DNA markers were used to determine zygosity. Twin pairs who were the same on five or more markers were classified as monozygotic (error rate of less than 1/40,000). Twin pairs who differed on one or more markers were classified as dizygotic. These pairs were divided into 11 groups based on their sex composition: male pairs, female pairs, and mixed sex pairs, which were ordered male first and female second. The number of pairs in each of the 11 groups with nonmissing values for sex-typed behavior and attitudes scale were 99 monozygotic (MZ), 95 dizygotic (DZ), 232 full-sibling (FS), 68 half-sibling (HS) male pairs; 108 MZ, 95 DZ, 261 FS, 63 HS female pairs; and 73 DZ, 160 FS, 48 HS mixed-sex pairs (1,301 total).

The average age of first siblings was 16.7; second siblings averaged 16.8 years. Of the sibling pairs, 54% reported their racial group as White, 23% African American, 14% Hispanic, 7% Asian, and 2% American Indian or other.

Construction of a Measure to Distinguish Gender: The Probability of Being a Boy (PRBOY)

We developed a measure of sex typicality of behaviors and attitudes for adolescents in Wave II of the Add Health study using an approach similar in nature to work described by Lippa and Connelly (1990). The main objective was to construct a sex-typed behaviors and attitudes score from the probability that an adolescent is male (or female) on the basis of participants' responses to a set of questions. It is similar to the approach used by Lippa and colleagues in that it retrofits a gender score to preexisting data based on differential responses of males and females but it differs in application. Instead of using probabilities of being male or female derived from larger psychological inventories, such as occupational preference inventories or the CPI (e.g., Lippa & Connelly, 1990; Lippa & Hershberger, 1999), our approach uses individual items and their contributions to the probabilities of being a boy. This approach is less demanding of data and allows the construction of gender scales on a wider range of existing data sets.

In applying this technique, we first selected a broad set of Wave II items from varying behavioral and attitudinal domains that showed sex differences in a response that could be attributed to individual preferences or behaviors. Care was taken not to select items where social restrictions presented males' and females' differential opportunities to participate (e.g., playing baseball). Preliminary analysis using the core sample of Wave II identified 21 questions that were useful in discriminating boys from girls. Stepwise logistic regression was used to select a subset of these questions that significantly contributed to predicting the log odds of being a boy. Table 1 provides the variable names, content of items, minimum and maximum scores, and difference scores (differences in mean response for boys and girls expressed as a fraction of the standard deviation of the boy). When examining the signs preceding difference scores in Table 1, bear in mind that the direction of these scores is affected by the direction of the responses for each item. For example, response patterns of items PHYSICFT, EMOTION, LIKESELF, LIVNOTHK, SENSITIV, TAKERISK, and UPSETPRB ranged from 1 (*strongly agree*) to 5 (*strongly disagree*). Accordingly, the difference score for EMOTION (.44) does not indicate that females are less likely than males to agree that they are emotional. In contrast, they are more likely to do so. The unsigned average of Table 1's difference scores was .37, indicating that responses to the individual items, although distin-

TABLE 1: Difference of Means for Boys and Girls Expressed as Fraction of Boys' Standard Deviation for Variables Used in Sex-Typicality Scale (PRBOY) for Boys and Girls

<i>Variable Name</i>	<i>Content of Items</i>	<i>Direction of Response Pattern</i>	<i>Minimum Value</i>	<i>Maximum Value</i>	<i>(Boy – Girl M) / (Boy SD)</i>	<i>Boys' SD</i>
CRYALOT	Frequency of crying	1 to 4; 4 = <i>every day</i>	0	4	-1.21	0.43
MOODY	Frequency of moodiness	1 to 4; 4 = <i>every day</i>	0	4	-0.47	0.87
NO_APPET	Frequency of poor appetite	1 to 4; 4 = <i>every day</i>	0	4	-0.38	0.74
HONEST	Honestly answered questions	1 to 4; 4 = <i>completely honest</i>	1	4	-0.21	0.86
TRBATTEN	Trouble paying attention	1 to 4; 4 = <i>every day</i>	0	4	0.11	1.06
BOTHER	Bothered by things	1 to 4; 4 = <i>every day</i>	0	3	-0.34	0.63
PHYSICFT	How physically fit	1 to 5; 5 = <i>strongly disagree</i>	1	5	-0.48	0.82
BADFIGHT	Past 12 months, serious fighting	0 to 3; 3 = <i>five or more times</i>	0	3	0.27	0.60
AEROBICS	Frequency of exercising	0 to 3; 3 = <i>five or more times</i>	0	3	-0.06	1.06
SKATE	Frequency of roller-blading/cycling	0 to 3; 3 = <i>five or more times</i>	0	3	0.29	0.99
EMOTION	How emotional you are	1 to 5; 5 = <i>strongly disagree</i>	1	5	0.44	0.98
LIKESELF	Do you like yourself as you are	1 to 5; 5 = <i>strongly disagree</i>	1	5	-0.44	0.79
LIVNOTHK	Live without thought for future	1 to 5; 5 = <i>strongly disagree</i>	1	5	-0.26	1.11
SENSITIV	How sensitive to others' feelings	1 to 5; 5 = <i>strongly disagree</i>	1	5	0.32	0.80
TAKERISK	Do you like to take risks	1 to 5; 5 = <i>strongly disagree</i>	1	5	-0.31	1.01
UPSETPRB	Upset by difficult problems	1 to 5; 5 = <i>strongly disagree</i>	1	5	0.36	1.06

NOTE: $N = 3,988$ for boys and $N = 4,433$ for girls.

guishing between the sexes, were not dramatically different for males and females. The regression equation including these variables is as follows:

$$\begin{aligned} \log_e[p_i/(1 - p_i)] = & 2.2098 + \text{CRYALOT}*(-1.2525) \\ & + \text{MOODY}*(-0.2255) \\ & + \text{NO_APPET}*(-0.2022) \\ & + \text{HONEST}*(-0.2435) \\ & + \text{TRBATTEN}*(0.3109) \\ & + \text{BOTHER}*(-0.1194) \\ & + \text{PHYSICFT}*(-0.3676) \\ & + \text{BADFIGHT}*(0.6654) \\ & + \text{AEROBICS}*(-0.1331) \\ & + \text{SKATE}*(0.3056) \\ & + \text{EMOTION}*(0.1217) \\ & + \text{LIKESELF}*(-0.2042) \\ & + \text{LIVNOTHK}*(-0.2171) \\ & + \text{SENSITIV}*(0.3175) \\ & + \text{TAKERISK}*(-0.1762) \\ & + \text{UPSETPRB}*(0.2025) \end{aligned}$$

where p_i is the probability of being a boy for the i^{th} adolescent. By solving for p_i , individual probabilities were computed for each adolescent and used as the measure of sex typicality. The resulting values range from 0 to 1, with 0 denoting the most female-like and 1 the most male-like score. The coefficients in the above formula provide the specific contribution of each variable to the log odds of being a boy while holding the contributions of other variables in the equation constant. One implication of this is that intrascale collinearity, which increases internal measures of reliability such as Cronbach's alpha, is

limited by design. Measures of internal reliability, such as Cronbach's alpha, are appropriate measures of reliability for scales using multiple items to assess a single domain. The approach used to construct PRBOY draws on measures from multiple domains that each contribute to the classification of individuals as males and females and is linked to our definition of gender not being limited to one or a few domains. Increasing the number of domains assessed by the scale increases PRBOY's ability to correctly classify adolescents by sex and provides explicit recognition that sex differences exist across domains of behavior and attitudes. Accordingly, a more relevant issue is whether the full model functions as intended. To determine if the PRBOY correctly classifies adolescents by sex, we used receiver operating characteristic (ROC) analysis. ROC measures of accuracy have been used in medical testing, information retrieval, weather prediction, (Swets, 1988), psychology (Swets, 1973), and epidemiology (Erdreich & Lee, 1981). To construct an ROC curve, we computed the probability of a true positive (probability that adolescent is a boy when truly a boy) versus the probability of a false positive (probability that adolescent is a boy when truly a girl) using various cut-points between 0 and 1. The area under this curve measures the probability of a correct ranking. It has been shown that this area measure is the same quantity that is estimated by the nonparametric Wilcoxon statistic (Hanley & McNeil, 1982). A value of 0.5 would indicate correct classification is the same as chance. Values of area under ROC curves range from 0.90 to 0.98 for diagnosis from applications of CT and chest x-ray films, 0.80 to 0.90 for mammography, and 0.75 to 0.90 for weather

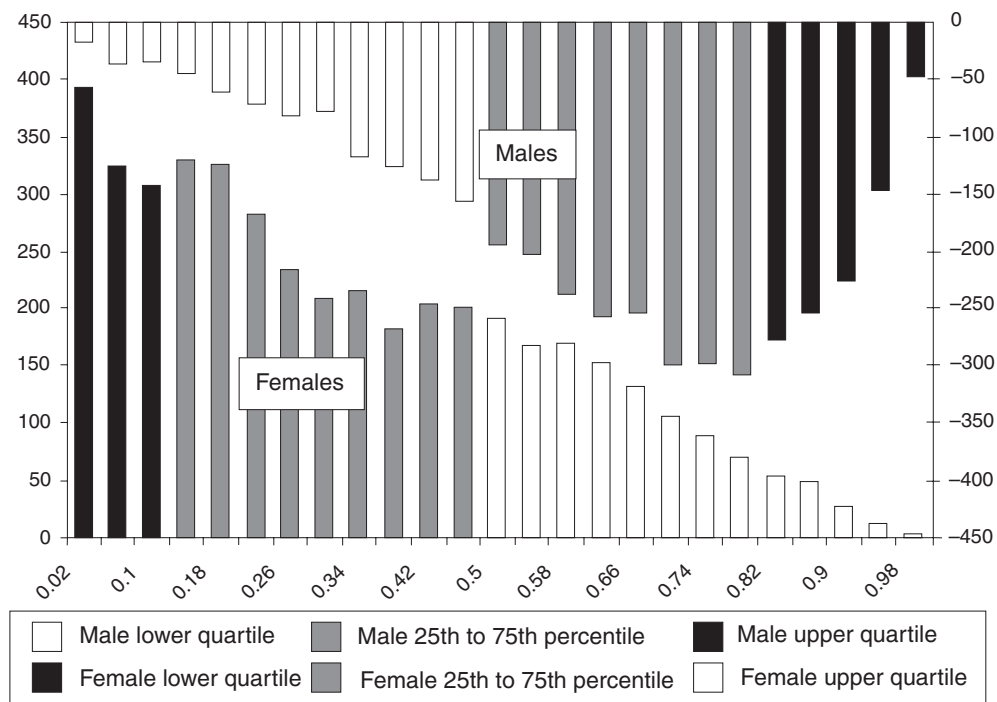


Figure 1 Sex-specific distributions for the Probability of Being a Boy (PRBOY) Scale.

prediction (Swets, 1988). Calculated on the core sample ($N = 8,421$), which was used to derive the formula for the PRBOY, analysis showed our model correctly discriminates between a randomly chosen boy and girl 81.7% of the time. To determine if this ROC was inflated due to being calculated on the same sample used to derive the PRBOY, the ROC was recalculated on the noncore sample respondents with complete PRBOY data ($N = 4,789$). The ROC as calculated on this sample was 79.6.

We experimented with the effects of deleting each variable in the final regression equation and computed the change in the percentage of boys and girls correctly classified by sex. On average, omitting one variable reduced the percentage correctly classified from 82% to 81%, with the most influential variable (CRYSALOT) reducing the percentage classified correctly to 78%.

The distribution of gender scores is bimodal, as shown in Figure 1. The upper quartile of the girls overlap the range where most of the boys score; similarly, the lower quartile of the boys overlaps the range where most of the girls score. The mean score is 0.63 for boys and 0.33 for girls. Standard deviations are 0.217 for boys and 0.235 for girls. The highest probability for both individual males and females is greater than 0.99, and the lowest probability for each sex is less than 0.01. It is interesting to note that the distribution of male scores is not a mirror image of the female scores. The girls' distribution

grows heaviest at the extreme left of the figure, indicating a high proportion of females with very female-like scores. In contrast, there is a scarcity of boys in the area of the histogram indicating the highest possible male-typical scores.

RESULTS

Preliminary Results: Sibling Pair Correlations for Sex-Typed Behaviors and Attitudes

The intrapair correlations for males on the PRBOY measure were .30 for MZ, .20 for DZ, .08 for FS, and .02 for HS. The generally small magnitude of these correlations reveals that most of the variance in sex-typed behaviors and attitudes is not shared between siblings, suggesting an important role for nonshared environmental influences and little if any role for shared environmental influences. The reduction in magnitude from highly related siblings (.30 for MZ) to less related siblings (.02 for HS), however, suggests that genetics influence sex-typed behaviors and attitudes for males. For female siblings, the correlations were .46 MZ, .22 DZ, .15 FS, and .20 HS. Although several correlations were larger in magnitude than their male counterparts, the overall pattern shown by these correlations appears to be similar to the corresponding values for males. The correlations for mixed-sex pairs also showed relatively low within-pair

TABLE 2: Fit Results for Genetic and Environmental Contributions to Sex-Typed Behaviors and Model Comparisons for Testing Sex-Typed Behavior and Attitude Hypotheses

Model Number	Parameters			Goodness-of-Fit Criteria					Model Comparisons			
	A	C	E	χ^2	p	df	AIC	RMSEA	Models	χ^2 Difference	df	Result
1	≠	≠	≠	9.544	.99	27	-44.456	.003				
2	≠	≠	=	9.495	1.00	28	-46.505	.003	1 vs. 2	0.049	1	accept
3	=	≠	=	13.490	0.99	29	-44.510	.008	2 vs. 3	3.995	1	reject
4	≠	=	=	9.569	1.00	29	-48.431	.003	2 vs. 4	0.074	1	accept
5	=	=	=	26.138	.668	30	-33.118	.027	4 vs. 5	16.569	1	reject
6	≠	Φ	=	9.569	1.00	30	-50.431	.002	4 vs. 6	0.000	1	accept
7	♀	=	=	93.848	.000	30	33.848	.075	4 vs. 7	84.279	1	reject
8	♂	=	=	266.407	.000	30	206.407	.103	4 vs. 8	5.336	1	reject

NOTE: A = additive genetic, C = shared environmental, E = nonshared environmental, AIC = Aikake's Information Criterion, and RMSEA = Root Mean Square Error Approximation. Φ signifies that the parameter was not estimated at all, ♂ signifies the parameter was estimated for males only, and ♀ signifies the parameter was estimated for females only.

similarity; DZ and FS were correlated .14 and .12. The HS PRBOY scores, correlated .27, were more similar than expected.

Alternative Models of Sex-Typed Behaviors and Attitudes

Table 2 provides information on the parameters specified for each model and the resulting goodness-of-fit statistics. For each model, the constraints imposed on parameters A, C, and E are provided. For example, in Model 1, parameters A, C, and E are allowed to be freely estimated across males and females, as indicated by the unequal symbol. Also provided are the χ^2 , degrees of freedom (*df*), Aikake's Information Criterion (AIC), and the Root Mean Square Error Approximation (RMSEA) for each model. The AIC is a measure of model fit that takes the parsimony of the model into account, making it ideal for comparisons of fit between nested models. The RMSEA is an approximation of the average of the discrepancies between the observed and implied covariance matrices (Loehlin, 1992). Table 2 also provides information on specific comparisons for model hypotheses, likelihood ratio tests, and outcomes for each hypothesis tested. Model 1, freely estimating additive genetic (A), shared environmental (C), and nonshared environmental (E) influences on variance in sex-typed behaviors across sexes, fit the data very well, $\chi^2(27) = 9.544$, $p = .999$. Constraining E to be equal across males and females in Model 2 did not significantly increase chi-square, $\chi^2(28) = 9.495$, $p = 1.00$, $\chi^2_{\text{diff}}(1) = 0.049$, revealing that nonshared environmental influences (E) were equivalent across sexes. Because Model 2 was more parsimonious (greater degrees of freedom) than Model 1, the hypothesis that nonshared environmental influences were equivalent across sexes was accepted (see Table 2).

Model 3 added the additional constraint that additive genetic influences were equivalent across sexes. Model 3 fit the data well, $\chi^2(29) = 13.490$, $p = .99$, but its chi-square was significantly greater than that of Model 2, $\chi^2_{m3} - \chi^2_{m2}(1) = 3.995$, leading to the rejection of the hypothesis that additive genetic influences were the same for males and females. Model 4 set shared as well as nonshared environmental influences to be equal across the sexes. The fit of Model 4 was exceptionally good, $\chi^2(29) = 9.569$, $p = 1.00$. Moreover, unlike Model 3, its fit was not significantly worse than Model 2, $\chi^2_{m4} - \chi^2_{m2}(1) = 0.074$, leading to the adoption of Model 4 as the comparison model and the acceptance of the hypothesis that both shared and nonshared environmental influences were the same for males and females. Model 5 added an equality constraint for additive genetic influences across sexes to the existing cross-sex constraints (E and C) of Model 4. Unsurprisingly, this constraint significantly reduced model fit, $\chi^2_{m5} - \chi^2_{m4}(1) = 16.569$. Accordingly, the hypothesis that additive genetic influences were the same for males and females was rejected.

Model 4's parameter estimates for shared environmental influences on male and female variance in sex-typed behaviors were near zero. To test the hypothesis that shared environmental influences were not significant for either sex, the C parameter was dropped entirely from Model 6. The resulting fit of Model 6 was identical to Model 4 with one more degree of freedom, $\chi^2(30) = 9.569$, $p = 1.00$, indicating that shared environmental influences were not important for explaining variance in sex-typed behaviors. Models 7 and 8 dropped additive genetic influences from females (Model 7) and from males (Model 8). The fit of both of these models was significantly worse than that of either Model 4 and Model 6 (see Table 3), indicating that genetic influences were important for both sexes. Taken together, these compar-

isons led to the acceptance of Model 6 as the best model. The conclusion that Model 6 was the best model is corroborated by the alternative fit statistics. Compared to those of other models, both the AIC (-50.431) and the RMSEA (.002) of Model 6 were smallest. Based on the extremely good fit of this Model 6—an RMSEA of .002 indicates that differences between estimated and observed covariance matrices is nearly zero—it was decided not to fit less parsimonious models with additional parameters to estimate nonadditive genetic influences.

Using parameter estimates drawn from Model 6, estimates of genetic and environmental influences were calculated. For males, 25% of the variance in sex-typed behaviors was accounted for by genetic influences and 75% by nonshared environmental influences and measurement error. For females, 38% of the variance in sex-typed behavior was accounted for by genetic influences and the remaining 62% by nonshared environmental influences and measurement error.

DISCUSSION

The degree that adolescents' behaviors are sex-typical appears to be affected by both genetic and environmental influences. Siblings were moderately similar for sex-typical behaviors and attitudes. This similarity, however, was better explained by genetic influences than by shared environmental influences. This leads to the same conclusion reached by Lippa and Hershberger (1999), who found that the majority of environmental influences on a gender diagnosticity factor contributed to by sex differences in responses to occupational preferences, everyday activities, and the California Psychological Inventory (CPI), as well as the Femininity subscale of the CPI, was associated with nonshared, rather than shared, environment influences. These findings conflict with what Mitchell et al. (1989) described as simple models of family transmission. If, as Bem suggested, between-family variations in gender socialization were responsible for individual differences in gendered attitudes, then shared family environment should make significant contributions to variance in sex-typed behaviors and attitudes. Mitchell et al. demonstrated this is not the case for M and F scales. It is apparently also not the case for sex-typicality of behaviors and attitudes. The findings do not mean that parents do not participate in the instruction and the modeling of sex-appropriate play and attitudes (Caldera et al., 1989; Condry & Condry, 1976; Polmureau et al., 1990), and it does not suggest that there are not differences between families in the gender socialization environment (Fagot & Leinbach, 1989). It does, however, suggest that these between-family differences do not have systematic effects on the sex-typicality

of adolescents who are exposed to them. Parental socialization efforts may be focused differentially on one child or another depending on the behaviors and attitudes expressed by the children in the household. This differential socialization could play a critical role in the etiology of sex-typical behaviors and attitudes, but it would not increase the aggregate level of similarity among siblings in a household and therefore cannot contribute to shared environmental influences.

In contrast to shared environmental influences, genetic influences accounted for significant amounts of variance in sex-typicality: 38% among females and 25% among males. These estimates support the perspective that inherited differences between individuals make important contributions to variance in sex-typed behaviors and attitudes. The difference in magnitude of genetic effects between the sexes was significant. The greater heritability for females than males suggests that different processes affect the variance of sex-typed behaviors and attitudes for the two sexes. Process differences are also evinced by the failure of opposite-sex siblings' intracorrelations across zygosity levels to reveal genetic influences. If males' and females' sex-typical behaviors were influenced by the same genes, then the patterns of correlation among opposite-sex siblings would indicate genetic influences (i.e., higher correlations among more highly related sibling pairs). In this case, the correlations across levels of zygosity among opposite-sex siblings did not demonstrate genetic influences. This suggests that although additive genetics make significant contributions to male and female sex typicality, the specific genes involved are not the same across the sexes. It is also interesting to note that compared to males ($SD = 0.217$), females showed slightly greater sex-typicality variance ($SD = 0.235$). It may be the case that greater, or at least different, social constraints for males than females play a role in reducing the expression of additive genetic influences for males' sex-typical behaviors and attitudes.

It also should be noted that the estimated heritabilities ($h^2 = .25$ and $.38$) appear to be sufficient in magnitude to explain the covariance between parental attitudes and behaviors and their offsprings' attitudes and behaviors that socialization researchers interpret as the product of social transmission (e.g., Barber & Eccles, 1992; Bronstein et al., 1996). The remainder of the variance, 62% for females and 75% for males, was associated with nonshared environmental influences. Therefore, although the largest contributions to male and female variance in sex-typed behaviors and attitudes are environmental influences, they are those influences not shared by siblings.

The substantial variance accounted for by nonshared environmental influences among both males and females suggests a large role for differential (i.e., idiosyncratic) experiences of siblings in the etiology of sex-typical behaviors for both males and females. This may be due to differential experiences of siblings with peers and teachers as well as other adults. It may be that these outside-the-home influences have the most important impact on the sex typicality of adolescents because of their particular relevance to adolescents, who at this developmental stage are organizing more of their lives around schools and peers than around parents (see Rowe et al., 1994). As mentioned above, differential parental treatment also may contribute to nonshared influences.

Although the items considered for inclusion in the scale were chosen to represent multiple domains of adolescent behavior, a possible limitation of this study is that the initial selection of items considered for inclusion in the sex-typicality measure was not chosen based on a specific sampling strategy. Accordingly, it cannot be guaranteed that they are representative of male and female behaviors and attitudes. It also should be acknowledged that the method of specifying a sex-typicality scale based on a nationally representative sample of adolescents restricts the generalizability of this study's behavioral genetic findings. We believe, however, that this limitation is outweighed by the advantages of the scale's sensitivity to the culture, context, and age cohort of the sample (see Lippa, 1995). Moreover, because this scale construction technique can be used to create similar scales on any data set with males and females, our conclusions can be readily evaluated using most other genetically informative data sets.

REFERENCES

- Baker, L. A., & Daniels, D. (1990). Nonshared environmental influences and personality differences in adult twins. *Journal of Personality and Social Psychology*, 58, 103-110.
- Barber, B. L., & Eccles, J. S. (1992). Long-term influences of divorce and single parenting on adolescent family- and work-related values, behaviors, and aspirations. *Psychological Bulletin*, 111, 108-126.
- Bearman, P. S., Jones, J., & Udry, J. R. (1998). *The national longitudinal study of adolescent health: Research design*. Retrieved on January 15, 1999, from the World Wide Web: <http://www.cpc.unc.edu/projects/addhealth/design.html>
- Bem, S. L. (1979). Theory and measurement of androgeny: A reply to the Pedhazur-Tetenbaum and Locksley-Colton critiques. *Journal of Personality and Social Psychology Bulletin*, 37, 1047-1054.
- Bem, S. L. (1981). Gender schema theory: A cognitive account of gender typing. *Psychological Review*, 88, 354-464.
- Block, J. H. (1973). Conceptions of sex role. Some cross cultural and longitudinal perspectives. *American Psychologist*, 28, 512-526.
- Bronstein, P., Briones, M., Brooks, T., & Cowen, B. (1996). Gender and family factors as predictors of late adolescent emotional expressiveness and adjustment. *Sex Roles*, 34, 739-765.
- Caldera, Y. M., Huston, A. C., & O'Brien, M. (1989). Social interactions and play patterns of parents and toddlers with feminine, masculine and neutral toys. *Child Development*, 60, 70-76.
- Collaer, M. L., & Hines, M. (1995). Human behavioral sex differences: A role for gonadal hormones during early development. *Psychological Bulletin*, 118, 55-107.
- Condry, J., & Condry, S. (1976). Sex differences: A study in the eye of the beholder. *Child Development*, 47, 812-819.
- Daniels, D. (1987). Differential experiences of children in the same family as predictors of adolescent sibling personality differences. *Journal of Personality and Social Psychology*, 51, 339-346.
- Ehrhardt, A. A., & Meyer-Bahlburg, H.F.L. (1981). Effects of prenatal sex hormones on gender related behavior. *Science*, 211, 1312-1317.
- Erdreich, L. S., & Lee, E. T. (1981). Use of relative operating characteristic analysis in epidemiology. *American Journal of Epidemiology*, 114, 649-662.
- Fagot, B. I., & Leinbach, M. D. (1989). The young child's gender schema: Environmental input, internal organization. *Child Development*, 60, 663-672.
- Goy, R. (1970). Experimental control of psychosexuality. *Philosophical Transactions for the Royal Society of London*, 159, 149-162.
- Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a Receiver Operating Characteristic (ROC) curve. *Radiology*, 143, 29-36.
- Huston, A. C. (1983). Sex-typing. In M. A. Hetherington (Ed.), *Handbook of child psychology: Vol. 4. Socialization, personality, and social development*. New York: John Wiley.
- Lippa, R. (1991). Some psychometric characteristics of gender diagnosticity measures: Reliability, validity, consistency across domains, and relationship to the big five. *Journal of Psychology and Social Psychology*, 61, 1000-1011.
- Lippa, R. (1995). Do sex differences define gender-related individual differences within sexes? Evidence from three studies. *Personality and Social Psychology Bulletin*, 21, 349-355.
- Lippa, R., & Connelly, S. (1990). Gender diagnosticity: A new Bayesian approach to gender-related individual differences. *Journal of Personality and Social Psychology*, 59, 1051-1065.
- Lippa, R., & Hershberger, S. (1999). Genetic and environmental influences on individual differences in masculinity, femininity, and gender diagnosticity: Analyzing data from a classic twin study. *Journal of Personality*, 67, 127-155.
- Loehlin, J. C. (1985). Fitting heredity-environment models jointly to twin and adoption data from the California Psychological Inventory. *Behavior Genetics*, 15, 199-221.
- Loehlin, J. C. (1992). *Latent variable models: An introduction to factor, path, and structural analysis* (2nd ed.). London: Lawrence Erlbaum.
- Lynton, H., & Romney, D. M. (1991). Parents differential socialization of boys and girls: A meta-analyses. *Psychological Bulletin*, 109, 267-296.
- Maccoby, E. E. (1988). Gender as a social category. *Developmental Psychology*, 24, 755-765.
- Maccoby, E. E. (1998). *The two sexes: Growing up apart, coming together*. Cambridge, MA: Harvard University Press.
- McBride-Chang, C., & Jacklin, C. N. (1993). Early play arousal, sex-typed play, and activity levels as precursors to later rough and tumble play. *Early Education and Development*, 4, 99-108.
- Mitchell, J. E., Baker, L. A., & Jacklin, C. N. (1989). Masculinity and femininity in children: Genetic and Environmental Factors. *Child Development*, 60, 1475-1485.
- Plomin, R. (1994). *Genetics and experience: The interplay between nature and nurture*. Thousand Oaks, CA: Sage.
- Plomin, R., & Daniels, D. (1987). Why are children in the same family so different from each other? *Behavioral and Brain Sciences*, 10, 1-16.
- Polmearau, A., Bolduc, D., Malcuit, G., & Cossette, L. (1990). Pink or blue: Environmental stereotypes in the first 2 years of life. *Gender Roles*, 22, 359-367.
- Reinisch, J. M. (1977). Prenatal exposure of human fetuses to synthetic progestin and estrogen affects personality. *Nature*, 266, 561-562.

- Reiss, D., Plomin, R., Hetherington, E. M., Rovine, N., Tryon, A., & Hagen, M. S. (1994). The separate world of teenage siblings: An introduction to the study of nonshared environment and adolescent development. In E. M. Hetherington, D. Reiss, & R. Plomin (Eds.), *Separate social worlds of siblings: The impact of nonshared environments of development* (pp. 62-109). Hillsdale, NJ: Lawrence Erlbaum.
- Rowe, D. C., Woulbroun, E. J., & Gully, B. L. (1994). Peers and friends as nonshared environmental influences. In E. M. Hetherington, D. Reiss, & R. Plomin (Eds.), *Separate social worlds of siblings: The impact of nonshared environments of development* (pp. 159-174). Hillsdale, NJ: Lawrence Erlbaum.
- Spence, J. T., & Helmreich, R. L. (1979). The many faces of androgyny: Reply to Locksley & Colten. *Journal of Personality and Social Psychology*, 37, 1032-1046.
- Swets, J. A. (1973). The Relative Operating Characteristics in psychology. *Science*, 182, 990-1000.
- Swets, J. A. (1988). Measuring the accuracy of diagnostic systems. *Science*, 240, 1285-1293.
- Terman, L. M., & Miles, C. C. (1936). *Sex and personality: Studies in masculinity and femininity*. New York: Russell and Russell.
- Udry, J. R. (2000). Biological limits of gender construction. *American Sociological Review*, 65, 443-457.

Received June 9, 2000

Revision accepted November 13, 2000

On-Line Versus Memory-Based Processing: The Role of “Need to Evaluate” in Person Perception

Zakary L. Tormala
Richard E. Petty
Ohio State University

Two studies are reported examining individual differences in the need to evaluate as a determinant of memory-based versus on-line information processing. In each study, participants read statements describing the behaviors of a hypothetical target person and reported their attitudes toward this person. Consistent with expectations, high need to evaluate individuals formed attitudes in a spontaneous, on-line fashion, whereas low need to evaluate individuals formed them in a less spontaneous, more memory-based fashion. This conclusion was supported by two kinds of evidence: attitude-recall valence relationships (Experiments 1 and 2) and response latency measures (Experiment 2). These results suggest that evaluative responding in the domain of person perception is less pervasive than concluded in prior research.

A long-standing assumption in social psychology has been that evaluation is a fundamental part of human nature (e.g., Markus & Zajonc, 1985; Osgood, Suci, & Tannenbaum, 1957). Considerable empirical evidence has been mounted in support of this view. For instance, Fazio, Sanbonmatsu, Powell, and Kardes (1986) found that attitudes can be activated automatically given the mere presence of an attitude object, particularly when the attitude is strong. Bargh and colleagues (Bargh, Chaiken, Gollwitzer, & Pratto, 1992; Bargh, Chaiken, Raymond, & Hymes, 1996) furthered this argument, noting that automatic attitude activation can occur for virtually all attitude objects. Although there exist differences in the conditions that foster such activation (e.g., Glaser & Banaji, 1999), researchers clearly agree that automatic evaluation is a pervasive phenomenon.

Additional empirical support for this position comes from the domain of person perception. Spontaneous trait inference researchers (see Uleman, Newman, & Moskowitz, 1996, for a review) have shown that even when impressions (i.e., attitudes or evaluative beliefs

about people) do not already exist in memory, they will be formed spontaneously whenever one encounters behavioral information describing a person. In fact, Uleman et al. (1996) argue that we form trait inferences “as naturally as we extract oxygen from the air” (p. 212). Work on on-line and memory-based impression formation also suggests that on-line evaluative processing of other people is the norm, although task constraints can disrupt it (e.g., Bargh & Thein, 1985; Hastie & Park, 1986; Lichtenstein & Srull, 1987). Hastie and Park (1986), for example, noted that they had difficulty finding any evidence of memory-based judgments until they “realized the importance of preventing subjects from making spontaneous on-line judgments” (p. 265, emphasis added).

Individual Differences in Evaluative Responding

Despite the apparent consensus regarding the extent to which people engage in evaluative responding—especially with respect to other people—work on the need to evaluate (Jarvis & Petty, 1996) challenges the notion that evaluative responding is so pervasive. The need to evaluate refers to the chronic tendency to engage in evaluative responding. In several studies demonstrating the predic-

Authors' Note: This article is based on a master's thesis submitted to the Ohio State University Graduate School by the first author under the supervision of the second. We thank thesis committee members Marilyn Brewer and William von Hippel and members of the 1997-1999 Groups for Attitudes and Persuasion at Ohio State University for helpful comments regarding this research. This research was supported in part by a National Science Foundation (NSF) predoctoral fellowship awarded to the first author and NSF Grant SBR 95-20854 awarded to the second author. Correspondence regarding this article can be addressed to either author at the Department of Psychology, Ohio State University, 1885 Neil Avenue Mall, Columbus, OH 43210-1222; e-mail: tormala.3@osu.edu or petty.1@osu.edu.

PSPB, Vol. 27 No. 12, December 2001 1599-1612
© 2001 by the Society for Personality and Social Psychology, Inc.

tive utility of this construct, Jarvis and Petty found evidence for individual differences in the extent to which people form evaluations. For example, when describing events from their previous day, people who scored highly on the Need to Evaluate Scale spontaneously generated a higher proportion of evaluative thoughts (e.g., "some of the songs at the concert were really good") than those who scored on the low end of the scale. Individuals low in their need to evaluate had a greater proportion of non-evaluative thoughts when describing their previous day (e.g., "woke up at 8:00 a.m. and took my dog out first thing"). Jarvis and Petty found that this relationship also emerged in people's responses to unfamiliar abstract paintings.

Research by Hermans and Eelen (2001) suggests that these differences also extend to the domain of automatic evaluative responding. Hermans and Eelen used a priming procedure in which positive or negative words (e.g., friendly, mean) preceded target words that were evaluatively congruent or incongruent with the primes (see Fazio et al., 1986). Participants were instructed to attend to the second word and evaluate it as quickly as possible as "positive" or "negative." They found that high need to evaluate individuals responded more quickly to evaluatively congruent than evaluatively incongruent target words. For those low in need to evaluate, however, there was no difference. Hermans and Eelen (2001) argued that high need to evaluate individuals have stronger object-evaluation associations due to their chronic evaluative responding. Further evidence for this view comes from research reported by Petty and Jarvis (1996), in which people high and low in the need to evaluate reported their attitudes toward nearly 100 familiar attitude objects (e.g., butterfly, landlords, spinach; see Bargh et al., 1992). The names of these objects were flashed on a computer screen and participants reported their evaluations on a dichotomous good/bad scale. Petty and Jarvis found those high in the need to evaluate reported their attitudes more quickly than those low in the need to evaluate.

This research demonstrates that there are individual differences in the general tendency to engage in evaluation; however, work on the need to evaluate has not explicitly examined individual differences in the evaluation of people. Because humans are social animals, evaluation of other people might be so fundamental that individual differences in the general propensity to evaluate would not moderate it. That is, even if we accept the notion that constant evaluation of objects and issues is not the rule, we are not in a position to reject the idea that evaluations of other people are more pervasive for everyone. Thus, one goal of the present research is to determine if individual differences in need to evaluate could moderate evaluative responding toward people.

Interestingly, a considerable body of research has demonstrated that individual differences do influence other aspects of person perception. The related constructs of personal need for structure (Neuberg & Newsom, 1993) and need for cognitive closure (Webster & Kruglanski, 1994), for example, appear to be important in this regard. Individuals high in need for structure (or closure) have been found to engage in more stereotyping (e.g., Neuberg & Newsom, 1993), form more spontaneous trait inferences (Moskowitz, 1993), demonstrate greater correspondence bias (Webster & Kruglanski, 1994), and be more susceptible to primacy effects in impression formation (Webster & Kruglanski, 1994) than those low in need for structure (or closure). These findings lend support to the notion that there may be detectable individual differences in evaluative responding toward people and also point to the desirability of distinguishing need to evaluate effects from those based on need for structure or closure (see Experiment 2).

The current research also has a second goal. Prior work has clearly shown that relative to high need to evaluate individuals, people low in need to evaluate show little evidence of evaluative thinking. Presumably, however, evaluation is sometimes necessary even for low need to evaluate individuals. For instance, when participating in psychology experiments (as when participating in real life), both low and high need to evaluate individuals are called on to express opinions, and both do express them. An important question, then, follows: If low need to evaluate individuals are not evaluating information actively as it is encountered (i.e., in an on-line fashion), on what basis are they forming their opinions? Work by Hastie and Park (1986) suggests that if opinions have not been formed on-line, they will be formed in a memory-based fashion when a judgment is required. Because no research has been conducted to explore the processes by which low need to evaluate individuals form attitudes or opinions when they are requested, we sought to examine these processes. We argue that whereas high need to evaluate individuals are likely to engage in on-line evaluation, those low in need to evaluate are likely to rely more on memory-based evaluative responding when the situation demands an evaluation.

The distinction between memory-based and on-line evaluation has important implications. For instance, the type of evaluative responding in which people engage may subject them to distinct types of biases. Anderson and Hubert (1963) found that although overall impressions are often susceptible to primacy effects, information recall is more susceptible to recency effects. Thus, to the extent that memory-based evaluations rely on information recall and on-line evaluations do not (e.g., Chartrand & Bargh, 1996; Hastie & Park, 1986; Lichtenstein & Srull, 1987; Mackie & Asuncion, 1990),

these differences might map onto recency and primacy effects, respectively. Moreover, considerable work on stereotyping indicates that situations inducing memory-based responding (e.g., cognitive load) also elicit increased reliance on stereotypes in both memory and judgment (e.g., Bodenhausen & Lichtenstein, 1987; Dijksterhuis & van Knippenberg, 1995; Macrae, Hewstone, & Griffiths, 1993). Thus, documenting individual differences in memory-based versus on-line evaluative responding is a potentially important task.

Assessing On-Line Versus Memory-Based Evaluative Responding

In prior research, two measures of on-line versus memory-based attitudes have been used: attitude-recall valence correlations and response latencies. The most common technique has been to assess the relationship between attitudes and the valence of information recalled (e.g., Hastie & Park, 1986; Lichtenstein & Srull, 1987). This technique is based on the notion that on-line attitudes are spontaneously formed and revised as information is received, whereas memory-based attitudes rely to a greater extent on the retrieval of information from long-term memory and the subsequent evaluation of that information. Because on-line attitudes are relatively less dependent on memory, they tend to have weak or nonsignificant correlations with the valence of information recalled. Because memory-based attitudes depend on memory by definition, however, they tend to have higher correlations with the valence of information recalled. In short, then, greater attitude-recall valence relationships are obtained for memory-based than for on-line attitudes.

On-line and memory-based attitude differences also have been assessed using response latency measures (e.g., Lingle & Ostrom, 1979; Mackie & Asuncion, 1990). Mackie and Asuncion (1990), for instance, found that on-line processing instructions produced shorter response latencies (i.e., faster responses) on attitude measures than did instructions designed to produce memory-based processing. The logic here is that memory-based attitudes take more time to report because the individual must think back to the information that can be recalled, evaluate this information, and then compute an attitude on the spot. Because on-line attitudes have presumably already been formed during information acquisition, they require no extra computation, making them much faster to report.

EXPERIMENT 1

The primary objective of Experiment 1 was to demonstrate that individual differences in the need to evaluate could determine memory-based versus on-line evaluations of people. All participants were exposed to behav-

ioral information about a person and reported their attitudes toward the person and recalled as much information about him or her as they could. We expected to find stronger attitude-recall valence correspondence for low need to evaluate individuals (indicative of memory-based attitude formation) than for high need to evaluate individuals (indicative of spontaneous, on-line attitude formation). We also included a common experimental manipulation of memory-based versus on-line processing sets (see Hastie & Park, 1986) to provide a replication of prior research and examine the possibility of an interaction between situational and individual difference determinants of this processing distinction. A significant interaction could suggest that low need to evaluate individuals tend to engage in little on-line evaluation but that when instructed to do so, they readily change their processing style. On the other hand, the absence of an interaction also would be informative. Although an on-line instructional set should certainly cause even low need to evaluate individuals to consider the evaluative implications of each sentence as it is received, it is possible that they would still not compute their overall, global evaluation of the person until asked to do so. Thus, the global evaluation would still be memory dependent.

In addition, this study sought to determine the extent to which a related individual difference, the need for cognition (Cacioppo & Petty, 1982), could predict memory-based versus on-line attitude formation. Need for cognition refers to the tendency to engage in and enjoy effortful processing and has been found to be moderately correlated with the need to evaluate (Petty & Jarvis, 1996). Thus, we felt it was important to rule this construct out as an alternative explanation for any need to evaluate effects. Moreover, there was some possibility that need for cognition might possess independent predictive potential in this domain. In a persuasion paradigm, Haugtvedt and Petty (1992) found significant attitude-recall correlations for low but not high need for cognition individuals. Because the need to evaluate has been shown to be uniquely predictive of the extent of evaluative thinking (Jarvis & Petty, 1996), however, we expected to rule need for cognition out as an alternative explanation for need to evaluate effects.¹

Method

PARTICIPANTS AND DESIGN

Participants included 140 undergraduates (65 men and 75 women) from the Ohio State University who participated in partial fulfillment of a course requirement. Participants were randomly assigned to instructional set conditions (memory-based vs. on-line) and then completed the Need to Evaluate and Need for Cognition Scales.

PROCEDURE

Participants were seated in a room with 10 cubicles partitioned such that visual contact between them was not permitted. They received instructions designed to induce either a memory-based or an on-line evaluation set. Following these instructions, all participants read the same 20 sentences about a hypothetical person named Ron. Of these 20 sentences, 10 described good behaviors and 10 described bad behaviors.² Two of the good sentences were as follows: "Ron lent money to a friend" and "As a result of his increasing interest in working with young children, Ron volunteered at the local youth center during his free time." Two of the bad sentences were as follows: "Ron stole \$20" and "Ron did not like his neighbor's dog and one afternoon he kicked him as he was walking from his car to the front door of his house." The sentences were presented in a fixed random order across participants and were varied in terms of length and complexity to be consistent with the memory-based instructional set cover story. After rating each of the 20 sentences according to the instructions they received, participants completed questionnaires containing attitude and recall measures as well as the individual difference scales.

INDEPENDENT VARIABLES

Instructional set. Participants were randomly assigned to either the on-line or memory-based set condition. Instructions were adapted from those used by previous researchers (e.g., Hastie & Park, 1986). In the on-line set condition, participants were instructed to read the sentences about Ron and try to form an impression of him. They were told to focus on the kind of person they think Ron might be and then to rate each of the sentences on the degree to which they implied that Ron was either very likable or not at all likable. Participants in this condition also were told that they would be asked a series of questions regarding the attitudes they had formed toward Ron. They then read each sentence and rated it on a 1 to 7 scale of likability.

In the memory-based set condition, participants were instructed to read the sentences about Ron and to focus on the degree to which they were simple or complex. Participants were not told that they would be asked any questions regarding their impressions of Ron as a person. Instead, they were told to focus their attention on the simplicity or complexity of each sentence and were advised to consider things such as the number of verbs or adjectives contained within each sentence. Participants rated each sentence using a 1 to 7 scale ranging from *very simple* to *very complex*. This manipulation was designed to distract participants from on-line evaluation.

Need to evaluate. Following the dependent measures, participants completed the Need to Evaluate Scale

containing 16 items, such as "I form opinions about everything" and "I enjoy strongly liking and disliking new things" (Jarvis & Petty, 1996). Participants responded to each item on a 5-point scale anchored at *extremely uncharacteristic* and *extremely characteristic*. The items on the Need to Evaluate Scale proved highly consistent ($\alpha = .86$); therefore, responses were summed to form one overall index for each participant. The range of scores was 32 to 80 (possible range is 16 to 80) and the median score was 50. Scores were not affected by the instructional set manipulation.

Need for cognition. Participants also completed the 18-item version of the Need for Cognition Scale (Cacioppo, Petty, & Kao, 1984). This scale contains items such as "I prefer complex to simple problems" and "Thinking is not my idea of fun" (reverse-scored). Participants responded to each item on a 5-point scale anchored at *extremely uncharacteristic* and *extremely characteristic*. Again, due to high internal consistency ($\alpha = .89$), scores were summed for each participant. The range of scores was 28 to 87 (possible range is 18 to 90) and the median score was 60. Scores were not affected by instructional set.

DEPENDENT MEASURES

Attitude index. Immediately following the 20 statements about Ron, participants' attitudes toward him were assessed using six 7-point attitude scales. These scales had the following anchors: *bad-good*, *would like him very much-would not like him at all* (reverse-scored), *negative-positive*, *honest-dishonest* (reverse-scored), *caring-uncaring* (reversed-scored), and *stupid-intelligent*. Scales were preceded by questions such as "In general, how good or bad a person do you think Ron might be?" "How much do you think you would like Ron?" and so on. The scales all ranged from 1 to 7, with 1 reflecting a *very negative attitude toward Ron* and 7 reflecting a *very positive attitude toward Ron* (after reverse-scoring the appropriate items). Responses to these scales were averaged to form one overall attitude index ($\alpha = .63$).³

Recall valence index. Following the attitude scales, participants engaged in a free recall task. Participants were asked to write down as many points as they could remember about Ron. They were told not to worry about exact wording, spelling, or grammar but to focus on the main idea of each point. They were given up to 5 minutes to recall as much as they could. Recall was later coded by a judge (blind to experimental condition and individual difference scores) as positive or negative toward Ron and these ratings were used for the recall index. In computing this index, only accurate recall was included. Each item recalled was considered accurate if it captured the gist of one of the behavioral statements.

For the recall valence index, a difference score was computed for each participant using the judge's rating.

Specifically, the number of negative points remembered about Ron was subtracted from the number of positive points remembered about Ron. Positive numbers thus indicated greater positivity in recall and negative numbers indicated greater negativity in recall. This index reflected our assumption that attitudes formed in the present experiment would be based on the explicit consideration of both positive and negative behaviors. Furthermore, it was based on previously used cognitive response valence indices computed in the same fashion (e.g., Mackie, 1987; Mackie & Asuncion, 1990; see also Petty, Priester, & Wegener, 1994).

To determine if participants' subjective assessments of the evaluative content of their own recall would produce different relationships with their attitudes than our *a priori* assessments of the evaluative implications of each behavior, we also asked participants to rate the sentences they listed. Following the recall task, participants were asked to return to each point recalled and rate it with a "+," "-", or "0," depending on whether they thought it was positive, negative, or neutral toward Ron. Unfortunately, a substantial number of participants (64% of the total sample) incorrectly completed this part of the experiment by returning to the original stimuli and rating those statements instead of their own recall. We did, however, compare the judge's ratings to the participants' ratings for those participants who completed this part of the experiment correctly. The ratings were not identical, suggesting that in some cases, participants recalled the gist of a sentence correctly but put their own evaluative spin on it; but they were highly correlated ($r = .83, p < .001$) and produced consistent results in subsequent analyses.

Results

Prior to analysis, we standardized all variables and measures and then followed the recommendation of Cohen and Cohen (1983) and removed outliers from the attitude and recall valence data that were at least three standard deviations above or below their respective means. For the preliminary analyses, only the specific outlying scores were eliminated; the corresponding participants' other data were retained for additional analyses. For our primary analysis of attitude-recall valence relationships, however, outlying scores on either measure resulted in the removal of a participant's entire data, because both attitude and recall valence scores were required to make the data relevant to our concerns. In other words, if an individual had an outlying score on either the attitude or recall valence index, the individual was eliminated from analysis because attitude-recall relationships cannot be determined without both attitude and recall data for each participant. Of importance, outliers were determined across all participants and were

equally likely to be found in any condition. This analysis led to the removal of data from just 4 participants (2.9% of total), resulting in a total N of 136 for our primary analysis.

In all regression analyses, we again followed the recommendation of Cohen and Cohen (1983) and used a hierarchical approach in which a sequence of regression analyses of increasing complexity was run. Within each analysis, only the highest order terms were interpreted. In other words, we first tested only the main effects. In the second analysis, we included main effects and two-way interactions, interpreting only the two-way interactions, and so forth.

PRELIMINARY ANALYSES

In preliminary analyses, we examined the relationships between our independent variables and amount of recall and overall valence of recall. We first analyzed amount of recall and found a significant main effect for instructional set ($\beta = .62, p < .001$) such that participants given an on-line instructional set recalled significantly more information than participants given a memory-based instructional set, perhaps because the memory-based task was more taxing than the on-line task. However, amount of recall was not predicted by need to evaluate ($\beta = -.10, p = .18$), need for cognition ($\beta = -.07, p = .38$), or participant gender ($\beta = -.11, p = .14$), or by any of the interactions between these variables ($ps > .19$).

We also analyzed the overall valence of recall. This analysis revealed a marginally significant main effect for participant gender ($\beta = -.17, p = .06$) such that women showed slightly less positive recall than men. Valence of recall was not predicted by need to evaluate ($\beta = -.03, p = .76$), need for cognition ($\beta = -.12, p = .18$), or instructional set ($\beta = .08, p = .34$), or by any of the interactions between these variables.

Also, consistent with prior research, need to evaluate and need for cognition were moderately correlated ($r = .28, p = .001$). Thus, it was important to assess the degree to which the need for cognition could account for any need to evaluate effects obtained.

ATTITUDE-RECALL VALENCE RELATIONSHIPS

Of primary interest in the present study was the extent to which the need to evaluate predicted on-line versus memory-based attitude formation as judged by the relationship between attitudes and recall valence. The attitude-recall valence relationships were analyzed using a multiple regression predicting attitudes, where recall valence, need to evaluate, and need for cognition were treated as continuous predictor variables and instructional set and participant gender were dummy coded. Participant gender, instructional set, recall valence, need to evaluate, need for cognition, and all of the interaction terms were submitted hierarchically as predictors

of the attitude index (see Cohen & Cohen, 1983). Predictors were considered relevant to the present concerns if they contained the recall term. For example, the effect of need to evaluate on the attitude-recall valence relationship was represented by the interaction (i.e., cross-product) between recall valence and need to evaluate in the regression analysis predicting attitudes. The instructional set effect on the attitude-recall valence relationship was represented by the interaction between recall valence and the instructional set variable. Other effects were similarly represented.

In this analysis, we found that the overall attitude-recall valence relationship was positive and marginally significant ($\beta = .16, p = .07$), suggesting that the more positive information they recalled, the more favorable participants' attitudes were toward Ron. However, there was no relationship between attitudes and need to evaluate ($\beta = -.09, p = .32$), need for cognition ($\beta = .07, p = .46$), participant gender ($\beta = .06, p = .47$), or instructional set ($\beta = .13, p = .14$).

More pertinent to our primary interests, however, were the interaction terms between recall valence and the other predictors. As expected, two significant interactions emerged. First, a Recall Valence \times Instructional Set interaction ($\beta = -.61, p < .05$) indicated that we replicated the basic effect of instructional set from prior research. As illustrated in the top panel of Figure 1, there was a stronger relationship between attitudes and recall valence under memory-based set conditions ($r = .29, p < .02$) than under on-line set conditions ($r = .02, p = .89$). Second, and more important, we found that the predicted interaction between recall valence and need to evaluate was significant as well ($\beta = -.72, p < .05$). As shown in the bottom panel of Figure 1, the relationship between attitudes and recall valence was greater for individuals low in need to evaluate ($r = .27, p < .03$) than for those high in need to evaluate ($r = .03, p = .83$), determined by a median split for illustrative purposes. Neither the Recall Valence \times Need for Cognition interaction ($\beta = -.33, p = .38$) nor the Recall Valence \times Participant Gender interaction ($\beta = .15, p = .62$) approached significance. Furthermore, none of the higher order interactions approached significance ($ps > .22$).⁴ Of particular interest was the absence of a three-way interaction between recall valence, need to evaluate, and instructional set ($\beta = .14, p = .87$). Under memory-based set conditions, the attitude-recall valence correlation was significant for low ($r = .48, p < .01$) but not high need to evaluate individuals ($r = .07, ns$). Moreover, although the correlation was not significant for either group under on-line set conditions, it also tended to be greater for low ($r = .11$) than for high need to evaluate individuals ($r = .01$).

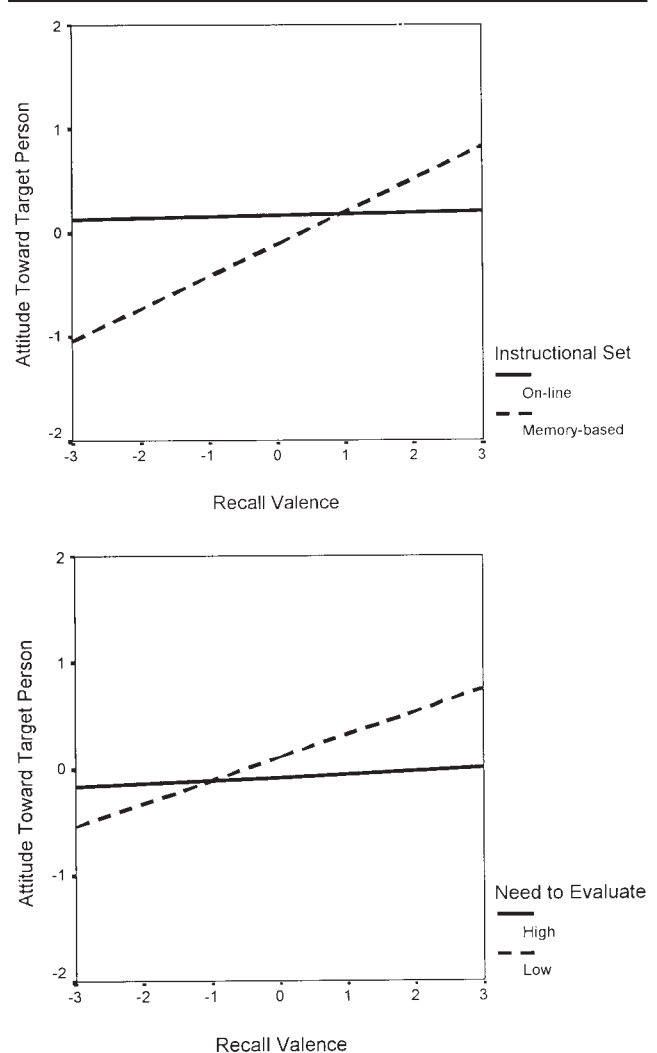


Figure 1 Top panel: Effects of instructional set on the relationship between attitudes and recall valence (Experiment 1). Bottom panel: Effects of need to evaluate on the relationship between attitudes and recall valence (Experiment 1).

Discussion

Our first experiment provided initial evidence that the need to evaluate predicts the extent of on-line versus memory-based attitude formation in person perception. Using attitude-recall valence correspondence as the criterion, Experiment 1 demonstrated that individuals high in need to evaluate engage in on-line attitude formation, whereas those low in need to evaluate engage in less on-line and more memory-based attitude formation. This finding suggests that there are individual differences in spontaneous evaluative processing of other people. High need to evaluate individuals evaluate during the acquisition of person information, whereas low need to evaluate individuals engage in less evaluation during

information acquisition and therefore depend more on the information they can recall when an evaluation question is posed. Furthermore, the findings in Experiment 1 suggest that need to evaluate is uniquely predictive of on-line versus memory-based evaluations, not being accounted for by its relationship with need for cognition. Interestingly, although analysis of the simple correlations revealed a nonsignificant attitude-recall valence relationship for low need to evaluate individuals in the on-line set condition, the interaction between need to evaluate, instructional set, and recall valence did not even approach significance in the regression analysis, which is the more reliable analysis because it did not involve a median split. Thus, it appears that an evaluative set (on-line instructions) can make low need to evaluate individuals a little more evaluative but that they are still relatively more reliant on memory than individuals who are higher in the need to evaluate. As mentioned earlier, it is possible that evaluative instructions cause low need to evaluate individuals to evaluate each piece of person information as it is received but not to actively synthesize this information into a global evaluation. Perhaps only when an evaluation is required does this synthesis occur. Thus, judgments remain somewhat dependent on memory.

EXPERIMENT 2

The purpose of Experiment 2 was to conceptually replicate and extend the findings from Experiment 1. Thus, in Experiment 2, we used a highly similar methodology with a few key differences. Specifically, a new hypothetical person ("Marie") was created, all of the behavioral statements were changed, and the instructional set manipulation was altered slightly. These changes were made to enhance the generality of our findings. Most important, however, Experiment 2 was conducted on computers to allow for the collection of response latency data. We sought to determine the extent to which need to evaluate could predict response latencies on attitude items in an impression formation paradigm. It was predicted that low need to evaluate individuals would take more time to formulate their evaluative responses (because they were computing them on the basis of the valence of information they could recall) than high need to evaluate individuals (who had presumably already formed them on-line).

Finally, in Experiment 2, we included the Personal Need for Structure Scale (Neuberg & Newsom, 1993), which assesses the general preference for cognitive simplicity. As mentioned earlier, need for structure has been shown to be important to other aspects of person perception. Thus, although uncorrelated with the need to evaluate in past research ($r = .03$) (Petty & Jarvis, 1996), we felt it was possible that the personal need for structure

also might predict on-line versus memory-based evaluations. Because high need for structure individuals desire simplicity and closure (see Kruglanski et al., 1997), one possibility was that they would be more likely to form quick, on-line evaluations than low need for structure individuals, who are more tolerant of ambiguity and might wait until all the information was presented before forming an evaluation. We included this scale to test this possibility.

Method

PARTICIPANTS AND DESIGN

Participants included 165 undergraduates (67 men and 98 women) from the Ohio State University who participated in Experiment 2 to help meet a course requirement. Participants were randomly assigned to instructional set conditions and then completed the Need to Evaluate Scale, the Personal Need for Structure Scale, and the Need for Cognition Scale.

PROCEDURE

Participants were seated in a room with eight partitioned computers and read instructions displayed on their monitors. Participants read 20 sentences (pre-tested and selected using the same procedure as in the first experiment; see Note 2) about a hypothetical person named Marie. In total, there were 10 positive and 10 negative sentences, presented in a different random order for each participant. Examples of sentences with positive and negative implications for Marie, respectively, are as follows: "Marie bought groceries for her elderly neighbor during the snowstorm" and "When Marie scraped the side of the other car in the parking lot, she drove away without leaving a note." After reading each of the 20 sentences, participants completed attitude and recall measures and the individual difference inventories.

INDEPENDENT VARIABLES

Instructional set. Participants were randomly assigned to either the on-line set condition or the memory-based set condition. In the on-line set condition, participants were given instructions that were identical to those used in Experiment 1. In the memory-based set condition, participants read instructions that were adapted from Mackie and Asuncion (1990); that is, they were told to read the sentences about Marie and to focus on the degree to which they were dynamic or not dynamic. The meaning of the word *dynamic* was left ambiguous but participants were told that in making this assessment they should pay attention to the number and type of verbs contained in each sentence. Participants in this condition were not told that they would be asked any questions regarding their impressions of Marie as a person. They

rated each sentence using a 1 to 7 scale, ranging from *not at all dynamic* to *very dynamic*.

Need to evaluate. Following the dependent measures, all participants completed the Need to Evaluate Scale. Items on this scale showed adequate consistency ($\alpha = .78$); therefore, responses to each item were summed to form one overall index for each participant. Scores ranged from 26 to 70 ($Mdn = 51$) and were unaffected by instructional set.

Personal need for structure. Participants then completed the Personal Need for Structure Scale (Neuberg & Newsom, 1993). This scale contains items such as "I don't like situations that are uncertain" and "I enjoy having a clear and structured mode of life." Participants responded to each item on a 6-point scale anchored at *strongly disagree* and *strongly agree*. Items were consistent ($\alpha = .72$); therefore, we summed responses to form an overall index. Scores ranged from 23 to 60 (possible range is 11 to 66) and the median score was 42. Need for structure scores also were unaffected by instructional set.

Need for cognition. Participants also completed the 18-item version of the Need for Cognition Scale. The reliability of items on this scale was quite high ($\alpha = .89$); therefore, we again summed them to form a single overall index. Scores ranged from 25 to 89 ($Mdn = 57$) and were unaffected by instructional set.

DEPENDENT MEASURES

Attitude index. Participants' attitudes toward Marie were assessed using seven 7-point attitude scales. These scales had the following anchors: *bad-good*, *negative-positive*, *would not like her at all-would like her very much*, *unfavorable-favorable*, *honest-dishonest* (reverse-scored), *caring-uncaring* (reverse-scored), and *stupid-intelligent*. Question wording for each item closely paralleled the wording used in the first study. Responses to these scales were averaged ($\alpha = .73$) to form an overall attitude index.

Recall valence index. Following the attitude measure, participants engaged in a free recall task that was identical to that used in Experiment 1, except that they were not asked to rate the valence of the information they recalled. We then used a judge's rating to compute a recall valence index in the same way as in Experiment 1. Of importance, a second judge also coded recall for 58 randomly selected participants. These additional ratings were used to assess interrater reliability and were found to correlate highly with the ratings of the first judge ($r = .94$, $p < .001$). Thus, the ratings of the first judge were deemed reliable and were used for all analyses.

Response latency. Reaction times for attitude responses also were collected. Because a key distinction between memory-based and on-line evaluation is the point at which an attitude is formed, the response latency index

was represented by the reaction time to the very first attitude item alone (i.e., the *good-bad* item). This item was placed first because of its global evaluative nature (i.e., it is not linked to any particular belief or feeling). Focusing our analysis on the first item was based on the notion that once a global evaluation is formed, it is stored and thus quicker to report (e.g., Lingle & Ostrom, 1979; see also Fazio, Chen, McDonel, & Sherman, 1982). Because the first attitude item required everyone to report an attitude, whether they had already formed one (i.e., on-line) or whether they formed one right then from memory, response latency differences should have disappeared (or become attenuated) after that item given that all participants had now produced evaluations.

Results

We again standardized all variables and measures prior to analysis. We then removed outliers from the attitude and recall valence data following the same criteria as in the first experiment. In total, data from 7 participants (3.6% of sample) were removed from the primary analysis of attitude-recall valence relationships, producing a final N of 158 for that analysis. Analyses in Experiment 2 were then conducted using the same hierarchical regression approach as in Experiment 1.

PRELIMINARY ANALYSES

We first examined the relationships between our independent variables and amount of recall. As in Experiment 1, there was a main effect for instructional set on the number of items of information recalled ($\beta = .44$, $p < .001$). Participants given the on-line instructional set recalled more information than those given the memory-based instructional set. There were no main effects for participant gender ($\beta = .11$, $p = .13$), need to evaluate ($\beta = -.06$, $p = .40$), need for cognition ($\beta = .04$, $p = .59$), or personal need for structure ($\beta = -.02$, $p = .84$). There was one significant interaction between participant gender and need for cognition ($\beta = .48$, $p = .05$) such that for women there tended to be a positive relationship between need for cognition and the amount of information recalled ($\beta = .18$, $p = .07$), whereas for men there was no relationship ($\beta = -.13$, $p = .31$). No further interactions were reliable ($ps > .20$).

We also examined the overall valence of recall. This analysis revealed one significant effect—a main effect for need to evaluate ($\beta = -.16$, $p < .05$)—indicating an inverse relationship between need to evaluate and positivity of recall. There were no main effects for participant gender ($\beta = .13$, $p = .11$), instructional set ($\beta = .08$, $p = .30$), need for cognition ($\beta = -.08$, $p = .36$), or personal need for structure ($\beta = .03$, $p = .69$), and there were no interactions between any of the predictor variables ($ps > .11$).

Consistent with past research (Petty & Jarvis, 1996), need to evaluate was positively correlated with need for cognition ($r = .28, p < .01$) and uncorrelated with need for structure ($r = -.01$), whereas need for cognition was negatively correlated with need for structure ($r = -.19, p = .02$).

ATTITUDE-RECALL VALENCE RELATIONSHIPS

We found that the overall attitude-recall valence relationship was marginally significant ($\beta = .13, p < .10$), again indicating a positive relationship between attitudes and recall valence. There was no relationship, however, between global attitudes and need to evaluate ($\beta = -.06, p = .45$), need for cognition ($\beta = .02, p = .85$), need for structure ($\beta = .10, p = .20$), or participant gender ($\beta = -.06, p = .42$). Attitudes were affected by instructional set ($\beta = .40, p < .001$), such that participants with an on-line instructional set had more favorable attitudes ($M = .41$) than those with a memory-based instructional set ($M = -.40$). Perhaps the on-line set was more enjoyable and thus produced positive affect that was transferred to the target person (see Schwarz & Clore, 1983).

Our primary concerns were with the interactions between recall valence and the other predictors. Most important, we replicated the first experiment and found a significant interaction between recall valence and need to evaluate ($\beta = -.62, p < .05$). As shown in Figure 2, the relationship between attitudes and recall valence was significant for individuals low in need to evaluate ($r = .31, p < .03$) but not for individuals high in need to evaluate ($r = .04, p = .76$), determined by a median split for illustration. Unlike Experiment 1, however, there was not a significant interaction between recall valence and instructional set ($\beta = -.17, p = .44$), although simple correlations were in the right direction and revealed that the attitude-recall valence relationship was significant in the memory-based set condition ($r = .24, p < .04$) but not the on-line set condition ($r = .06, p = .59$). There were no significant interactions between recall valence and gender ($\beta = .09, p = .70$), need for cognition ($\beta = .21, p = .49$), or need for structure ($\beta = -.18, p = .57$). Furthermore, there was only one significant higher order interaction—the three-way interaction between recall valence, instructional set, and gender ($\beta = 1.11, p < .02$). This interaction suggested that although men demonstrated the predicted interaction between instructional set and recall valence ($\beta = -.75, p = .02$), women did not ($\beta = .29, p = .32$). No further interactions were significant ($ps > .18$). As in Experiment 1, there was no three-way interaction between recall valence, need to evaluate, and instructional set ($\beta = .51, p = .56$). Under memory-based set conditions, the attitude-recall valence correlation was signif-

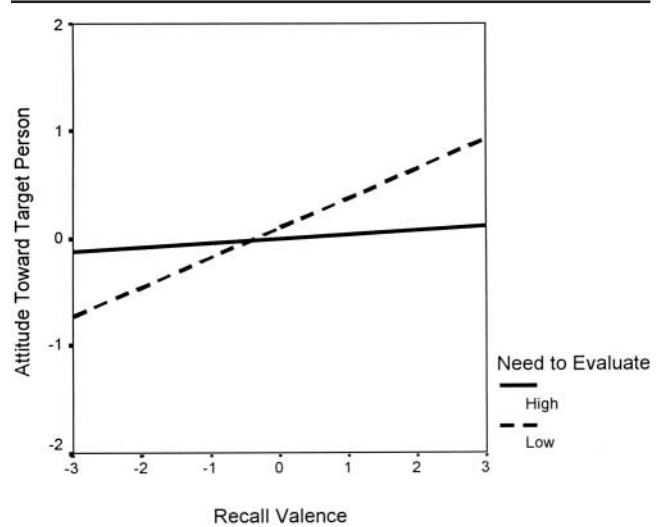


Figure 2 Effects of need to evaluate on the relationship between attitudes and recall valence (Experiment 2).

icant for low ($r = .45, p < .02$) but not high need to evaluate individuals ($r = .09, ns$). Furthermore, although the correlation was not significant for either group, low and high need to evaluate individuals were similarly spread under on-line set conditions ($rs = .21$ and $-.11$, respectively).

RESPONSE LATENCIES

Reaction times were expected to be slower for individuals engaging in relatively little on-line (and thus more memory-based) evaluative processing than for individuals engaging in extensive on-line evaluative processing. In other words, we predicted main effects for both need to evaluate and instructional set. Because outlying data in the previous analysis pertained to attitude and recall ratings per se, and not response latencies, we reanalyzed for outlying scores that were at least three standard deviations above or below the mean response latency across all participants. This analysis again resulted in the removal of seven (different) participants from the final analysis. Following the recommendation of Fazio (1990), we then performed a logarithmic transformation on reaction times to correct for any remaining skewness in their distribution.

These data were submitted to a hierarchical regression, predicting reaction times. Standardized need to evaluate, need for cognition, and need for structure scores served as continuous predictor variables, whereas instructional set and participant gender were dummy coded. No recall term was included in this analysis. As expected, this analysis revealed two significant effects. A

significant effect for instructional set ($\beta = -.17, p < .04$) indicated that individuals in the on-line set condition were significantly faster to report their attitudes ($M = 6.49$ seconds) than individuals in the memory-based set condition ($M = 7.60$ seconds). More important, a significant effect for need to evaluate ($\beta = -.17, p = .04$) indicated that as need to evaluate increased, response time decreased; that is, high need to evaluate individuals were faster to report their attitudes ($M = 6.36$ s) than those low in need to evaluate ($M = 7.75$ s), divided by a median split for illustrative purposes. Consistent with our other findings, reaction times were not predicted by need for cognition ($\beta = .05, p = .60$) or need for structure ($\beta = .02, p = .77$). There was, however, a marginally significant main effect for gender ($\beta = -.13, p < .10$), suggesting that women reported their attitudes more quickly than did men. Also consistent with our other findings, there was no interaction between need to evaluate and instructional set ($\beta = -.05, p = .86$). No other interactions were significant.⁶

As a final analysis, we averaged the reaction times to the remaining attitude items (i.e., beyond the first item), conducted a logarithmic transformation on these scores, and submitted the data to the same analysis. As expected, neither instructional set ($\beta = -.09, p = .26$) nor need to evaluate ($\beta = -.12, p = .14$) predicted this reaction time index. Similarly, this index was not predicted by need for structure ($\beta = -.12, p = .15$) or by gender ($\beta = .05, p = .55$). All other effects closely paralleled those from the analysis of the first item alone.

Discussion

Using both attitude-recall valence relationships and response latency measures, Experiment 2 replicated and extended the major findings from Experiment 1. We found a stronger attitude-recall valence relationship for low than for high need to evaluate individuals and also that low need to evaluate individuals were slower to report their attitudes than were high need to evaluate individuals. Both findings were consistent with the notion that low need to evaluate individuals engage in reduced on-line evaluative processing and thus relatively more memory-based evaluative processing. Of course, a bit of caution is warranted here because our reaction time index was based on responses to a single self-report item. Although we had reason to form the index from responses to the first item alone, future research using more controlled reaction time methodology, and perhaps a more reliable index, would be useful.

As in Experiment 1, we found no evidence of an interaction between instructional set and need to evaluate for either the attitude-recall data or the response latency data, despite changing the operationalization of the memory-based evaluative set. The absence of this inter-

action provides further evidence that even when provided with evaluation instructions (i.e., on-line set), the global attitudes of low need to evaluate individuals are still somewhat more dependent on memory than the attitudes of high need to evaluate individuals. Under such conditions, low need to evaluate individuals may evaluate the implications of each behavior individually without synthesizing, updating, and revising these "mini-assessments" on-line. Again, it is interesting to note that our analysis of the simple correlations did reveal an attenuation of the attitude-recall valence relationship for low need to evaluate individuals under on-line set conditions, consistent with the idea that evaluative instructions do cause them to engage in somewhat more on-line evaluation. We hesitate to put too much weight on this finding, however, given that it was based on a median split and inconsistent with the regression analysis.

It is worth addressing the possibility that our reaction time results could have stemmed from a tendency of high need to evaluate individuals to just respond more quickly in general; that is, it is conceivable that their higher motivation to evaluate makes them more likely to make quick decisions of all kinds, regardless of the extent to which they have evaluated something on-line. We believe, however, that the present data are inconsistent with this possibility. Specifically, the faster response times of high need to evaluate individuals did not extend beyond the first attitude item. If high need to evaluate individuals were simply faster in general than low need to evaluate individuals, this difference would be expected to hold over all items. Thus, this explanation does not appear to provide a viable alternative account for our findings.

GENERAL DISCUSSION

Research on the need to evaluate construct (Hermans & Eelen, 2001; Jarvis & Petty, 1996; Petty & Jarvis, 1996) suggests that meaningful individual differences exist in the extent to which people form evaluations. The primary goals of the present research were twofold: First, we sought to extend the work on the need to evaluate to the domain of person perception and demonstrate that individual differences exist even in the tendency to evaluate other individuals—a domain in which on-line evaluative responding has appeared to be particularly pervasive (e.g., Hastie & Park, 1986; Uleman et al., 1996). Second, we sought to apply this notion to the literature on on-line versus memory-based judgment formation, which has just begun to consider the role of individual differences in motivation (e.g., McConnell, 2001). In the present research, we accomplished each of these goals. Although people high in the need to evaluate engage in considerable on-line evaluation, people low in the need to evaluate are more prone to making global evaluative judg-

ments only when an evaluative question is posed and thus are relatively more dependent on the information they can recall at the time the judgment is required.

Curiously, we did not replicate the finding of the Instructional Set \times Recall Valence interaction in Experiment 2. To further test the reliability of this effect, we combined the data from the two experiments to achieve maximum power and reanalyzed the attitude-recall valence data. Because the two experiments differed in time (i.e., they were conducted in separate academic terms), operationalization of the memory-based set, gender of target person, and the specific behavioral sentences used, this analysis provided a valuable test of the generality of our findings. Of importance, we explicitly included an "experiment" factor (dummy coded) to account for possible differences across experiments. We then conducted a hierarchical regression, predicting attitudes and including the predictors that appeared in both studies (i.e., recall, need to evaluate, need for cognition, instructional set, gender, and experiment).

Analysis revealed the same main effects for instructional set ($\beta = .28, p < .001$) and recall valence ($\beta = .11, p < .06$) that appeared in each experiment. Interestingly, there was also a marginally significant interaction between recall valence and experiment ($\beta = .37, p < .09$), indicating that the overall attitude-recall valence relationship was slightly stronger in the first experiment than in the second. Aside from this one marginal interaction, there were no significant effects at any level involving the experiment factor ($ps > .17$), need for cognition ($ps > .12$), or participant gender ($ps > .41$). There was, however, a significant interaction between instructional set and recall valence ($\beta = -.34, p < .05$). This interaction replicated the basic set effect from prior literature in that the relationship between attitudes and recall valence was significant under memory-based set conditions ($r = .23, p < .01$) but not under on-line set conditions ($r = .05, p = .51$). The interaction between need to evaluate and recall valence also was significant ($\beta = -.59, p < .02$) and indicated that the relationship between attitudes and recall valence was significant for low need to evaluate individuals ($r = .31, p = .001$) but not for high need to evaluate individuals ($r = .05, p = .61$). These findings suggest that across experiments, both instructional set and need to evaluate moderated the attitude-recall relationship. Notably, there was no three-way interaction between need to evaluate, instructional set, and recall valence ($\beta = -.15, p = .83$), consistent with the results from each individual experiment. Again, although it is reasonable to speculate that under some conditions low need to evaluate individuals could be motivated to engage in more extensive on-line evaluation and integration of social stimuli (e.g., undergraduates evaluating a real-life dating partner), the present research clearly

suggests that they have a tendency to wait to form global evaluations of other people until asked to do so, even under evaluation instructions.

Why Does Need to Evaluate Matter?

As these analyses illustrate, both instructional set and the need to evaluate moderate the attitude-recall valence relationship. As explained by Srull and Wyer (1986), processing goals are the key to memory-based versus on-line evaluative responding. Although previous operationalizations of evaluative versus nonevaluative processing goals have come strictly in the form of instructional set manipulations, the need to evaluate also describes differences in this regard. When processing goals are evaluative in nature, as with an on-line set or high need to evaluate, on-line evaluation is more likely. When processing goals are nonevaluative, as with a memory-based set or low need to evaluate, on-line evaluation is less likely and memory-based evaluation is more likely. Similar to past research in this area, the present research does not address the specific processes by which on-line and memory-based attitude formation occur, but it does raise important questions. For example, on what information do high need to evaluate individuals base their attitudes? They appear to be forming attitudes on-line, but does that mean memory plays absolutely no role in the formation of their attitudes? Equally important, what exactly does it mean that low need to evaluate individuals engage in memory-based attitude formation?

HIGH NEED TO EVALUATE

Srull and Wyer (1989) outlined a series of processes they thought were relevant to on-line evaluations. They argued that when people have evaluative processing goals, they seek an evaluative organization of information in memory. Thus, they spontaneously encode behaviors as evaluative traits. These traits become unified as a single trait concept that is revised and updated as subsequent behaviors are encoded. This trait concept or representation gets stored separately in memory; so, when asked to report their attitude, people simply retrieve the summary evaluation they have already made. As a result, the evaluation is independent of the individual behaviors that can be recalled. This explanation is speculative but most likely approximates the general type of processing in which high need to evaluate individuals engage. At the very least, they appear to be integrating information into an evaluation as they receive it, thus reporting their attitudes quickly without scanning long-term memory for behaviors.

LOW NEED TO EVALUATE

According to Srull and Wyer (1989), when an evaluative processing goal is not in place, people do not spontaneously evaluate while information is presented

but instead rely on memory for information to form their attitudes when they are asked to report them. Under these conditions, the evaluation of a person is contingent on the specific behaviors stored in memory; that is, people who have not already formed an evaluation recall what they can about the person and then evaluate this information. If primarily positive behaviors are recalled, one's evaluation of the person will be favorable. If more negative behaviors are recalled, one's evaluation of the person will be unfavorable. Of course, this does not imply that there is no on-line evaluation whatsoever but just that there is a relatively greater reliance on memory for raw information. Because low need to evaluate individuals tend not to have an evaluative processing goal (Jarvis & Petty, 1996), a memory-based process more likely describes the evaluations they report.

As described earlier, it remains possible that low need to evaluate individuals are forming evaluations on-line (e.g., evaluations of each behavior) but for some reason, they are not synthesizing them into a single global evaluation until a judgment is required. In fact, our analysis of simple correlations suggested that evaluative instructions did produce at least some increase in on-line evaluative responding for these individuals. Perhaps because they do not enjoy engaging in evaluative responding, low need to evaluate individuals simply do less of it and thus fail to synthesize on-line evaluations. Alternatively, low need to evaluate individuals might form evaluations on-line but have less confidence in those evaluations, thus relying on them to a lesser degree. Other recent research (Petty, Briñol, & Tormala, *in press*) has indicated that people are less willing to rely on their evaluative thoughts in forming attitudes when they lack confidence in the validity of those thoughts. If low need to evaluate individuals lacked confidence in their on-line evaluations (perhaps because they have less practice forming them), it would be sensible for these individuals to refrain from forming global evaluative assessments until it is necessary to do so.

Another possibility is that the attitude-recall valence relationships in the present research suggest that low need to evaluate individuals are somehow more rational and thus retrieve information from memory to support their attitudes once they are formed; that is, it is possible that low need to evaluate individuals recall information selectively to justify the evaluations they form. The fact that our findings generalized to the response latency measure, however, reduces the likelihood of this possibility because latencies were recorded before participants even knew they would be asked to recall any information. Furthermore, there is no *a priori* reason to think that low need to evaluate individuals would be more rational than high need to evaluate individuals, particularly given the positive correlation between need to eval-

uate and need for cognition (which may map more accurately onto differences in rationality; see Cacioppo, Petty, Feinstein, & Jarvis, 1996). In any case, future research might vary the order of the attitude and recall measures to test the notion that recall is used as justification by low need to evaluate individuals.

New Questions and Future Directions

While answering some important questions, the present research also raises new ones. For example, is it possible that a more realistic situation in person perception would elicit more pervasive on-line evaluative responding? Perhaps in more lifelike or interpersonal situations, we are more generally inclined to spontaneously evaluate other people. The present studies were somewhat removed from the reality of our everyday experiences with others. They were, however, modeled after a paradigm common in this literature, and need to evaluate proved to be equally predictive across instructional set conditions. Future research using more real-world situations with others would be useful to address these questions. It would certainly be reasonable to find that some situations were so involving for all participants that need to evaluate would no longer moderate on-line versus memory-based attitude formation.

Another interesting path for future research would be to include additional measures that might further identify the different types of processes involved in the attitude formation strategies of individuals varying in the need to evaluate. Thought listings, for instance, might shed light on the nature of on-line attitude formation. According to Greenwald (1968), recall tends not to matter as much for attitudes when a person is actively generating arguments or evaluative thoughts. Thought listings or cognitive responses (see Petty, Ostrom, & Brock, 1981) could reveal the extent to which people are elaborating, counterarguing, accepting, or thinking favorable thoughts in response to the behaviors of another person. These types of cognitive responses, rather than recall, may be correlated with attitudes for people engaging in on-line attitude formation. Indeed, a great deal of previous research suggests that they are (Mackie & Asuncion, 1990; Petty & Cacioppo, 1986).

The present research also may have implications for areas of research not directly falling under the attitudes rubric. For example, our findings may be relevant to the stereotyping literature. As noted earlier, using both cognitive load and processing goal manipulations, a number of researchers have found that memory favors stereotype-consistent information under memory-based processing conditions (e.g., Dijksterhuis & van Knippenberg, 1995; Macrae et al., 1993; Stangor & Duan, 1991). Similarly, it has been found that judgments are more stereotypic (and correlated with memory)

under such circumstances (Bodenhausen & Lichtenstein, 1987; Dijksterhuis & van Knippenberg, 1995; Macrae et al., 1993). Interestingly, it may follow that the need to evaluate would predict stereotypic recall and the stereotypicality of judgments based on that recall. In particular, we would expect to find low need to evaluate individuals relying on stereotypes to a greater degree than high need to evaluate individuals. Future research might explore this possibility.

NOTES

1. It is important to note that Haugtvedt and Petty's (1992) findings do not necessarily speak directly to the present concerns. This is because the persuasive message used in their research contained only favorable (positive) information about the attitude object. Therefore, it is possible that the attitude-recall correlations were based as much on the mere number of arguments recalled as on the evaluative implications (i.e., valence) of those arguments. Counting the number of arguments is a cue-based process (Petty & Cacioppo, 1984) that does not necessarily imply either on-line or memory-based attitude formation. Indeed, such counting could occur during message presentation, whereby an individual keeps track of the approximate number of arguments as they are received, or after message presentation, whereby an individual counts up the arguments after they have all been received. In neither case would attitudes have been computed by a consideration of the valence of information stored in memory.

2. Of importance, the sentences (designed to be similar to those used in previous research in this area) were selected through pretesting, where they were identified from a slightly larger pool of sentences as similarly extreme. That is, overall, positive and negative sentences differed from neutrality to the same degree.

3. Due to their somewhat low reliability, we also conducted a factor analysis on the attitude items. We found that although there were two factors, most items still loaded onto the first. Using only those items as the attitude index do not change the results.

4. We also tested the effects of need for cognition in a model in which need for cognition was the only individual difference variable included. In this model as well, need for cognition was not a significant predictor of the attitude-recall relationship. Furthermore, we conducted two additional hierarchical regression analyses using different recall indices: one using positive recall alone and one using negative recall alone. No significant effects emerged from the analysis using positive recall as the recall index. When negative recall was used, however, one significant effect emerged: Participants in the on-line set condition reported more favorable attitudes than those in the memory-based set condition ($\beta = .25, p = .02$). No other effects were significant.

5. Consistent with the primary analysis, need for cognition and personal need for structure also failed to predict the attitude-recall relationship in separate regression models where they were the only individual difference variables entered. Furthermore, when positive recall alone was substituted as the recall index in the primary model (i.e., the model with all variables included), we found significant effects for instructional set ($\beta = .41, p < .001$) and the interaction between recall valence, instructional set, and participant gender ($\beta = 1.16, p < .02$), consistent with those in the main analysis. No other effects were significant. When negative recall alone was treated as the recall index, we found the same effect for instructional set ($\beta = .46, p < .001$), a marginally significant negative relationship between attitudes and recall ($\beta = -.14, p = .08$; indicating that as negative recall decreased, attitudes became more favorable), and a marginally significant interaction between recall and need to evaluate ($\beta = .60, p = .09$), suggesting that the attitude-recall relationship was stronger for low than high need to evaluate individuals. Given the apparent asymmetry between positive and negative recall in this experiment and Experiment 1, it could be useful in future research to address their independence in this context, perhaps by analyzing them as separate predictors.

6. The interaction between need for structure and instructional set approached significance ($\beta = .49, p = .08$) and indicated that the predicted effect for instructional set was found for individuals low in need for structure but not for those who scored more highly on this scale. Reaction times also were submitted to separate regression models in which need to evaluate, need for cognition, and personal need for structure were the only individual difference variables included. In these models as well, these variables were not significant predictors of reaction times.

REFERENCES

- Anderson, N., & Hubert, S. (1963). Effects of concomitant verbal recall on order effects in personality impression formation. *Journal of Verbal Learning and Behavior*, 2, 379-391.
- Bargh, J. A., Chaiken, S., Govender, R., & Pratto, F. (1992). The generality of the automatic attitude activation effect. *Journal of Personality and Social Psychology*, 62, 893-912.
- Bargh, J. A., Chaiken, S., Raymond, P., & Hymes, C. (1996). The automatic evaluation effect: Unconscious automatic attitude activation with a pronunciation task. *Journal of Experimental Social Psychology*, 32, 104-128.
- Bargh, J. A., & Thein, R. D. (1985). Individual construct accessibility, person memory, and the recall-judgment link: The case of information overload. *Journal of Personality and Social Psychology*, 49(5), 1129-1146.
- Bodenhausen, G. V., & Lichtenstein, M. (1987). Social stereotypes and information-processing strategies: The impact of task complexity. *Journal of Personality and Social Psychology*, 52(5), 871-880.
- Cacioppo, J. T., & 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., & Jarvis, W. B. G. (1996). Individual differences in cognitive motivation: The life and times of people varying in need for cognition. *Psychological Bulletin*, 119, 197-253.
- Cacioppo, J. T., Petty, R. E., & Kao, C. F. (1984). The efficient assessment of need for cognition. *Journal of Personality Assessment*, 48(3), 306-307.
- Chartrand, T. L., & 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(3), 464-478.
- Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum.
- Dijksterhuis, A., & van Knippenberg, A. (1995). Memory for stereotype-consistent and stereotype-inconsistent information as a function of processing pace. *European Journal of Social Psychology*, 25, 689-693.
- Fazio, R. H. (1990). A practical guide to the use of response latency in social psychological research. In C. Hendrick & M. S. Clark (Eds.), *Research methods in personality and social psychology* (pp. 74-97). Newbury Park, CA: Sage.
- Fazio, R. H., Chen, J., McDonel, E. C., & Sherman, S. J. (1982). Attitude accessibility, attitude-behavior consistency, and the strength of the object-evaluation association. *Journal of Personality and Social Psychology*, 18, 339-357.
- Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, 50, 229-238.
- Glaser, J., & Banaji, M. R. (1999). When fair is foul and foul is fair: Reverse priming in automatic evaluation. *Journal of Personality and Social Psychology*, 77, 669-687.
- Greenwald, A. G. (1968). Cognitive learning, cognitive response to persuasion, and attitude change. In A. G. Greenwald, T. C. Brock, & T. M. Ostrom (Eds.), *Psychological foundations of attitudes* (pp. 147-170). San Diego, CA: Academic Press.
- Hastie, R., & Park, B. (1986). The relationship between memory and judgment depends on whether the judgment task is memory-based or on-line. *Psychological Review*, 93(3), 258-268.
- Haugtvedt, C. P., & 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.
- Hermans, D., DeHouwer, J., & Eelen, P. (2001). A time course analysis of the affective priming effect. *Cognition and Motivation*, 15, 143-165.
- Jarvis, W.B.G., & Petty, R. E. (1996). The need to evaluate. *Journal of Personality and Social Psychology*, 70(1), 172-194.
- Kruglanski, A. W., Adash, M. N., DeGrada, E., Mannetti, L., Pierro, A., & Webster, D. M. (1997). Psychological theory testing versus psychometric nay-saying: Comment on Neuberg et al.'s (1997) critique of the need for closure scale. *Journal of Personality and Social Psychology*, 73, 1005-1016.
- Lichtenstein, M., & Srull, T. K. (1987). Processing objectives as a determinant of the relationship between recall and judgment. *Journal of Experimental Social Psychology*, 23, 93-118.
- Lingle, J. H., & Ostrom, T. M. (1979). Retrieval selectivity in memory-based impression formation. *Journal of Personality and Social Psychology*, 37(2), 180-194.
- Mackie, D. M. (1987). Systematic and nonsystematic processing of majority and minority persuasive communications. *Journal of Personality and Social Psychology*, 53, 41-52.
- Mackie, D. M., & Asuncion, A. G. (1990). On-line and memory-based modification of attitudes: Determinants of message recall-attitude change correspondence. *Journal of Personality and Social Psychology*, 59(1), 5-16.
- Macrae, C. N., Hewstone, M., & Griffiths, R. J. (1993). Processing load and memory for stereotype-based information. *European Journal of Social Psychology*, 23, 77-87.
- Markus, H., & Zajonc, R. B. (1985). The cognitive perspective in social psychology. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology* (3rd ed., Vol. 1, pp. 137-230). New York: Random House.
- McConnell, A. R. (2001). Implicit theories: Consequences for social judgments of individuals. *Journal of Experimental Social Psychology*, 37, 215-227.
- Moskowitz, G. B. (1993). Individual differences in social categorization: The influence of personal need for structure on spontaneous trait inferences. *Journal of Personality and Social Psychology*, 65, 132-142.
- Neuberg, S. L., & Newsom, J. T. (1993). Personal need for structure: Individual differences in the desire for simple structure. *Journal of Personality and Social Psychology*, 65, 113-131.
- Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning*. Urbana: University of Illinois Press.
- Petty, R.E., Briñol, P., & Tormala, Z. L. (in press). Thought confidence as a determinant of persuasion: The self-validation hypothesis. *Journal of Personality and Social Psychology*.
- Petty, R. E., & Cacioppo, J. T. (1984). The effects of involvement on response to argument quantity and quality: Central and peripheral routes to persuasion. *Journal of Personality and Social Psychology*, 46, 69-81.
- Petty, R. E., & Cacioppo, J. T. (1986). The Elaboration Likelihood Model of persuasion. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 19, pp. 123-205). New York: Academic Press.
- Petty, R. E., & Jarvis, B. G. (1996). An individual differences perspective on assessing cognitive processes. In N. Schwarz & S. Sudman (Eds.), *Answering questions: Methodology for determining cognitive and communicative processes in survey research* (pp. 221-257). San Francisco: Jossey-Bass.
- Petty, R. E., Ostrom, T. M., & Brock, T. C. (1981). Historical foundations of the cognitive response approach to attitudes and persuasion. In R. E. Petty, T. M. Ostrom, & T. C. Brock (Eds.), *Cognitive responses in persuasion* (pp. 5-29). Hillsdale, NJ: Lawrence Erlbaum.
- Petty, R. E., Priester, J. R., & Wegener, D. T. (1994). Cognitive processes in attitude change. In R. S. Wyer & T. K. Srull (Eds.), *Handbook of social cognition* (2nd ed., Vol. 2, pp. 69-142). Hillsdale, NJ: Lawrence Erlbaum.
- Schwarz, N., & Clore, G. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology*, 45, 513-523.
- Srull, T. K., & Wyer, R. S., Jr. (1986). The role of chronic and temporary goals in social information processing. In R. M. Sorrentino & E. T. Higgins (Eds.), *Motivation and cognition: Foundations of social behavior* (pp. 503-549). New York: Guilford.
- Srull, T. K., & Wyer, R. S., Jr. (1989). Person memory and judgment. *Psychological Review*, 96(1), 58-83.
- Stangor, C., & Duan, C. (1991). Effects of multiple task demands upon memory for information about social groups. *Journal of Experimental Social Psychology*, 27, 357-378.
- Uleman, J. S., Newman, L. S., & Moskowitz, G. B. (1996). People as flexible interpreters: Evidence and issues from spontaneous trait inference. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 28, pp. 211-279). San Diego, CA: Academic Press.
- Webster, D. M., & Kruglanski, A. W. (1994). Individual differences in need for cognitive closure. *Journal of Personality and Social Psychology*, 67, 1049-1062.

Received July 24, 2000

Revision accepted November 4, 2000

How One Cause Discounts or Augments Another: A Connectionist Account of Causal Competition

Frank Van Overwalle

Dirk Van Rooy

Vrije Universiteit Brussel, Belgium

The authors investigated the degree of discounting and augmentation of a target cause by an alternative cause given a varying number of observations on the alternative cause while holding its degree of covariation constant. Two experiments showed that more observations of the alternative cause resulted in greater discounting or augmentation of a target cause. This sample size effect cannot be explained by current attribution theories based on statistical notions or belief updating but can be accounted for by a connectionist framework. In addition, the authors found that the sample size effect was stronger when the information was presented in a sequential trial-by-trial format as opposed to a summarized format but found no effect of information order. Possible extensions of statistical models with confidence weights that take account of sample size were considered and simulated but none accommodated the data as well as connectionist models.

How does a perceiver learn which one among multiple factors caused an event? According to Kelley (1967, 1971), perceivers take into account not only how a potential cause covaries with the event but also how this cause competes with rival explanations. The principle of covariation prescribes that an event is attributed to a cause with which it covaries over time (Kelley, 1967). Despite the central place accorded to this principle in attribution theory, Kelley (1971) also argued that the covariation principle in itself is insufficient to explain how perceivers select between alternative causes. For instance, although both speeding and bad weather may covary with a car accident, perceivers often discredit or discount the effect of one cause in favor of the other. To account for competition of multiple possible causes that goes beyond the covariation principle, Kelley (1971) proposed two complementary principles of discounting and augmentation.

The discounting principle specifies that “the role of a given cause in producing a given effect is discounted if other plausible causes are also present” (Kelley, 1971, p. 8). One of the most common examples of discounting in social cognition is when internal attributions to the actor are discounted given evidence on the potent influence of external pressures. The opposite tendency is described in the augmentation principle. This principle specifies that “if for a given effect, both a plausible inhibitory cause and a plausible facilitatory cause are present, the role of the facilitative cause in producing the effect will be judged greater” (Kelley, 1971, p. 12). For instance, a person’s success will be more strongly attributed to internal capacities when the task was hard rather than easy (for an overview, see McClure, 1998).

Although Kelley (1971) initially reserved the terms *discounting* and *augmentation* to describe competition between alternative explanations based on prior knowledge and minimal causal information (see also Morris & Larrick, 1995; Read & Montoya, 1999), other researchers applied these terms more broadly to denote causal competition also during causal induction, that is, in conjunction with covariation information (e.g., Hansen & Hall, 1985; Kruglanski, Schwartz, Maides, & Hamel, 1978; Rosenfield & Stephan, 1977; Van Overwalle & Van Rooy, 1998; Wells & Ronis, 1982). In line with this latter approach, the present studies focus on the effect of dis-

Authors’ Note: This research was in part supported by the Belgian Foundation of Scientific Research–Flanders (F.W.O.) under Grants 8.0192.95 and G.0128.97. We are grateful to Bert Timmermans for his comments on an earlier version of this article. Address correspondence to Frank Van Overwalle, Department of Psychology, Vrije Universiteit Brussel, Pleinlaan 2, B-1050 Brussels, Belgium; e-mail: frank.vanoverwalle@vub.ac.be.

PSPB, Vol. 27 No. 12, December 2001 1613-1626

© 2001 by the Society for Personality and Social Psychology, Inc.

counting and augmentation during covariation detection and causal induction.

Specifically, the aim of the present article is to evaluate current attribution models that provide computational accounts of the principles of discounting and augmentation in relation to the principle of covariation. Some of these computational theories are based on statistical principles (Anderson & Sheu, 1995; Cheng & Holyoak, 1995; Fales & Wasserman, 1992; Försterling, 1989), others on an anchoring and adjustment analogy (Busemeyer, 1991; Hogarth & Einhorn, 1992), whereas still others are inspired on associative or connectionist principles (Read & Montoya, 1999; Van Overwalle, 1998; Van Overwalle & Van Rooy, 1998).

To test the validity of these models, we induced changes in the discounting and augmentation of a target cause by varying only the number of observations of the competing (or alternative) cause while keeping its degree of covariation constant. Research has amply documented that people make more pronounced causal judgments after receiving more covariation information (Baker, Berber, & Vallée-Tourangeau, 1989; Försterling, 1992; Shanks, 1985, 1987, 1995; Shanks, Lopez, Darby, & Dickinson, 1996; Van Overwalle & Van Rooy, 2000). The specific hypothesis tested in this article is that as the number of observations for a competing explanation goes up, a target explanation will be more discounted or augmented (Hansen & Hall, 1985; Wells & Ronis, 1982). For instance, when there is growing evidence indicating that a successful task was easy, then the person's ability will be more discounted. Conversely, the greater the evidence that the task was hard, the more the person's ability will be augmented.

Why do we test a hypothesis that seems so intuitively plausible and straightforward? The reason is that, surprisingly, all prominent attribution models in social psychology that use a statistical formulation (e.g., Anderson & Sheu, 1995; Busemeyer, 1991; Cheng & Holyoak, 1995; Fales & Wasserman, 1992; Försterling, 1989; Hogarth & Einhorn, 1992) actually contradict this hypothesis and predict no variation in competition due to sample size differences. Only the connectionist approach makes this prediction (Read & Montoya, 1999; Van Overwalle, 1998; Van Overwalle & Van Rooy, 1998). This has far-reaching implications on how we view the causal attribution process. In contrast to statistical models that describe attribution as a complex and laborious application of statistical rules, connectionist models assume that causality is often an implicit process outside awareness and that only the outcome of that process is open to consciousness (Smith & DeCoster, 1999). This latter view seems to describe much better how causes are learned and understood in the hustle of everyday social life. Before turning to this connectionist framework, we

begin by explaining briefly why existing statistical and adjustment models fail to predict variations due to sample size.

STATISTICAL MODELS

The majority of attribution models in social psychology are rooted in the idea that people are intuitive statisticians who apply a kind of statistical rule to estimate causality. One of the most popular rules was formalized in the probabilistic contrast model of Cheng and Novick (1990, 1992; Cheng & Holyoak, 1995) and "requires that people . . . estimate and compare proportions" (p. 549). As pointed out by Cheng and Novick (1992), the probabilistic contrast approach is computationally superior to similar statistical formulations that make anomalous predictions in some cases (Einhorn & Hogarth, 1986; Morris & Larrick, 1995; Schustack & Sternberg, 1981). The model accounts for discounting and augmentation by assuming that probabilities are calculated conditional on the presence or absence of the competing cause (Cheng & Holyoak, 1995). This is analogous to experimental designs, where extraneous variables are kept constant across conditions. Other valuable models that can account for discounting and augmentation in addition to covariation are Försterling's (1992) formulation of Kelley's ANOVA model that is based on the statistical ANOVA parameter of effect size (i.e., η^2), and the Bayesian model proposed by Fales and Wasserman (1992).

However, none of these statistical models can account for differences in discounting and augmentation when the degree of covariation is kept constant and only the number of observations for the alternative cause is varied. The main reason is that these models are all based on proportions between frequencies of the causes and the effects rather than on the raw frequencies themselves. Consequently, because the degree of covariation is kept constant and, in addition, because some critical cause-effect frequencies are set to zero to induce straightforward discounting or augmentation, these proportions remain unchanged and thus so do the predicted causal estimates (see Appendix A for a mathematical demonstration). Thus, although these statistical models may adequately describe attribution judgments at asymptote (i.e., given a sufficient large size), they clearly fail to account for the gradual increase or decrease in judgments while learning is still going on and asymptote has not been reached.

JUDGMENT UPDATING MODELS

Recently, a number of models have been proposed that avoid these limitations of earlier statistical models by incorporating some form of anchoring and updating

rule (the step-by-step belief-adjustment model of Hogarth & Einhorn, 1992; the serial averaging strategy of Busemeyer, 1991), which makes them sensitive to sample size. However, a serious restriction of these proposals is that the adjustment rules involve only a single cause and do not take into account the influence of alternative causes, which is necessary to predict discounting and augmentation. Therefore, we will ignore these models and immediately turn to the connectionist approach.

CONNECTIONIST MODELS

Recently, several authors moved away from statistical approaches and proposed an entirely different, connectionist framework that is inspired by basic neurological properties of the human brain (Read & Montoya, 1999; Smith, 1996; Van Overwalle, 1998; see also Allan, 1993; McClelland & Rumelhart, 1988; Shanks, 1995). We claim that unlike earlier models, adaptive connectionist models with an error-correcting learning algorithm predict differences in discounting and augmentation given changes in sample size of the competing cause.

To explain the properties of adaptive networks, we focus on one of the simplest architectures: the "feedforward" model. In the feedforward model, causal estimates are represented by the weight of the connections between input nodes representing causes and an output node representing the effect. Activation in the network typically runs from causes to effect (hence the name feedforward). Whenever a cause is present, its input node is activated and this activation is then automatically propagated to the output node in proportion to the weight of the connection (i.e., the causal strength at that moment). All input activations received at the output node are linearly summed to determine the output activation, which reflects the effect predicted by the network on the basis of the causal input given.

A key feature of adaptive connectionist networks is that the connection weights are adapted in response to information on new co-occurrences between causes and effect, using a learning algorithm. The learning algorithm we focus on here is the delta algorithm (McClelland & Rumelhart, 1988). The goal of this algorithm is to reduce the error between the mental representation of the effect as predicted by (the output node of) the network and the actual effect. This error is reduced by adjusting the weights of the cause-effect connections. When the occurrence of the effect is underestimated, the weights are adjusted upward; when the occurrence of the effect is overestimated, adjustments are made downward. Thus, as Dennett (1998) noted, this learning algorithm mimics important aspects of reasoning and intelligence, because "intelligence is . . . for improving the fidelity of your expectations about what is going to happen next" (p. 184).

The delta learning algorithm has been applied in many investigations on human categorization and causality (for reviews, see Allan, 1993; Shanks, 1995; Van Overwalle & Van Rooy, 1998) and is formally identical to Rescorla and Wagner's (1972) associative model. Although it does not tally frequencies or compute statistical probabilities, it forces the weights to converge to the probabilistic norm (Cheng & Novick, 1990) after a sufficient number of observations (Chapman & Robbins, 1990; Van Overwalle, 1996). Thus, the feedforward network respects the statistical principle of covariation.

Most important, the delta algorithm can easily account for our manipulation, in which the sample size of the competing cause is increased. The reason is that the network's output is computed on the summed activation of all causes present. Thus, adjustments are driven by competition for predictive strength. To illustrate, Figure 1 depicts the weights of two tennis partners, Theo (the target explanation T) and Xavier (the alternative explanation X), who together win a doubles game.

The top panel illustrates discounting. Consider a network that is provided with information that Xavier won either five single games (large size) or one single game (small size). As can be seen, after five games, Xavier's weight is much stronger than after one game. Next, when the network learns the new information that Xavier and Theo win their double game, competition between the two explanations arises. If Xavier takes already a great part of the available weight, there is less weight left for Theo to gain. This results in stronger discounting of Theo's weight than if Xavier takes a smaller part of the available weight.

The bottom panel illustrates the reverse mechanism of augmentation, where the network receives the same information, except that Xavier first lost five or one single game so that he acquired inhibitory weight. As can be seen, to reduce the inhibitory impact of Xavier in winning the double game, augmentation of Theo's contribution is much stronger when Xavier lost five times rather than only once.

DESIGN AND HYPOTHESES

The design of the experiments reported here was modeled after the example illustrated above. In a first experiment, we induced competition between two causal explanations (e.g., the contribution of Theo versus Xavier in winning a game) and manipulated the strength of the competing cause by varying how often this cause was presented alone: either one time (small size) or five times (large size). Thus, for instance, participants learned that Theo's tennis partner, Xavier, won five single games in the large size condition as opposed to only one game in the small size condition. However, this manipulation confounds size with consistency infor-

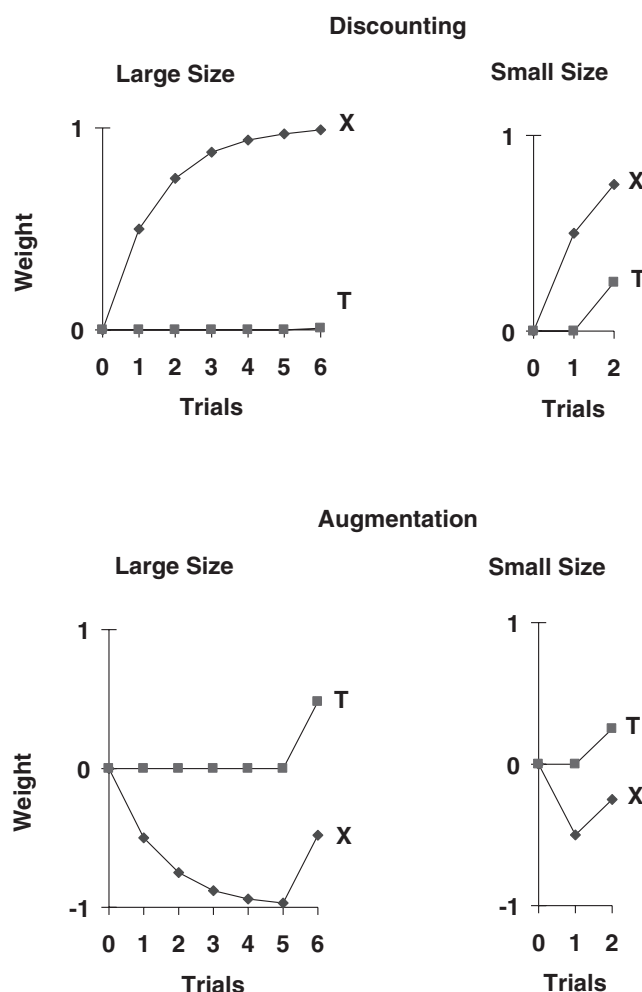


Figure 1 A feedforward simulation of competition between a target cause T (e.g., Theo) and a competing cause X (e.g., Xavier), given positive (+) or negative (–) outcomes.

NOTE: For discounting, there were five (large size) or one (small size) X+ trials, indicating that the competing cause (Xavier) won his single games five times or one time, respectively. This was followed by one compound TX+ trial, indicating that Theo and Xavier together won their doubles game. For augmentation, there were five (large size) or one (small size) X– trials, indicating that the competing cause (Xavier) lost one or five single game(s), followed by the same compound TX+ trial, indicating that Theo and Xavier together won their doubles game. The figure illustrates forward competition. In backward competition, the compound TX+ trial, indicating that both players won their doubles game, comes first.

mation (or covariation with time, see Kelley, 1967). For instance, the information that Xavier won five times also increases the consistency of Xavier's performance on different occasions. To assess whether the size effect is not solely due to increased consistency, in a second experiment, increasing size was manipulated by proving information that the target cause competed against many alternative actors or stimuli (each on a single occasion) rather than against one alternative on several occasions

(as in the first experiment). This should rule out consistency information as an alternative explanation for the size effect. Consistent with the connectionist framework, we predict for the two experiments that a large as opposed to small size will result in a greater facilitatory or inhibitory strength of the competing cause, leading to, respectively, more discounting and more augmentation of the target cause.

We also explored two additional questions. One question was whether the order in which the competing and target causes are presented would influence the predicted size effects. The competing cause was either presented before the target (forward order) or after the target (backward order). In associative research with animals (where discounting and augmentation are analogous to blocking and superconditioning, respectively), the typical finding is that only forward competition effects occur (Kamin, 1968; Rescorla, 1969). However, more recent studies with humans indicate that forward and backward competition are generally equally effective (Shanks, 1985; Van Hamme, 1994; Williams & Docking, 1995; Williams, Sagness, & McPhee, 1994). Therefore, we expect that information order will have little effect on our size manipulation.

Another exploratory question was whether our predictions would apply on different presentation formats. In associative research, covariation information is typically presented in a sequential trial-by-trial format in which each occurrence of one or more causes is described in short sentences at successive trials. This format seems to reflect people's incidental learning during everyday life and also is consistent with a connectionist approach where input information is received and processed on-line (i.e., after each piece of information is received). Conversely, in social research, covariation information is typically presented in a prepacked format where the occurrences of causes are summarized in one single sentence. This format seems to capture people's verbal interactions with one another but is less suitable for sequential processing as required by connectionist models. Consequently, we anticipated that the size effect would be stronger in a sequential format than in a summary format.

EXPERIMENT 1: COMPETING AGAINST ONE ALTERNATIVE

Method

PARTICIPANTS

Participants were 106 male and female students from the Vrije Universiteit Brussel who participated for a partial course requirement. They were tested in groups of one to five. Approximately half of the participants

received the sequential format ($n = 51$), whereas the other half received the summary format ($n = 55$).

MATERIAL

The overall design of the experiment involved one between-subjects Format factor (sequential or summarized) and three within-subject factors, including Order (forward or backward), Type (discounting or augmentation), and Size (small or large). We created two stories for each combination of the three within-subjects factors, resulting in 16 stories. The stories were counterbalanced across participants between Size and Order. The target and competing causes involved either actors or stimuli (objects or persons). To make sure that the actors and stimuli would be seen as causally related to the effect (which is necessary for discounting and augmentation to occur, see McClure, 1998), the stories involving actors used action verbs (which tend to imply the actor as the cause) and the stories involving stimuli used state verbs (which tend to imply the stimulus as the cause; see Rudolph & Försterling, 1997).

Size. Each story consisted of five trials (large size) or one trial (small size) in which a competing cause was present alone. In addition and regardless of size, there were another five compound trials in which both the target and competing cause were present together.

Order. In the forward order, the trials describing the sole presence of the competing cause were given first, whereas in the backward order, these competing trials were given last.

Type. To induce discounting, both target and competing trials were followed by the same outcome. In contrast, to induce augmentation, the outcome of the competing cause alone was opposite to the focal outcome when also the target cause was present. To make sure that the participants would encode this opposite information correctly, the semantic negation of an outcome was always indicated in capitals.

The manipulation of these three within-subjects factors is illustrated in the next example with “An” as the discounted target actor and “Elena” as the discounting competing actor. The example illustrates forward order (i.e., competing cause first), with the large size phrasing given between straight brackets. This gives for the summary presentation format:

Elena passed the first [five] selection round[s] in single scull, and An and Elena passed the five selection rounds in double scull.

The sequential format of this example is illustrated with 10 consecutive trials between slashes. The trials between parentheses were omitted for the small size:

Elena passed the first [/ second / third / fourth / fifth] selection round in single scull;

An and Elena passed the first / second / third / fourth / fifth selection round in double scull.

Similarly, the augmentation manipulation of a stimulus is illustrated below in a backward order (i.e., competing cause last) given a summary format, with “mackerel” as the augmented target stimulus and “salmon” as the augmenting competing stimulus:

Ella and four other women felt sick after eating mackerel and salmon but five [one] other women [woman] did NOT feel sick after eating salmon.

PROCEDURE

Participants were seated in front of an IBM-compatible PC and the experiment was monitored by MEL software. Instructions appeared on the screen and the use of the rating scale was practiced. Participants read 16 stories, which appeared in a different random order for each subject. In the summary format, each story was presented during one single trial. In the sequential format, each story was presented during several consecutive trials. The computer randomized for each subject the order in which specific actors or stimuli appeared, with the provision that forward and backward order was not affected.

After reading each story, participants had to rate the causal influence of the target factor and the competing factor: *something special about [actor or stimulus]*. They rated the causal influence of each factor on the outcome using an 11-point rating scale ranging from 0 (*absolutely no influence*) to 100 (*very strong influence*), with midpoint 50 (*partial influence*). This question phrasing is standard in most attribution research on the influence of covariation information (e.g., Cheng & Novick, 1990; Försterling, 1989). For example, in the story with An and Elena as actors, participants rated the causal influence of *something special about An* and *something special about Elena*. Similarly, in the story with mackerel and salmon as stimuli, participants rated the causal influence of *something special about mackerel* and *something special about salmon*.

Results

Because we made predictions irrespective of whether the causes involved an actor or stimulus, the ratings were collapsed across actor and stimulus. This was justified by a multivariate analysis of variance (MANOVA) with Measure (actor or stimulus), Order (forward or backward), Type (discounting or augmentation), and Size (small or large), which revealed that Measure did not interact with Size or with any higher order interaction including Size. However, because we found the expected difference between presentation formats, as the triple interaction

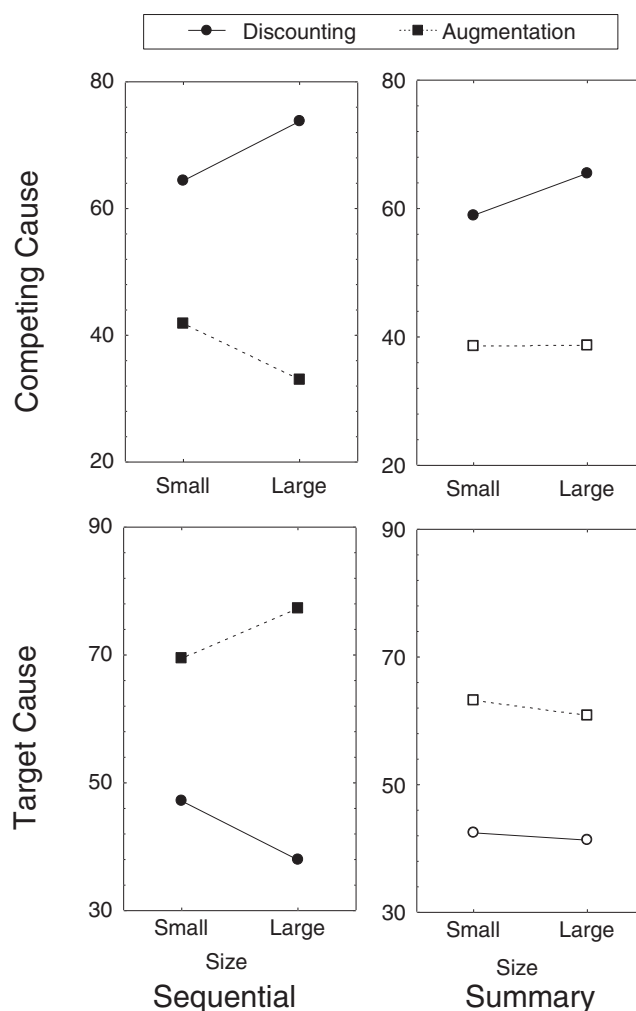


Figure 2 Experiment 1: Causal ratings in function of size, type, and format.

NOTE: Black data points denote a significant size effect ($p < .01$).

between Format, Type, and Size was significant for the target and competing ratings, $F(1, 104) = 6.58$ – 17.44 , $p < .05$, we will report these results separately.

SEQUENTIAL FORMAT

Figure 2 (left panel) depicts the average ratings for the target and competing causes. We begin with the competing cause. Consistent with our sample size hypothesis, we found a significant interaction between Type and Size, $F(1, 50) = 30.31$, $p < .0001$. Simple effect tests revealed that in the discounting condition, the (facilitatory) competing cause acquired more strength when size was large rather than small, $F(1, 50) = 22.03$, $p < .0001$. The predicted reversed pattern was found in the augmentation condition, where the (inhibitory) competing cause acquired less strength when size was large rather than small, $F(1, 50) = 11.74$, $p < .01$.

Our critical hypothesis concerned the target cause, which should be more discounted and more augmented given a larger sample size. This prediction also was supported because the critical interaction between Type and Size was significant, $F(1, 50) = 29.76$, $p < .0001$. Simple effect tests revealed that when size was large as opposed to small, the target cause was more discounted, $F(1, 50) = 13.66$, $p < .001$, and more augmented, $F(1, 50) = 11.57$, $p < .01$.¹ In addition, our prediction that order would have little effect on our size manipulation was confirmed, because the interaction between Order and Size was not significant, $F < 1$.

SUMMARY FORMAT

Figure 2 (right panel) depicts the average ratings for the target and competing ratings. For the competing rating, the expected interaction between Type and Size approached significance, $F(1, 54) = 3.99$, $p = .051$. Simple effect tests revealed that although the size manipulation led to the predicted increase of the (facilitatory) competing cause for discounting, $F(1, 54) = 13.19$, $p < .001$, it failed to show any change for augmentation, $F < 1$, *ns*. More important, for the target cause, neither the expected interaction nor any of the simple effect tests reached significance, F s < 1 , *ns*. Thus, contrary to our hypothesis, size did not increase discounting or augmentation of the target cause. Consistent with our prediction, however, the interaction between Order and Size again failed to reach significance for the target cause, $F < 1$.

Discussion

The most important finding of this experiment is that, at least in the sequential format, the sample size of the competing cause substantially influenced the perceived causality of the target cause. As more evidence on the competing cause was accumulated, the target cause was more discounted or augmented. Overall, these results lend some support to a connectionist account of causal competition. However, the size effect did not appear in the summary format. This is consistent with the connectionist conception that information uptake and encoding is an incremental process that works best on the basis of trial-by-trial information.

In addition, we found that the order in which the competing cause was presented (before or after the target causes) did not influence this size effect. This suggests that most often both forward and backward competition are effective in social explanation, consistent with associative research on human learning (e.g., Shanks, 1985; Van Hamme, 1994; Williams & Docking, 1995; Williams et al., 1994).

EXPERIMENT 2: COMPETING AGAINST MANY ALTERNATIVES

A limitation of the first experiment, mentioned earlier, is that by increasing the number of observations of the competing cause, we also increased its perceived consistency (or covariation with time). Although Försterling (1989) actually documented that high consistency decreased attributions to the actor or the stimulus (whereas our results show the opposite pattern), we wanted to rule out consistency as an explanation for the results in the first experiment. Therefore, in a second experiment, the target did not compete against one single actor or stimulus but rather against one (small size) or five different actors or stimuli (large size). Because individual actors and stimuli appear only once in this design, consistency is kept constant and thus ruled out as a potential alternative explanation. Moreover, by competing against many alternatives rather than a single alternative repeated several times, the predicted competition effects might be stronger in this experiment.

We make essentially the same predictions as in the previous experiment. To capture the idea of competition by several alternative actors or stimuli, we borrowed Weiner's (1985) terminology and referred to them as the *external* or *general* context, respectively (see also Van Overwalle, 1997). Thus, an explanation in terms of a target actor must compete against an alternative explanation in terms of an external context. For instance, we expect that people will attribute high sales figures to a (target) clerk's capacities when there is only one other clerk with equally high sales. Conversely, when there are five other clerks with high sales, attributions to alternative causes such as external conditions (e.g., convenient location of the shop) are more likely. We make the same prediction for target explanations in terms of a specific stimulus (e.g., salmon) as opposed to competing explanations in terms of a general context (e.g., a restaurant's food quality).

Method

PARTICIPANTS

Participants were 115 male and female students from the Vrije Universiteit Brussel who participated for a partial course requirement. They were tested in groups of one to five. Approximately half of the participants received the sequential presentation format ($n = 64$) and the other half received the summary format ($n = 51$).

MATERIAL AND PROCEDURE

The overall design and procedure was similar to the previous experiment, with the following modifications. We developed 16 novel stories in which the competing causes involved different comparison actors or stimuli rather than a single competing actor or stimulus. To

make things easier, the focal event (in which the target is present, together with the external or general context) was presented only once.

The next example illustrates the discounting manipulation of Annie as target actor in a forward order (with sentence parts describing the large size between straight brackets). For the summary and sequential format respectively, this gives the following:

One [Five] other salesgirl[s] and Annie attained high sales figures for perfumes

Sandra [/ Daniela / Katharina / Sabrina / Katrien] /
Annie attained high sales figures for perfumes

Likewise, the augmentation manipulation of Danny as target stimulus is illustrated below in a backward order:

Jos respected his peer student Danny but he did NOT respect one [five] other peer student[s].

Jos respected his peer student, Danny;

Jos did NOT respect his peer student, Luis [/ Gaston
/ Ruben / Ricardo / David]

To assess the causal influence of the target causes, similar to Experiment 1, participants rated the causal influence of *something special about [actor or stimulus]*. To assess the causal influence of the external or general context, they rated the causal influence of *something external (outside [actor])* or *something general (plays a role with many [stimulus category])*.

Results

The same preliminary MANOVA was performed as in the previous experiment and again indicated that the ratings of the actor and stimulus could be safely collapsed.² Although there were no significant interactions between Format and Size, $F_s(1, 113) < 2.25$, ns , to remain consistent with the previous experiment, we will analyze the presentation formats separately.

SEQUENTIAL FORMAT

Figure 3 (left panel) depicts the average context and target ratings. As can be seen, the results depict the same predicted pattern as in the sequential format of the previous experiment. This was confirmed by the significant interaction between Type and Size on the target and context ratings, $F(1, 63) = 28.66\text{--}43.06$, $p < .0001$. Simple effect tests showed that given a larger size, the competing context was rated stronger in the discounting condition, $F(1, 63) = 10.35$, $p < .0001$, and weaker in the augmentation condition, $F(1, 63) = 26.33$, $p < .0001$. More important, given a larger size, the target cause was more discounted, $F(1, 63) = 20.93$, $p < .01$, and more augmented, $F(1, 63) = 30.32$, $p < .0001$. In addition, as predicted,

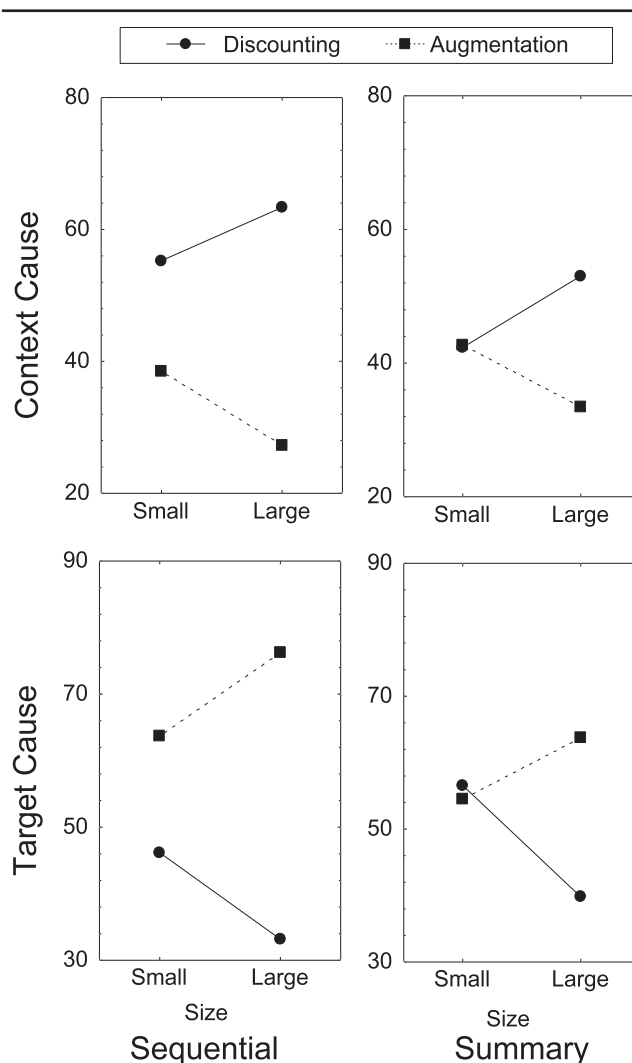


Figure 3 Experiment 2: Causal ratings in function of size, type, and format.

NOTE: Black data points denote a significant size effect ($p < .01$).

there was no significant interaction between Order and Size on the target, $F < 1$.

SUMMARY FORMAT

Figure 3 (right panel) depicts the average target and context ratings. As we had hoped for, the hypothesized size effect now also appeared in the summary presentation format, although competition was generally weaker than in the sequential format (as can be seen, for instance, from the ratings of the discounting and augmentation conditions that did not differ given a small size). This was confirmed by a significant Type \times Size interaction on the target and context ratings, $F(1, 50) = 22.60$ – 27.54 , $p < .0001$. Simple effect tests showed that given a larger size, the context was rated higher given dis-

counting, $F(1, 50) = 12.64$, $p < .001$, and lower given augmentation, $F(1, 50) = 11.15$, $p < .01$. More important, given a larger size, the target was more discounted, $F(1, 50) = 32.06$, $p < .0001$, and more augmented, $F(1, 50) = 9.91$, $p < .01$. As predicted, there was again no significant interaction between Order and Size on the target rating, $F < 1$.

Discussion

The findings of this experiment confirmed that as the number of the comparison cases increased, the target cause was more discounted or augmented. Because the frequency of each individual comparison cause was kept constant, consistency is ruled out as an explanation for these results. Similar to the previous experiment, there was no effect of order.

Perhaps the more interesting finding of this experiment is that this sample size effect was now also observed in the summary format. The most plausible explanation for the more robust size effect in the summary format is that the target competed against five different actors or stimuli rather than only one repeated over five trials as in the first experiment. Alternatively, it is also possible that the use of broad contextual categories to measure the comparison causes may have allowed for a clearer separation between target and competing causes, leading to increased competition and a more reliable effect of sample size.

MODEL SIMULATIONS

The present experiments demonstrate that there was more discounting and augmentation after increasing the number of competing cases. To evaluate how closely a connectionist formulation can predict our data, we ran a feedforward simulation and correlated the simulated results with the observed attribution ratings. To incorporate the finding that both forward and backward competition are effective, we implemented Van Hamme and Wasserman's (1994) suggestion that absent causes are encoded with a negative activation (for more details on the model specification, see Appendix B).

As a way of comparison, we did the same for the most prominent statistical models in social psychology: the probabilistic contrast model (Cheng & Holyoak, 1995) and the ANOVA model (Försterling, 1992). However, it would be of little interest to reiterate the fact that these models fail to be sensitive to sample size. Therefore, we extended these existing models with additional parameters that take into account sample size to evaluate whether these extensions would be sufficient to account for the observed size effect.

Specifically, we followed the suggestion by Cheng and Holyoak (1995) that "confidence in the assessment of a contrast is presumed to increase monotonically with the

number of cases observed" (p. 273). That is, we weighted the major theoretical variables of the models (conditional probabilities: Cheng & Novick, 1990; sum of squares: Försterling, 1989) in proportion to the number of observations available. We allowed two such different confidence weights for frequencies that involved the presence and absence of the target cause (ω_t and ω_x , respectively). This procedure is identical to the one used recently by Lober and Shanks (2000, p. 207) and parallels that of the connectionist models, where we also allowed two different learning rate parameters for target and comparison factors (ϵ_t and ϵ_x , respectively). More technical details on the model specifications are given in Appendix B.

Method

The models were run using exactly the same order of trials and blocks as in the experiments. The connectionist model was updated after each trial. We sought the best overall fitting parameters of each model by searching for the maximum correlation between simulated and observed data across all conditions, given all admissible parameter values (see Gluck & Bower, 1988; Nosofsky, Kruschke, & McKinley, 1992). We did not attempt to find common best-fit parameters for all data sets because the presentation formats and stories differed too much between conditions and experiments. The best-fit parameter values were generally quite robust, and small deviations of 0.10 in the values decreased the maximum correlations only minimally.

To evaluate the performance of the models with respect to stronger discounting and augmentation given an increased sample size, we then computed separate correlations between observed and simulated data within each discounting and augmentation condition (using the same overall best-fitting model parameters). Next, we averaged these correlations across target and competing causes as well as across forward and backward order. Table 1 depicts the results. An average correlation of +1 reflects the predicted size effect in all cases, a correlation of 0 indicates that the sample size effect is absent in all cases, and a correlation of -1 indicates that the predicted sample size effect is reversed in all cases.

Results and Discussion

As can be seen in Table 1, the fit for discounting and augmentation was generally poor for the statistical models because they showed many zero correlations indicating that they failed to predict the observed sample size effect. To illustrate, the probabilistic contrast model predicts that, for the present augmentation design, the probability of a target cause is always 1 and for a competing cause always 0 (see Equation A4), so that an increasing confidence weight has no effect at all on these esti-

TABLE 1: Fits of the Models to the Data

Model	Experiment 1		Experiment 2	
	Seq	Sum	Seq	Sum
Weighted ANOVA				
Discounting	.000	.000	.000	.000
Augmentation	.000	.500	.000	.000
	$\omega_t > .00$	$\omega_t > .06$	$\omega_t > .00$	$\omega_t > .00$
	$\omega_x = .00$	$\omega_x = .70$	$\omega_x = .00$	$\omega_x = .00$
Weighted probabilistic				
Discounting	1.000	1.000	1.000	1.000
Augmentation	.000	.000	.000	.000
	$\omega_t = .98$	$\omega_t = .20$	$\omega_t = 1.00$	$\omega_t = .60$
	$\omega_x = .72$	$\omega_x = .82$	$\omega_x = .60$	$\omega_x = .10$
Feedforward connectionist				
Discounting	1.000	1.000	1.000	1.000
Augmentation	1.000	-.500	1.000	1.000
	$\epsilon_t = .95$	$\epsilon_t = .81$	$\epsilon_t = .42$	$\epsilon_t = .68$
	$\epsilon_x = .50$	$\epsilon_x = .72$	$\epsilon_x = .20$	$\epsilon_x = .20$

NOTE: Cell entries are correlations. Seq = sequential format, sum = summarized format, ω_t = weight for target factors, ω_x = weight for competing factor, ϵ_t = learning rate for target factors, and ϵ_x = learning rate for competing factors. The activation level for absent causes in the feedforward model was set to $\alpha_n = -.5$.

mates, resulting in correlations of 0. In contrast, the connectionist model was capable of simulating most observed size effects. The correlations were all +1 (except in the summary format of Experiment 1 where some conditions showed a nonsignificant opposite size effect). To emphasize that the obtained results are not due to the feedforward architecture used here but rather to the delta learning mechanism, we also conducted simulations with a recurrent network (Read & Montoya, 1999).³ As would be expected, the results for the discounting and augmentation conditions were exactly the same as for the feedforward network.

We explored other solutions to improve the performance of the statistical models, but they all failed. Measures such as the *F* value or the *F* test in Försterling's ANOVA model that incorporate the number of observations (i.e., by the degrees of freedom) fail because in our discounting and augmentation designs, the error variance was always zero (see Appendix A). Other possibilities are to add a sort of anchoring and adjustment notion to statistical models. This is exactly what Hogarth and Einhorn (1992) proposed in their step-by-step belief-adjustment model and Bussemeyer (1991) in his serial averaging strategy. However, the proposed adjustment rules are mathematically identical to a simplified version of the delta algorithm, one that deals with only one cause at a time (see Wasserman, Kao, Van Hamme, Katagiri, & Young, 1996). Hence, these extensions cannot account for competition between alternative explanations. Interestingly, if one would extend these adjustment rules to

include more causes, they would become identical to the delta algorithm of connectionist models.

GENERAL DISCUSSION

The general finding of the present experiments was that increasing the sample size of a competing cause reliably increased discounting and augmentation of a target cause. This provides support for adaptive connectionist models with an error-correcting delta algorithm because they were the only models that made this prediction. These connectionist models suggest that causal learning on the basis of relevant information is an on-line incremental process in which estimates of causality are continuously updated on-line.

Our experiments have, however, a shortcoming in that they did not include a baseline condition in which competition by an alternative explanation was absent. Hence, we cannot be sure in the small size conditions whether differences in causal estimates were due to discounting or augmentation, or both. However, it is difficult to escape the interpretation that discounting and augmentation took place in the large size conditions where the difference between estimates became even greater.

The experiments further confirmed that forward and backward competition are equally effective, in line with earlier associative learning research (Shanks, 1985; Van Hamme, 1994; Williams & Docking, 1995; Williams et al., 1994). In addition, they demonstrated that information provided in a summary sentence has much weaker size and competitive effects. A possible explanation is that summary information is often encoded more shallowly than sequential information and so leads to weaker effects. Alternatively, according to a connectionist view that allows only for trial-by-trial input, summary information must be converted for sequential input before it can be appropriately encoded. Perhaps this conversion is far from perfect and so results in less sample size and competition effects. We return to this issue later in the discussion.

SAMPLE SIZE ADJUSTMENTS IN STATISTICAL MODELS?

Because statistical models do not incorporate a learning mechanism that allows adjusting causal estimates when novel information comes in (Anderson & Sheu, 1995; Cheng & Holyoak, 1995; Fales & Wasserman, 1992; Försterling, 1989), sample size effects are particularly problematic for these models. In the simulations, we demonstrated that many attempts to extend these models with supplementary theoretical notions (e.g., adding confidence weights or an anchoring-and-adjustment mechanism) failed to salvage these models from this limitation.

One might argue that sample size does not so much affect causal judgments but rather the confidence with which these judgments are given. In line with this reasoning, one might suggest that facilitatory information increases confidence over trials, whereas inhibitory information decreases confidence. However, this argument simply shifts the burden of proof from causal judgments to confidence judgments. Moreover, it does not explain why the reverse confidence effects should occur for the discounted and augmented causes. Cheng (1997) provided an account of why perceivers become increasingly uncertain of the causal status of the discounted cause, but her reasoning cannot explain why they should become more certain in the case of augmentation. We see no way to solve these logical quandaries of the confidence notion within the boundaries of the existing statistical theories. If our participants had been given the possibility to express their subjective confidence in the causal ratings, perhaps this might have solved some questions, but it would have left unanswered most of the reservations that we raised.

Finally, it might be argued that the scope of statistical models is explaining judgments only at the asymptotic level (What is computed?) rather than at a pre-asymptotic level that keeps track of the development of judgment (How is it computed?). However, the fact that the connectionist approach accommodates both levels makes that argument questionable. Therefore, we tend to conclude that connectionist models may complement or even replace these earlier models by providing a low-level description of the attribution process.

IMPLICATIONS FOR SOCIAL EXPLANATION

What are the implications of a connectionist perspective in social explanation? We believe that the most important contribution is theoretical parsimony. The delta error-correcting algorithm provides a unifying mechanism that incorporates not only Kelley's (1967, 1971) principles of covariation and competition but also the statistical principle of sample size. Moreover, a connectionist model not only explains how people incrementally adjust their causal estimates (cf. delta-algorithm) but can also address how people make a quick causal judgment in situations where they do not receive a stream of new causal information. In such cases, contextual cues about the appropriate place, time, and content spread their activation automatically to a host of potential causal candidates in memory. The cause that is most strongly connected with the outcome (because of strong weights built up during prior learning) will then be selected as the most likely cause. Thus, a connectionist model is relevant for both causal learning from covariation information as well as for snapshot judgments that are based on memorized experiences

from the past. For instance, after learning that Xavier won five tennis tournaments, there is a strong connection in memory between Xavier and success in tennis. Later judgment can invoke this memory trace to make predictions on Xavier's future successes and that of his opponents.

A connectionist approach also has applied value. First, it assumes that activation spreading and weight adjustment occur at a relatively implicit and autonomous level and that only the outcome of these processes is available for conscious introspection (Smith & DeCoster, 1999). Hence, people do not need to learn complex and elaborate rules for causal induction like statistical approaches assume. Rather, they infer causality quite often even when they are not fully aware of it, much like young infants and children learn how to understand and predict their environment and how to control it. A second important implication is that biases such as sample size should not be viewed as errors of the mind but rather as a window on how the mind works. Whereas statistical models often see biases as errors against logical norms, connectionist models actually see them as illustrations of human information processing. For instance, a connectionist approach can explain other well-known biases and sample size effects such as illusory correlation, group homogeneity, and group polarization and the impact of increasing or alternative information on impression formation and attitude change (Van Overwalle, Labiouse, & French, 2000).

One question that remains largely unsolved, however, is the relationship between simple connectionist adjustments and higher level "cognitive" reasoning and language use that is typical of social interaction. In this respect, it is quite interesting that the data from the summary format generally paralleled those of the sequential format (although in a weaker form and not always for all effects). Shanks (1991; Lober & Shanks, 2000) argued that people's experience with causal acquisition in real situations gave rise to causal metabeliefs, which may be readily applied in the interpretation of verbal summary sentences. Because these metabeliefs developed from observed situations, they may show similar but weakened effects of covariation, competition, and sample size at the surface. Alternatively, Van Overwalle and Van Rooy (1998) speculated that verbal summary information is unconsciously decomposed in the form of dummy exemplars or mental models (cf. Johnson-Laird, 1983), which are then sequentially analyzed by a connectionist processor. Occasional differences between trial-by-trial and summary formats are explained by an imperfect conversion of summary information, a type of error that has been amply documented in research on deductive reasoning. Which of these two proposed mechanisms actu-

ally underlies causal induction from verbal summaries is a question for future research.

APPENDIX A

This appendix demonstrates that existing statistical models are incapable of showing a size effect of the competing cause X on the target cause T. The information given to the participants can be represented by a standard contingency table with four relevant frequencies denoted by a - d : when the target is present and the outcome occurs (a) or not (b) and when the target is absent and the outcome occurs (c) or not (d).

A1. Probabilistic Contrast Formulation

According to the probabilistic contrast formulation, when two or more potential causes are present, the genuine causal influence of a target T is measured *conditional* on the presence or the absence of the other competing factor X (Cheng & Holyoak, 1995). Hence, the causal influence of a target factor T on the outcome O is measured by ΔP_T , or the contrast between two conditional proportions:

$$\Delta P_T = P(O|T.X) - P(O|\sim T.X), \quad (A1)$$

where O represents the outcome, T the target cause, X the competing cause, and a tilde denotes their absence. Because the competing factor X is always present in our designs, the ΔP formulation can be reduced to the following:

$$\Delta P_T = P(O|T) - P(O|\sim T) = [a/(a+b)] - [c/(c+d)]. \quad (A2)$$

In the designs of all experiments, $b = 0$, whereas for discounting, $d = 0$ and for augmentation, $c = 0$. Hence, in the case of discounting, we can simplify the ΔP formulation as follows:

$$\Delta P_T = [a/a] - [c/c] = 0, \quad (A3)$$

and in the case of augmentation:

$$\Delta P_T = [a/a] - [0/(0+d)] = 1. \quad (A4)$$

This indicates that the discounted target will always attain zero causal strength, whereas the augmented target will always receive full causal strength regardless of the frequencies of the competing X. It can be easily shown that in more recent versions of probabilistic theory (Cheng, 1997), the same predictions are made for augmentation, whereas discounting is undetermined (i.e., division by zero).

A2. ANOVA Formulation

The ANOVA formulation defines causal strength as an analog to the effect size η^2 of a standard ANOVA (Försterling, 1992), which is given by the following:

$$\eta^2 = SS_{\text{between}} / SS_{\text{total}} = 1 - [SS_{\text{within}} / SS_{\text{total}}]. \quad (A5)$$

If the presence of the outcome is indicated by 1, and its absence by 0, then SS_{within} can be expressed as follows:

$$SS_{\text{within}} = a[a/(a+b) - 1]^2 + b[a/(a+b)]^2 + c[c/(c+d) - 1]^2 + d[c/(c+d)]^2. \quad (\text{A6})$$

For discounting where $b = d = 0$, this formulation can be simplified as follows:

$$SS_{\text{within}} = a[a/a - 1]^2 + 0 + c[c/c - 1]^2 + 0 = 0. \quad (\text{A7})$$

For augmentation where $b = c = 0$, this reduces to the following:

$$SS_{\text{within}} = a[a/a - 1]^2 + 0 + 0 + 0 = 0, \quad (\text{A8})$$

showing that in both cases, SS_{within} equals zero. In addition, SS_{total} can be written as follows:

$$SS_{\text{total}} = (a+c) [(a+c)/(a+b+c+d) - 1]^2 + (b+d) [(a+c)/(a+b+c+d)]^2. \quad (\text{A9})$$

It can be easily shown that for discounting, $SS_{\text{total}} = 0$, so that η^2 in Equation A5 is undetermined, whereas for augmentation, $SS_{\text{total}} \neq 0$, so that $\eta^2 = 1$, irrespective of the frequencies of the competing X.

A3. Bayesian Formulation

The argument that Bayesian reasoning is not influenced by sample size is based on the assumption—shared by most models—that perceivers evaluate to what extent a causal hypothesis can explain all or most of the available data. However, this assumption can be relaxed by assuming that perceivers simultaneously entertain a number of causal hypotheses with mutually exclusive degrees of causal probability—for instance, 11 hypotheses reflecting a probability of 0, 0.1, 0.2, 0.3, and so forth to 1. In this extended case, the model is sensitive to sample size (Fales & Wasserman, 1992). Because this assumption is psychologically very implausible as it puts a great burden on human information processing, and because it is not required by any of the other models discussed, this extension is not considered here.

APPENDIX B

This appendix discusses how the models were specified for the simulations and how the statistical models were extended to account for sample size. Note that for simulating Experiment 2, the comparison causes were represented by a single context factor that was assumed to be always present in the same way as the competing causes in Experiment 1. Hence, the causal structure was essentially identical for all experiments.

B1. Weighted Probabilistic Model

To make the probabilistic model sensitive to the number of observations, we weighted each of the conditional probabilities P (see Appendix A1) with a freely estimated proportion (between 0% and 100%) of the frequencies involved. That is, if ω_t denotes a proportion of the frequencies when cause T is present and if ω_x denotes a proportion of the frequencies when cause T is absent, then the strength of a target cause may be formalized as follows:

$$\Delta P_T = \omega_t P(O|T) - \omega_x P(O|\sim T), \quad (\text{B1})$$

and, likewise, the strength of a competing cause is as follows:

$$\Delta P_X = \omega_x P(O|\sim T). \quad (\text{B2})$$

B2. Weighted ANOVA Model

To make Försterling's (1989) model sensitive to the number of observations, we weighted η^2 in Appendix A2 with a freely estimated proportion of the frequencies involved, that is, with the same ω_t and ω_x as defined above. Thus, for the target cause, this becomes as follows:

$$\eta_T^2 = [\omega_t SS_{\text{between}}] / [(\omega_t + \omega_x) SS_{\text{total}}] \quad (\text{B3})$$

and, similarly, for the competing cause as follows:

$$\eta_X^2 = [\omega_x SS_{\text{within}}] / [(\omega_t + \omega_x) SS_{\text{total}}]. \quad (\text{B4})$$

B3. Feedforward Model

The feedforward architecture consists of two input nodes representing the target cause and the competing cause and an output node representing the outcome. The input nodes are connected to the output node via weighted, unidirectional links. When a cause is present at a trial, its input node is activated to the default level 1; when a cause is missing at a trial, its input node is activated to a negative value α_n , which can be freely estimated between 0 and -1 (Van Hamme & Wasserman, 1994). However, to keep the number of free parameters in all models equal, this parameter was arbitrarily set at an intermediate value of $\alpha_n = -.5$. The positive activation of all input nodes is spread automatically to the output node in proportion to the weights of the links and then linearly summed to represent the output activation.

After each trial, the weights of the links are incrementally adjusted by reducing the error between the output activation (the outcome predicted by the network) and a teaching activation (the actual outcome), which is 1 when the focal outcome is present and -1 when absent (this coding assumes that an absent outcome was interpreted as opposite to the focal outcome, which is plausible for most of our stories). This adjustment is mathematically expressed by the following delta algorithm (McClelland & Rumelhart, 1988, p. 87):

$$\Delta w = \varepsilon (a_t - a_o) a_i, \quad (\text{B5})$$

where ε is the learning rate (freely estimated between 0 and 1) and a_t , a_o , and a_i denote, respectively, the teaching, output, and input activations. We assumed that there were separate learning rates for target factors and competing factors, denoted respectively by ε_t and ε_x .

NOTES

1. To verify that these sample size effects were not due to the within-subject nature of the experiment, we replicated the sequential format using a between-subjects design. To avoid a growing anticipation among participants that the number of trials would always be identical, as well as the experimental demand that even a minimal number of tri-

als (in the small size condition) is informative, we inserted filler stories with two and six trials. Consistent with our predictions, when size was large as opposed to small, the target was more discounted, $F(1, 75) = 3.78$, $p = .056$, and more augmented, $F(1, 45) = 5.24$, $p < .05$. Although these results are less reliable than in the present within-subjects experiment, they indicate that the sample size effect on competition survives under different experimental conditions.

2. Although most interactions with Measure and Size did not reach significance, there was a less interesting Measure \times Size \times Type interaction on the target rating, $F(1, 114) = 16.01$, $p < .001$, which indicated that the target stimulus was less augmented than the target actor given a large size.

3. The simulations of the recurrent model were run using the same specifications as the feedforward model, with the following additional recurrent parameters: $\text{istr} = \text{estr} = \text{decay} = 1$, using the linear activation rule with 1 internal processing cycle (McClelland & Rumelhart, 1988). These parameters make the recurrent model most similar to the feedforward model. In addition, we also ran the recurrent model as specified in earlier social research (Read & Montoya, 1999; Smith & DeCoster, 1999) with parameters $\text{istr} = \text{estr} = \text{decay} = .15$, using the non-linear activation rule with 10 internal processing cycles. The results given both sets of parameters were identical.

REFERENCES

- Allan, L. G. (1993). Human contingency judgments: Rule based or associative? *Psychological Bulletin*, 114, 435-448.
- Anderson, J. R., & Sheu, C.-F. (1995). Causal inferences as perceptual judgments. *Memory and Cognition*, 23, 510-524.
- Baker, A. G., Berber, M. W., & Valée-Tourangeau, F. (1989). Judgments of a 2×2 contingency table: Sequential processing and the learning curve. *Quarterly Journal of Experimental Psychology*, 41B, 65-97.
- Busemeyer, J. R. (1991). Intuitive statistical estimation. In N. Anderson (Ed.), *Contributions to information integration theory: Vol. 1. Cognition*. Hillsdale, NJ: Lawrence Erlbaum.
- Chapman, G. B., & Robbins, S. J. (1990). Cue interaction in human contingency judgment. *Memory and Cognition*, 18, 537-545.
- Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review*, 104, 367-405.
- Cheng, P. W., & Holyoak, K. J. (1995). Complex adaptive systems as intuitive statisticians: Causality, contingency, and prediction. In H. L. Roiblat & J.-A. Meyer (Eds.), *Comparative approaches to cognitive science* (pp. 271-302). Cambridge, MA: MIT Press.
- Cheng, P. W., & Novick, L. R. (1990). A probabilistic contrast model of causal induction. *Journal of Personality and Social Psychology*, 58, 545-567.
- Cheng, P. W., & Novick, L. R. (1992). Covariation in natural causal induction. *Psychological Review*, 99, 365-382.
- Dennett, D. C. (1998). *Brainchildren: Essays on designing minds*. London: Penguin.
- Einhorn, H. J., & Hogarth, R. M. (1986). Judging probable cause. *Psychological Bulletin*, 99, 3-19.
- Fales, E., & Wasserman, E. A. (1992). Causal knowledge: What can psychology teach philosophers? *Journal of Mind and Behavior*, 13, 1-28.
- Försterling, F. (1989). Models of covariation and attribution: How do they relate to the analogy of analysis of variance? *Journal of Personality and Social Psychology*, 57, 615-625.
- Försterling, F. (1992). The Kelley model as an analysis of variance analogy: How far can it be taken? *Journal of Experimental Social Psychology*, 28, 475-490.
- Gluck, M. A., & Bower, G. H. (1988). From conditioning to category learning: An adaptive network model. *Journal of Experimental Psychology: General*, 117, 227-247.
- Hansen, R. D., & Hall, C. A. (1985). Discounting and augmenting facilitative and inhibitory forces: the winner takes all. *Journal of Personality and Social Psychology*, 49, 1482-1493.
- Hogarth, R. M., & Einhorn, H. J. (1992). Order effects in belief updating: The belief-adjustment model. *Cognitive Psychology*, 24, 1-55.
- Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge, UK: Cambridge University Press.
- Kamin, L. J. (1968). Attention-like processes in classical conditioning. In M. R. Jones (Ed.), *Miami symposium on the prediction of behavior: Aversive stimulation* (pp. 9-31). Miami: University of Miami Press.
- Kelley, H. H. (1967). Attribution 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, & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior* (pp. 1-26). Morristown, NJ: General Learning Press.
- Kruglanski, A. W., Schwartz, S. M., Maides, S., & Hamel, I. Z. (1978). Covariation, discounting, and augmentation: Towards a clarification of attributional principles. *Journal of Personality*, 76, 176-189.
- Lober, K., & Shanks, D. R. (2000). Causal induction based on causal power? Critique of Cheng (1997). *Psychological Review*, 107, 195-212.
- McClelland, J. M., & Rumelhart, D. E. (1988). *Explorations in parallel distributed processing: A handbook of models, programs and exercises*. Cambridge, MA: Bradford.
- McClure, J. (1998). Discounting of causes of behavior: Are two reasons better than one? *Journal of Personality and Social Psychology*, 74, 7-20.
- Morris, M. W., & Larrick, R. P. (1995). When one cause casts doubt on another: A normative analysis of Discounting in causal attribution. *Psychological Review*, 102, 331-355.
- Nosofsky, R. M., Kruschke, J. K., & McKinley, S. C. (1992). Combining exemplar-based category representations and connectionist learning rules. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 18, 211-233.
- Read, S. J., & Montoya, J. A. (1999). An autoassociative model of causal reasoning and causal learning: Reply to Van Overwalle's critique of Read and Marcus-Newhall (1993). *Journal of Personality and Social Psychology*, 76, 728-742.
- Rescorla, R. A. (1969). Conditioned inhibition of fear. In W. K. Honig & N. J. Mackintosh (Eds.), *Fundamental issues in associative learning*. Halifax, UK: Dalhousie University Press.
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64-98). New York: Appleton-Century-Crofts.
- Rosenfield, D., & Stephan, W. G. (1977). When discounting fails: An unexpected finding. *Memory and Cognition*, 5, 97-102.
- Rudolph, U., & Försterling, F. (1997). The psychological causality implicit in verbs: A review. *Psychological Bulletin*, 121, 192-218.
- Schustack, M. W., & Sternberg, R. J. (1981). Evaluation of evidence in causal inference. *Journal of Experimental Psychology: General*, 110, 101-120.
- Shanks, D. R. (1985). Forward and backward blocking in human contingency judgment. *Quarterly Journal of Experimental Psychology*, 37b, 1-21.
- Shanks, D. R. (1987). Acquisition functions in contingency judgment. *Learning and Motivation*, 18, 147-166.
- Shanks, D. R. (1991). On similarities between causal judgments in experienced and described situations. *Psychological Science*, 2, 341-350.
- Shanks, D. R. (1995). Is human learning rational? *Quarterly Journal of Experimental Psychology*, 48a, 257-279.
- Shanks, D. R., Lopez, F. J., Darby, R. J., & Dickinson, A. (1996). Distinguishing associative and probabilistic contrast theories of human contingency judgment. In D. R. Shanks, K. J. Holyoak, & D. L. Medin (Eds.), *The psychology of learning and motivation* (Vol. 34, pp. 265-311). New York: Academic Press.
- Smith, E. R. (1996). What do connectionism and social psychology offer each other? *Journal of Personality and Social Psychology*, 70, 893-912.

- Smith, E. R., & DeCoster, J. (1999). *Heuristic-systematic and other dual process models in social and cognitive psychology: An integration and connectionist interpretation*. Unpublished manuscript.
- Van Hamme, L. J. (1994). *Associative and statistical accounts of cue competition in causality judgments*. Unpublished doctoral dissertation, University of Iowa, Iowa City.
- Van Hamme, L. J., & Wasserman, E. A. (1994). Cue competition in causality judgments: The role of nonpresentation of compound stimulus elements. *Learning and Motivation*, 25, 127-151.
- Van Overwalle, F. (1996). The relationship between the Rescorla-Wagner associative model and the probabilistic joint model of causality. *Psychologica Belgica*, 36, 171-192.
- Van Overwalle, F. (1997). A test of the Joint model of causal attribution. *European Journal of Social Psychology*, 27, 221-236.
- Van Overwalle, F. (1998). Causal explanation as constraint satisfaction: A critique and a feedforward connectionist alternative. *Journal of Personality and Social Psychology*, 74, 312-328.
- Van Overwalle, F., Labiouse, C., & French, R. (2000). *Connectionist exploration in social cognition*. Unpublished manuscript.
- Van Overwalle, F., & Van Rooy, D. (1998). A connectionist approach to causal attribution. In S. J. Read & L. C. Miller (Eds.), *Connectionist models of social reasoning and social behavior* (pp. 142-171). Hillsdale, NJ: Lawrence Erlbaum.
- Van Overwalle, F., & Van Rooy, D. (2000). *When more observations are better than less: A connectionist account of the acquisition of causal strength*. Manuscript submitted for publication.
- Wasserman, E. A., Kao, S.-F., Van Hamme, L., Katagiri, M., & Young, M. E. (1996). Causation and association. *The Psychology of Learning and Motivation*, 34, 207-264.
- Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review*, 92, 548-573.
- Wells, G. L., & Ronis, D. L. (1982). Discounting and augmentation: Is there something special about the number of causes? *Personality and Social Psychology Bulletin*, 8, 566-572.
- Williams, D. A., & Docking, G. L. (1995). Associative and normative accounts of negative transfer. *Quarterly Journal of Experimental Psychology*, 48a, 976-998.
- Williams, D. A., Sagness, K. E., & McPhee, J. E. (1994). Configural and elemental strategies in predictive learning. *Journal of Experimental Psychology, Learning, Memory and Cognition*, 20, 694-709.

Received July 24, 2000

Revision accepted January 23, 2001

Pessimistic Bias in Comparative Evaluations: A Case of Perceived Vulnerability to the Effects of Negative Life Events

Hart Blanton

State University of New York at Albany

Danny Axsom

Virginia Polytechnic Institute and State University

Kimberly P. McClive

State University of New York at Albany

Simani Price

American Institutes for Research

This study tested the hypothesis that unrealistic pessimism characterizes comparative estimates of coping ability. Participants rated their ability to adjust to a range of negative life events in comparison to the abilities of other same-sex students at their college. Most coping estimates showed signs of unrealistic pessimism, in that students rated their own abilities as worse than those of other same-sex students. Analyses indicated that this effect was due, in part, to the presence of an egocentric bias and, in part, to the absence of a self-enhancement bias. First, pessimism appeared to arise because participants paid more attention to the difficulties that they would have coping with severe misfortunes than they paid to the difficulties that others would have. Second, pessimism appeared to arise because participants were not motivated to enhance their coping appraisals, given that they were optimistic that they would not experience these events in the future.

When people think about their future, they often do so through rose-colored glasses. For example, in a widely cited study by Weinstein (1980), undergraduates who were asked to consider a range of future events reported that relative to other same-sex students at their university, they were less likely to experience such negative outcomes as contracting a venereal disease or divorcing a few years after marriage. Subsequent research using diverse populations (including community samples) has replicated this basic finding (see Klein & Weinstein, 1997). Of course, many of the people in these studies will

not contract venereal diseases and many of those who marry will not get divorced. It is often impossible to say whether a given individual's beliefs about the future are unduly optimistic. At the group level, however, an optimistic bias is typically inferred if the average response among group members is to see the self as better than the group's average. Using this convention, the evidence for unrealistic optimism is so compelling that Taylor and Brown (1988) tabbed it as one of three major illusions characteristic of normally functioning people (along with overly positive views of the self and exaggerated perceptions of control).

Nevertheless, optimism about the future is not unbridled. It is certainly the case that people do at times worry about their futures. We argue that the current social psychological literature, which focuses primarily on the common sources of optimism, has not adequately explored common sources of pessimism. Moreover, we suggest that such pessimism can arise from the very judgment processes that are thought to cause optimism in

Authors' Note: We would like to thank Jim Jaccard and Bill Klein for comments and assistance on earlier drafts. We would also like to thank Jerry Suls for his careful and constructive editing and the two reviewers, Justin Kruger and a second anonymous reviewer, for their thoughtful comments. Correspondence concerning this article should be addressed to Hart Blanton, Department of Psychology, Social Science 112, State University of New York, Albany, NY 12222; e-mail: hblanton@csc.albany.edu.

PSPB, Vol. 27 No. 12, December 2001 1627-1636

© 2001 by the Society for Personality and Social Psychology, Inc.

other judgment domains. We argue that the process of comparative evaluation, which leads to unrealistic optimism about the likelihood of experiencing negative life events, will often lead to pessimism about one's ability to adjust to negative life events. To understand why this might occur, it is necessary first to understand why people feel uniquely invulnerable when they reflect on the likelihood that negative events will occur.

*The Judgment Processes
Leading to Unique Invulnerability*

When individuals make comparative estimates of their likelihood of experiencing negative life events, they must go through a series of cognitive maneuvers to arrive at a final judgment. Consistent demonstrations that people underestimate their risk of experiencing misfortune relative to others suggest that the default judgment process is flawed in some way. Past research has pointed to two primary flaws that contribute to perceptions of unique invulnerability. First, there is an egocentric bias, which is the tendency for individuals to focus greater attention on the self when making these evaluations than on the "other" who is the comparison standard (e.g., "other same-sex students"). Second, there is a self-enhancement bias, which is the tendency for individuals to focus primarily on positive information about the self as opposed to negative information. These two tendencies combine to create unrealistic optimism about the likelihood of experiencing negative life events.

Egocentric bias. The egocentric bias has been illustrated in many ways. The most straightforward demonstration is found in study participants' statements about how they arrive at comparative judgments. Participants answer these questions almost entirely by reflecting on factors affecting their own risk and then very rarely, and only in passing, do they mention factors affecting others (e.g., Kruger, 1999; Weinstein & Lachendro, 1982). Interestingly, when attention is forced away from the self and onto others, the optimistic bias is attenuated (Hoorens, 1995; Weinstein & Lachendro, 1982; see also Klar & Giladi, 1997). Other evidence of the egocentric bias comes from correlational studies. Various researchers have found that study participants' comparative evaluations do not covary with the attributes of other people in their lives (e.g., Blanton, Buunk, Gibbons, & Kuyper, 1999; Diener & Fujita, 1997; Klar & Giladi, 1999). This suggests that people do not reflect on this information when they make comparative estimates or if they do, that they give this information relatively little weight when forming a judgment. Because these ratings appear to be made with only modest consideration of others, Wood (1996) has even argued that questions asking people to

make comparative evaluations should not be viewed as measures of social comparison.

Self-enhancement. The egocentric nature of comparative ratings can help explain why these judgments are systematically biased, but it cannot explain why they are biased toward optimism. To understand why this occurs, it is necessary to consider the type of information about the self that comes to mind when comparative evaluations are made. In this regard, it appears that the self-enhancement bias influences information accessibility such that mostly positive information is retrieved. To the extent that individuals focus mostly on positive information about the self—information suggesting to them that their personal risk is low—then failure to reflect in detail on similar information about others will result in a distorted view of the self as uniquely invulnerable.

There are two primary reasons for predicting that people will focus on positive information when assessing their comparative risk. First, recall that unrealistically positive views of the self and exaggerated perceptions of personal control also are characteristic of normal functioning (Taylor & Brown, 1988). Mostly positive information about the self should therefore come to mind during personal risk estimation, just as mostly positive information comes to mind during various other self-evaluations. The second reason that positive information comes to mind has to do specifically with the nature of risk estimates. In truth, the average college student will not experience many of the negative life events that have been studied. College students will not typically trip and break bones, have heart attacks before age 40, or attempt suicide (all life events studied by Weinstein, 1980). Thus, college students who reflect accurately on their own low probability of experiencing such events, but who then fail to notice that this is true of college students in general, will feel uniquely invulnerable with respect to the likelihood of experiencing misfortune.

*Egocentric Processing That
Produces Unrealistic Pessimism*

Although the egocentric nature of comparative judgments leads to unrealistically positive evaluations when attention is drawn to positive information about the self, it leads to unrealistically negative comparative evaluations when attention is drawn to negative information about the self. This was demonstrated in a recent set of studies by Kruger (1999). He had participants make comparative evaluations of their abilities for some tasks that were relatively easy (e.g., using a computer mouse correctly) and for some that were relatively hard (e.g., computer programming). He found a tendency to view the self as above average at the easy tasks but below average at the difficult tasks. Thus, it is possible to make people feel uniquely inadequate by asking them to make

comparative evaluations about their negative attributes (even when these negative attributes are common).

In this regard, we argue that people are likely to reflect on mostly negative information when they reflect on their ability to adjust to or cope with negative life events. Judgments about adjustment are different from judgments about probability because, in this case, the desired conclusions actually seem less tenable than the undesired conclusions. Whereas most college students will not contract HIV, for instance, most college students would expect to feel devastated by the news that they had contracted it. Through the egocentric analysis of these difficulties, college students should predict unique devastation in response to HIV. Moreover, these feelings of unique devastation should be greater than those that they would have felt were they to have considered less severe events. This prediction is based on the notion that the more severe the life event, the more people will think about their own inability to cope. This view parallels Kruger (1999), who found that the more difficult the performance task, the more people will think about their own inability to succeed. Thus, when a negative life event is severe, such that most people consider it difficult in absolute terms to cope, the egocentric bias should lead people to perceive a unique inability to cope. In contrast, when a negative life event is mild, such that most people would consider it easy to cope, the egocentric bias should lead people to perceive either a unique capacity to cope or a capacity that is roughly similar to the capacities of comparable others. These hypotheses are tested in the current study.

Motivated Processing That Diminishes Unrealistic Pessimism

If people are pessimistic about their abilities to cope with severe misfortune, this would seem to suggest that they are not motivated to enhance self-evaluations about their coping abilities. In short, coping would appear to be one of the unusual life domains in which people are comfortable embracing their own shortcomings. There may be times, however, when the motivation to believe in one's coping ability will be high, leading to self-enhancement. This might occur, for instance, for situations in which a perceived inability to cope would cause people to doubt their abilities to effectively manage their emotions in the future. In support of this prediction, research has shown that people who are currently experiencing misfortune often do report that they are coping with their situation better than similar others in the same situation (Wood, Taylor, & Lichtman, 1985; also see Crosby, Pufall, Snyder, O'Connell, & Whalen, 1989; Taylor, 1989; Wood & VanderZee, 1997). If the coping appraisals of actual victims reflect a motive to feel in control of one's own emotions when this ability is needed,

then people may also shore up their coping appraisals when a misfortune is only anticipated. This prediction is consistent with the pragmatic tradition in social cognition (see Taylor, 1998) or with the view of the individual as a "motivated tactician" (Fiske & Taylor, 1991). It suggests that people are typically pessimistic about how they would cope with future misfortunes but that they override these tendencies when a feature of the judgment context motivates them to feel otherwise. Thus, egocentric processing may lead people to believe that they are uniquely vulnerable to the effects of severe life events when these events remain in the abstract. As the possibility of actually experiencing these events becomes real, however, people should resist these conclusions. This hypothesis will be tested in the current study.

CURRENT STUDY

Overview

A sample of 92 participants from two large state universities (State University of New York at Albany and Virginia Tech) completed a questionnaire assessing beliefs related to 20 life events. The 20 events (listed in Table 1) were chosen based on pretesting so that they varied along perceptions of severity and perceived likelihood. After completing the questionnaire, participants were given a full debriefing explaining the focal hypotheses.

Ratings

Comparative adjustment. Comparative adjustment ratings were assessed with the question, "How well would you adjust to [event] compared to other same-sex college students?" To help participants interpret this question, they were first told, "We define people who are *completely adjusted* as those who have returned to good mental health, who are able to carry out their normal daily responsibilities effectively, and who are not overcome by their emotions about the negative life event." Ratings were made on 21-point Likert scales, with endpoints of -10 (*much worse adjusted*) and +10 (*much better adjusted*) and a midpoint value of 0 (*same*).

Severity. Severity of each event was assessed by having participants first consider that "some events are devastating to an individual whereas others are merely nuisances." Then, they were asked to think of each event and rate "How severe do you think it would be for someone to experience [event]?" Responses were made on an 11-point scale with endpoints of 0 (*not at all severe*) and 10 (*extremely severe*).

Comparative likelihood. The relative likelihood of experiencing an event was adapted from Weinstein (1980). Participants rated their likelihood of experiencing each event in comparison to other same-sex students at their

university. These ratings were made on 21-point Likert scales, with endpoints of -10 (*much less likely*) and $+10$ (*much more likely*) and a midpoint value of 0 (*same*).

Absolute ratings. To test for the presence of egocentric processing, participants made separate ratings of their own absolute ability to adjust to each event (absolute self) and others' absolute ability to adjust to each event (absolute other). To minimize demand characteristics that might cause participants to try to equate their absolute adjustment ratings with their comparative adjustment ratings, we reversed the direction of the response scale on the absolute ratings. Participants were asked, "How hard would it be for you [other same-sex students] at [participant's college] to adjust with [event]?" Ratings were made on 11-point scales, with endpoints of 0 (*not at all hard*) and 10 (*extremely hard*). For ease of presentation, responses on the absolute measures were reverse coded so that the response metric on both would be scaled with high scores indicating ease of adjustment.

Social desirability. An ancillary measure, social desirability, was assessed to determine if any of the observed effects could be attributed to a self-presentational bias. Participants rated the degree to which they thought it would be psychologically desirable to recover quickly from each of the 20 events. Specifically, they were asked, "How psychologically unhealthy would it be to recover quickly after [event]?" with scale endpoints of 0 (*not at all unhealthy*) and 10 (*extremely unhealthy*).

Question Order

To address potential artifacts that might result from a fixed questionnaire format, two versions of the questionnaire were created. In version one, the comparative coping and adjustment ratings were presented prior to the absolute ratings, whereas this order was reversed in version two. Also in version 1, comparative adjustment ratings occurred prior to comparative likelihood questions and the absolute ratings for others occurred before the absolute ratings for the self, whereas these orders were reversed in version 2. With both versions, the two event moderators were completed at the end of the questionnaire, with severity rated first in version 1 and social desirability rated first in version 2.

RESULTS

Exploratory Analyses

All analyses reported in the sections that follow were preceded by a set of analyses that tested for significant effects of the questionnaire version and for effects of the participant's university affiliation. A small set of higher

order interactions was uncovered. These effects were small enough in magnitude that all of the effects reported below were significant and in the same direction, regardless of the questionnaire version or the university affiliation. We thus do not elaborate these effects any further in the results below.

Event-Level Analyses

Averaging across the 20 events, the sample's mean comparative adjustment ratings were significantly more negative than a neutral point of zero ($M = -2.29$, $SD = 2.41$, $t = -9.12$, $p < .01$). This indicated a pessimistic bias with regard to coping ability. Ratings varied from event to event, however, as shown in the left-most column of Table 1. We found a significant tendency toward optimism for 1 of the 20 events: "You trip and break a bone" ($M = 1.12$, $SD = 3.19$); difference from zero, $t(91) = 3.37$, $p < .001$. We found neither pessimism nor optimism for 4 events. Most important, we found a significant tendency toward pessimism for 15 of the 20 events. The event at the extreme end of pessimism was as follows: "While cleaning a gun, you accidentally shoot a friend" ($M = -6.63$, $SD = 3.83$); difference from zero, $t(91) = -16.59$, $p < .001$.

In contrast to the ratings for comparative adjustment, ratings for comparative likelihood tended toward optimism. Averaging across the 20 events, the overall comparative likelihood score was significantly more negative than a neutral point of zero ($M = -2.98$, $SD = 1.71$), $t(91) = 16.71$, $p < .01$. This indicates that participants thought they would be less likely than others to experience these events, replicating Weinstein (1980). As shown in Table 1, responses on this measure did not show the same range of variability across events as the adjustment ratings, because perceptions of unique invulnerability were observed with 19 of the 20 events. The one exception to this was "falling and breaking a bone," for which the mean response ($M = -.28$, $SD = 2.87$) did not differ from the neutral point of zero, $t(91) < 1$.

In sum, the current sample showed both a pessimistic bias with respect to the ability to cope with negative life events and an optimistic bias with respect to the likelihood of experiencing negative life events. As a first pass to determine if these two biases were related, we inspected the zero-order correlation between the set of 20 sample means for comparative adjustment and the set of 20 sample means for comparative likelihood. Because these two sets of means each have a sampling error that should be taken into account in any significance test, and because each set of means was generated from the same sample participants, the correlation estimate did not meet the assumptions necessary to run standard sig-

TABLE 1: Comparative Coping Ability and Event Likelihood

	Comparative Adjustment		Comparative Likelihood	
	M	SD	M	SD
You trip and break a bone	1.12*	3.19	-0.28	2.87
You are injured in an auto accident	0.58	3.38	-0.62*	2.48
You are injured in a flood	0.11	3.01	-2.77**	3.82
You buy a car that turns out to be a lemon	0.02	3.51	-2.14**	3.68
You are diagnosed with gum problems	-0.28	3.08	-1.63**	3.78
You are injured in an earthquake	-0.71*	3.48	-3.27**	4.10
You are fired from a job	-0.87*	3.49	-2.88**	3.33
You are mugged	-0.89*	3.83	-1.33**	3.05
You fail to find a job 6 months after college graduation	-1.46**	4.34	-2.54**	3.88
You contract a treatable venereal disease	-2.28**	4.18	-5.62**	3.62
You drop out of college	-2.47**	3.80	-5.67**	3.39
You experience a heart attack before age 40	-2.50**	3.70	-2.19**	3.93
You contract lung cancer	-2.89**	4.30	-2.71**	5.18
You get a divorce a few years after marriage	-3.34**	3.85	-3.20**	3.95
You become sterile	-3.63**	4.75	-1.79**	3.16
You are kicked out of college for an honor code violation	-4.29**	3.72	-4.84**	4.04
You contract AIDS	-4.92**	4.54	-5.95**	3.76
A loved one is murdered	-5.01**	4.88	-1.87**	3.15
Your spouse dies in an auto accident	-5.42**	3.88	-0.80**	2.33
You accidentally shoot a friend while cleaning a gun	-6.63**	3.83	-7.22**	3.63
Total	-2.29**	2.41	-2.98**	1.71

NOTE: Comparative adjustment ratings range from -10 to +10. Negative numbers indicate degrees of pessimism (coping worse than comparable others) and positive numbers indicate degrees of optimism (coping better than comparable others). Comparative likelihood ratings range from -10 to +10. Negative numbers indicate degrees of optimism (less likely than comparable others to experience the event) and positive numbers indicate degrees of pessimism (more likely than comparable others to experience the event). *p* values indicate scores differ significantly from neutral points of zero for each rating.

p* < .05. *p* < .01.

nificance testing. We thus do not present *p* values with this analysis. Nevertheless, the degree of association between these two measures suggested that there was a strong and positive relationship between comparative likelihood and comparative adjustment biases, $r(18) = .52$. Specifically, the more an event evoked the perception of unique vulnerability with respect to adjustment, the more it evoked the perception of unique invulnerability with respect to likelihood. This lends support to the hypothesized pragmatic engagement of self-enhancement, in which people reserve their pessimism about their coping ability for those events that they perceive as unlikely. It is possible, however, that the observed

relationship occurred because the sample held a collective misperception that the events that they themselves would experience were less severe than the events that most other people would experience. This possibility will be addressed in some of the analyses that follow.

Individual-Level Analyses

Between-individual effects. To map out the decision process underlying comparative adjustment ratings, we conducted a second set of analyses on the 20 event ratings. In these, we treated the individual as the unit of analysis instead of the event. As a first pass, we computed 20 correlation matrices of the key variables for each event. These 20 tables were then averaged together to gain a general sense of the average between-individual covariation on key variables. Because the correlation coefficients were averaged across 20 events, and because each of the 20 correlation matrices were based on the responses of 92 participants, all parameter estimates reached statistical significance, *z* values > 3.00, *p* < .01.

The resulting correlation matrix is presented in Table 2, where a few notable relationships can be observed. Consistent with analyses at the event level, comparative adjustment ratings were positively related to comparative likelihood ratings, although this relationship was smaller than what was found at the event level, $r = 0.15$, $z = 6.36$, $p < .001$. The reason this association diminished at the between-individual level in comparison to the between-event level may be that the variability within events was diminished relative to the variability between events. For instance, perceptions of event likelihood varied somewhat in degrees of optimism across the 20 events, but most participants were extremely optimistic that they would not contract HIV. This restriction in range would reduce the degree of association that likelihood ratings can have with other ratings. Despite this problem, some of the values in Table 1 suggested strong associations. In particular, comparative adjustment estimates were inversely related to event severity, $r = -.40$, $z = 17.95$, $p < .001$. This lends support for the egocentric nature of the comparative adjustment. It indicates that the participants who were most pessimistic about their abilities to cope were the ones who viewed the events as especially severe. In further support of the egocentric mechanism, comparative adjustment was positively associated with both the absolute self-ratings, $r = .45$, $z = 20.58$, $p < .001$, and absolute other ratings, $r = .27$, $z = 11.58$, $p < .001$. This same pattern was found by Klar and Gilaldi (1999, p. 588). As they noted, if comparative adjustment ratings reflected an actual comparison between self and other, one would most typically expect them to be inversely correlated with the absolute ratings for other—not positively associated. However, positive

TABLE 2: Average Correlation Matrix Across the 20 Life Events

	2	3	4	5	M	SD
1. Comparative adjustment	0.15	0.45	0.27	-0.40	-2.29	2.41
2. Comparative likelihood	—	0.14	0.07	-0.12	-2.98	1.71
3. Absolute self	—	—	0.57	-0.67	5.89	1.26
4. Absolute other	—	—	—	-0.56	5.81	1.20
5. Event severity	—	—	—	—	6.03	1.12

NOTE: Values are the average correlation coefficients for the focal variables across the 20 negative life events. Comparative adjustment and comparative likelihood were rated on -10 to +10 scales and the remaining variables were rated on 0 to 10 scales. All coefficients are significant at z value > 3.00 , $p < .01$.

associations might be observed at the level of zero-order correlation in some cases as a result of the shared variance between the absolute ratings for self and absolute ratings for other, $r = .57$, $z = 27.40$, $p < .001$. Moreover, the hypothesized inverse relations have been found in other studies (Kruger, 1999). The observed pattern is thus only suggestive. To better assess the case for an egocentric mechanism, we moved to analyses that investigated the within-individual associations between ratings for the 20 events.

Within-individual effects. To investigate how ratings varied between events but within individuals, we conducted a set of within-subject regressions. To perform these analyses, it was necessary first to reconfigure the data set into a pooled-time series or "stacked subjects" configuration (Sayrs, 1989). In this configuration, the responses from each participant were broken into 20 separate lines of data. To remove any effects of participant on these 20 lines of data, a modification of pooled time-series analysis called least squares dummy variable regression was used (Jaccard & Wan, 1993; Sayrs, 1989). This was accomplished by generating a set of 91 dummy variables representing the 92 participants. These codes were entered into all regression equations described below to remove effects attributable to the participant. Because the results from such analyses would be adversely affected by variations in sample size, missing values (representing no more than 2% to 3% of the cases on a given variable) were imputed using the expectation-maximization method of Little and Schenker (1995). Because the nonstandardized regression coefficients do not provide easy interpretation using this method, results from all analyses using this procedure are presented in terms of standardized coefficients.¹

First, to test for the effects consistent with the proposed egocentric and self-enhancing mechanisms, we regressed comparative adjustment onto comparative likelihood and event severity. Consistent with the findings at the event level, there was a significant effect of severity, $\beta = -.77$, $t(1746) = 17.23$, $p < .001$, and a signifi-

cant effect of comparative likelihood, $\beta = .15$, $t(1746) = 4.58$, $p < .001$. The severity effect indicated that participants were the most pessimistic about their ability to cope with the events that they rated as the most severe. The comparative likelihood effect indicated that participants were the most pessimistic about their ability to cope with the events that they felt the most optimistic about avoiding in the future. Because this latter effect controls for the effect of severity, it is difficult to argue that it occurred because participants thought they would be more likely than others to experience the less severe events. This thus resolves the concern expressed in the event-level analyses.

The next set of analyses provided a formal test of the egocentric nature of comparative adjustment ratings. This was done by regressing comparative adjustment onto the absolute self-ratings and the absolute-other ratings. Consistent with previous research showing that comparative ratings are driven by self-assessment more than other-assessment (e.g., Blanton et al., 1999; Diener & Fujita, 1997; Klar & Giladi, 1999; Kruger, 1999), only the absolute ratings for self were found to have significant influence on the comparative adjustment ratings, $\beta = 0.54$, $t(1746) = 11.54$, $p < .001$. The nature of this effect was consistent with past research. The more participants thought that it would be hard for them to cope in absolute terms, the more they thought it would be hard for them to cope in relative terms. In contrast, the absolute rating for others had only a small and nonsignificant effect on comparative adjustment, $\beta = -.05$, $t(1746) = 1.07$, *ns*. Combined, these two findings suggest that participants' comparative adjustment ratings were the result of an egocentric analysis of the difficulties they thought they would encounter while trying to adjust.

We then conducted two analyses to determine if the egocentric analysis of difficulties was responsible for the effects of event severity and comparative likelihood on comparative adjustment ratings. This was done by first regressing absolute self-ratings on severity and comparative likelihood. Consistent with an egocentric account of each effect, this revealed both a significant effect of severity, $\beta = -.82$, $t(1746) = 35.79$, $p < .001$; and a small but significant effect of comparative likelihood, $\beta = 0.05$, $t(1746) = 2.61$, $p < .01$. The nature of these two effects combined was that as the severity of an event increased and as the likelihood of an event decreased, the absolute estimate of coping ability decreased. It should be noted, however, that the effect of comparative choice was a small effect and one that just barely made conventional levels of significance. Nevertheless, to determine if these two effects mediated the effect of absolute adjustment ratings on comparative adjustment ratings, we regressed comparative adjustment on severity, comparative likelihood, and the absolute self-rating. Mediation would be

demonstrated for the independent effects of severity and comparative likelihood if the predictive power of either was diminished when the absolute ratings for self were entered.

The results did not support a full mediation account for either effect. Controlling for the effects of absolute self-ratings, $\beta = .51$, $t(1745) = 6.42$, $p < .01$, participants still thought they would adjust worse in comparison to others for events they rated as severe, $\beta = -.36$, $t(1745) = 4.55$, $p < .01$, and as unlikely, $\beta = .13$, $t(1745) = 4.01$, $p < .01$. These findings argue against full mediation because neither the effect of severity nor the effect of comparative likelihood became nonsignificant after controlling for the absolute self-ratings. However, the drop in the magnitude of the severity effect, from $\beta = -.77$ and $t(1746) = 17.23$ when absolute self-ratings were not controlled to $\beta = -.26$ and $t(1745) = 4.55$ when comparative likelihood was controlled, suggested that the severity effect was at least partially mediated by the egocentric nature of comparative ratings. To determine if the data supported this interpretation, a commonality analysis was conducted to establish whether a significant proportion of the severity effect could be attributed to variance it shared with the absolute self-ratings. This revealed that fully 94% of the variance in comparative adjustment that could be attributed to event severity also was shared with the absolute self-ratings. This was a significant amount of variance, $t(1745) = 12.59$, $p < .01$. Thus, the regression analyses were consistent with a mediational account. However, the relatively small proportion of the severity effect that was not shared with the absolute self-ratings was also significant, $t(1745) = 3.28$, $p < .01$. Thus, the regression analyses were consistent with a partial mediation model for event severity. When a commonality analysis was conducted on the comparative likelihood effect, only 1.8% of this effect was attributable to variance that it shared with the absolute self-ratings. This was not a significant amount of variance, $F < 1$. Thus, there was little evidence that the effect of comparative likelihood was mediated by the egocentric nature of comparative adjustment ratings.

In summary, the within-subject analyses indicated that event severity had a large influence on comparative adjustment estimates and that this effect was partially mediated by the egocentric nature of comparative ratings. In contrast, comparative likelihood had a moderate influence on comparative adjustment estimates, but this effect did not appear to be mediated by the egocentric nature of comparative ratings. The finding with event severity is consistent with Kruger (1999) and suggests that event severity (or coping "difficulty") diminishes comparative adjustment ratings by diminishing absolute adjustment estimates for the self. Because we did not find full mediation, however, event severity may have

influenced comparative adjustment ratings directly as well and in ways that were not investigated by Kruger (1999). Alternatively, our single-item rating of event severity may have lacked the statistical reliability that was needed to establish full mediation. In contrast, findings with comparative likelihood suggested that comparative likelihood estimates influenced comparative adjustment ratings through a mechanism or a set of mechanisms that differed from the egocentric mechanism that was investigated by Kruger (1999).

Ancillary Analyses

We conducted a set of ancillary analyses to address a potential third variable confound and to address a possible artifact of the design. First, we reran the above analyses controlling for the social desirability variable. When comparative adjustment was regressed onto social desirability, this variable provided significant prediction, $\beta = -.72$, $t(1747) = 15.88$, $p < .01$. The nature of this effect was that participants thought they would have more difficulty adjusting for events in comparison to others when they thought it was healthier to take time recovering. When regression analyses also controlled for the effects of comparative adjustment, event severity, comparative likelihood, and absolute abilities for the self, the effect of social desirability remained significant, although the effect size diminished, $\beta = -.18$, $t(1744) = 2.69$, $p < .01$. Importantly, however, all of the above effects of the focal variables remained. In particular, severity and comparative likelihood continued to predict comparative adjustment in the predicted directions. Moreover, the absolute ratings for self partially mediated the effect of severity but not the effect of comparative likelihood. Finally, there was no evidence that the effect of social desirability on comparative adjustment was mediated by egocentric processing.

A second concern in the study related to experimenter demand. We were concerned that by having participants make a set of highly similar ratings, we might have forced associations between the variables that would not have otherwise existed. To address this, we used three independent samples to generate estimates for the focal variables and then investigated the associations between these variables, treating event as the level of analysis. Comparative adjustment estimates were rated by a new sample of 278 participants ($n = 151$ women, $n = 127$ men). Severity and then social desirability estimates were rated by a different sample of 243 participants ($n = 125$ women, $n = 118$ men). Finally, comparative likelihood estimates were taken from a subsample of the participants analyzed in the original study. This subsample was the group who completed the version of the questionnaire that presented the comparative likelihood estimates as the first rating in the packet. Thus, we

were able to investigate the degree of association between ratings at the level of the event, with minimal concerns about experimenter demand forcing associations. There were two primary shortcomings with using a design of this nature. First, it did not allow for inspection of responses at the individual level of analysis; therefore, we could not test for egocentric processing. Second, analyses at the event level did not meet the assumptions necessary for conducting significance testing. Nevertheless, we were able to use these ratings to compare the magnitude of associations at the event level with those found in the original study to ensure compatibility of effect sizes. When we regressed the 20 means for comparative adjustment on the 20 means for severity, comparative likelihood, and social desirability, we found a large effect of severity, $\beta = -.94$, a moderate effect of comparative likelihood, $\beta = .26$, and a small effect of social desirability, $\beta = -.14$. These three effects were in the same direction and were comparable in size to what we had found in the larger study. This indicated that the primary results did not arise due to associations forced by the questionnaire design.

DISCUSSION

Overview

These findings indicate that people's beliefs about their ability to adjust to misfortune do not follow as a simple extension from the findings of optimism with respect to personal risk assessment. Although participants were generally optimistic about their risk level, thinking that they would be less likely than others to experience misfortunes, they were simultaneously pessimistic about their coping abilities, thinking that they would be less able than others to cope with misfortunes if they were to occur. This latter tendency toward pessimism did not apply with all events, however. Pessimism was greatest for events that were consensually viewed as severe. This moderating role of event severity was predicted on the basis of research pointing to the egocentric nature of comparative judgments (e.g., Kruger, 1999). We hypothesized that participants would make comparative adjustment ratings by focusing primarily on the difficulties they themselves would experience while paying relatively little attention to the difficulties others would experience. Consistent with this interpretation, we found that participants' comparative adjustment estimates were predicted by their ratings of their own coping ability but not by their ratings of others' coping ability. Moreover, when participants' ratings of their own coping ability were controlled, the effect of event severity was diminished to a significant degree. It should be noted, however, that this analysis only offered support for a partial mediation model. Thus, an account of the

effect of severity that focuses on just the egocentric nature of comparative ratings is not sufficient. Future research will need to determine if severity leads to more pessimistic appraisals of one's coping ability through other mechanisms or if full mediation can be achieved by improving the psychometric properties of the relevant measures.

Although we thought people would be open in general to admitting their frailties when they reflected on their coping abilities, we also believed there would be contexts in which they would be motivated to deny such vulnerabilities. Based on past research showing that victims of misfortune report that they are coping better than others (Wood et al., 1985; also see Crosby et al., 1989; Taylor, 1989; Wood & VanderZee, 1997), we predicted that nonvictims would downplay their coping difficulties for events that they thought they might experience in the future. For the most part, participants were optimistic that negative events would not occur to them. However, for those events that they were the least optimistic about avoiding in the future, they were also the least pessimistic about adjusting. It is possible that this self-enhancing tendency could actually yield optimism about adjustment when event likelihood appears certain.

Interestingly, we found little evidence that this self-enhancement tendency occurs as a result of egocentric processing. This suggests that when people feel that misfortune may strike, they do not augment their assessments of their relative coping ability by shoring up their perceptions of their ability to cope in absolute terms. Instead, they appear to move directly to the comparative estimate. This pattern of results possibly points to the emotional significance of comparative adjustment estimates relative to absolute estimates (Diener & Fujita, 1997; Klein, 1997). When people think misfortune is imminent, their first defense may be a social one in which they first bolster their perceptions that they can cope better than others. Perceptions that coping will be "easy" in absolute terms may only follow afterward.

Limitations

As with any study, the results of the present investigation must be interpreted in light of methodological constraints associated with data collection. The research relied on self-reports and used a cross-sectional design. These are always causes for caution. Self-report is somewhat less of a concern in the current study than it is in many because the criterion variable of interest was related to a subjective evaluation. Moreover, ancillary analyses suggested that the associations observed were not manufactured by the questionnaire format and did not result from social desirability concerns. Nevertheless, both tendencies could have influenced responses in

ways that we failed to operationalize with our questionnaire. Even if questionnaire format and social desirability were not influencing responses, other concerns remain due to the cross-sectional nature of the study. Many of the observed relationships may have been due to unanticipated third-variable influences or to reverse causality effects. The current findings would thus benefit from replications in which participants are randomly assigned to conditions and in which relevant self-perceptions are experimentally manipulated. For instance, research could determine if increases in the perceived likelihood and decreases in perceived severity lead to diminished pessimism about coping. Despite these limitations, the findings in this study are consistent with current theory and offer strong evidence of a type of pessimism that has received little empirical attention.

Implications for Persuasive Communication

The observed pessimistic bias, if robust, may offer guidance for constructing messages to change behavior. Past research has documented that it is difficult to make people feel at risk of experiencing negative life events (Weinstein & Klein, 1995). In contrast, the current studies suggest that it may be possible to target feelings of vulnerability related to adjustment without triggering message rejection. As an example of how this finding might be applied, our research suggests that health campaigns designed to increase condom use on a college campus should not try to convince students that their sexual practices put them at risk for contracting AIDS. In general, college students do not feel at risk of experiencing such outcomes and most would probably resist efforts to change these perceptions (Gerrard, Gibbons, & Warner, 1991; Gerrard & Warner, 1994; Smith, Gerrard, & Gibbons, 1997). The alternative suggested by the current studies is to construct messages that communicate to college students that were they to contract AIDS, they would be ill-equipped for managing this crisis. An intervention using this strategy would tap into a feeling of vulnerability that appears common and may be more effective at motivating behavior change.

It is interesting to note that an intervention such as the one we propose is expected to influence behavior by reinforcing a perception that is generally erroneous. In truth, students who contract AIDS will, on average, be no more or less able to adjust than the average student. In fact, half of the students who contract AIDS will adjust better than the other half. Although it may be disagreeable to build interventions that take advantage of a perceptual bias, it is the goal of health communication to change behavior before victimization occurs. Toward this end, it would be preferable to reinforce a bias held by nonvictims than to have many discover their own coping competence by encountering misfortune firsthand.

Before interventions such as this are implemented, however, it is important to conduct research to determine whether they will lead to the predicted effects. It is quite possible, for instance, that people will resist communications that suggest that their current actions will lead to outcomes they cannot handle. Recall that our participants thought that they would cope better the more likely the event. This suggests that steps to increase the perceived likelihood of an event may increase the perceived ability to cope. The reverse may occur as well. It is possible that reinforcing a perceived inability to cope will lead to a decreased perception in an event's likelihood. This raises an interesting possibility that is worth pursuing in future research. Beliefs about likelihood and beliefs about adjustment may compensate for one another in a dynamic fashion to minimize fear (see Boney-McCoy, Gibbons, & Gerrard, 1999; Gibbons, Eggleston, & Benthin, 1997). If compensation of this sort does occur, interventions that are oriented toward changing behavior will face a dual, and challenging, task. They will need to shatter an illusion of unique invulnerability with respect to event likelihood while simultaneously reinforcing an illusion of unique vulnerability with respect to coping and adjustment.

NOTE

1. The primary concern that must be addressed when using this analytic technique is that there may be nonconstant error variance or "dependencies" occurring within individuals. Likely dependencies for the current analyses fall into two broad categories. First, there may be dependencies due to the spatial ordering of different events. To test for this, the Durbin-Watson test of spatial dependencies was performed on each version of the questionnaire. This did not uncover any signs of dependence and so no corrections were made to address this. Second, there may be dependencies across the 30 events if clusters of events share variance. This would suggest that they should not be treated as independent units of analysis within individuals. To test for this possibility, we identified clusters of events that shared variance before conducting regression analyses by factor-analyzing responses on the criterion variable. We then tested if these factor groupings resulted in dependencies within individuals. This procedure did reveal dependencies that would inflate significance levels. We thus corrected for these by inflating the standard error terms, following the procedures of Deane (2000). All effects we report using this analytic procedure are thus unbiased with respect to the shared variance within these event clusters.

REFERENCES

- Blanton, H., Buunk, B. P., Gibbons, F. X., & 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.
- Boney-McCoy, S., Gibbons, F. X., & Gerrard, M. (1999). Self-esteem, compensatory self-enhancement, and the consideration of health risk. *Personality and Social Psychology Bulletin*, 25(8), 954-965.
- Crosby, F., Pufall, A., Snyder, R. C., O'Connell, M., & Whalen, P. (1989). The denial of personal disadvantage among you, me and all the other ostriches. In M. Crawford & M. Gentry (Eds.), *Gender and thought* (pp. 79-99). New York: Springer-Verlag.

- Deane, G. D. (2000). "Check all that apply . . .": A simple method for correcting observational dependence in multiple response data (Working Paper 2000-4). Albany, NY: Center for Social and Demographic Analysis, State University of New York at Albany.
- Diener, E., & Fujita, F. (1997). Social comparisons and subjective well-being. In B. P. Buunk & F. X. Gibbons (Eds.), *Health, coping, and well-being: Perspectives from social comparison theory* (pp. 329-358). Mahwah, NJ: Lawrence Erlbaum.
- Fiske, S. T., & Taylor, S. E. (1991). *Social cognition* (2nd ed.). New York: McGraw-Hill.
- Gerrard, M., Gibbons, F. X., & Warner, T. D. (1991). Effects of reviewing risk-relevant behavior on perceived vulnerability among women marines. *Health Psychology, 10*(3), 173-179.
- Gerrard, M., & Warner, T. D. (1994). Comparison of Marine and college women's HIV/AIDS-relevant sexual behaviors. *Journal of Applied Social Psychology, 24*(11), 959-980.
- Gibbons, F. X., Eggleston, T. J., & Benthin, A. C. (1997). Cognitive reactions to smoking relapse: The reciprocal relation between dissonance and self-esteem. *Journal of Personality and Social Psychology, 72*(1), 184-195.
- Hoorens, V. (1995). Self-favoring biases, self-presentation, and the self-other asymmetry in social comparison. *Journal of Personality, 63*(4), 793-817.
- Jaccard, J., & Wan, C. K. (1993). Statistical analysis of temporal data with many observations: Issues for behavioral medicine data. *Annals of Behavioral Medicine, 15*, 41-50.
- Klar, Y., & Giladi, E. E. (1997). No one in my group can be below the group's average: A robust positivity bias in favor of anonymous peers. *Journal of Personality and Social Psychology, 73*(5), 885-901.
- Klar, Y., & Giladi, E. E. (1999). Are most people happier than their peers, or are they just happy? *Personality and Social Psychology Bulletin, 25*(5), 585-594.
- Klein, W. M. (1997). Objective standards are not enough: Affective self-evaluative, and behavioral responses to social comparison information. *Journal of Personality and Social Psychology, 72*(4), 763-774.
- Klein, W. M., & Weinstein, N. D. (1997). Social comparison and unrealistic optimism about personal risk. In B. P. Buunk & F. X. Gibbons (Eds.), *Health, coping, and well-being: Perspectives from social comparison theory* (pp. 25-61). Mahwah, NJ: Lawrence Erlbaum.
- Kruger, J. (1999). Lake Woebegone be gone! The "below-average effect" and the egocentric nature of comparative ability judgments. *Journal of Personality and Social Psychology, 77*(2), 221-232.
- Little, R. J., & Schenker, N. (1995). Missing data. In G. Arminger, C. Clogg, & M. Sobel (Eds.), *Handbook of statistical modeling for the social and behavioral sciences*. New York: Plenum.
- Sayrs, L. W. (1989). *Pooled time series analysis*. Newbury Park, CA: Sage.
- Smith, G. E., Gerrard, M., & Gibbons, F. X. (1997). Self-esteem and the relation between risk behavior and perceptions of vulnerability to unplanned pregnancy in college women. *Health Psychology, 16*(2), 137-146.
- Taylor, S. E. (1989). *Positive illusions*. New York: Basic Books.
- Taylor, S. E. (1998). The social being in social psychology. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (Vol. 1, 4th ed., pp. 58-95). New York: McGraw-Hill.
- Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin, 103*, 193-210.
- Weinstein, N. D. (1980). Unrealistic optimism about future life events. *Journal of Personality and Social Psychology, 39*, 806-820.
- Weinstein, N. D., & Klein, W. M. (1995). Resistance of personal risk perceptions to debiasing interventions. *Health Psychology, 14*(2), 132-140.
- Weinstein, N. D., & Lachendro, E. (1982). Egocentrism as a source of unrealistic optimism. *Personality and Social Psychology Bulletin, 8*, 195-200.
- 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. E., & Lichtman, R. R. (1985). Social comparison in adjustment to breast cancer. *Journal of Personality and Social Psychology, 49*, 1169-1183.
- Wood, J. V., & VanderZee, K. (1997). Social comparisons among cancer patients: Under what conditions are comparisons upward and downward? In B. P. Buunk & F. X. Gibbons (Eds.), *Health, coping, and well-being: Perspectives from social comparison theory* (pp. 299-328). Mahwah, NJ: Lawrence Erlbaum.

Received August 9, 2000

Revision accepted January 14, 2001

Social Anxiety and the Cued Activation of Relational Knowledge

Mark W. Baldwin

McGill University

Kelley J. Main

University of Manitoba

A cued activation procedure was used to examine the hypothesis that social anxiety involves an expectation of being rejected or evaluated negatively by others, combined with a concern about impression management. Participants underwent a conditioning procedure in which distinctive computer tones were paired with thoughts of social rejection and acceptance, respectively. In a pilot study, a lexical decision task established that when these tone cues were played later, they differentially activated expectations of rejection. In the main study, female participants interacted with a male confederate while one of the tones, or a control tone, sounded repeatedly in the background. Several indicators of social anxiety showed an interaction between level of public self-consciousness and the nature of the tone played. High-self-conscious individuals tended to be affected by the cues, whereas low-self-conscious people were not affected.

Social anxiety in one of its forms, such as shyness or dating anxiety, is a common experience in interpersonal relationships: we feel awkward, we stammer, we blush. Occasionally, we may come to avoid social situations that could lead to embarrassment, and this social avoidance can produce loneliness and other relationship difficulties.

A fear of negative interpersonal evaluation has been identified as the central precipitating factor in social anxiety (Edelmann, 1992; Ingram & Kendall, 1987; Winton, Clark, & Edelmann, 1995). In their well-known model of social anxiety, Schlenker and Leary (1982; also Leary & Kowalski, 1995) characterized this fear as arising from two factors: (a) The individual anticipates making an unwanted impression on others, and (b) is very concerned about or focused on the impression he or she is making.

Previous research has already fleshed out this model with some of the elements contributing to the fear of

negative evaluation, such as the perception of others as rejecting and hypercritical (e.g., Hartmann, 1983; Wallace & Alden, 1991). We sought to build on previous work by using a novel experimental paradigm to examine the effects of knowledge activation on social anxiety during an initial encounter. Our guiding assumption was that the expectation of negative evaluation results from negative memories and knowledge structures becoming activated and influencing the anticipation and interpretation of the current interaction. When a woman meets a man for the first time, for example, what autobiographical memories resonate with the current context? What interpersonal scripts (e.g., "If I say something foolish, he will dismiss me?") influence—even implicitly, outside of conscious awareness—the interpretation of ongoing experience? What images of social events (e.g., being teased or criticized) pop into mind so easily that they seem inevitable outcomes of the upcoming interaction? And, the focus of this article, "What determines which memories and knowledge structures get activated?"

The Cued Activation of Relational Knowledge

Priming research has demonstrated the possibility of directly activating relational schemas representing social acceptance and rejection. Various forms of priming manipulation, including guided visualizations (Baldwin &

Authors' Note: This research was supported by a grant from the Social Sciences and Humanities Research Council of Canada. We thank Jacquie Vorauer and Grainne Fitzsimons for their assistance in collecting data, Jason Ediger for serving as the confederate, and Jacquie Vorauer and Steven Wright for their helpful comments on earlier drafts. Correspondence may be addressed to the first author at the Department of Psychology, McGill University, 1205 Dr. Penfield Avenue, Montreal, Quebec, Canada, H3A 1B1; e-mail: mbaldwin@ego.psych.mcgill.ca.

PSPB, Vol. 27 No. 12, December 2001 1637-1647

© 2001 by the Society for Personality and Social Psychology, Inc.

Holmes, 1987; Baldwin, Keelan, Fehr, Enns & Koh-Rangarajoo, 1996; Baldwin & Sinclair, 1996), presentations of significant others' names (Baldwin, 1994), and subliminal exposures of a significant other's face (Baldwin, Carrell, & Lopez, 1990), have produced self-evaluative and interpersonal effects. In one study, for example, graduate students who were subliminally primed with their department chair's scowling face were more self-critical of their own research performance than were their unprimed counterparts (Baldwin et al., 1990). In another study, people visualizing significant others who made them feel avoidant, anxious, or secure were later differentially motivated to meet prospective dating partners with similar attachment characteristics (Baldwin et al., 1996).

Although the activation of relational information often occurs in just this way—a telephone call early in the morning from a critical acquaintance certainly can make one feel less than secure for the rest of the day—during actual encounters, evaluative expectancies can be triggered by more indirect cues. In theory, any minimal cue, if it becomes associated with specific interpersonal experiences, should be able to activate relational knowledge. For example, a woman who was repeatedly criticized as a child by her piano teacher might find that the mere sight, or thought, of a piano makes her somewhat uneasy and insecure. Conversely, a man routinely enjoying warm interactions with a loved one while dining on ethnic cuisine might develop a positive association to that particular comfort food. As Bargh and Ferguson (2000) have argued, many social cognitive structures and processes are set in motion by environmental cues; this principle surely applies to the activation of relational schemas.

Testing this mechanism in the lab would involve experimentally creating a new association between a relational schema and a neutral cue, such that presentation of the neutral cue later serves to activate the relational schema. A few recent studies have demonstrated that this is possible: In three studies by Baldwin, Granzberg, Pippus, and Pritchard (2001) using the same manipulation as in the current studies, participants completed a bogus computerized questionnaire that gave them 10 trials of approval (a row of smiling faces) and disapproval (a row of frowning faces) feedback, paired respectively with two distinctive computer-generated tones. Later, while participants completed a package of dependent measures, a computer on the other side of the room repeatedly emitted one of the tones. Women's ratings of their self-esteem at that moment were lower if the tone being played was the one that had once signaled rejection.

Although much can be learned by studying how cued activation mechanisms influence self-evaluative thoughts

in the context of an experimental task, the most appropriate situation for studying social anxiety is during an uncomfortable social encounter. Would this kind of cued activation procedure have an impact on people's thoughts, feelings, and behavior during a social interaction? After all, there are so many other influences and distractions: the give and take of the conversation, the comments the other person makes, the actual characteristics of the interaction partner. If knowledge activation is a key element in social anxiety, however, perception of the ongoing interaction should be shaped to some degree by whatever knowledge is activated at the moment, even if it is triggered by so minimal a cue as a tone sounding in the background. As a strong test of the knowledge activation premise, therefore, we administered the cued activation manipulation during an initial interaction between a woman and a man and included measures of mood, self-esteem, and interpersonal behavior.

Self-Consciousness and the Concern With Impression Management

Schlenker and Leary's (1982) model of social anxiety holds that the activation of a negative or unwanted evaluative expectancy is only one of two factors contributing to social anxiety; the other is impression motivation, a concern with creating a certain impression on others. It is logical that people who are highly focused on how they are evaluated by an interaction partner should be the most influenced by activated relational knowledge about acceptance and rejection. Indeed, there is a substantial literature (see Wicklund, 1975, for a review) demonstrating that when people focus their attention on themselves, this increases their evaluative concerns, making them more responsive to evaluative standards and feedback of various kinds. Consequently, social anxiety researchers have operationalized the impression motivation factor as high and low scores on the individual difference measure of public self-consciousness (Fenigstein, Scheier, & Buss, 1975) and have indeed found a strong link to social anxiety. With items such as "I'm concerned with the way I present myself" and "I usually worry about making a good impression," people scoring highly on this measure tend to report an increased awareness of how they are regarded by others and attach considerable importance to another person's impressions of them. Public self-consciousness (hereinafter referred to simply as self-consciousness) has been shown to correlate with social anxiety in several studies (e.g., Hope & Heimberg, 1988; Leary & Kowalski, 1995).

As would be predicted on the basis of the two-factor model of social anxiety, the combination of both high-self-focus and negative social expectancies tends to produce the most pronounced evaluative distress and social

anxiety. In an early study, Fenigstein (1979) used the scale to designate female undergraduate participants as high- or low-self-conscious and had them undergo a social interaction where they were snubbed by two experimental confederates. He found that women high in self-consciousness displayed a more negative reaction to the social rejection and greater desire to avoid further interaction with the confederates than did those low in self-consciousness. Similar results have been reported by Alden, Teschuk, and Tee (1992) and Burzio, Merluzzi, and Pryor (1986). The current research also builds on some previous studies of self-evaluations in testing situations, in which visualization primes of accepting and critical relational schemas were found to have more impact where self-awareness also was induced experimentally (Baldwin, 1994, Study 2; Baldwin & Holmes, 1987, Study 2). In our study of social interaction, therefore, we chose to compare the reactions of high- versus low-self-conscious individuals to the cued activation of relational information.

The Current Studies

In both studies, a conditioning procedure was used to associate expectations of acceptance and rejection with different computer-generated tones. In the pilot study, the effects of this procedure were examined using a lexical decision task to establish that later presentation of the conditioned tones activated thoughts of acceptance and rejection. In the main study, the same procedure was used to activate relational knowledge during an initial, potentially anxiety-producing meeting between a woman and a man. We sought to test whether a knowledge activation procedure would have an impact in this context.

PILOT STUDY

In a pilot study, we examined the cued activation procedure using a lexical decision task (Meyer & Schvaneveldt, 1971) to establish whether the conditioned tones would indeed activate acceptance and rejection information. In previous research using the lexical decision task (Baldwin & Sinclair, 1996), participants performed word/nonword judgments on a range of targets, including words representing interpersonal acceptance (e.g., liked, accepted) and rejection (e.g., criticized, rejected). This task is thought to reveal the accessibility of different schemas or constructs: Words that correspond to cognitively accessible content produce shorter reaction times because the individual is quicker to recognize them as words. A sequential-priming version of the task builds on the principle of spreading activation to assess the associations the individual perceives among different constructs. In the studies by Baldwin and Sinclair, each trial began with a prime word representing either

success or failure. Consistent with a view of low self-esteem as resulting from the perception that acceptance from others is conditional or contingent on successful performances (e.g., Rogers, 1959), low-self-esteem individuals were quicker to recognize rejection targets when primed with failure and quicker to recognize acceptance targets when primed with success.

Baldwin and Meunier (1999) extended this research by examining cued activation. Cues were established during a brief conditioning phase in which participants visualized either a contingently or noncontingently accepting significant other while a computer repeatedly emitted a distinctive sequence of tones. When the tone sequence was played again later during the lexical decision task, reaction times reflected the activated interpersonal context. For example, when the tone that had been paired with a contingent relationship was played, people were quicker to recognize rejection targets when primed with failure or acceptance targets when primed with success.

In the current pilot study, we used a modified version of this task to test for the direct associations we were trying to establish between a conditioned stimulus and the anticipation of acceptance or rejection. We first used a conditioning paradigm to create cues (distinctive tone sequences) for either rejection or acceptance. Then, in a lexical decision task, participants heard one cue or the other and performed word-nonword judgments on acceptance or rejection targets. We predicted an interaction effect whereby the conditioned stimulus for acceptance (hereinafter the CS-acceptance) would facilitate reaction times (RTs) to acceptance words (compared to the CS-rejection) but slow down RTs to rejection words. This would demonstrate the cued activation of interpersonal expectancies.

METHOD

Participants

Thirteen McGill University undergraduate students (7 women, 6 men, with a median age of 21 years) participated and received \$8 (Cdn.) in compensation. Data from 1 female participant were dropped from analyses because she made errors on more than 20% of the word trials.

Procedure

Participants were informed by the female experimenter that they were in a study about cognitive styles and attitudes that would involve a number of different tasks. First was the conditioning procedure, developed by Baldwin et al. (2001), which consisted of an attitude questionnaire on a computer. This questionnaire presented a series of benign multiple choice questions (e.g.,

"What is your favorite flavor of ice cream?"). Participants were told that the questions had been pretested with other students, asking them which answers they would like someone to give, and the goal of the present study was to see if the participants' opinions actually lined up with these socially desirable answers. Earlier research by Baldwin et al. showed that participants find this a plausible, compelling exercise. As they responded to questions, they were given bogus feedback on every third trial: Feedback consisted of a 1-second presentation of a row of two male and two female faces smiling in approval, to indicate that their answers were matching the ideal, or a row of frowning faces, to indicate that their answers were not matching the ideal. The feedback was in fact given in a fixed random order, unrelated to the participants' responses. Each time feedback was given it was signaled 1.5 seconds in advance by one of two distinctive 1-second tone sequences: either a high-pitched doorbell sound or a low-pitched sequence of tones that increased in pitch. Thus, after the 60-item questionnaire, which took approximately 15 minutes to complete, the participants had received 10 acceptance trials signaled by one tone sequence, the CS-acceptance, and 10 rejection trials signaled by the other tone sequence, the CS-rejection (tones were counterbalanced across participants). Following this task, all participants performed a 2-minute distractor task, which consisted of a paper-and-pencil word-search puzzle.

Participants then performed a 96-trial lexical decision task in which they made word/nonword judgments of a series of targets. On each trial, the computer played one of the tone sequences and immediately after presented a letter string that was either a word or a nonword. The participant responded by pressing one of two keys as quickly as possible to indicate whether the target string was a word or nonword. This task was divided into three blocks. In one block, the presentation of the target string on each trial was signaled by the CS-acceptance, in another block the CS-rejection, and in another block a novel tone sequence (CS-control). The tones were characterized as orienting cues to help them stay focused on the task. The order of blocks was counterbalanced across participants. Within each block, there were four acceptance target words and four rejection target words. There were also four positive, but noninterpersonal, words (e.g., tranquil, amuse) and four negative but noninterpersonal words (e.g., slavery, decay) included as control stimuli. The 16 word trials were interspersed randomly with 16 nonword trials. All targets were presented only once to each participant; the targets assigned to each block were randomly selected for each participant from a longer list of targets (see Baldwin & Sinclair, 1996, for a more detailed description of the lexical decision procedure).

Following the lexical decision task, participants filled out a battery of questionnaires that were administered for exploratory purposes. They were then debriefed, paid, and thanked for their participation. Debriefing confirmed that although participants were aware of the contingency between the tones and the social feedback during the conditioning phase of the study, they soon realized that during the lexical decision task, each tone was equally likely to be followed by either an acceptance or rejection target word and so they were not suspicious about the use of the tones as signals during the task.

RESULTS AND DISCUSSION

Participants gave an incorrect response or took longer than the 2 seconds allowed on an average of 1.63 (out of 48) word trials. These error trials were discarded and mean RTs for each stimulus condition (i.e., CS-acceptance, acceptance words; CS-acceptance, rejection words; and so on) were calculated based on correct responses.

The interpersonal targets were analyzed in a 2 (tone condition: CS-acceptance vs. CS-rejection) \times 2 (target words: acceptance vs. rejection) within-participants ANOVA. As predicted, the only significant effect was the interaction between the CS condition and the nature of the target words, $F(1, 11) = 7.69$, $p = .02$ (note that this effect was unchanged if RTs¹ in the corresponding CS-control condition, or those for noninterpersonal targets, were included as covariates, and this interaction effect was not significant in a similar analysis of the noninterpersonal targets, $F < 1.5$). As predicted, a planned comparison showed that rejection words were identified 90 milliseconds more quickly after the presentation of the CS-rejection ($M = 697.05$, $SD = 131.02$) than after the presentation of the CS-acceptance ($M = 787.77$, $SD = 154.61$), $t(11) = 2.86$, $p < .01$, one tailed.² Contrary to predictions, although the means were 9 ms in the anticipated direction, acceptance targets were not significantly more quickly identified after the CS-acceptance ($M = 705.51$, $SD = 171.40$) than after the CS-rejection ($M = 714.65$, $SD = 176.04$), $t < 1$.

As predicted, then, the lexical decision task revealed the cued activation of interpersonal knowledge as a result of a brief conditioning procedure. Other social psychological work on human conditioning (e.g., Krosnick, Betz, Jussim, & Lynn, 1992; Murphy & Zajonc, 1993) and affective associations (e.g., Fazio, Sanbonmatsu, Powell, & Kardes, 1986; Greenwald, McGhee, & Schwartz, 1998) often has been limited to examining the overall valence of affective responses. The lexical decision task allows the researcher to present target stimuli that are more specific in their social content, in this case relating to rejection and acceptance; it was only on these targets, and not the noninterpersonal

targets, that effects were observed. Since Tolman's (1932; see also Bolles, 1972; Dickinson, 1989) seminal article, associative learning in the type of conditioning paradigm used here often has been characterized as the generation of expectancies about which type of experiences tend to follow certain events. This analysis fits well with our general assumption that activation of various cues can spread to representations of social outcomes (see Baldwin & Sinclair, 1996, for further discussion). In this study, the effect was accounted for primarily by rejection targets, indicating that what was being learned was the presence or absence of a negative outcome (i.e., rejection) rather than the presence or absence of a positive outcome (i.e., acceptance). This finding is consistent with the notion that anxiety—in this case, social anxiety—is the emotional response to the anticipation of a negative outcome—in this case, social rejection (e.g., Higgins, 1987).

MAIN STUDY

In the main study, we examined the impact of the cued activation of interpersonal knowledge on an ongoing interaction. Past research has shown that an interaction with a self-confident member of the other sex can serve as a powerful manipulation of social anxiety (Burgio et al., 1986; Melchoir & Cheek, 1990). Because of a gender difference observed in some previous cued-activation research (Baldwin, Granzberg, Pippus, & Pritchard, 2001), and also consistent with earlier social anxiety research focusing on women (e.g., Alden et al., 1992; DePaulo, Epstein, & LeMay, 1990), we studied the impact of the cues while female participants conversed with a male confederate. We hypothesized that those women who interacted while the cue for acceptance played would rate themselves as less anxious than those who interacted while the cue for rejection played. Based on Schlenker and Leary's (1982) two-factor model, however, we also anticipated that the activation effect would be qualified by an interaction between activation condition and level of premeasured self-consciousness: We predicted that highly self-conscious participants would be most affected by the evaluative cues.

Undergraduate women first underwent the conditioning procedure described in Study 1. Shortly thereafter, they had a 5-minute interaction with a male confederate who acted rather cool and aloof while one of the CS tones played in the background. We predicted that participants would be more comfortable if the CS-acceptance was playing and more anxious if the CS-rejection was playing. This effect was expected to be especially pronounced for high-self-conscious individuals, who tend to be chronically focused on themselves and the impression they are making on others.

METHOD

Participants

Fifty-six female introductory psychology students at the University of Manitoba served as participants in the experiment, receiving course credit for their participation. In a series of mass-testing sessions approximately 4 months earlier in the term, they completed the Self-Consciousness Scale (Fenigstein et al., 1975) in a questionnaire packet that included other measures not used in this study. Data were discarded from 3 participants who did not fill out the pretest questionnaire.

Procedure

Participants were given the same instructions, and underwent the same conditioning procedure, as described in the pilot study. After this conditioning phase, the participant moved to a different desk on the other side of the room and completed a 5-minute filler task of word puzzles. During this time, the female experimenter worked at the computer, ostensibly developing a new program.

Participants were then informed that they would engage in a brief conversation with a male experimenter, with the only restriction on their conversation being that they were not allowed to talk about the experiment. The experimenter then left the room to summon him. After a short period of time, the computer she had been working on began repeatedly, at 5-second intervals, emitting one of three tone sequences: one of the two tones from the bogus attitude questionnaire or else a third novel tone as a control condition. These three conditions represent the activation manipulation.

The male experimenter, who was actually a well-dressed senior undergraduate confederate, then arrived to carry on the 5-minute conversation. Following the procedure used by Stopa and Clark (1993; see also Alden et al., 1992; Burgio et al., 1986) to induce social anxiety, the confederate was instructed to behave in a reserved, but not unfriendly, way, allowing uncomfortable pauses to occur and not introducing new topics of conversation unless there was a pause of longer than 30 seconds. During the conversation, only a few participants mentioned the tones from the computer on the other side of the room; the confederate remarked that the experimenter must have been working on it. Most participants said later that they noticed the tones but tried to ignore them in order to focus on keeping the conversation going. As intended, participants generally reported that the conversation was rather uncomfortable.

Following the conversation, participants were asked to complete a number of dependent measures. The first measure was a mood scale. This 43-item scale included 20 adjectives from the Positive and Negative Affect

Schedule (Watson, Clark, & Tellegen, 1988), 15 self-evaluative mood items used by McFarland and Ross (1982), and 8 items drawn from the anxiety subscale of the Multiple Affect Adjective Check List (Zuckerman & Lubin, 1965). The second measure was the 23-item Behaviour Checklist (Stopa & Clark, 1993), which asked participants to rate their behavior during the interaction on items such as nervous, confident, blushing, self-assured, and awkward. Next, participants completed the State Self-Esteem Scale, which contains items such as, "I feel inferior to others at this moment" and "I feel good about myself" (Heatherton & Polivy, 1991). Finally, participants completed a reflected-appraisal measure devised by Ryan, Plant, and Kuczkowski (1991). They rated how they thought the confederate would rate them on 14 adjectives (interesting, active, stupid, unfriendly, happy, passive, boring, intelligent, sad, friendly, shy, confident, insecure, and outgoing). After participants completed these questionnaires, they were debriefed by the female experimenter and thanked for their participation. No participant expressed suspicion about the presence of the tones during the interaction.

The confederate also completed two questionnaires following each interaction with participants. First, he completed a parallel form of the Behavior Checklist (Stopa & Clark, 1993), rating the participant's behavior during the interaction. Next, he completed the same reflected appraisal measure that participants completed (Ryan et al., 1991), except the confederate completed this scale according to how he saw the participant. The confederate was blind throughout to which tone sequence had been paired with which type of feedback for each participant.

RESULTS

Consistent with the analytic approach used in other social anxiety studies (e.g., Burgio et al., 1986; DePaulo et al., 1990) and to allow comparability with experimental manipulations of self-focus (e.g., Alden et al., 1992; Burgio et al., 1986; Fenigstein, 1979), participants were designated as high or low in self-consciousness following a median split procedure.

Initial analyses revealed that the self-report measures of mood, behavior, self-esteem, and reflected appraisals were highly intercorrelated, average pairwise $r = .70$. A self-report index was therefore calculated by first coding all measures such that higher numbers represented positive ratings (i.e., meaning less anxious, more poised, higher in state self-esteem, and with a more positive expectation of the confederate's appraisal) and then standardizing and summing across measures. This index ($\alpha = .86$) was then analyzed using a 3 (condition: control,

CS-acceptance, or CS-rejection) \times 2 (self-consciousness: high vs. low) analysis of variance.

The activation manipulation affected participants' self-reported comfort during the conversation, as shown by a significant main effect for activation condition, $F(2, 47) = 4.23, p < .05$. Self-consciousness also played a role in social anxiety reactions because, not surprisingly, high-self-conscious individuals reported less comfort during this awkward interaction than their low-self-conscious counterparts, $F(1, 47) = 12.25, p < .001$. These main effects were qualified by the predicted interaction between activation condition and self-consciousness, $F(2, 47) = 6.30, p < .01$. Planned comparisons showed that, as depicted in Figure 1, high-self-conscious individuals reported less comfort if the CS-rejection tone was playing ($M = -4.77, SD = 2.64$), compared with the control condition ($M = -1.97, SD = 2.37$), $t(47) = 1.86, p < .05$. High-self-conscious individuals also reported greater comfort if the CS-acceptance tone was playing ($M = 1.87, SD = 2.13$), compared with controls, $t(47) = 2.84, p < .01$; indeed, in this condition, their ratings were at least as positive as those of low-self-conscious individuals. For their part, low-self-conscious individuals were not affected by the activation conditions, $t_s < 1, ns$, reporting fairly high levels of comfort across the board. This critical interaction effect, whereby high-self-conscious individuals were affected by the activation manipulation but low-self-conscious individuals were not, also was significant across the four univariate analyses (see Table 1).

Confederate's Ratings

We were interested in whether participants' insecurity would be apparent to their interaction partner or would be limited to internal states. The two sets of ratings filled out by the confederate were highly correlated, $r(52) = .87, p < .001$, and therefore were combined into a single index. There were no main effects on this index; however, the interaction term was marginally significant, $F(2, 48) = 2.60, p = .085$. As can be seen in Table 1, the interaction was significant for the confederate's ratings of how poised the participant's interactive behaviors were. Comparisons of cell means showed no significant differences on either the index or the univariate behavior ratings. The significant interaction effect, therefore, was not attributable to one specific comparison of an experimental group with its control condition. The overall pattern, however, was generally similar to that of the participants' self-ratings, and the two indices were significantly correlated, $r(51) = .58, p < .001$. Thus, the effects of cued activation and self-consciousness on participants' social anxiety levels were evident in their behavior, as observed by their interaction partner, as well as in their internal state.

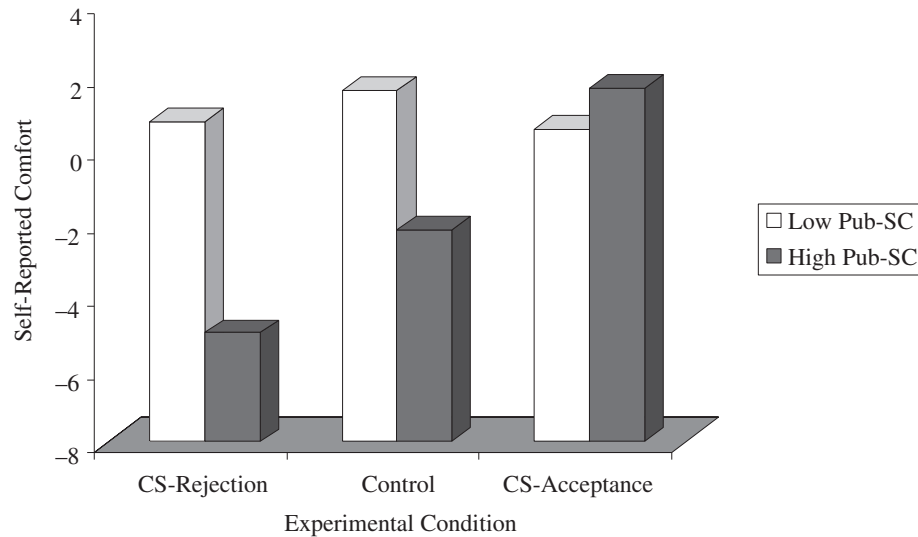


Figure 1 Index of self-reported comfort by activation condition and public self-consciousness.

NOTE: CS-Rejection = conditioned stimulus for rejection, CS-Acceptance = conditioned stimulus for acceptance, Low Pub-SC = low public self-conscious, and High Pub-SC = high public self-conscious.

TABLE 1: Self-Reports and Confederate Ratings as a Function of Activation Condition and Level of Self-Consciousness

Measure	Activation Condition						F Values		
	CS-Rejection		Control		CS-Acceptance				
	Lo-SC	Hi-SC	Lo-SC	Hi-SC	Lo-SC	Hi-SC	Condition	Self-Consciousness	Interaction
Self-reports									
Mood									
<i>M</i>	153.91	126.83	157.10	146.75	153.13	158.70	4.23**	6.45**	5.56***
<i>SD</i>	17.38	11.25	15.14	13.48	13.99	10.52			
Behavior									
<i>M</i>	112.82	66.67	117.03	88.22	116.21	114.61	3.65**	11.91***	3.17**
<i>SD</i>	25.40	13.26	28.78	25.67	28.37	26.77			
Self-esteem									
<i>M</i>	76.36	58.67	81.40	65.00	74.63	74.22	1.77	15.07****	3.63**
<i>SD</i>	10.48	13.41	6.15	12.97	9.29	10.89			
Reflected appraisals									
<i>M</i>	68.91	51.33	70.20	56.78	66.76	74.89	2.53*	4.24**	4.58**
<i>SD</i>	16.19	10.17	14.40	8.71	13.66	12.60			
Confederate ratings									
Behavior									
<i>M</i>	113.64	95.83	113.60	101.33	97.38	113.10	.02	.57	3.17**
<i>SD</i>	23.70	11.11	22.21	13.69	20.15	26.10			
Evaluative appraisals									
<i>M</i>	71.57	63.50	71.00	66.78	63.13	71.70	.03	.07	1.81
<i>SD</i>	14.78	9.05	12.05	9.81	15.64	15.89			

NOTE: Higher numbers represent more positive scores on all measures. CS-Rejection = conditioned stimulus for rejection, CS-Acceptance = conditioned stimulus for acceptance, Lo-SC = low self-conscious, and Hi-SC = high self-conscious.

* $p < .01$. ** $p < .05$. *** $p < .01$. **** $p < .001$.

DISCUSSION

The results across a number of measures demonstrated the impact of knowledge activation during interpersonal interactions. Activated relational knowledge and level of self-consciousness combined together to

determine levels of social anxiety, as predicted on the basis of Schlenker and Leary's (1982) two-factor model. In particular, high-self-conscious individuals rated their behavior and mood more negatively in the CS-rejection condition and more positively in the CS-acceptance condition, with the control condition in between; low-self-conscious individuals showed little if any impact of the

activation manipulation, reporting minimal anxiety irrespective of experimental condition.

For high-self-conscious individuals, then, the mere sound of a computer tone on the other side of the room had an impact on how they felt and acted when meeting a stranger. For these participants, concerned with the impression they were making, application of activated relational knowledge to their current situation had an effect on their mood, self-esteem, and poise during an initial encounter. Thus, as proposed by Schlenker and Leary (1982), social anxiety was most evident for people who tended to be focused on the impression they were making and had negative expectancies about how they would be received by their interaction partner. By contrast, in the condition where the manipulation activated positive expectancies, social anxiety was absent. As in previous research using experimental manipulations of self-focus, there was little impact of the activation manipulation on low-self-conscious individuals. These people are not inclined to focus on or be concerned about the impression they are making on others and therefore did not show changes in their typically low level of anxiety regardless of activation condition.

Although the planned contrasts revealed effects for both the CS-rejection and CS-acceptance conditions, compared to the control condition, casual examination of the univariate means in Table 1 indicates that the effect was somewhat more robust for the CS-acceptance. When one takes into account the finding of the pilot study that the lexical decision effects were mostly observed on rejection targets, it suggests the intriguing possibility that when people are in an anxiety-producing situation, they are most strongly influenced by cues signaling safety from negative outcomes (see, e.g., Seligman & Binik, 1977). A similar idea has been suggested in the adult attachment literature by Mikulincer and Arad (1999) in their discussion of the "secure base" function served by attachment working models. Future research is required to determine conditions in which acceptance or rejection cues will have more impact.

One caveat is in order regarding the generalizability of these findings to men as well as women. We elected to study women in this experiment partly because in some of our previous test-anxiety research with the conditioning procedure (Baldwin et al., 2001), we have found a gender difference such that whereas women showed straightforward activation effects, men occasionally did not or even showed opposite effects indicating defensiveness. Although the possibility therefore exists that our current findings are not relevant to social anxiety in men, we suspect that gender differences may have to do with the content of social evaluations rather than the basic processes involved. That is, one study by Baldwin et al. found that the gender difference in the test-anxiety

studies was due to different emphases placed by men and women on agency and performance as a source of self, and presumably social, evaluation. Therefore, in the current social interaction situation, we expect that the conditioning procedure would have similar effects on men and women so long as the social feedback was based on the kinds of self-aspects (e.g., attitudes, values, performances) that the participant believes typically lead to acceptance or rejection. The impact of specific social expectancies on social anxiety is a topic deserving of additional research.

An alternative account of our findings follows from previous research and clinical observation that has indicated that trait anxiety is correlated with the speed of acquisition and generalization of punishment expectancies (e.g., Eysenck, 1965; Wenar, 1954; Zinbarg & Mohlman, 1998). Perhaps, then, only the highly self-conscious individuals learned the contingency, which is why only they showed the impact of the cue. This likely does not apply to the current findings, however. The Self-Consciousness Scale (Fenigstein et al., 1975) has two other subscales in addition to the public self-consciousness measure studied here. Exploratory analyses based on both private self-consciousness and chronic social anxiety—the latter being the most direct indicator of anxiety—showed some main effects on the dependent measures (e.g., chronically socially anxious people reported lower state self-esteem) but no significant interaction effects involving the conditioning manipulation. Thus, only public self-consciousness, the indicator of impression motivation, interacted with the activation of social expectancies.

GENERAL DISCUSSION

The results of the two studies attest to the value of taking a social cognitive view of social anxiety and relational phenomena in general. The fact that the tones produced any effects at all, particularly on social anxiety and interactive behavior, demonstrates a number of important points. First, it demonstrates the profound tendency of people to attend to and learn the contingencies of interpersonal acceptance and rejection (e.g., Safran, 1990; Sullivan, 1953). The tone sequences were associated to acceptance and rejection through just 10 trials each, and counterbalancing ensured that there was nothing about the tones per se that could account for the findings.

Second, the results illustrate the powerful impact of activated relational knowledge. Often, problems such as social anxiety, as well as related characteristics such as low self-esteem and insecure attachment, are considered to be rather immutable traits that are learned early in life and inevitably expressed in all later relationships. Although individual differences undoubtedly provide a

backdrop of chronic tendencies, the current findings along with related results in the self-esteem and attachment domains have shown that the momentary, temporary accessibility of relational knowledge can have a profound effect on perceptions, feelings, and interpersonal behaviors. As mentioned earlier, we know that direct priming manipulations involving reminding an individual of a critical or accepting significant other can produce shifts in self-evaluation and interaction intentions (e.g., Baldwin et al., 1996; Baldwin & Holmes, 1987; Baldwin & Sinclair, 1996). Another line of research by Andersen and colleagues (see Andersen & Berk, 1998) on a social cognitive interpretation of transference has shown that activation of a schema often occurs because a stranger shares some features with a significant other that serve as cues or triggers for the structure. Based on these minimal cues, participants often go beyond the information given to assume that the new person will have a range of other characteristics associated with the significant other: They even expect this new person to be accepting or rejecting, or pleasant versus unpleasant to interact with, to the same degree as the significant other. The current findings show that the trigger for such activation need not be a feature of the new person but rather can be a relatively incidental environmental cue. In this case, a computer tone activated relational expectancies to influence people's mood and self-esteem, their reflected appraisals of how they thought their interaction partner saw them, and even the way their partner did in fact see them. These kinds of effects of activated structures are commonplace in the social cognitive literature and are highly consistent with the cardinal features of social anxiety, such as the perceptions of ambiguous feedback as rejecting (e.g., Pozo, Carver, Wellens, & Scheier, 1991) or beliefs that others hold unreachable standards (e.g., Wallace & Alden, 1991). The finding that such social expectancies can be activated by minimal environmental cues underscores the value of taking a social cognitive approach to studying the mechanisms underlying both temporary and chronic tendencies to experience social anxiety (see, e.g., Baldwin & Fergusson, *in press*).

Finally, social anxiety causes significant interpersonal distress for a large population, and our findings are relevant to the issue of change in relational knowledge activation. We assume (see, e.g., Baldwin et al., 1996; Baldwin & Fergusson, *in press*) that people generally have multiple models of different kinds of relationships and interactions, which can be cued by all manner of triggers. Therapy often is directed at detecting and modifying those triggers. The treatment of social anxiety has varied extensively over the decades, but some combination of cognitive and behavioral therapies is often used.

A key ingredient in most therapies is exposure (Edelmann, 1992; Heimberg & Barlow, 1991; Taylor & Arnou, 1988), in which clients gradually experience each of several feared situations in the absence of aversive consequences. These treatments are partially based on the notion of redefining the cues in the situation that become associated with anxiety, as well as modifying the specific cognitions that commonly occur in anxiety-provoking situations. Recent work on the mechanisms of exposure therapies and extinction (see, e.g., Bouton, 1991), however, has shown that it is not enough to learn a new way of seeing things; for example, social interactions do not always lead to failure or social mistakes or awkwardness do not always lead to rejection and humiliation. What also must happen in a successful intervention is that the newly learned, positive structures are more likely to be activated during a social interaction rather than the old, dysfunctional structures. The important factor, then, is activation—what script or relational schema gets activated at the time of performance. Thus, as Brewin (1989) pointed out, research into therapeutic change needs to focus on changes in knowledge activation and the cues that can trigger positive rather than negative expectations. The current paradigm could prove quite useful in this task. We would not advocate a direct application to therapy: Social success would hardly be facilitated by the periodic sound of a doorbell tone emanating from a handheld computer. Techniques could be developed, however, to try to reconfigure the relational knowledge that is activated by particular situations or particular sensations (see, e.g., Baldwin & Fergusson, *in press*).

In the broader scheme of things, the current research is in part a response to calls for a truly social cognition that focuses on cognition about interpersonal relationships rather than social objects such as self and other in isolation (Baldwin, 1992, 1995; Fiske & Haslam, 1996). It also fits with the agenda of pursuing more experimental rather than correlational research in the field of close relationships (e.g., Clark & Reis, 1988). The results demonstrate that it is possible to manipulate social cognitive mechanisms in a relatively subtle manner via environmental cues to produce fairly dramatic emotional and even interpersonal effects. If our interactions and relationships are shaped to this degree by shifts in knowledge activation, future research should examine in detail the principles that determine which social knowledge structures become cued for which individuals in which situations.

NOTES

1. The 2-s cutoff for reaction times was used to avoid skewness in the data, which is common if reaction times are not limited in this way. To

control for any remaining skewness, the critical analyses were redone following a log transformation of the raw reaction times (RTs). The interaction was slightly weaker, $F(1, 11) = 4.16, p = .066$, but the critical contrast comparing rejection targets in the two priming conditions remained significant, $t(11) = 2.60, p < .02$.

2. All planned comparisons between means are one-tailed tests.

REFERENCES

- Alden, L. E., Teschuk, M., & Tee, K. (1992). Public self-awareness and withdrawal from social interactions. *Cognitive Therapy and Research, 16*, 249-267.
- Andersen, S. M., & Berk, J. S. (1998). Transference in everyday experience: Implications of experimental research for relevant clinical phenomena. *Review of General Psychology, 2*, 81-120.
- Baldwin, M. W. (1992). Relational schemas and the processing of social information. *Psychological Bulletin, 112*, 461-484.
- Baldwin, M. W. (1994). Primed relational schemas as a source of self-evaluative reaction. *Journal of Social and Clinical Psychology, 13*, 380-403.
- Baldwin, M. W. (1995). Relational schemas and cognition in close relationships. *Journal of Social and Personal Relationships, 12*, 547-552.
- Baldwin, M. W., Carrell, S. E., & Lopez, D. F. (1990). Priming relationship schemas: My advisor and the pope are watching me from the back of my mind. *Journal of Experimental Social Psychology, 26*, 435-454.
- Baldwin, M. W., & Fergusson, P. (in press). Relational schemas: The activation of interpersonal knowledge structures in social anxiety. In R. Crozier & L. Alden (Eds.), *International handbook of social anxiety*. Sussex, UK: Wiley.
- Baldwin, M. W., Granzberg, A., Pippus, L., & Pritchard, E. T. (2001). *Cued activation of relational schemas: Self-evaluation and gender effects*. Manuscript submitted for publication.
- Baldwin, M. W., & Holmes, J. C. (1987). Salient private audiences and awareness of the self. *Journal of Personality and Social Psychology, 52*, 1087-1098.
- Baldwin, M. W., Keelan, J. P. R., Fehr, B., Enns, V., & Koh-Rangarajoo, E. (1996). Social cognitive conceptualization of attachment working models: Availability and accessibility effects. *Journal of Personality and Social Psychology, 71*, 94-104.
- Baldwin, M. W., & Meunier, J. (1999). The cued activation of attachment relational schemas. *Social Cognition, 17*, 209-227.
- Baldwin, M. W., & Sinclair, L. (1996). Self-esteem and "if . . . then" contingencies of interpersonal acceptance. *Journal of Personality and Social Psychology, 71*, 1130-1141.
- Bargh, J. A., & Ferguson, M. J. (2000). Beyond behaviorism: On the automaticity of higher mental processes. *Psychological Bulletin, 126*, 925-945.
- Bolles, R. C. (1972). Reinforcement, expectancy, and learning. *Psychological Review, 95*, 394-409.
- Bouton, M. E. (1991). A contextual analysis of fear extinction. In P. R. Martin (Ed.), *Handbook of behavior therapy and psychological science* (pp. 435-453). New York: Pergamon.
- Brewin, C. R. (1989). Cognitive change processes in psychotherapy. *Psychological Review, 96*, 379-394.
- Burgio, K. L., Merluzzi, T. V., & Pryor, J. B. (1986). Effects of performance expectancy and self-focused attention on social interaction. *Journal of Personality and Social Psychology, 50*, 1216-1221.
- Clark, M. S., & Reis, H. T. (1988). Interpersonal processes in close relationships. *Annual Review of Psychology, 39*, 609-672.
- DePaulo, B. M., Epstein, J. A., & LeMay, C. S. (1990). Responses of the socially anxious to the prospect of interpersonal evaluation. *Journal of Personality, 58*, 623-640.
- Dickinson, A. (1989). Expectancy theory in animal conditioning. In S. Klein & R. R. Mowrer (Eds.), *Contemporary learning theories: Pavlovian conditioning and the status of traditional learning theory* (pp. 279-308). Hillsdale, NJ: Lawrence Erlbaum.
- Edelmann, R. J. (1992). *Anxiety: Theory, research and intervention in psychology*. West Sussex, UK: Wiley.
- Eysenck, H. J. (1965). Extraversion and the acquisition of eye blink and GSR conditioned responses. *Psychological Bulletin, 63*, 258-270.
- Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On automatic activation of attitudes. *Journal of Personality and Social Psychology, 50*, 229-238.
- Fenigstein, A. (1979). Self-consciousness, self-attention, and social interaction. *Journal of Personality and Social Psychology, 37*, 75-86.
- Fenigstein, A., Scheier, M. F., & Buss, A. H. (1975). Public and private self-consciousness: Assessment and theory. *Journal of Consulting and Clinical Psychology, 43*, 522-527.
- Fiske, A. P., & Haslam, N. (1996). Social cognition is thinking about relationships. *Current Directions in Psychological Science, 143*-148.
- Greenwald, A. G., McGhee, D. E., Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality & Social Psychology, 74*, 1464-1480.
- Hartmann, L. M. (1983). A metacognitive model of social anxiety: Implications for treatment. *Clinical Psychological Review, 3*, 435-456.
- Heatherton, T. F., & Polivy, J. (1991). Development and validation of a scale for measuring state self-esteem. *Journal of Personality and Social Psychology, 60*, 895-910.
- Heimberg, R. G., & Barlow, D. H. (1991). New developments in cognitive-behavioral therapy for social phobia. *Journal of Clinical Psychiatry, 52*, 21-29.
- Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review, 94*, 319-340.
- Hope, D. A., & Heimberg, R. G. (1988). Public and private self-consciousness and social phobia. *Journal of Personality Assessment, 52*, 626-639.
- Ingram, R. E., & Kendall, P. C. (1987). The cognitive side of anxiety. *Cognitive Therapy and Research, 11*, 523-536.
- Krosnick, J. A., Betz, A. L., Jussim, L. J., & Lynn, A. R. (1992). Subliminal conditioning of attitudes. *Personality and Social Psychology Bulletin, 18*, 152-162.
- Leary, M. R., & Kowalski, R. M. (1995). *Social anxiety*. New York: Guilford.
- McFarland, C., & Ross, M. (1982). Impact of causal attributions on affective reactions to success and failure. *Journal of Personality and Social Psychology, 43*, 937-946.
- Melchoir, L. A., & Cheek, J. M. (1990). Shyness and anxious self-presentation during a social interaction [Special issue: Communication, cognition, and anxiety]. *Journal of Social Behavior and Personality, 5*, 117-130.
- Meyer, D., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology, 90*, 227-234.
- Mikulincer, M., & Arad, D. (1999). Attachment working models and cognitive openness in close relationships: A test of chronic and temporary accessibility effects. *Journal of Personality & Social Psychology, 77*, 710-725.
- Murphy, S. T., & Zajonc, R. B. (1993). Affect, cognition, and awareness: Affective priming with optimal and suboptimal stimulus exposures. *Journal of Personality and Social Psychology, 64*, 723-739.
- Pozo, C., Carver, C. S., Wellens, A. R., & Scheier, M. F. (1991). Social anxiety and social perception: Construing others' reactions to the self. *Personality and Social Psychology Bulletin, 17*, 355-362.
- Rogers, C. R. (1959). Therapy, personality and interpersonal relationships. In S. Koch (Ed.), *Psychology: A study of a science* (Vol. 3). Toronto, Canada: McGraw-Hill.
- Ryan, R. M., Plant, R. W., & Kuczkowski, R. J. (1991). Relation of self-projection processes to perform, emotion, and memory in a controlled interaction setting. *Personality and Social Psychology Bulletin, 17*, 427-434.
- Safran, J. D. (1990). Towards a refinement of cognitive therapy in light of interpersonal theory: I. Theory. *Clinical Psychology Review, 10*, 87-103.
- Schlenker, B. R., & Leary, M. R. (1982). Social anxiety and self-presentation: A conceptualization and model. *Psychological Bulletin, 92*, 641-669.

- Seligman, M., & Binik, Y. (1977). The safety signal hypothesis. In H. Davis & H. Hurwitz (Eds.), *Pavlovian-operant interactions* (pp. 165-187). Hillsdale, NJ: Lawrence Erlbaum.
- Stopa, L., & Clark, D. N. (1993). Cognitive processes in social phobia. *Behavior Research and Therapy*, 31, 255-267.
- Sullivan, H. S. (1953). *The interpersonal theory of psychiatry*. New York: Norton.
- Taylor, C. B., & Arnow, B. (1988). *The nature and treatment of anxiety disorders*. New York: Free Press.
- Tolman, E. C. (1932). *Purposive behavior in animals and men*. New York: Appleton-Century-Crofts.
- Wallace, S. T., & Alden, L. E. (1991). A comparison of social standards and perceived ability in anxious and nonanxious men. *Cognitive Therapy and Research*, 15, 237-254.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS Scale. *Journal of Personality and Social Psychology*, 54, 1063-1070.
- Wenar, C. (1954). Reaction time as a function of manifest anxiety and stimulus intensity. *Journal of Abnormal and Social Psychology*, 49, 335-340.
- Wicklund, R. A. (1975). Objective self-awareness. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 8, pp. 233-275). New York: Academic Press.
- Winton, E. C., Clark, D., & Edelmann, R. J. (1995). Social anxiety, fear of negative evaluation and the detection of negative emotion in others. *Behavior Research and Therapy*, 33, 193-196.
- Zinbarg, R. E., & Mohlman, J. (1998). Individual differences in the acquisition of affectively valenced associations. *Journal of Personality and Social Psychology*, 74, 1024-1040.
- Zuckerman, M., & Lubin, B. (1965). *The multiple affect adjective check list*. San Diego, CA: Educational and Industrial Testing Service.

Received August 9, 2000

Revision accepted January 14, 2001

Lessons From the Past: Do People Learn From Experience That Emotional Reactions Are Short-Lived?

Timothy D. Wilson

Jay Meyers

University of Virginia

Daniel T. Gilbert

Harvard University

Do people learn from experience that emotional reactions to events are often short-lived? Two studies indicate that it depends on whether the events are positive or negative. People who received positive or negative feedback on a test were not as happy or unhappy as they would have predicted. People in the positive feedback condition did not learn from this experience when making predictions about their reactions to future positive events. People in the negative feedback condition moderated their predictions about their reactions to future negative events, but this may not have been a result of learning. Rather, participants denigrated the test as a way of making themselves feel better and, when predicting future reactions, brought to mind this reconstruction of the test and inferred that doing poorly on it again would not make them very unhappy. Experience with a negative event (but not with a positive event) may improve the accuracy of one's affective forecasts, but the extent to which people learn from their affective forecasting errors may be limited.

“I’ll be elated for days if we beat the Rams,” a sports fan thinks, even though he or she should know better. Sporting events typically do not have long-lasting effects on people’s happiness. Wilson, Wheatley, Meyers, Gilbert, and Axson (2000), for example, assessed college football fans’ happiness 1 to 3 days after their favorite teams won or lost a game and found that the outcomes of the games had no detectable effect on their happiness; yet, the fans predicted that the outcome of the games would influence their happiness for several days.

This *durability bias*—the tendency to overestimate the duration of one’s emotional reactions to future events—has been found in a variety of settings with a variety of

populations (Gilbert, Pinel, Wilson, Blumberg, & Wheatley, 1998; Wilson et al., 2000). Curiously, it occurs in domains in which people have considerable past experience. Wilson et al.’s (2000) participants had watched many football games, and yet they overpredicted the impact of future ones on their happiness. Similarly, people expect new television sets, upgraded computers, and fancy cars to cause lasting pleasure when their previous wide-screen televisions, power notebooks, and luxury sedans did not. The purpose of the present studies was to investigate why people fail to generalize from their past experiences when making affective forecasts and the conditions under which they will generalize, at least to some extent.¹

On the face of it, it seems easy to learn from past emotional experiences. People simply need to recall how they felt after a similar event in the past, such as the fact that the thrill of their new 27-inch television lasted only a couple of weeks, and apply that knowledge to the future, predicting that the thrill of owning a new 32-inch model will probably wear off quickly as well. We believe, however, that generalization of this sort can be quite difficult.

Authors’ Note: We gratefully acknowledge the support of the National Institute of Mental Health (Research Grant No. RO1-MH56075) to Daniel T. Gilbert and Timothy D. Wilson. Correspondence concerning this article may be sent to Timothy D. Wilson, Department of Psychology, University of Virginia, P.O. Box 400400, Charlottesville, VA 22904-4400; e-mail: tdw@virginia.edu.

PSPB, Vol. 27 No. 12, December 2001 1648-1661

© 2001 by the Society for Personality and Social Psychology, Inc.

*Necessary Conditions to Learn
From Past Emotional Experiences*

To learn from their past emotional experiences, people would have to satisfy three conditions. First, they would need to make an effort to compare their past experiences with future ones, rather than simply thinking about a future event in isolation with no reference to the past (mental effort criterion). Buehler and McFarland (2001), for example, found that when people were induced to focus on a future event without thinking about similar past events, they made less accurate affective forecasts. Second, if people do make the effort to consult the past, they need to decide which of their many past experiences are most applicable to the future event they are considering (applicability criterion). Third, even if people recognize which events are applicable and make an effort to think about them, they need to recall or reconstruct accurately how they felt after those events (accuracy criterion).

We suggest that people often fail to satisfy one or more of these conditions. People often opt for simple mental strategies over effortful ones, suggesting that in many situations, they might fail to satisfy the mental effort criterion (e.g., Gilbert, 1991; Petty & Cacioppo, 1986). If people do think about the past, they might not recognize which events are most relevant to future ones, failing to satisfy the applicability criterion. Suppose that football fans are wondering how long they will be happy after their favorite team beats the Rams next month. Is the best predictor how happy they were after their team beat the Rams 2 years ago, how happy they were after their team beat the Lions last week, or how happy they were after their bowling team beat the Elks Lodge last night? No two events are identical and it is not always obvious which past events are most relevant to one's future reactions (Higgins, 1996; Koehler, 1996).

Even if people decide that a past event is applicable, they might not recall how happy they were after this event, failing to satisfy the accuracy criterion. There is evidence that people's memory for emotional reactions is poor. Christianson and Safer (1996) went so far as to conclude that "There are apparently no published studies in which a group of subjects has accurately recalled the intensity and/or frequency of their previously recorded emotions" (p. 235). People have been shown to have poor memory for the intensity or frequency of specific emotions such as anger and sadness (Levine, 1997), pain (Erskine, Morley, & Pierce, 1990), positive and negative moods (Thomas & Diener, 1990), and attitudes (Ross, 1989; Ross & Newby-Clark, 1998).

It thus seems unlikely that people will typically satisfy all three conditions necessary to learn from their past emotional experiences. There is reason to believe, however, that people learn different lessons from positive

and negative events. We argue that people often do not learn from positive events because they fail to meet one or more of the criteria for learning to occur. In contrast, people sometimes appear to learn from negative events, not necessarily because they meet the necessary criteria but because negative experiences change their construal of the event in ways that moderate their predictions for the future.

Failing to Learn From Positive Events

One reason people fail to generalize from positive events is that as time goes by, they fail to remember that these reactions were short-lived. Meyers, Wilson, and Gilbert (2000) and Mitchell, Thompson, Peterson, and Cronk (1997) found evidence for a retrospective durability bias, whereby people overestimated how happy they were after positive emotional events in the past. In one study, people interested in politics were surveyed at three points in time: a few weeks before the 1996 presidential election, right after the election, and 3 months later. Democrats showed a strong durability bias before the election, predicting that they would be much happier following Bill Clinton's victory than they in fact were. They also showed a retrospective durability bias, whereby 3 months after the election they recalled being happier following Bill Clinton's victory than they in fact had been. Similarly, participants in studies by Mitchell et al. (1997) predicted that enjoyable activities such as a bicycle trip would be significantly more enjoyable than they in fact were, and they remembered that the activities were more enjoyable than they in fact were.

Meyers et al. (2000) suggested that there is a common mechanism responsible for both the retrospective and prospective durability bias, namely, *focalism*, whereby people think too much about the event in question and fail to consider the consequences of other events that are likely to occur. When predicting how they will feel in the future after an emotional event, people think too much about that event and too little about other events that will influence their thoughts and feelings (Wilson et al., 2000). Similarly, when people try to reconstruct how happy they were in the past, they appear to focus too much on the event and not enough on other events that occurred at the time, thereby overestimating the emotional impact of the event. This retrospective durability bias helps to explain why people do not learn from their past experiences when predicting the future. One of the mechanisms that produces the durability bias (focalism) appears to operate in prospect and retrospect.²

Sometimes, of course, people think about the future right after experiencing an emotional event, when their reaction to it is fresh in their minds. Under these conditions, the accuracy criterion is likely to be satisfied; people can easily recall that their emotional reaction was rel-

actively short-lived ("Hmm, the new 27-inch television has not made me happy for as long as I expected") and thus might generalize from this fact when forecasting their future reactions ("I guess a new 32-inch model wouldn't make me happy for very long either").

Even if they can recall how they felt in the past, however, people still might fail to satisfy the mental effort criterion; that is, they might not focus on their past feelings and apply this knowledge to their predictions about the future. There is evidence that when people experience neutral or positive events, they are not very motivated to analyze the causes of their feelings (Hastie, 1984; Wong & Weiner, 1981). Generalizing from these findings, we hypothesized that when people experience a neutral or positive event and then are asked to predict how they will feel after similar events in the future, they will not engage in an effortful recall and application of their past reactions. Instead, they will focus on the future event and assume that it will have a long-lasting effect ("Wow, a new 32-inch TV, that would be great").

To test these hypotheses, we randomly assigned people to receive positive or no feedback on a test of social aptitude and then asked them to predict how happy they would be if they did well on similar future tests. Consistent with prior research on the durability bias, we expected that this feedback would not influence people as much as they expected. Because people made forecasts about future events soon after experiencing similar ones, they should be able to recall accurately that these events did not influence them very much. Even so, we hypothesized that people who received positive feedback would not generalize from this experience and would predict that similar positive events in the future would make them quite happy.

We do not mean to suggest that people will never recall and apply how they felt after past positive events. If the event that was experienced or the one to be predicted is extremely consequential or unexpected, people might well go to the effort of recalling and applying their past reactions. In the present studies, however, people were not expected to experience extremely discrepant reactions, and the events they predicted were not extremely consequential. Under these conditions, we suggest, people might not engage in the necessary mental effort to learn from their recent positive experiences.

Reconstructing Negative Events

In contrast, we predicted that after experiencing a negative event, such as getting a poor grade on a test, people would moderate their forecasts about how similar events would influence them in the future. One reason for such generalization might be that people genuinely learn from the fact that they did not have as negative a reaction as they expected and apply this

knowledge to the future. Negative experiences often have less effect on people's happiness than they expect; in fact, we have found that this absence of a negative effect is more unexpected to people than the absence of a positive effect (Meyers et al., 2000). Consequently, people might be more likely to think about why they did not feel as negatively as they expected, leading them to meet the necessary criteria to learn from experience: They think about their reaction to doing poorly on a test (mental effort); they recall accurately that they did not feel very upset (accuracy); and they apply this knowledge to the future, predicting that they will not feel upset if they do poorly on similar tests in the future (applicability).

Although people may well go through these steps under some circumstances, we suggest that there is a simpler, less effortful way that receiving negative feedback can moderate people's predictions about their future reactions. Such generalization could occur not because people satisfy the conditions necessary for genuine learning but because the negative experience triggers a reconstrual of the event. It is well known that people possess a powerful "psychological immune system" that speeds recovery from negative experiences (e.g., Festinger, 1957; Taylor, 1991). People rationalize and reconstrue negative events in ways that dampen the pain they cause. For rationalization to be most effective, it is helpful that people not be aware that they are doing it (Gilbert et al., 1998). One consequence of such immune neglect—the tendency to be unaware of the operation of one's own psychological immune system—is the durability bias. Because people do not recognize how much they will transform negative events psychologically in ways that ameliorate negative feelings, they overestimate how long it will take them to recover from the events (Gilbert et al., 1998).

We suggest that after a negative event has occurred, people's reconstrual of it also will influence their forecasts about how they will feel after similar, future events. That is, instead of consulting and learning from their prior reactions, people simply might bring to mind their new, benign construal of the experience and predict that such a mild event would not influence them very much. In other words, people might moderate their predictions not because they meet the criteria necessary to learn from experience but because their view of the test has changed.

To test this reconstrual hypothesis, we randomly assigned some participants to receive negative feedback on the test of social aptitude and asked them to predict how happy they would be if they did poorly on similar tests in the future. We hypothesized that people would rationalize their poor performance such that it did not make them as unhappy as they would have predicted

and, further, that this rationalization would cause them to predict that a future negative performance on the test would not make them very unhappy either.

We should note that the learning-from-experience and reconstrual hypotheses are not mutually exclusive. People might predict that a future negative performance will have little impact because (a) they remember that they did not feel badly after doing poorly the first time and apply this knowledge to the future and (b) their view of the test has changed, leading them to predict that doing poorly on an invalid test will not affect them very much. The goal of the present studies was to test the latter reconstrual view to see if this relatively simple process can account for changes in people's predictions about future negative events. We will return to the learning-by-experience hypothesis in the General Discussion and discuss evidence relevant to it.

In addition to asking people how happy they would be after experiencing similar events in the future, we asked them to predict how they would feel if they received positive or negative feedback in situations that were different from the test they had just taken (e.g., overhearing a neighbor complain that they were insensitive to other people). We predicted that the extent to which people who received negative feedback would generalize to these different situations would depend on the specificity of their rationalizations about their poor performance on the test. If people limit their rationalizations to devaluing the specific test, then they should generalize only to similar situations (e.g., taking similar tests in the future). If they make more general rationalizations, such as devaluing the test and deciding that social aptitude is not a very important skill, then they should generalize to both the similar and dissimilar future negative events (e.g., the neighbor's negative comment), predicting that all of them will have less impact than they will for those people who did not receive negative feedback.

STUDY 1: GENERALIZING FROM POSITIVE AND NEGATIVE EXPERIENCES

Method

OVERVIEW

People received positive, negative, or no feedback on a test of social aptitude and then predicted how happy they would be after receiving negative or positive feedback in a variety of future situations. Some of these situations were very similar to taking the social aptitude test, whereas others were quite different. The study thus employed a 3 (feedback: positive, negative, none) \times 2 (similarity of future situations: similar vs. dissimilar) \times 2 (valence of future events, positive vs. negative) design, with repeated measures on the last two factors. We hypothesized that people in the negative feedback con-

dition would predict that they would be less affected by future negative events than people in the positive and no feedback conditions predicted they would be. No differences between conditions were expected for future positive events; people in the positive feedback condition should say that they would be as happy after these events as people in the negative and no feedback conditions.

PARTICIPANTS

The participants were 128 undergraduates enrolled in psychology courses at the University of Virginia who received course credit for their participation. Before participating, people completed the Beck Depression Inventory (BDI) and a baseline measure of happiness. To avoid the possibility that negative feedback would be distressing to people who were depressed, people who rated their current happiness as 3 or less or scored 11 or more on the BDI were not included in the study. We ran these people in the no feedback condition so as not to single them out but did not include their data in the analyses. Twenty-two participants met these criteria. The final sample thus consisted of 106 participants (88 women and 18 men).

PROCEDURE

Individual participants were asked to take a test called the Virginia Social Aptitude Scale (V-SAS), which was ostensibly a newly developed test of "social aptitude . . . the ability to perceive and interpret the social behavior of others." We deliberately made the description of this trait somewhat ambiguous to make it easy for people who did poorly on the test to decide that it was not a very important skill. The experimenter mentioned that because many students had expressed an interest in knowing how well they had done on the test, they would receive feedback in the form of a letter grade. They were given a sheet that ascribed the following descriptors and percentiles to grades A through D: "A" or "Outstanding," 90th-100th percentile; "B" or "Good," 70th-89th percentile; "C" or "Fair," 50th-69th percentile; "D" or "Below Average," 30th-49th percentile.

Participants rated their current level of happiness ("In general, how happy would you say you are today, compared to how happy you are ON AVERAGE?" 1 = *below average happiness*, 5 = *average happiness*, 9 = *above average happiness*) and some filler personality scales (e.g., the Rosenberg Self-Esteem Scale). They then took the V-SAS, which involved looking at 40 black-and-white photocopied pictures of faces and guessing which emotion each person was expressing from a list of eight emotions and emotion blends. The faces, which were from Ekman and Friesen (1975), were pretested on an independent group of students who were asked to identify the emotion without the benefit of multiple-choice options. Based on the pilot results, we constructed options for

each face that were difficult. After 10 minutes, the experimenter collected the test and scored it in a different room.

Feedback manipulation. In the negative feedback condition, the experimenter scored the test by marking 19 of the 40 items incorrect and assigned it a grade of D. Based on pilot testing, the items that generated the most inconsistent responses were marked wrong. In the positive feedback condition, the experimenter marked 5 of the 40 items incorrect and assigned it a grade of A. In both conditions, she gave the participant the graded answer sheet and the description of the percentiles for each grade in a folder and said that she had to go make a photocopy of the next questionnaire. Participants in the no feedback condition were told that they would receive their grade at the end of the study.

DEPENDENT MEASURES

Happiness measures. Participants rated their happiness 5 minutes after receiving their feedback on the same scale as they had rated their baseline happiness. Most participants then filled out an indirect measure of their mood—the Associated Reasoning Scale (Mayer & Hanson, 1995). (We added this scale after the study had begun; 74 of the 106 participants completed it.) This scale asks people to rate the probability that positive and negative events will occur, such as an improvement in the economy or the divorce of a married couple, and to choose a sample member of various categories (e.g., “a type of worker”) from a list of positive and negative examples (e.g., conscientious vs. lazy). Previous studies have found that responses to this scale correlate reasonably well with self-reported mood and are sensitive to experimental manipulations of mood (e.g., Mayer & Hanson, 1995).

Before making their predictions about how happy they would be in future situations, it was important that participants’ current mood be equivalent in each condition (to avoid the possibility that their current moods would color their predictions about the future). We asked all participants to read articles that were intended to neutralize any lingering effects of the feedback. They were told that future versions of the V-SAS might include examining people’s reactions to written stimuli and were asked to read two articles and answer questions related to them. The first article was a consumer report about compact disc players, whereas the second article was a piece by Dave Barry called “Consumers From Mars” that discussed advertising and marketing in a humorous manner. After reading each article, participants answered filler questions about how well written, enjoyable, interesting, informative, and entertaining the article was and the extent to which it kept their attention. All participants then received a questionnaire on which they

again indicated their current happiness on the same happiness scale as before. This question was answered approximately 15 minutes after taking the V-SAS test. Next, participants filled out a second form of the Associated Reasoning Scale. The forms of this scale that people completed first and second were counterbalanced.

Prediction questions. Participants then received the main dependent measures, on which they predicted how happy they would be after receiving positive and negative feedback on a variety of tests and in other situations. They did so on the same 9-point scales on which they had rated their actual happiness earlier, where 1 = *below average happiness*, 5 = *average happiness*, and 9 = *above average happiness*. Participants were told to imagine the events occurred in 3 weeks and to judge each separately and independently of the others. They also were told that the abilities measured on the tests were not related to one another and that performance on one test was not predictive of performance on the other tests.

There were five similar events that involved taking another version of the V-SAS. The first test was described as identical to the one people had just taken except that it used a different set of black and white photographs. The next four were all subtests of the V-SAS that varied on at least one dimension from the subtest people had taken: one used high-quality color photos instead of black-and-white photos, one involved judging people’s personalities (instead of their emotions) from a set of color photos, one involved judging people’s emotions from recordings of their tone of voice instead of photographs, and one involved judging people’s personality from recordings of their tone of voice. All five were clearly labeled as subtests of the V-SAS. People predicted how happy they would be 5 minutes after receiving an A and a D on each test.

The four dissimilar events involved receiving positive or negative feedback in situations other than taking the V-SAS: overhearing a stranger in a restaurant comment how “sensitive to others” you are (positive) or overhearing the stranger complain how “insensitive of others you are” (negative); being told by a teaching assistant that your work shows great creativity (positive) or that your work lacks creativity (negative); hearing that your neighbor told your landlord that you are “kind and sensitive to everyone around you” or that your neighbor complained that you were “loud and insensitive to those around you”; and getting an A on a Shape Manipulation Test of Creativity (in which the task is to reconfigure shapes to solve various problems) or getting a D on the Shape Manipulation Test of Creativity. In each case, people rated how happy they would be 5 minutes after the positive and negative version of the event. Note that each of these events was similar in at least one way to the V-SAS; for example, some involved feedback on how sensi-

tive people were (the same trait measured by the V-SAS), whereas others involved feedback on tests of other abilities or in other academic situations. However, unlike the similar events, none involved taking a subtest of the V-SAS.

Rationalization measure. Participants then answered several questions designed to assess how much they had rationalized a poor performance on the V-SAS, including how valid the test was (1 = *not very valid*, 9 = *extremely valid*), its fairness (1 = *not very fair*, 9 = *extremely fair*), how easy it was to concentrate while taking the test (1 = *very difficult*, 9 = *very easy*), how easy it is to measure social aptitude (1 = *very difficult*, 9 = *very easy*), how important psychology experiments are (1 = *not very important*, 9 = *extremely important*), how much something unusual (such as how tired they were) influenced their test performance (1 = *not at all*, 9 = *very much*), their opinions of multiple-choice tests (1 = *very poor test format*, 9 = *very good test format*), and their physical health (1 = *not so great*, 9 = *just fine*). They also rated the importance of the following outcomes and abilities (all using a scale that ranged from 1 = *not very important* to 9 = *extremely important*): the importance of doing well on the face-reading test, face-reading ability, being a good face reader in the work place, being a good face reader when dealing with people, and being a good face reader in friendships. All participants were then fully debriefed.

Results and Discussion

People's initial level of happiness in the negative feedback, no feedback, and positive feedback conditions were 5.38, 6.00, and 5.82 (SD s = 1.06, 1.17, 1.24), respectively, $F(2, 103) = 2.79$, $p = .07$. The mean in the positive feedback condition did not differ significantly from the mean in the negative feedback condition, $F(1, 103) = 2.55$, $p = .11$. To control for individual differences in initial happiness, we subtracted people's baseline happiness reports from their happiness reports after taking the test.

EFFECTS OF FEEDBACK ON ACTUAL HAPPINESS

People reported their happiness 5 and 15 minutes after taking the test. At 5 minutes, there was a tendency for people in the positive feedback condition to be happier than people in the no feedback and negative feedback conditions (M s = .09, -.17, and -.19, respectively; SD s = .38, .70, .62, respectively). However, the main effect of feedback did not reach significance, $F(2, 103) = 2.41$, $p = .10$. At 15 minutes, the means were .12, -.22, and .00, respectively (SD s = .70, .72, .67, respectively), and again, the main effect of feedback was not significant, $F(2, 103) = 2.19$, $p = .12$. A 3 (feedback: negative, none, positive) \times 2 (time: 5 vs. 15 minutes) mixed-model ANOVA did not reveal any significant effects: main effect

of feedback, $F(2, 103) = 2.29$, $p = .11$; Feedback \times Time interaction, $F(2, 103) = 2.19$, $p = .12$.³

People's responses on the Associated Reasoning Scale (Mayer & Hanson, 1995) also indicated that the feedback had little, if any, effect on people's mood. There was neither a main effect of feedback nor a Feedback \times Time interaction on this measure, $F(2, 71) \leq 1$. There was a nearly significant effect of time, $F(1, 71) = 3.39$, $p = .07$, reflecting the fact that people's judgments were more positive at 5 than 15 minutes (M s = 23.16 and 21.39, SD s = 6.89 and 7.15). The absence of a feedback effect on the Associated Reasoning Scale is, of course, a null effect that should be interpreted cautiously, especially because there was a reduced sample size on this measure. It is worth noting, however, that other studies have found that this scale correlates reasonably well with self-reported mood and is sensitive to experimental manipulations of mood (e.g., Mayer & Hanson, 1995). In our study, the scales correlated significantly (albeit modestly) with self-reported happiness averaged across 5 and 15 minutes, $r(72) = .26$, $p = .02$.

THE PREDICTION DURABILITY BIAS

We could test for a prospective durability bias in the no feedback condition because these participants predicted how they would feel about getting an A or a D on the same test they had taken that day (with new photographs) without having received any prior feedback. As hypothesized, they predicted that they would be relatively happy if they got an A ($M = .78$, $SD = 1.31$), which was significantly higher than positive experiencers reported feeling after actually getting an A ($M = .09$, $SD = .38$), $t(67) = 2.89$, $p = .005$. Similarly, they predicted that they would be unhappy if they received a D ($M = -2.31$, $SD = 1.62$), which was significantly lower than negative experiencers reported feeling after actually getting a D ($M = -.19$, $SD = .62$), $t(71) = 7.42$, $p < .001$. No feedback participants, of course, were making predictions for a somewhat different situation (taking the test in the future for the second time) than people in the positive and negative feedback condition actually experienced. It is thus worth noting that the magnitude of their apparent durability bias was nearly identical to that found by Meyers et al. (2000) among forecasters who made predictions for the same situation that people in the positive and negative feedback conditions experienced. These results suggest that people's emotional reactions to doing well or poorly on the test did not last as long as they would have predicted.

RATIONALIZATION

An iterative principal axis factor analysis of the 13 items on the rationalization questionnaire, with a varimax rotation, yielded a sensible three-factor solution (with eigenvalues > 1). One factor consisted of four

items about the specific test people had taken that day (e.g., “How valid do you think the face-reading test of empathy is?”), a second factor consisted of four items about the general importance of social aptitude in different settings (e.g., “How important is being a good face reader when dealing with people?”), and a third factor consisted of three items about people’s physical and psychological state when they took the test (e.g., “How easy was it for you to concentrate while taking the test?”). Two items (how easy it is to measure social aptitude and their opinions of multiple-choice tests) did not load strongly on any of the three factors and were dropped from the analyses.⁴ People’s responses were reverse scored such that high numbers revealed more rationalization (e.g., more negative ratings of the validity of the test).

A series of ANOVAs revealed that negative experiencers had higher scores on each of the rationalization indices. They rated the specific test more negatively ($M = 5.43$, $SD = 1.11$) than people in the no feedback and positive feedback conditions ($M_s = 4.27$ and 4.42 , $SD_s = 1.32$ and 1.38), $F(2, 103) = 8.98$, $p < .001$. They said that social aptitude was a less important skill than did people in the no feedback and positive feedback conditions ($M_s = 2.78$ vs. 2.26 and 2.20 ; $SD_s = 1.32$, 1.10 , $.93$, respectively), $F(2, 103) = 2.83$, $p = .06$, and they also said that they were in a worse psychological and physical state while taking the test than did people in the no feedback and positive feedback conditions ($M_s = 4.01$ vs. 3.61 and 2.93 ; $SD_s = 1.87$, 1.64 , 1.58 , respectively), $F(2, 103) = 3.55$, $p = .03$. These results suggest that negative experiencers rationalized their poor performance broadly, derogating the specific test they had taken, downplaying the importance of social aptitude as a trait, and claiming that they were not in an optimal physical or psychological state while taking the test.

PREDICTIONS ABOUT REACTIONS TO FUTURE POSITIVE AND NEGATIVE EVENTS

People’s predictions for the five similar future events (i.e., the five V-SAS subtests) were highly correlated ($\alpha = .93$), as were their predictions for the four dissimilar similar events, ($\alpha = .90$). Therefore, we averaged their ratings within each category. As seen in Figure 1, our hypotheses were largely confirmed. First, as hypothesized, people who received positive feedback did not generalize at all to the future experiences. Their affective forecasts were very similar to the forecasts made by no feedback participants; a 2 (actual feedback: positive vs. no feedback) \times 2 (anticipated future feedback: positive vs. negative) \times 2 (situations: similar vs. dissimilar) ANOVA did not reveal any significant main effects or interactions involving actual feedback, $F_s(1, 67) < 1$. The only significant effects were an unsurprising main effect of anticipated future feedback, $F(1, 67) = 533.45$, $p <$

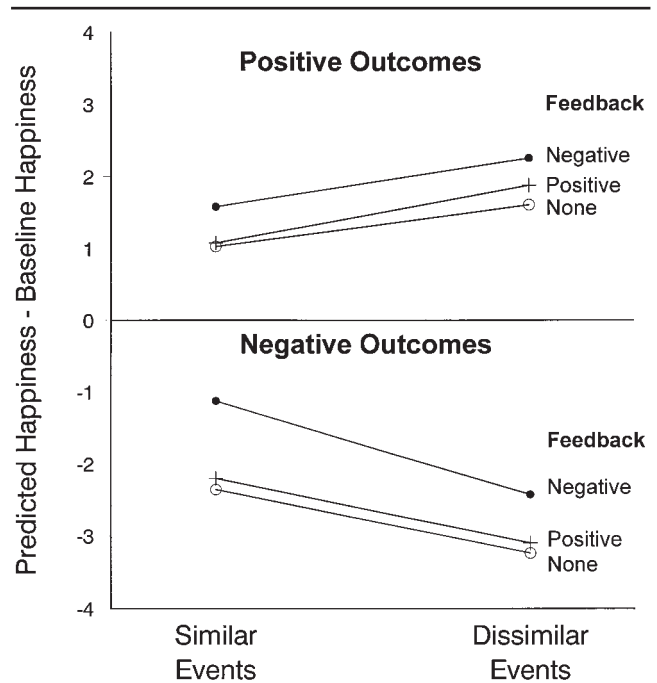


Figure 1 Study 1: Predicted happiness following similar and dissimilar future events.

.001, reflecting the fact that people predicted they would be happier if they received positive feedback in the future situations than if they received negative feedback, and an Anticipated Future Feedback \times Situations interaction, $F(1, 67) = 129.20$, $p < .001$, reflecting the fact that people predicted that the similar test situations would have less impact (e.g., less of a difference between receiving positive vs. negative feedback) than the dissimilar situations would (see Figure 1).

In short, as hypothesized, receiving positive feedback on the V-SAS test had no detectable effect on people’s predictions about how happy they would be in similar or dissimilar situations in the future. Another way of illustrating this is to examine people’s predictions about how they would feel if they got an A on the same test that they had just taken (this was the first situation they were asked to predict). Even though they had reported that they were only .09 ($SD = .38$) points above baseline 5 minutes after receiving an A on the test, they predicted that they would be .97 ($SD = 1.45$) points higher than their baseline level of happiness 5 minutes after getting an A on the same test in the future, $t(32) = 3.90$, $p < .001$.

Also as hypothesized, negative experiencers predicted that future negative feedback would not make them as unhappy as people in the no feedback condition said it would. A 2 (actual feedback: received a D vs. no feedback) \times 2 (anticipated future feedback: positive vs. negative) \times 2 (situations: similar vs. dissimilar) ANOVA revealed a nearly significant three-way interaction, $F(1,$

71) = 3.33, $p = .07$, which we decomposed by conducting separate 2 (actual feedback: received a D vs. no feedback) \times 2 (situations: similar vs. not similar) ANOVAs for predicted reactions to future positive versus negative events. For negative future events, there was a highly significant effect of actual feedback, $F(1, 71) = 10.66$, $p = .002$, reflecting the fact that people in the no feedback condition predicted that they would feel unhappier after future negative feedback than did negative experiencers ($M_s = -2.80$ vs. -1.78 ; $SD_s = 1.47, 1.19$). As expected, negative experiencers moderated their forecasts more for similar than dissimilar situations, as reflected by a significant Actual Feedback \times Situations interaction, $F(1, 71) = 3.90$, $p = .05$. However, the difference between the negative feedback and feedback conditions was significant for both the similar and dissimilar situations, $F_s(1, 71) = 13.47$ and 6.20 , $p_s < .001$ and $.03$, respectively. Thus, whereas negative experiencers moderated their forecasts more for similar than dissimilar future events, this moderation was significant in both cases.

As seen in Figure 1, negative experiencers did not generalize completely from their lack of reaction to doing poorly on the test. Their predicted happiness 5 minutes after getting a D on the same test in the future was $-.84$ ($SD = 1.26$), which was significantly lower than how happy they reported being 5 minutes after getting a D, $-.19$ ($SD = .62$), $t(36) = 3.82$, $p < .005$. Nonetheless, their predictions about how happy they would be after future negative events were more moderate than the predictions in the no feedback and positive feedback conditions, indicating that some generalization occurred.

For positive future events, people who received negative feedback made more positive predictions than did no feedback people; main effect of actual feedback, $F(1, 71) = 4.63$, $p = .04$. This finding is similar to a contrast effect that Meyers et al. (2000) found, whereby negative experiencers said they would have felt better if they had received an A on the test than positive experiencers reported feeling. After a negative experience, positive ones seem even sweeter. The Actual Feedback \times Situations interaction was not significant, $F(1, 71) < 1$.

MEDIATION ANALYSES

As hypothesized, people who received negative feedback were more likely to rationalize their performance by derogating the test, downplaying social sensitivity as a skill, and reporting that they were in a worse physical and psychological state. Also as hypothesized, they predicted that future negative feedback would not make them as unhappy as people in the other conditions said it would. The next step was to see if the changes in rationalization mediated the changes in predictions. First, we averaged the three rationalization scales (ratings of the test,

importance of social aptitude, ratings of current state) to create an overall index of rationalization. Second, because the effect of negative feedback changed people's predictions for similar and dissimilar negative situations (see Figure 1), we averaged the predicted happiness ratings over both types of situations to create an index of predicted happiness. Third, because we predicted that positive feedback would have no effect on rationalization or predictions about future negative situations, we combined the positive and no feedback group into a control condition (dummy-coded 0) and compared them to the negative feedback group (dummy-coded 1). We then conducted a mediation analysis on these variables (see Kenny, Kashy, & Bolger, 1998, for details). The effect of negative feedback on the mediator, rationalization, was significant ($\beta = .38$, $SE = .09$, $p < .001$). The relationship between rationalization and predicted happiness also was significant, after controlling for feedback ($\beta = .22$, $SE = .10$, $p = .03$). The direct effect of negative feedback on predicted happiness was significant ($\beta = .32$, $SE = .09$, $p = .001$), but was reduced significantly when the effect of rationalization was controlled ($\beta = .23$, $SE = .10$, $p = .02$, $z = 2.00$, $p < .05$). These results meet all of the conditions necessary to demonstrate that changes in rationalization mediated the effects of feedback on predicted happiness.

In sum, negative experiencers generalized to new, negative situations, whereas positive experiencers did not generalize to new, positive situations. Interestingly, negative experiencers generalized even to situations that were not very similar to the test they had just taken, such as overhearing a stranger say they were not very sensitive. People appear to have rationalized the negative feedback rather broadly, inferring not only that the V-SAS was an invalid test but that the trait of social aptitude was not very important. As a result, they predicted that receiving a critical comment about their sensitivity would not make them as unhappy as people in the positive feedback and no feedback conditions predicted it would.

If this interpretation is correct, then it should be possible to limit the scope of people's generalizations by narrowing the kinds of rationalizations they make for their poor performance. We tested this hypothesis in Study 2 by attempting to get some participants in the negative feedback condition to narrow their rationalizations to the invalidity of the V-SAS test. We predicted that these participants would moderate their predictions about how unhappy they would be after receiving negative feedback on future V-SAS tests but not their predictions about how unhappy they would be after receiving negative feedback in the dissimilar situations (e.g., overhearing a negative comment from a stranger).

STUDY 2: NARROWING THE SCOPE
OF PEOPLE'S RATIONALIZATIONS*Method**OVERVIEW*

We replicated the no feedback and negative feedback conditions in Study 1 with the addition of a manipulation designed to narrow the scope of people's rationalizations for negative feedback. The study thus employed a 2 (feedback: negative or none) \times 2 (rationalization: specific vs. undirected) \times 2 (similarity of future situations: similar vs. dissimilar) design. We expected to replicate Study 1 in the undirected rationalization condition; negative experiencers should moderate their forecasts for how unhappy they would be in both similar and dissimilar future situations. Negative experiencers in the specific rationalization condition were expected to moderate their forecasts only for how unhappy they would be in the similar situations.

PARTICIPANTS

The participants were 127 undergraduates enrolled in psychology courses at the University of Virginia who received course credit for their participation. Of these, 111 met the criteria for the study of scoring 10 or less on the BDI (74 women and 37 men).

PROCEDURE

The negative and no feedback conditions of Study 1 were replicated exactly except for the following changes: After they took the V-SAS, half of the participants were randomly assigned to the specific rationalization condition. They received a questionnaire whose purpose was supposedly to obtain critical feedback on the V-SAS. Its true purpose was to suggest to people in the negative feedback condition that their poor performance was due to the problems with the test. The first set of questions suggested that the poor quality of the photographs might have made it difficult to judge people's emotions. The instructions noted that there were two versions of the test: one that used "small, black-and-white photocopies of photographs that are somewhat grainy" and one that used "high-quality, 5- \times 7-inch color prints." The first question asked participants which set of photographs they had received (in fact, everyone had received the black-and-white photos). The second question asked how easy it was to detect people's emotions from the pictures they received. The third question noted that "some people have complained that it is especially difficult to detect blends of two or more emotions in the low-quality, black-and-white pictures" and asked people how easy it was for them to detect blends. People responded on 9-point scales (1 = *very difficult*, 9 = *very easy*). Three additional questions focused people's attention on other pos-

sible problems with the test; namely, it was too short, it used a multiple-choice format, and social aptitude is something that cannot be measured on a test. People responded to these questions on 9-point scales (1 = *disagree*, 9 = *agree*). People's ratings on these measures were reverse coded such that high scores reflected negative evaluations of the test. Participants were given 5 minutes to fill out the rationalization questionnaire. Those in the undirected rationalization condition sat quietly for 5 minutes.

Dependent measures. The dependent measures were identical to Study 1 except that we did not include the Associated Reasoning Scale. The prediction items were the same as in Study 1 except that people only rated the negative outcomes (e.g., getting a D on the tests or overhearing negative remarks). As a measure of how people construed the new situations, people also rated how valid each test or situation was (1 = *not very valid*, 9 = *extremely valid*) and how important the test or situation would be to them (1 = *not very important*, 9 = *extremely important*). People then were asked to recall what grade they had received on the V-SAS (all participants recalled their grade accurately), completed the same rationalization questionnaire as in Study 1, and were fully debriefed.

Results and Discussion

The baseline happiness ratings in the four cells of the Feedback \times Rationalization design did not differ significantly, $F_s(1, 106) < 1$. As in Study 1, we subtracted people's baseline happiness ratings from their happiness ratings 5 and 15 minutes after taking the test.

EFFECTS OF FEEDBACK ON ACTUAL HAPPINESS

Once again, the negative feedback had little or no effect on people's happiness, as indicated by the absence of any significant effects in a 2 (feedback: none vs. negative) \times 2 (rationalization: specific vs. undirected) \times 2 (time: happiness ratings 5 vs. 15 minutes after the test) mixed-model ANOVA, $F_s(1, 104) < 1.70$, $p_s > .19$. At 5 minutes, the means were $-.07$ and $-.04$ ($SD_s = .60, .76$) in the no feedback and negative feedback conditions; at 15 minutes, the means were $-.05$ and $.02$ ($SD_s = .84, .84$).

*PREDICTIONS ABOUT FUTURE REACTIONS TO
SIMILAR AND DISSIMILAR NEGATIVE EVENTS*

As in Study 1, we averaged people's happiness predictions for the five tests that were similar to the V-SAS they had just taken ($\alpha = .95$) and for the four situations that were dissimilar to the V-SAS ($\alpha = .91$), and analyzed these predictions with a 2 (feedback: none vs. negative) \times 2 (rationalization: specific vs. undirected) \times 2 (similarity: similar vs. different future events) mixed-model ANOVA. As hypothesized, the three-way interaction was significant, $F(1, 106) = 4.09$, $p = .046$.

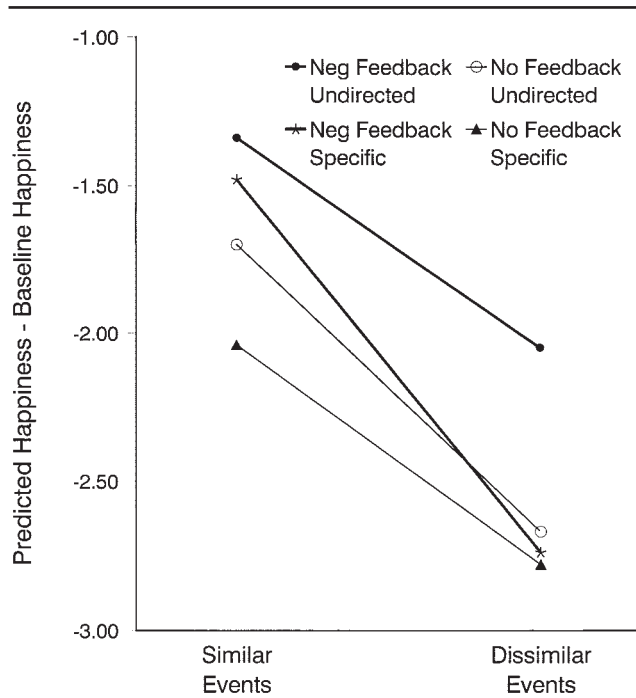


Figure 2 Study 2: Predicted happiness following similar and dissimilar future events.

As seen in Figure 2, the results in the undirected rationalization condition were similar to the results of Study 1. Negative experiencers again made more moderate predictions for both similar and dissimilar situations, compared to people in the no feedback condition. Although a 2 (actual feedback: negative vs. no feedback) \times (future event: similar vs. dissimilar) ANOVA in the undirected rationalization condition revealed that the main effect of feedback was not significant, $F(1, 55) = 1.44$, $p = .24$, this main effect was significant when combined with the identical conditions in Study 1 ($z = 3.06$, $p = .002$). The reliability of this effect did not differ significantly across studies, $\chi^2(1) = 1.91$, $p = .17$.

As hypothesized, people in the negative feedback-specific rationalization condition appear to have made more limited generalizations. Compared to their counterparts in the no feedback condition, they predicted that similar negative experiences would not make them as unhappy but that dissimilar negative experiences would make them just as unhappy. The Feedback \times Similarity of Event interaction was nearly significant in the specific rationalization condition, $F(1, 51) = 3.81$, $p = .057$.

RATIONALIZATION

In the specific rationalization condition, people evaluated the test right after taking it. If the manipulation was successful, then negative experiencers should have

taken this opportunity to be more critical of the test than people who did not receive any feedback as a means of rationalizing their poor performance. This was the case; negative experiencers had significantly more negative responses on these questions than people who received no feedback ($M_s = 6.86$ vs. 5.83 , $SD_s = 1.22$ and $.99$), $F(1, 50) = 11.31$, $p < .001$. In other words, the manipulation did not make all participants more critical of the test. Rather, people who received negative feedback were more likely to follow the lead of the questions and rate the test negatively.

At the end of the study, participants in all conditions completed the rationalization questionnaire used in Study 1. A factor analysis yielded a very similar solution (see Note 4). As in Study 1, negative experiencers rated the specific test more negatively ($M_s = 6.13$ vs. 4.62 , $SD_s = 1.35$ and 1.28), $F(1, 105) = 35.84$, $p < .001$; they tended to say that social aptitude was a less important skill ($M_s = 2.89$ vs. 2.44 , $SD_s = 1.57$ and 1.29), $F(1, 105) = 2.59$, $p = .11$, and that they were in a worse psychological and physical state while taking the test ($M_s = 3.55$ vs. 2.97 , $SD_s = 1.57$ and 1.23), $F(1, 105) = 4.57$, $p = .04$.

There were no significant main effects or interactions involving the rationalization manipulation on these indices. As hypothesized, however, negative experiencers in the undirected and specific rationalization conditions both derogated the specific test, but only those in the undirected condition derogated social aptitude as a skill. On the latter measure, negative experiencers in the undirected condition derogated social aptitude as a skill somewhat more than people in the no feedback condition ($M_s = 6.91$ vs. 7.57 , $SD_s = 1.83$ and 1.30), $F(1, 106) = 2.94$, $p = .09$ (this replicates the results of Study 1). In the specific rationalization condition, there was less of a tendency among negative feedback participants to derogate social aptitude as a skill ($M_s = 7.32$ vs. 7.57 , $SD_s = 1.22$ and 1.29), $F(1, 106) < 1$. A planned contrast that weighted the negative feedback/undirected rationalization -3 , and the other means 1 , was nearly significant, $F(1, 105) = 3.21$, $p = .07$.

People also rated the validity and importance of each of the future situations about which they made predictions. Because the validity and importance ratings were highly correlated ($.86$), we averaged them. Consistent with the hypothesis that negative experiencers rationalized their poor performance, this group rated the similar situations (the V-SAS tests) as less valid and important than did people in the no feedback condition ($M_s = 3.18$ vs. 3.90 , $SD_s = 1.26$ and 1.09). They also rated the dissimilar situations as less valid and important than did people in the no feedback condition ($M_s = 4.68$ vs. 4.88 , $SD_s = 1.33$ and 1.15), but to a lesser degree, as reflected by a significant Feedback \times Similarity of Events interaction, $F(1, 107) = 4.72$, $p = .03$.

The three-way interaction with the rationalization factor was not significant, $F(1, 107) < 1$. As predicted, however, there was some evidence that people in the negative feedback/undirected rationalization condition were more likely to rate the dissimilar situations as invalid and unimportant ($M = 4.40$, $SD = 1.33$) than did people in any of the other three conditions; means in the no feedback/undirected, no feedback/specific, and negative feedback/specific conditions were 4.87, 4.89, and 4.98, respectively ($SDs = .95, 1.35, 1.29$, respectively), $F(1, 107) = 3.57$, $p = .06$. This result is consistent with the idea that people in the negative feedback/undirected condition rationalized their poor performance rather broadly, leading them to rate even the dissimilar situations as less valid and important than people in the other conditions.

MEDIATION ANALYSES

As in Study 1, we tested whether rationalization mediated the effects of negative feedback on predicted happiness. First, we examined whether the mediation results of Study 1 replicated by performing the identical analysis in the undirected rationalization condition (which was equivalent to the same condition as run in Study 1). The results closely paralleled those of Study 1. The effect of negative feedback on rationalization was significant ($\beta = .54$, $SE = .11$, $p < .001$). The relationship between rationalization and predicted happiness also was significant, after controlling for feedback ($\beta = .37$, $SE = .15$, $p = .02$). The direct effect of negative feedback on predicted happiness ($\beta = .19$, $SE = .13$, $p = .16$) was reduced significantly when the effect of rationalization was controlled ($\beta = -.01$, $SE = .15$, ns , $z = 2.22$, $p < .05$). The only part of this analysis that did not replicate Study 1 is that the effect of negative feedback on predicted happiness was not significant. As mentioned earlier, however, this effect is significant when averaged across Studies 1 and 2.

Because we manipulated how people rationalized, there were more specific mediation hypotheses that we could test. First, we predicted that negative experiencers in the undirected and specific rationalization conditions would both moderate their predictions about how they would feel after doing poorly on similar tests because both had rationalized their poor performance by derogating these tests. To test this prediction, we dummy coded negative experiencers in both rationalization conditions 0 and no feedback participants 1 and performed mediation analyses of their predictions about similar events (the left-hand side of Figure 2). As predicted, the effect of negative feedback on rationalization was significant ($\beta = .43$, $SE = .09$, $p < .001$), as was the relationship between rationalization and predicted happiness, after controlling for feedback ($\beta = .29$, $SE = .10$, $p < .01$). Furthermore, the direct effect of negative feedback on pre-

dicted happiness ($\beta = .17$, $SE = .10$) was reduced significantly when the effect of rationalization was controlled ($\beta = .04$, $SE = .10$, ns , $z = 2.51$, $p < .05$).

When making predictions about dissimilar events, negative experiencers in the specific rationalization condition were expected to behave similarly to no feedback participants because they had denigrated the test rather than focusing on more general aspects of the situation. To test this hypothesis, we combined this group with no feedback participants (dummy coded 0) and compared them with negative experiencers in the undirected rationalization condition in an analyses of their predictions about dissimilar events (the right-hand side of Figure 2). As predicted, the effect of negative feedback on rationalization was significant ($\beta = .30$, $SE = .09$, $p < .001$), as was the relationship between rationalization and predicted happiness, after controlling for feedback ($\beta = .29$, $SE = .10$, $p < .005$). Furthermore, the direct effect of negative feedback on predicted happiness ($\beta = .21$, $SE = .10$) was reduced significantly when the effect of rationalization was controlled ($\beta = .12$, $SE = .10$, ns , $z = 2.29$, $p < .05$).

GENERAL DISCUSSION

Summary of Studies 1 and 2

When left to their own devices (i.e., when we did not attempt to direct people's rationalizations), negative experiencers predicted that future negative feedback would not make them as unhappy as did people who received no feedback ($z = 3.06$, $p = .002$) (averaged across studies). Negative experiencers generalized more to similar than dissimilar situations, as indicated by a reliable Feedback \times Similarity of Situations interaction ($z = 2.01$, $p = .04$). Nonetheless, the feedback effect was significant for both similar and dissimilar situations ($zs = 3.10$ and 2.65 , $ps = .002$ and $.008$).

The undirected negative experiencers rationalized their poor performance rather broadly. They devalued the specific test they had taken more than did people in the no feedback condition ($z = 5.86$, $p < .001$) and said that social aptitude was a less important skill ($z = 2.57$, $p = .01$). There was also a trend for negative experiencers to say that they were in a worse psychological and physical state while taking the test ($z = 1.83$, $p = .07$). This broad pattern of rationalization apparently led negative experiencers to believe that future negative feedback would not affect them as much in several different situations. Study 2 showed that people who were induced to make more specific rationalizations about the causes of their poor performance generalized more narrowly when making affective forecasts about the future. Furthermore, mediation analyses in both studies were consistent with the hypothesis that changes in rationaliza-

tion mediated the effects of negative feedback on predicted happiness.⁵

Thus, there is good evidence that negative experiencers moderated their forecasts about their reactions to future negative events because of the way in which they rationalized their poor performance on the test, consistent with our reconstrual hypothesis. As mentioned earlier, however, it is possible that there was also some genuine learning from experience; that is, the reconstrual hypothesis and learning-by-experience hypotheses could both be true. In addition to basing their forecasts on their reconstrual of the test, people might also, at least to some extent, have based their forecasts on their memory that the previous test did not influence them very much. If so, then there should be a correlation between people's actual happiness after the test and their predicted happiness to future negative events, more so than in the positive or no feedback conditions. This is because the actual reactions of people who received negative feedback are diagnostic for how they might feel after future negative feedback, whereas the actual reactions of people who received positive or no feedback are not.

The correlational evidence for the learning-from-experience hypothesis was mixed. In Study 1, the correlation between people's actual happiness 5 minutes after the test and their predicted happiness 5 minutes after future negative feedback was .62 in the negative feedback condition ($df = 35$, $p < .05$) versus .31 and .22 in the no feedback and positive feedback conditions, respectively ($dfs = 34$ and 31 , ns). The former correlation was significantly higher than the average of the latter two ($z = 2.12$, $p < .05$). However, this evidence was weaker when the covariances were examined, which are less susceptible to condition differences in means and standard deviations. Furthermore, this finding did not replicate in Study 2, where the correlation was actually higher in the no feedback than the negative feedback condition (.45 vs. .24, $dfs = 53$ and 51), albeit nonsignificantly so ($z = 1.17$, ns). (There was no positive feedback condition in Study 2.⁶) Thus, the evidence for the learning-from-experience hypothesis is, at best, weak. We can conclude that learning from experience is not a necessary condition for negative experiencers to moderate their forecasts because there was evidence that they rationalized their poor performance and that this rationalization mediated their predictions of their future happiness.

In contrast, people did not generalize at all from the positive experience of doing well on the test. Even though receiving positive feedback had a minimal impact on people's happiness, people predicted that doing well on an identical test in the future would make them quite happy, as happy as people in the no feedback condition said it would. The reason people failed to gen-

eralize from positive experiences, we suggest, is that they did not make the effort to consult their recent experiences with the V-SAS test when making their forecasts; that is, they failed to satisfy the mental effort criterion by making "top-of-the head" forecasts about how happy they would be in the future situations. Consistent with this view, there were no significant between-condition differences in Study 1 in the correlations between actual happiness 5 minutes after the test and how they predicted they would feel 5 minutes after future positive feedback ($rs = .38$, $.10$, and $.20$ in the positive feedback, no feedback, and negative feedback conditions, respectively; $dfs = 31$, 34 , and 35 , respectively; positive feedback vs. average of other two conditions, $z = 1.14$, ns). Again, the covariances showed even less of a pattern; they were $.20$, $.07$, $.15$ in the positive feedback, no feedback, and negative feedback conditions, respectively.

Admittedly, our evidence for the hypothesis that people in the positive feedback condition did not consult their earlier experience and instead used a top-of-the-head strategy is indirect. Our strategy was to satisfy two of the three conditions necessary to learn from experience (accuracy of recall and applicability) and to infer that if people still failed to generalize, they must not have met the third criterion (mental effort). Accuracy of recall was satisfied by asking people to make predictions soon after experiencing the event; presumably, people could remember how they had felt 15 minutes earlier. Applicability was satisfied by asking people to forecast their feelings for very similar events in the future, one of which was identical to the test they had just taken. The fact that people still failed to generalize from their recent experiences suggests that they did not engage in the mental effort necessary to compare their recent experiences with the future ones.

We do not mean to suggest that people will never engage in such mental comparison or that they will never learn from positive experiences. In the present studies, people may have used a top-of-the head strategy because the future events they were considering were hypothetical ones. When making more consequential forecasts about real events, people might expend the mental effort necessary to retrieve their previous experiences from memory and apply them to their forecasts about similar events in the future. Even if they did, however, they might not always succeed in recalling their past reactions accurately. In everyday life, people often think about future events (e.g., next month's vacation) long after their experience with similar events (last summer's vacation). Meyers et al. (2000) and Mitchell et al. (1997) found that after time passes, people exaggerate how happy they were during and after positive events. Thus, whereas the three conditions necessary to learn from positive experiences may sometimes be met, we suspect

they typically are not in everyday life. Such a conclusion explains why the durability bias for positive events has been found in many settings (Gilbert et al., 1998; Wilson et al., 2000), even in ones in which people had ample past experience.

Limited Learning From Experience

The present studies seem to imply that people will be more accurate at predicting their emotional reactions to negative than positive events, at least for negative events they have experienced in the past. We believe there are a number of reasons, however, to doubt this view. First, we have found ample evidence for a durability bias for negative outcomes in previous studies, even for negative events that people had experienced before and could have learned from (e.g., a loss by one's favorite football team; Gilbert et al., 1998; Wilson et al., 2000). Second, even though negative experiencers in Studies 1 and 2 moderated their forecasts to some extent, they still predicted that the future negative events would make them quite unhappy.

The studies were not designed to test the accuracy of people's forecasts. We did not bring people back to the lab and give them another bad grade or arrange for them to be insulted by a neighbor. In our research on affective forecasting, however, we have been struck by how quickly people recover from similar negative events. In the present studies, we found no evidence that receiving negative feedback on the V-SAS test made people unhappy. People who received negative feedback were no less happy 5 minutes later than people who received no feedback ($z = -.08$, ns). In addition, there were not any differences on an indirect measure of mood—the Associated Reasoning Test (Study 1). Thus, the fact that negative experiencers in Studies 1 and 2 predicted that a poor performance on a test 2 weeks later would lower their happiness by between 1 and 2 points on the happiness scale (see Figures 1 and 2) makes us suspicious. It is possible, of course, that doing poorly on a V-SAS test a second time would make them substantially unhappier than doing poorly on it the first time. Given how resilient we have found people to be after several different kinds of negative experiences, however, we suspect that people in the negative feedback conditions were overestimating how unhappy they would be, albeit to a lesser extent than were people in the no feedback or positive feedback conditions.

In short, there seems to be a paradox: We found that people who experienced a negative event predicted that future negative occurrences would not influence their happiness as much as people who did not experience the negative event. However, in previous studies and in the present ones, we have found ample evidence for a strong durability bias for predictions for negative events. Why

do people appear to have learned from negative experiences but continue to commit the durability bias for negative events?

One possibility is that the magnitude of the durability bias for negative events is so large that it remains even after people have corrected it to some degree. People may have learned from their past negative experiences but not enough to completely avoid the durability bias. This interpretation explains why the bias continues to be as strong or stronger for negative events than for positive events. The two mechanisms known to produce the durability bias—focalism (viewing the event in a vacuum) and immune neglect (underestimating how much they will rationalize the event)—both operate on forecasts about negative events. Only one of these mechanisms (focalism) operates on forecasts about positive events, because people are not motivated to rationalize away positive experiences. Uncorrected forecasts, then, would be expected to be more extreme for negative events. Thus, even if people do learn from past negative experiences, the magnitude of the durability bias could still be quite large.

We close by noting that it may not always be advisable to correct the durability bias. Overestimating the duration of emotional reactions can serve as a motivator, energizing people to work for positive outcomes and avoid negative ones. We are not convinced, however, that the durability bias is an altogether good thing, given that many important decisions are based on forecasts about the duration of affective reactions, such as whom to marry, whether to become a lawyer or psychologist, and whether to plant daffodils in the garden. It is to people's advantage to know exactly which events really will cause lasting pleasure and which will not. If people knew that they will recover quickly if their favorite team loses the big game next week, they might spend less time obsessing about the game and more time planting daffodils.

NOTES

1. A durability bias, we should note, could result from at least two kinds of errors in prediction. People could overestimate the peak intensity of their initial emotional reaction, underestimate the rate at which they will recover from the reaction, or both. People could be wrong about how badly they will feel 30 minutes after getting a bad grade on a test, for example, because they overestimated how badly they felt right after getting the grade, underestimated how quickly they would recover from the negative feedback, or both. Although it is important to consider the conditions under which each of these errors occur, the purpose of the present studies was not to address this question but to investigate the conditions under which people learn about the durability bias more generally; that is, regardless of whether people overestimate initial intensity or underestimate recovery, after the event, they end up with a less intense emotional reaction than they predicted, and when and how they come to recognize this fact is the question of the present studies.

2. In both the Mitchell, Thompson, Peterson, and Cronk (1997) and Meyers, Wilson, and Gilbert (2000) studies, the retrospective durability bias was not as strong as the prospective one. Meyers et al. sug-

gested that focalism is probably weaker when recalling the past than when predicting the future. It is easier to "fill in the vacuum" of the past by recalling that other events influenced one's thoughts and feelings than to imagine the occurrence of future events.

3. Very similar results were found in both Studies 1 and 2 when baseline happiness was used as a covariate rather than performing an ANOVA on difference scores.

4. The same factor analysis was performed on the rationalization measures in Study 2 and a similar three-factor structure was found. Because exploratory factor analyses of small samples are unstable, we decided to keep only items that loaded .4 or more on their respective factor in both data sets. All items except the two mentioned in the text met this requirement.

5. These tests of significance averaged over the two studies using the method of adding t s (Rosenthal, 1978). In no case did the reliability of the effects differ significantly across studies, with one possible exception. There was a marginally significant tendency for the effects of negative feedback on predictions about similar situations to differ across studies, $\chi^2(1) = 3.42, p = .06$.

6. All of these correlations are between people's reported happiness 5 minutes after the test and their predicted happiness. The correlations between people's reported happiness 15 minutes after taking the test and their predicted happiness were very similar.

REFERENCES

- Buehler, R., & McFarland, C. (2001). Extremity bias in affective forecasting: The role of temporal focus. *Personality and Social Psychology Bulletin*, 27, 1480-1493.
- Christianson, S. A., & Safer, M. A. (1996). Emotional events and emotions in autobiographical memory. In D. Rubin (Ed.), *Remembering our past: Studies in autobiographical memory* (pp. 218-243). New York: Cambridge University Press.
- Ekman, P., & Friesen, W. V. (1975). *Unmasking the face*. Englewood Cliffs, NJ: Prentice Hall.
- Erskine, A., Morley, S., & Pierce, S. (1990). Memory for pain: A review. *Pain*, 41, 255-265.
- Festinger, L. (1957). *A theory of cognitive dissonance*. Palo Alto, CA: Stanford University Press.
- Gilbert, D. T. (1991). How mental systems believe. *American Psychologist*, 46, 107-119.
- Gilbert, D. T., Pinel, E. C., Wilson, T. D., Blumberg, S. J., & Wheatley, T. P. (1998). Immune neglect: A source of durability bias in affective forecasting. *Journal of Personality and Social Psychology*, 75, 617-638.
- Hastie, R. (1984). Causes and effects of causal attribution. *Journal of Personality and Social Psychology*, 46, 44-56.
- Higgins, E. T. (1996). Knowledge activation: Accessibility, applicability, and salience. In E. T. Higgins & A. Kruglanski (Eds.), *Social psychology: Handbook of basic principles*. New York: Guilford.
- Kenny, D. A., Kashy, D. A., & Bolger, N. (1998). Data analysis in social psychology. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 1, pp. 233-265). New York: McGraw-Hill.
- Koehler, J. J. (1996). The base rate fallacy reconsidered: Normative, descriptive, and methodological challenges. *Behavioral and Brain Sciences*, 19, 1-53.
- Levine, L. J. (1997). Reconstructing memory for emotions. *Journal of Experimental Psychology: General*, 126, 165-177.
- Mayer, J. D., & Hanson, E. (1995). Mood-congruent judgment over time. *Personality and Social Psychology Bulletin*, 21, 237-244.
- Meyers, J. M., Wilson, T. D., & Gilbert, D. T. (2000). [The accuracy of predicted and recollected happiness after emotional events]. Unpublished raw data.
- Mitchell, T. R., Thompson, L., Peterson, E., & Cronk, R. (1997). Temporal adjustments in the evaluation of events: The "rosy view." *Journal of Experimental Social Psychology*, 33, 421-448.
- Petty, R. E., & Cacioppo, J. T. (1986). *Communication and persuasion: Central and peripheral routes to attitude change*. New York: Springer-Verlag.
- Rosenthal, R. (1978). Combining the results of independent studies. *Psychological Bulletin*, 85, 185-193.
- Ross, M. (1989). Relation of implicit theories to the construction of personal histories. *Psychological Review*, 96, 341-357.
- Ross, M., & Newby-Clark, I. R. (1998). Construing the past and future. *Social Cognition*, 16, 133-150.
- Taylor, S. E. (1991). Asymmetrical effects of positive and negative events: The mobilization-minimization hypothesis. *Psychological Bulletin*, 110, 67-85.
- Thomas, D. L., & Diener, E. (1990). Memory accuracy in the recall of emotions. *Journal of Personality and Social Psychology*, 59, 291-297.
- Wilson, T. D., Wheatley, T. P., Meyers, J. M., Gilbert, D. T., & Axsom, D. (2000). Focalism: A source of durability bias in affective forecasting. *Journal of Personality and Social Psychology*, 78, 821-836.
- Wong, P. T. P., & Weiner, B. (1981). When people ask "why" questions, and the heuristics of attributional search. *Journal of Personality and Social Psychology*, 40, 650-663.

Received September 27, 2000

Revision accepted December 31, 2000

Incremental Commitment and Reciprocity in a Real-Time Public Goods Game

Robert Kurzban
Kevin McCabe
Vernon L. Smith
Bart J. Wilson
University of Arizona

Allowing players in public goods games to make small incremental commitments to contributing to the good might facilitate cooperation because it helps to prevent players from being “free ridden,” contributing more to the public good than other group members. Two experiments using a real-time version of the voluntary contribution mechanism were conducted to investigate the hypothesis that players are generally willing to contribute public goods conditional on beliefs that others are doing so at similar levels. Experiment 1 provided evidence that affording a strategy of commitment can increase the production of public goods. Experiment 2 provided evidence that most players are willing to contribute to the public good at a level at or slightly above the contribution of the lowest contributor in the group. Both experiments point to inequity aversion as an important element of play in public goods games.

Public goods have the property that once they are produced, any individual in a group can consume them, regardless of whether he or she contributed to the production of the good. A strictly rational agent should, in general, refuse to provision a public good because the agent can enjoy the benefit of the good without bearing the cost of provisioning it (Olson, 1965).¹ When public goods have large aggregate benefits relative to their costs, their production constitutes a social dilemma—a situation in which individually rational choices lead to socially deficient outcomes (e.g., Dawes, 1980)—because group members would be better off in aggregate if the good were produced but each individual member would prefer not to pay to produce it. However, public goods are produced both in the real world, as in contributions to public radio, and in the laboratory (see below).

Willingness to provision public goods has frequently been assessed experimentally with the voluntary contribution mechanism (VCM) (e.g., Isaac & Walker, 1988a). Typically, in these experiments, groups of between four and eight participants are faced with a decision to invest money provided to them by the experimenter into two accounts: a private account and a public (or group) account. Money placed in the private account is kept by the investing individual, whereas money placed in the public account is increased at some interest rate (> 1) and divided among all group members equally. This creates a social dilemma because each individual player maximizes earnings by investing everything in his or her private account, but everyone would be better off if all group members contributed to the public account (i.e., the public good).

In the large number of experiments using the VCM, participants are partially able to overcome the social dilemma, routinely contributing to the public good, much as they do in the real world (see Ledyard, 1995, for a review). Why people voluntarily contribute is an important and heavily debated issue. It is important because knowing why and how social dilemmas are solved can inform both our understanding of human social motives

Authors' Note: The writing of this article was supported by a fellowship from the International Foundation for Research in Experimental Economics and a grant from the Russell Sage Behavioral Economics Roundtable, both to the first author. We would like to thank Mark Dincecco for his valuable assistance with this project. Correspondence concerning this article should be addressed to Robert Kurzban, Department of Anthropology, 341 Haines Hall, University of California, Box 951553, Los Angeles, CA 90095-1553; e-mail: rkurzban@ucla.edu.

PSPB, Vol. 27 No. 12, December 2001 1662-1673
© 2001 by the Society for Personality and Social Psychology, Inc.

as well as how to structure social policy to facilitate the production of public goods.

An important clue to why people contribute in public goods games comes from the very consistent finding that in multi-round games, players' contributions begin at moderate levels but then decrease over time (e.g., Davis & Holt, 1993). Although some of this decrease in contributions can be attributed to participants' learning the incentive structure of the game, it is clear that learning alone does not account for this decline (Andreoni, 1988, 1995; Houser & Kurzban, *in press*; Isaac & Walker, 1988a).

Instead, Andreoni (1995), among others, has suggested that this decrease might be due to "frustrated attempts at kindness." This argument supposes that players are generally willing to contribute to the public good only if others are doing so at similar levels. Thus, participants who contributed more than the average amount that others did in one round will want to decrease contributions in subsequent rounds, whereas players contributing at or below the average will not change their contribution rates. If, in each round, some participants roll back their contributions while others keep their contributions constant, the inevitable result is a downward spiral to zero.

This argument, then, is that players do not want to contribute significantly more than others do, on average, which we refer to as "being free ridden." This explanation is similar to claims that participants are reluctant to contribute to public goods due to fear that others are not doing so as well (Chen, 1996; Chen & Komorita, 1994; Komorita & Parks, 1995; Yamagishi & Sato, 1986), although fear has been used in other senses, including the motivation not to have one's resources wasted (Rapoport & Eshed-Levy, 1989). One hypothesis, then, to explain the pattern of results in public goods games is that players would like to achieve the group-efficient outcome but are unwilling to risk contributing significantly more than others in their group to do so (Sugden, 1984). Thus, players may begin the game contributing at moderate levels, willing to risk a small amount of inequity in the hope that their contributions will be reciprocated, but decrease these contributions when they are not.

There is some evidence in favor of this view. First, players' reported expectations about other group members' contributions correlate well with their own actual contribution decisions across a number of experiments (Bornstein & Ben-Yossef, 1994; Braver & Barnett, 1974; Croson, 1998; Dawes, McTavish, & Shaklee, 1977; Komorita, Parks, & Hulbert, 1992; Messick et al., 1983; Wit & Wilke, 1992; Yamagishi & Sato, 1986), although the direction of causality of this relationship can of course be questioned. This suggests that players want to

contribute at the same level as others in their group, preferring neither to free ride nor to be free ridden.

Additional evidence comes from experiments investigating the impact of commitment² in public goods games. Chen and Komorita (1994) ran a series of studies in which participants submitted a pledge to contribute some fraction of their endowment during the subsequent phase of the game. In one condition, these pledges bound not only the player making the pledge, but also all other players in the player's group to the same amount. Making a pledge did not expose the player to being free ridden in this condition, and both pledges and contributions were quite high, up to 73% of players' endowments (see also Chen, 1996). In another condition, players' pledges applied only to themselves, meaning that an individual making a large pledge ran the risk of obligating himself or herself to a contribution greater than that of other players. In this condition, pledges and contributions were much smaller, 36% of players' endowments.

Taken together, these results suggest that commitment can facilitate public good production, but only when the mechanism of commitment does not expose players to being free ridden by the other members of the group. Thus, commitment is a means by which players can assure one another that they are not going to free ride on others' contributions, so that group members can contribute without fearing that they will be free ridden. However, people seem unwilling to use a commitment mechanism if doing so exposes them to being free ridden. This presents an interesting problem from the standpoint of eliciting contributions to public goods: People might be willing to match committed contributions of others but not to commit before others have done so.

This problem was described by Schelling (1960) in his discussion of two hypothetical parties who both want to contribute a large amount of money to the Red Cross, but only if the other does so as well. The solution Schelling suggested was to allow sequential commitments of small amounts by each individual, thus keeping a tight reign on inequality of contributions. So, in this scenario, one person contributes a small amount, which is then matched by the second person, and so on, allowing each person to risk only the amount of the incremental contributions rather than the whole sum (see also Admati & Perry, 1991; Osgood, 1962; Roberts & Sherratt, 1998).

The same problem applies when players in public goods environments are willing to cooperate only to the extent that everyone else is willing to do so. In turn, a similar solution is possible. What is needed is a mechanism by which players can commit to cooperating to some

small degree and observe other players' reciprocal contributions. This allows players to signal their commitment to provisioning the public good without exposing themselves to being free ridden by other group members who do not match their committed contributions.

To instantiate a mechanism that allows players to make "consecutive small contributions" (Schelling, 1960), we used the real-time VCM first developed by Dorsey (1992). In the real-time VCM, participants have some short amount of time in each round in which to update their contributions to the public good. Their actual contribution in a given round is equal to their contribution when the countdown clock reaches zero.

When players can adjust their contributions upward and downward during the round, information about others' contributions amounts to little more than cheap talk.³ However, similar to Dorsey (1992), we modified the mechanism by which contributions could be updated such that in some groups, players could increase their contribution to the public account in single token increments but were not allowed to decrease their contributions to the public account during the course of the round. This increase only (IO) mechanism can be construed as affording a commitment strategy—once a player has raised his or her contribution to the public good to a particular level, they are unable to reverse this decision, committing them to that level. This mechanism allows players to make small commitments to the public good while allowing them simultaneously to limit their commitments so that they can control the extent to which they expose themselves to being free ridden.

If the hypothesis is correct that players' willingness to provision public goods is a positive function of their ability to prevent themselves from being free ridden, then providing a mechanism of incremental commitment should increase contributions to a public good relative to the case in which incremental commitment is not possible. In Experiment 1, we used the IO mechanism to test this hypothesis, predicting that contributions in a condition in which players could only increase their contributions (commitment) would be higher than in a condition in which players could increase or decrease their contribution (cheap talk).

EXPERIMENT 1

Method

PARTICIPANTS

Fifty participants were recruited from the University of Arizona undergraduate community using the electronic recruitment system maintained by the Economic Science Laboratory. Each participant was told that he or

she would earn \$5 for showing up to the experimental session and could earn additional money depending on the decisions that he or she and other people in the experiment made during the experiment.

DESIGN

There were two conditions: one with the increase/decrease pledge mechanism, and one with the increase-only mechanism. Five groups of five participants were run in each condition.

PROCEDURE

The procedure was a standard public goods game that largely duplicated that used by Marwell and Ames (1979) with the real-time contribution mechanism developed by Dorsey (1992). Participants were given a time to report to the laboratory and either one or two groups of five were run in a given session, depending on the number of people available for that session.

After arriving, participants received their \$5 show-up payment and were assigned to one of the computers in the main laboratory area. Computers in this laboratory are separated by partitions so that players cannot see one another or any other player's computer screen. The entire experiment was conducted by computer.

Once all participants had arrived and were seated at a computer terminal, they read the instructions for playing the public goods game (Andreoni, 1995). These instructions appeared on players' computer screens and participants were allowed to proceed through them at their own pace. Any questions that arose were answered privately by the experimenter. The instructions indicated that the public goods game would continue for 10 rounds. Participants were informed that at the beginning of each round, they would be given an endowment of 50 tokens, that tokens could be invested in accounts that earned points which would be converted to cash and paid at the conclusion of the experimental session, and that they would receive the average amount that they earned over the course of the 10 rounds. The instructions informed them that they could divide their endowment (in units of whole tokens) any way they chose between the two accounts during each round and that they would earn the full value of each token that they put in their personal account as well as one third of the value of each token they and the other participants put in the group account. As part of the instructions, participants were given an opportunity to familiarize themselves with the interface they would be using to update their contributions during the round. The countdown clock, the information that they would see during the round, and the mechanism for updating their contribution (IO or ID) also were explained.

At the beginning of each round, the players' entire endowments were placed in the private account. In the IO condition, players could increase their contribution to the group account one unit at a time by clicking on a small button provided for this purpose. In the ID condition, two buttons were visible, one for increasing the contribution to the group account and one for increasing the contribution to the private account. Across conditions, players could see the current contribution levels of all five members of their group during the round, updated five times per second. The placement of the boxes was constant across all 10 rounds, although there was no way to know which information corresponded to which player in the room.

When all players indicated that they were ready, Round 1 began. The countdown clock was set to 90 seconds and counted down in increments of 1 second. Players could modify their contributions during the entire 90-second countdown. When the time for the round had elapsed, players were informed of the aggregate contribution to the group account and their total earnings in tokens for that round (the number of tokens in the personal account and one third of the tokens in the group account). When all players had indicated that they were ready to begin the next round, the countdown clock returned to 90 seconds and Round 2 began. Subsequent rounds proceeded similarly.

When Round 10 was complete, participants were asked to fill out a short questionnaire, which included a free-response section that asked participants to indicate how they had made their contribution decisions. After filling out the questionnaires, each participant was given a sealed envelope with his or her earnings and dismissed.

Results

We conducted a mixed-effects analysis for repeated measures (e.g., Longford, 1993). The two factors (pledge mechanism [ID, IO] and round) are modeled as (dichotomous zero-one) fixed effects, whereas the groups and the participants within each group are modeled as random effects. Because there could be substantial variation of the contributions across groups, and because contribution decisions are likely to be autocorrelated as participants learn over rounds, we generalized the error structure to include groupwise heteroskedastic variances and a first-order autoregressive (AR[1]) process for residuals by participant and estimated the model via maximum likelihood.⁴

This analysis revealed a main effect of round, $LR(9) = 59.47, p < .0001$, but no significant effect of pledge mechanism, $t(8) = 1.18, p = .27$.⁵ Of more interest is the two-way interaction of pledge mechanism and round, $LR(9) = 17.43, p < .05$. This interaction was driven by the drop-off in contributions in the ID condition compared with the

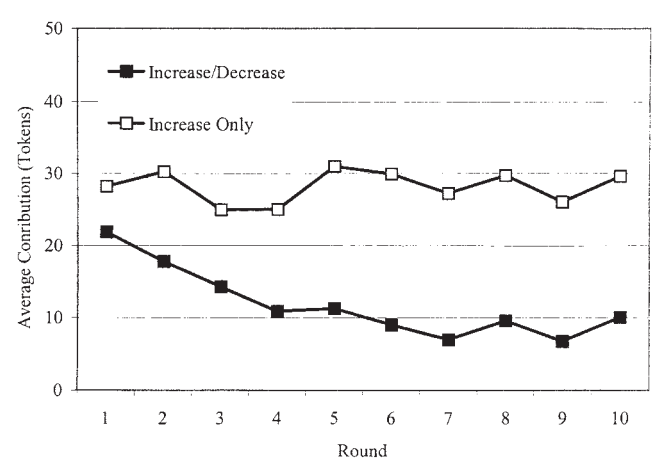


Figure 1 Contributions by condition in Experiment 1.

relatively stable contributions in the IO condition (see Figure 1).

EXPERIMENT 2

Experiment 1 suggests that the IO mechanism, which affords commitment, is effective in facilitating cooperation when participants have access to full information about others' current contributions. This is consistent with the idea that affording a strategy of commitment that allows players to limit their exposure to being free ridden can facilitate cooperation.

Nonetheless, even with the commitment mechanism, groups in Experiment 1 attained only moderate rates of contribution to the public good. There may be a number of reasons for this, but one possibility is that the environment allowed players to exploit those participants that overcommitted, contributing to the public good in substantial amounts during the round even when others did not. Some evidence exists that whereas conditional cooperation is generally reciprocated, unconditional cooperation tends to be exploited (e.g., Komorita, Hilty, & Parks, 1991). Thus, it is possible that participants who observed other group members making unilateral large contributions chose to free ride on these contributions, keeping their own allocation to the public account low because their fellow group members were generous without needing the incentive of reciprocal cooperation.

This suggests that a mechanism that simultaneously allows incremental commitment but prevents players from seeing others as exploitable might further increase contribution levels. One way to implement such a mechanism is to provide players within a group only the lowest current contribution to the public good. By providing players with only the lowest current contribution, they are prevented from observing the potential for free rid-

ing on overly cooperative players in a given round, thus hiding potential exploitative opportunities.

We make two assumptions in Experiment 2. First, we assume, as in Experiment 1, that players have a preference for achieving the group-level optimum outcome provided that they do not expose themselves to being free ridden in the process. Second, we assume that players are “sophisticated” (after Milgrom & Roberts, 1991) and know that others similarly do not want to be free ridden (see also Keser & van Winden, *in press*). This idea goes back at least as far as Pruitt and Kimmel (1977), whose “goal expectation” model held that cooperation was due to “*an expectation that the other will cooperate either immediately or in response to the actor’s cooperation*” (p. 375).⁶

Given these two assumptions, consider players who receive only information about the lowest current contribution to the group account. These players know that everyone in their group is currently contributing at least the value of the information that they observe. Furthermore, players using the IO mechanism will know that the other players are committed to these contributions. If players are sophisticated, they will know that keeping their contribution at this level will freeze this value and, importantly, dissuade others from contributions significantly above this value because they will not want to be free ridden by the player currently at the minimum. However, players will not know how much above this level others are currently contributing, obscuring opportunities for free riding.

In this condition, players can incur the relatively low cost of contributing one unit to ensure that the minimum information does not get stuck at its current level, inhibiting additional group cooperation. Note that raising one’s contribution above the minimum in this condition also reveals to a player whether others are tied with him or her; therefore, these marginal increases also can be construed as relatively inexpensive information gathering (Ward, 1989). If all players in a group increase their contribution gradually, just above the minimum value, this will lead to a kind of “ratchet effect,” with contributions increasing incrementally by small amounts over time.

As a comparison class for the low information condition, we also included a condition in which players receive only information about the highest current contribution. This condition has the same amount of information (one player’s contribution) but does allow players to observe the possibility of free riding off of another’s contribution and, critically, does not assure players that others are not free riding off of their own. If the hypothesis is correct that players withhold contributions out of fear that other group members are free riding, providing the highest information should lead to

less cooperation than when the lowest contribution information is provided.

To summarize, receiving the value of the lowest contribution means that participants can be sure that all members of the group are committed to at least the level of cooperation indicated by the current value. The IO mechanism combined with the lowest information treatment (IOL⁷) allows one to contribute small amounts, keeping a tether on the extent to which one can be free ridden. Thus, the IOL condition should be effective in eliciting contributions from players because it allows players to make small incremental contributions to the public good while monitoring whether other players are reciprocating, in much the way described by Schelling (1960). Thus, in Experiment 2, we predict that cooperation (contributions) in the IOL information condition will be high and sustainable compared to those in all other cells. A second prediction is that because it is relatively easy in the IOL condition to ensure that one’s contribution never strays far from that of others, there will be a close correspondence between players’ contributions and the information they receive.

Method

PARTICIPANTS

One hundred participants who had not taken part in Experiment 1 were recruited from the University of Arizona undergraduate community. Each participant was told that he or she would earn \$5 for showing up to the experimental session and could earn additional money depending on the course of the experiment. The amount that each participant actually earned depended on the decisions that he or she and the other participants in their group made during the experiment.

DESIGN

The experiment employed a 2 (contribution information: highest, lowest) \times 2 (pledge mechanism: ID, IO) factorial design. Five groups of five participants were run in each of the resulting four conditions.

PROCEDURE

The procedure was identical to that used in Experiment 1 with the following exception: Participants in half of the groups in this experiment could see the current contribution level of the highest current contributor to the group account, whereas participants in the other half of the groups could see the current contribution level of the lowest current contributor to the group account. Of course, in both cases, the determination of the highest or lowest contribution included the focal participant.

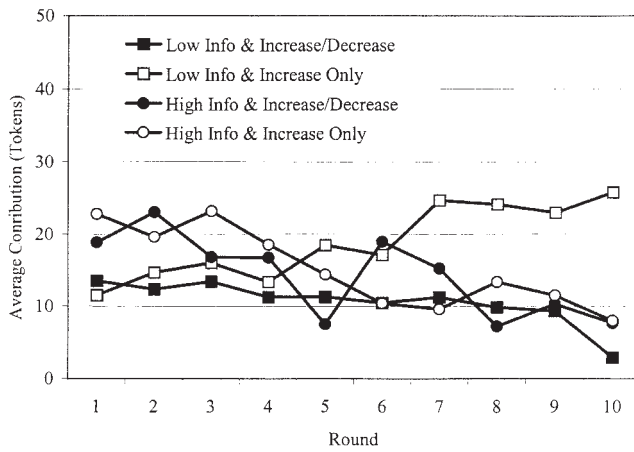


Figure 2 Contributions by condition in Experiment 2.

Results

CONTRIBUTION LEVELS

The primary dependent measure of interest was participants' final contribution at the end of each of the 10 rounds. We employed a mixed-effects model with repeated measures for our analysis. The 2×2 treatment effects (contribution information [lowest, highest] and pledge mechanism [ID, IO]) and round are modeled as dichotomous (zero-one) fixed effects, whereas the groups and the participants within each group are modeled as random effects. Because we expected a priori that (a) the variation of the contributions across groups will be heterogeneous and (b) a learning effect across the rounds may manifest itself as autocorrelation in the participants' decisions, we generalized the error structure to include groupwise heteroskedastic variances and a first-order autoregressive (AR[1]) process for residuals by participant.⁸

This analysis from the maximum likelihood estimation yielded no significant effect of contribution information, $t(16) = 1.03$, $p = .32$, and pledge mechanism, $t(16) = -0.46$, $p = .65$. The round effect, however, was significant, $LR(9) = 30.31$, $p < .0005$.

However, these null findings for the main treatment effects are qualified by highly significant two-way interactions between contribution information and round, $LR(9) = 29.42$, $p < .001$, and between pledge mechanism and round, $LR(9) = 41.69$, $p < .0001$. The contribution information and round interaction is driven by the observation that contributions in the low information conditions are, on average, relatively constant over the course of the game, whereas contributions in the highest information condition decrease over time. The final two-

TABLE 1: Results of the Lowest Information Only (IOL) Treatment From the Mixed-Effects Analysis of the Pledge Mechanism and Round Interaction on Contributions, Experiment 2

Variable	Coefficient	p
IOL (all rounds)	-2.59	.65
IOL \times Round 2	6.77	.0008
IOL \times Round 3	8.62	.0005
IOL \times Round 4	8.86	.0007
IOL \times Round 5	8.80	.0010
IOL \times Round 6	12.00	.0001
IOL \times Round 7	14.04	< .0001
IOL \times Round 8	17.04	< .0001
IOL \times Round 9	16.68	< .0001
IOL \times Round 10	21.17	< .0001

way interaction, contribution information and pledge mechanism, was not significant, $t(16) = 3.37$, $p = .70$.

These two-way interactions were themselves qualified by a significant three-way interaction among contribution information, pledge mechanism, and round, $LR(9) = 27.70$, $p < .005$. This interaction is driven by the observation that in the IOL condition, contributions increase over the course of the 10 rounds, whereas in the other three conditions, contributions fall off with time. Figure 2 displays the average contribution across all five groups for each condition over the course of the 10 rounds.

As a joint test for all rounds, a statistical test for the two- and three-way interactions does not reveal the sign or magnitude of the interactions, both of which are relevant to the main hypothesis of this article. The interaction by round in the IOL cell is of particular interest given that the combination of these treatments, by hypothesis, should lead to successful provisioning of the public good. Table 1 reports the coefficients of the specific interaction in this cell by round. Notice that except for the first round, the interaction is highly significant every round and that the magnitude of this effect on the contributions was much larger for Rounds 6 through 10 than for Rounds 2 through 5. With more experience within a session, the total group contributions in the non-IOL treatment are progressively falling, whereas the contributions in the IOL treatment remain relatively high and constant. Hence, the estimates of the IOL treatment effect increase relative to the non-IOL treatment.

The participant and group random effects control for one interesting aspect of the data: between-group variation. Again, the IOL groups are particularly noteworthy. In this condition, two of the groups' contribution levels look similar to those of groups in the other three conditions, with contributions starting off at moderate levels and decreasing toward zero over time. In contrast, in the three other groups in this condition, contributions tended to increase over the course of the game, reaching

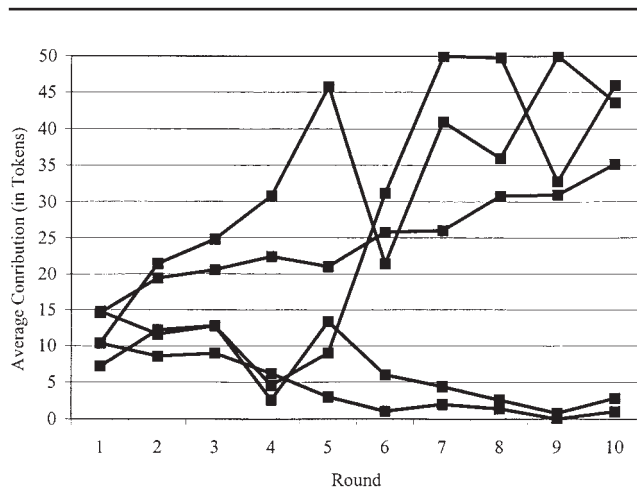


Figure 3 Contributions in the lowest information and increase only conditions in Experiment 2.

levels of more than 75% by the end (see Figure 3). We return to the question of the source of this between-group variation below.

RECIPROCITY

To try to evaluate the extent to which players' contributions were influenced by the information they were provided, we ran a regression of the players' actual contributions on the value of the information at the end of the round, with treatments entered as independent variables for all groups across all 10 rounds. Note that the information observed at the end of the round is really only a proxy for the players' expectations because the value could, in principle, have changed at the last moment in a round before a player had a chance to react. Also, because one of the five players in each group was the individual who actually set the highest or lowest value, including all players in these regressions would overestimate the strength of the relationship between the final value and players' contributions because 20% of the observations would necessarily be perfectly correlated. For this reason, we removed the player whose contribution matched the final information value in each round in every cell. If more than one player's contribution matched this value, only one of these tying players was removed. We ran the following regression:

$$\begin{aligned} \text{Contribution}_{it} = & a_0 + a_1 I_i + a_2 H_i + a_3 H_i * I_i + \beta_0 \text{Endvalue}_{it} \\ & + \beta_1 I_i * \text{Endvalue}_{it} + \beta_2 H_i * \text{Endvalue}_{it} \\ & + \beta_3 H_i * I_i * \text{Endvalue}_{it} + e_{it}, \end{aligned}$$

where $I=0$ for the ID condition and $I=1$ for the IO condition, $H=0$ for the low information condition and $H=1$ for the high information condition, and Endvalue refers

TABLE 2: Results of the Regression of Final Information Value on Contribution by Treatment, Experiment 2

Variable	Coefficient	p	η^2
Constant	5.96	< .0001	
Highest	-8.06	< .0001	0.02
Increase only	-0.97	.20	0.03
Highest \times Increase Only	1.27	.60	0.01
Endvalue	0.99	< .0001	0.47
Highest \times Endvalue	-0.60	< .0001	0.06
Increase Only \times Endvalue	-0.09	.06	0.00
Highest \times Increase Only \times Endvalue	0.20	< .05	0.00

NOTE: Refer to the text for the regression model. Number of observations = 800, $R^2 = .60$, and $s^2 = 83.92$. p values are based on the standard errors for a groupwise heteroskedastic model.

to the value of the information (highest or lowest) at the end of the round. Subscripts refer to player i s at time t , where t refers to Rounds 1 through 10. The results of this regression are summarized in Table 2.

Overall, there was a significant relationship between the information observed at the end of a round and players' contributions. This relationship held across conditions but the slope coefficient differed significantly depending on the information condition. In the lowest information condition, an increase of one token in the information the player observed led to an increase of roughly one token in actual contributions. In contrast, under the highest information treatment, the effect of an increase of one token in the information observed was an increase of roughly one half of a token.

Discussion

Experiment 2 provided strong support for the idea that players in public goods games are willing to contribute to the extent that they believe others are similarly willing to do so. Correlations between actual contributions and the information observed were relatively strong, particularly in the lowest information condition. When players could observe the lowest information, they could be certain that every other player was contributing at least the value of the current value that they were seeing. In the IO condition, players were committed to this contribution, encouraging reciprocal contributions.

In general, the establishment of high levels of contribution followed the ratchet pattern described by Schelling (1960). In the three groups in which high rates of cooperation were observed, players increased their contributions systematically over the course of the round to match the lowest information value and kept their contribution at roughly one token above this level. Figure 4 illustrates this pattern for one period of play for one group that achieved complete cooperation in the IO and low information condition.

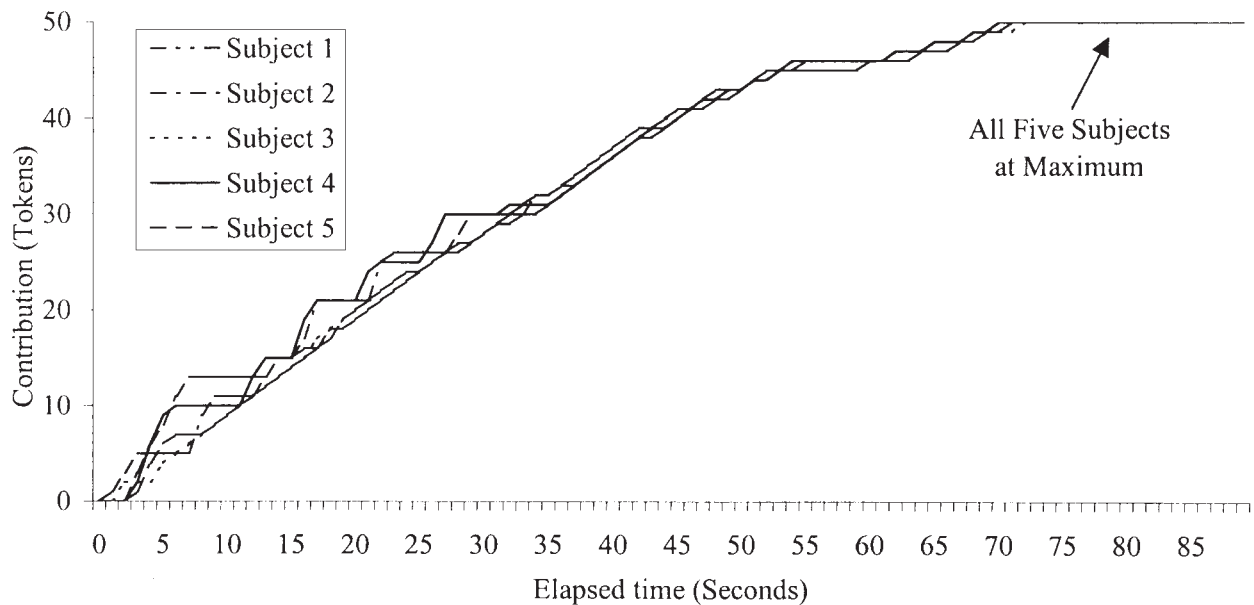


Figure 4 Contributions by all five players over time in a representative round in the lowest information and increase only conditions, Experiment 2.

Groups reached high but not perfect levels of contribution when one player unilaterally refused to increase his or her contribution above the current minimum, keeping the group stuck at that particular level. Why certain players chose to increase their contributions to a seemingly arbitrary point and then stop is not clear, because these players would almost certainly have been better off increasing their contributions because other players in their groups seemed to be using a strategy of keeping their contribution one token above the minimum.

The hypothesis that the low contribution and IO cell would lead to enhanced provisioning of the public good received mixed support. In this cell, three out of the five groups achieved contribution rates of between 60% and 100% during the latter rounds of the game, a respectable amount of within-group cooperation given consistent findings that cooperation rates drop off toward zero in latter rounds of most public goods games with repeated play (Davis & Holt, 1993). However, two of the groups in this cell resoundingly failed to achieve substantial rates of cooperation, with contributions sinking to less than 10% in the final rounds of the game (see Keser & van Winden, *in press*, for a similar result). Thus, the only statistical evidence for the IOL mechanism's effectiveness emerged in the context of its ability to increase levels of cooperation over time, in contrast to the other three mechanisms.

We cannot be certain about the source of this between-group variation. It is possible that idiosyncratic differences among experimental sessions could account for some of this variation, but we have no particular reason to believe this is the case. Another possibility is that there is some unmeasured individual difference variable among the players in these groups. If, indeed, players are using a kind of matching strategy, the low contribution information conditions are particularly sensitive to individual differences. Imagine that there is some small fraction of players in the population who simply choose to contribute zero in every round of a public goods game (strong free riders). In the high information conditions, these players are somewhat invisible, their presence indicated only in the end-of-round, aggregate contribution information, which reveals their reticence. In contrast, one of these players in the lowest information condition will be obvious to everyone in their group, because the minimum value will not budge from zero. Even if all players but one keep their contribution slightly above the minimum value, just one "zero player" will prevent the group from establishing mutual cooperation. Small numbers of strong free riders have been observed in other experiments (e.g., Andreoni, 1995; Croson, 1998), suggesting this phenomenon is not simply a function of the design of this particular experiment.

What is the origin of these strong free riders? We can only speculate at this point. Perhaps they are extremely competitive, playing zero to ensure that no one else in

their group earns more than they do—it is known that the individual difference variable “social value orientation” can have important effects on cooperation in other games (e.g., Kramer, McClintock, & Messick, 1986; McClintock & Liebrand, 1988; Van Lange & Visser, 1999; but see Parks, 1994). Perhaps they believe they are playing some optimal strategy, trained in game theory, and believing that equilibrium play in these games is zero. Indeed, in the free-response portion of the questionnaire, one player in one of the two groups that was unable to achieve cooperation in the IOL condition indicated that he or she was playing the dominant strategy. This player did contribute zero on 7 of the 10 rounds of the game, dooming the group to extremely low contribution levels. It seems possible that strong free riders were unable to understand what effect their playing zero would have on other players’ decisions. Additional work will be required to isolate any individual difference variable that may be at work in these games (for additional work on individual differences, see Kurzban & Houser, *in press*; Liebrand, 1984; Parks & Hulbert, 1995; Rapoport & Suleiman, 1993; Yamagishi, 1986).

GENERAL DISCUSSION

The results of the experiments reported here yield two primary findings. The first is that providing a mechanism of commitment in the public goods environment can be effective in eliciting cooperation from players, but only under particular conditions. Experiment 1 showed that the commitment mechanism was effective in sustaining cooperation over time when players had access to complete information about others’ contributions. The level of cooperation under these conditions did not show the typical pattern of decay over the course of the 10 rounds of play (e.g., Andreoni, 1988; Isaac & Walker, 1988b). In Experiment 2, three groups in the IOL condition were able to achieve extremely high rates of cooperation, particularly in the latter rounds of the game, which contrasted starkly with the other experimental conditions. The IO mechanism seems to be able to facilitate cooperation but also makes groups susceptible to strong free riders, whose presence scuttles attempts to cooperate. In contrast, when players saw only the highest contribution, providing a commitment mechanism had no significant effect on the level of cooperation.

The second finding is that participants in public goods games use reciprocal strategies but that the extent to which they do so depends on the nature of the information they have about other players’ contributions. In Experiment 2, there were close relationships between players’ actual contributions and the single piece of information that they had available to them about others’ contributions. The relationship between the infor-

mation that players observed and their own contributions was significantly weaker when they had access to the current highest contribution to the public good.

An additional finding is that there is some evidence that players are sophisticated in the sense that they believe others are playing some kind of reciprocal strategy as well. This is clear from the results of the IOL condition. If players played a simple matching strategy by which they set their allocation to the group account to the level that they observed, the minimum value would never change and cooperation could not be established. However, in three of the five groups, players set their allocations to the group account slightly above the level of the current minimum, ensuring that they did not cause the minimum value to get stuck, inhibiting further contributions from reciprocators.

There is evidence from two additional sources suggesting that players are sophisticated. The first is the free-response section of the questionnaires that participants filled out that asked them to indicate how they had made their contribution decisions. Many participants indicated that they themselves were using a reciprocal strategy (e.g., “If others put tokens in the group account, so did I”) and that they were contributing to elicit contributions from others (e.g., “I wanted my contributions to be matched”), suggesting that these participants believed others would also use some type of reciprocal strategy.

Second, when participants had the capability of increasing and decreasing their contributions, at least some players put large numbers of tokens in the group account during the course of the round and left them there up until the last few seconds of the game, when they removed them with some haste. This suggests that these players were trying to signal that they were going to contribute a large number of tokens to the group account in an attempt to induce others to do so.

From a theoretical standpoint, these results lend weight to the hypothesis that players in public goods games are motivated by a fear of being free ridden as well as by a desire to achieve high levels of cooperation within one’s group. This contrasts with classical economic models that assume that people have preferences over only their own payoffs but is consistent with recently proposed “inequity aversion” models that suggest that people have preferences over their own outcomes as well as the distribution of outcomes among other relevant agents (Dufwenberg & Kirchsteiger, 1998; Rabin, 1993). More specifically, people seem to dislike unequal outcomes but are particularly upset if they are on the short end of the unequal allocation (Fehr & Schmidt, 1999).

More concretely, in the context of public goods games, there seem to be two principles that explain a great deal of contribution behavior. The first is that play-

ers do not want to contribute more than other members of their group. The second is that if players believe that everyone is going to contribute in roughly equal amounts, they prefer that amount to be higher rather than lower. On this theory, the incremental IO mechanism allows players to limit their fear of exploitation because they can condition their own play on their observations of others' contributions, ensuring that they will not be the victim of a large and disadvantageous unequal outcome. The low information condition is particularly effective because it essentially allows group members to coordinate on high contribution levels.

If this analysis is correct, it suggests why obtaining cooperation in public goods games in which players make their contributions simultaneously is problematic. In the simultaneous game, in any given round, players do not know how much others are contributing when they make their own decision. Thus, any contribution one makes exposes the player to being free ridden by others who contribute less, leading players to make small contributions to avoid this unpleasant state of affairs. This suggests that moderate levels of contribution toward the beginning of the game are conservative attempts to establish high levels of cooperation but that the spiral downward during the course of multiple round games is the result of the failure of these attempts.

Also, to the extent that this model is correct, doubt is cast on explanations of contribution behavior that make reference to altruism or learning as important factors (e.g., Andreoni, 1990) and suggest instead that reciprocity is a key element (e.g., Croson, 1998; Komorita, Chan, & Parks, 1993; Komorita et al., 1992). Players seem to use their own contributions to elicit contributions from others (Pruitt & Kimmel, 1977) and are willing to expose themselves to small amounts of being free ridden to do so. Thus, reciprocity in public goods games needs to be understood not only in the context of responding in kind to others' contributions but using one's own contributions to elicit cooperation from others.

An important feature of the low information treatment is that it allows players to monitor to some extent every other member of the group. The low information indicates that every single group member is contributing at least at the indicated level. This suggests a role for perceptions of unanimity in group cooperation (Smith, 1991), but because unanimity per se was not manipulated in these experiments, further research will be needed to clarify when and if this is an important factor. For example, it is an open question how results of Experiment 2 would change if information about the second lowest contributor were provided (for computer simulation data that bear on this issue, see de Heus, 2000; Parks & Komorita, 1997). This is a potentially important area of research because it would shed light on the exact nature

of the fear of being free ridden; that is, are people reluctant to be free ridden by even one other player, or are they willing to tolerate a certain amount of free riding to establish cooperation among remaining group members?

As always, caution should be exercised in generalizing from these results. There are many different contexts in which individuals must decide how much to cooperate with one another, ranging from small work groups up to large scale phenomena such as provisioning public radio, and many of these contexts obviously differ in important ways from the stylized laboratory environment. Settings outside the lab might differ in the information that one has about what others are doing, the opportunity to signal one's commitment, the degree of interpersonal interaction, and so forth. However, the real-time mechanism does reflect the structure of many types of public goods environments, such as many fundraising drives, which often provide potential donors continuously updated information about how much money has been pledged up to that time (Dorsey, 1992). More generally, many cooperative activities occur in real time and, of course, most actions that we take in the real world cannot be undone or taken back once they are completed.

It is also important to note that it is not clear how specific the effects we observed are to potentially important factors such as group size, the per capita return of the public good, participant population, and so forth. In addition to determining the generality of these findings, an important goal for future research will be to develop techniques that are capable of distinguishing among the different kinds of reciprocal strategies participants might be using in public goods environments.

NOTES

1. We add "in general" because there are situations in which a rational agent will provision a public good, such as when the benefit of the good to the individual exceeds the cost of its production (see Olson, 1965, for a thorough discussion).

2. We will use the term *commitment* strictly to mean an action that is binding on the actor. Others have used the word in the sense of a spoken promise, which might or might not be broken (e.g., Kerr & Kaufman-Gilliland, 1994).

3. Cheap talk is communication that is costless and nonbinding. "Little more" is an important hedge because the talk is more "expensive" as the clock gets closer to zero. Because there are physical limits to how fast a player can remove tokens using our interface, a player with a very high contribution might not be able to decrease his or her contribution all the way to zero if only a few seconds are left, making contributions toward the end of a round more like commitments.

4. Likelihood ratio (LR) tests find that both of these specifications are significant: heteroskedastic group variances, $LR(19) = 131.96$, $p < .0001$, and AR(1) residuals, $LR(1) = 60.83$, $p < .0001$. These specifications improve the efficiency of the estimates. The LR test determines whether the difference in the maximized value of the likelihood function with the restriction is significantly different than the unrestricted maximum value of the likelihood function (see, e.g., Kennedy, 1992, p. 61).

5. To illustrate the improvement in efficiency, the standard error on the main effect of the pledge mechanism, without compensating for groupwise heteroskedasticity and AR(1) error terms, is 6.30, but with the corrections for nonspherical disturbances, the standard error is 6.13.

6. As an aside, it is interesting to note that evolutionary psychologists have predicted that people should be "sophisticated" in this sense across a variety of contexts. This derives from the fact that natural selection builds mechanisms that embody assumptions that reflect the stable, recurrent features of the environment in which a population evolves (Cosmides & Tooby, 1992). Thus, the important elements of human psychology that have been reliably present over evolutionary time should be embodied in people's assumptions about others.

7. For ease of exposition, treatment cells hereafter are referred to by a three-letter combination of the mechanism and the information participants observed (e.g., IOH is the increase only and high information condition).

8. Likelihood ratio tests find that both of these specifications are significant: heteroskedastic group variances, $LR(19) = 384.16$, $p < .0001$, and AR(1) residuals, $LR(1) = 110.12$, $p < .0001$. These specifications improve the efficiency of the estimates in Table 2.

REFERENCES

- Admati, A. R., & Perry, M. (1991). Joint projects without commitment. *Review of Economic Studies*, 58, 259-276.
- Andreoni, J. (1988). Why free ride? Strategies and learning in public goods experiments. *Journal of Public Economics*, 37, 291-304.
- Andreoni, J. (1990). Impure altruism and donations to public goods: A theory of warm glow giving. *Economic Journal*, 100, 464-477.
- Andreoni, J. (1995). Cooperation in public goods experiments: Kindness or confusion? *American Economic Review*, 85, 891-904.
- Bornstein, G., & Ben-Yossef, M. (1994). Cooperation in intergroup and single-group social dilemmas. *Journal of Experimental Social Psychology*, 30, 52-67.
- Braver, S. L., & Barnett, B. (1974). Perception of opponent's motives and cooperation in a mixed-motive game. *Journal of Conflict Resolution*, 18, 686-699.
- Chen, X. P. (1996). The group-based binding pledge as a solution to public goods problems. *Organizational Behavior and Human Decision Processes*, 66, 192-202.
- Chen, X. P., & Komorita, S. S. (1994). The effects of communication and commitment in a public goods social dilemma. *Organizational Behavior and Human Decision Processes*, 60, 367-386.
- Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In J. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind* (pp. 163-228). New York: Oxford University Press.
- Crosen, R.T.A. (1998). *Theories of altruism and reciprocity: Evidence from linear public goods games*. Unpublished manuscript, the University of Pennsylvania, Philadelphia.
- Davis, D. D., & Holt, C. A. (1993). *Experimental economics*. Princeton, NJ: Princeton University Press.
- Dawes, R. M. (1980). Social dilemmas. *Annual Review of Psychology*, 31, 169-193.
- Dawes, R. M., McTavish, J., & Shaklee, H. (1977). Behavior, communication, and assumptions about other people's behavior in a commons dilemma situation. *Journal of Personality and Social Psychology*, 35, 1-11.
- Dorsey, R. E. (1992). The voluntary contributions mechanism with real time revisions. *Public Choice*, 73, 261-282.
- Dufwenberg, M., & Kirchsteiger, G. (1998). *A theory of sequential reciprocity*. Unpublished manuscript.
- Fehr, E., & Schmidt, K. (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, 114, 817-868.
- de Heus, P. (2000). Do reciprocal strategies work in social dilemmas? *Personality and Social Psychology Review*, 4, 278-288.
- Houser, D., & Kurzban, R. (in press). Revisiting kindness in public goods games. *American Economic Review*.
- Isaac, R. M., & Walker, J. M. (1988a). Communication and free-riding behavior: The voluntary contribution mechanism. *Economic Inquiry*, 26, 585-608.
- Isaac, R. M., & Walker, J. M. (1988b). Group size effects in public goods provision: The voluntary contributions mechanism. *Quarterly Journal of Economics*, 103, 179-200.
- Kennedy, P. (1992). *A guide to econometrics* (3rd ed.). Cambridge, MA: MIT Press.
- Kerr, N. L., & Kaufman-Gilliland, C. M. (1994). Communication, commitment, and cooperation in social dilemmas. *Journal of Personality and Social Psychology*, 66, 513-529.
- Keser, C., & van Winden, F. (in press). Conditional cooperation and voluntary contributions to public goods. *Scandinavian Journal of Economics*.
- Komorita, S. S., Chan, D. K., & Parks, C. (1993). The effects of reward structure and reciprocity in social dilemmas. *Journal of Experimental Social Psychology*, 29, 252-267.
- Komorita, S. S., Hilty, J. A., & Parks, C. D. (1991). Reciprocity and cooperation in social dilemmas. *Journal of Conflict Resolution*, 35, 494-518.
- Komorita, S. S., & Parks, C. D. (1995). Interpersonal relations: Mixed-motive interaction. *Annual Review of Psychology*, 46, 183-207.
- Komorita, S. S., Parks, C. D., & Hulbert, L. G. (1992). Reciprocity and the induction of cooperation in social dilemmas. *Journal of Personality and Social Psychology*, 62, 607-617.
- Kramer, R. M., McClintock, C. G., & Messick, D. M. (1986). Social values and cooperative response to a simulated resource conservation crisis. *Journal of Personality*, 54, 576-592.
- Kurzban, R., & Houser, D. (in press). Individual differences and cooperation in a circular public goods game. *European Journal of Personality*.
- Ledyard, J. O. (1995). Public goods: A survey of experimental research. In J. Kagel & A. E. Roth (Eds.), *Handbook of experimental economics* (pp. 111-194). Princeton, NJ: Princeton University Press.
- Liebrand, W. B. (1984). The effect of social motives, communication and group size on behaviour in an N-person multi-stage mixed-motive game. *European Journal of Social Psychology*, 14, 239-264.
- Longford, N. T. (1993). *Random coefficient models*. New York: Oxford University Press.
- Marwell, G., & Ames, R. (1979). Experiments in the provision of public goods, I: Resources, interest, group size and the free rider problem. *American Journal of Sociology*, 84, 1335-1360.
- McClintock, C. G., & Liebrand, W.B.G. (1988). Role of interdependence structure, individual value orientation, and another's strategy in social decision making: A transformational analysis. *Journal of Personality and Social Psychology*, 55, 396-409.
- Messick, D. M., Wilke, H., Brewer, M. B., Kramer, R. M., Zemke, P. E., & Lui, L. (1983). Individual adaptations and structural change as solutions to social dilemmas. *Journal of Personality and Social Psychology*, 44, 294-309.
- Milgrom, P. R., & Roberts, J. (1991). Adaptive and sophisticated learning in normal form games. *Games and Economic Behavior*, 3, 82-100.
- Olson, M. (1965). *The logic of collective action*. Cambridge, MA: Harvard University Press.
- Osgood, C. E. (1962). *An alternative to war or surrender*. Urbana: University of Illinois Press.
- Parks, C. D. (1994). The predictive ability of social values in resource dilemmas and public goods games. *Personality and Social Psychology Bulletin*, 2, 431-438.
- Parks, C. D., & Hulbert, L. G. (1995). High and low trusters' responses to fear in a payoff matrix. *Journal of Conflict Resolution*, 39, 718-730.
- Parks, C. D., & Komorita, S. S. (1997). Reciprocal strategies for large groups. *Personality and Social Psychology Review*, 4, 314-322.
- Pruitt, D. G., & Kimmel, M. J. (1977). Twenty years of experimental gaming: Critique, synthesis, and suggestions for the future. *Annual Review of Psychology*, 28, 363-392.
- Rabin, M. (1993). Incorporating fairness into game theory and economics. *American Economic Review*, 83, 1281-1302.
- Rapoport, A., & Eshed-Levy, D. (1989). Provision of step-level public goods: Effects of greed and fear of being gyped. *Organizational Behavior & Human Decision Processes*, 44, 325-344.
- Rapoport, A., & Suleiman, R. (1993). Incremental contribution in step-level public goods games with asymmetric players. *Organizational Behavior and Human Decision Processes*, 55, 171-194.

- Roberts, G., & Sherratt, T. N. (1998). Development of cooperative relationships through increasing investment. *Nature*, 394, 175.
- Schelling, T. (1960). *The strategy of conflict*. Cambridge, MA: Harvard University Press.
- Smith, V. L. (1991). The principle of unanimity and voluntary consent in social choice. In V. L. Smith (Ed.), *Papers in experimental economics* (pp. 377-391). Cambridge, UK: Cambridge University Press.
- Sugden, R. (1984). Reciprocity: The supply of public goods through voluntary contributions. *Economic Journal*, 94, 772-787.
- Van Lange, P.A.M., & Visser, K. (1999). Locomotion in social dilemmas: How people adapt to cooperative, tit-for-tat, and noncooperative partners. *Journal of Personality and Social Psychology*, 77, 762-773.
- Ward, H. (1989). Testing the waters: Taking risks to gain reassurance in public goods games. *Journal of Conflict Resolution*, 33, 274-308.
- Wit, A. P., & Wilke, H.A.M. (1992). The effect of social categorization on cooperation in three types of social dilemmas. *Journal of Economic Psychology*, 13, 135-151.
- Yamagishi, T. (1986). The provision of a sanctioning system as a public good. *Journal of Personality and Social Psychology*, 51, 110-116.
- Yamagishi, T., & Sato, K. (1986). Motivational bases of the public goods problem. *Journal of Personality and Social Psychology*, 50, 67-73.

Received November 15, 2000

Revision accepted January 10, 2001

Goals, Culture, and Subjective Well-Being

Shigehiro Oishi

University of Minnesota

Ed Diener

University of Illinois, Urbana-Champaign

The present studies examined the role of independent and interdependent goal pursuits in the subjective well-being (SWB) of Asian and European American college students. In Study 1, the authors found that independent goal pursuit (i.e., goal pursuit for fun and enjoyment) increased the benefit of goal attainment on SWB among European Americans but not among Asian Americans. In Study 2, the authors found that interdependent goal pursuit (i.e., goal pursuit to please parents and friends) increased the benefit of goal attainment on the SWB of Asian Americans, whereas it did not increase the benefit of goal attainment on the SWB of European Americans. In Study 3, the authors found that whereas interdependent goal pursuit increased the benefit of goal attainment, independent goal pursuit did not increase the benefit of goal attainment among Japanese college students. Altogether, the present findings suggest that independent and interdependent goal pursuits result in divergent affective consequences across cultures.

From daily experiences, all of us must recognize the pervasive role of goals in our lives, because achieving a goal or failing to do so makes our everyday lives enjoyable or miserable. For example, breaking one's personal record in a 5K race, receiving a rejection letter from a journal editor, hosting a successful cocktail party, and giving a horrible lecture are all likely to, at least temporarily, influence a person's sense of well-being. Indeed, there is ample evidence that goal attainment is associated with positive emotional experience (Brunstein, 1993) and life satisfaction (Emmons, 1986; see Cantor & Blanton, 1996; Emmons, 1996, for review). But is goal attainment equally good for anyone? Recently, researchers found that the effect of goal attainment on well-being varies depending on individuals' motives (e.g., Brunstein, Schultheiss, & Graessman, 1998; Emmons, 1991; Oishi, Diener, Suh, & Lucas, 1999; Sagiv & Schwartz, 2000; Sheldon & Kasser, 1998). For instance, Sheldon

and Kasser (1998) found that goal attainment had a very positive effect for those who pursued their goals for intrinsic reasons (i.e., for the fun and enjoyment they provide) but did not have any positive effect for those who pursued their goals for extrinsic reasons. The question regarding the effect of goal attainment on well-being takes on additional importance in light of cultural variation in goal motivation (e.g., Heine, Lehman, Markus, & Kitayama, 1999; Heine, Takata, & Lehman, 2000; Iyengar & Lepper, 1999; Kitayama, Markus, Matsumoto, & Norasakkunkit, 1997; Markus & Kitayama, 1991, 1994; Triandis, 1995). That is, is the type of person who benefits most from goal attainment the same or different across cultures? The present article tackles this question from the cultural psychological perspective (e.g., Heine et al., 1999; Markus & Kitayama, 1994; Miller, 1999) and examines the role of culture in the link between goal attainment and well-being.

Goals and Culture

Goals have been central constructs in cross-cultural and cultural psychology (e.g., Schwartz, 1992; Triandis, 1995). Most notably, Triandis (1995) distinguished individualist cultures from collectivist cultures by the type of goals that people pursue. He argued that people in individualist cultures tend to pursue personal goals that reflect personal desires, wishes, and needs, whereas people in collectivist cultures tend to pursue communal

Authors' Note: We thank Professor Hideki Okabayashi for collecting data used in Study 3. We also thank Christie Napa Scollon and Sumie Okazaki for their helpful comments on earlier versions of this article. Correspondence concerning this article should be addressed to Shigehiro Oishi, 75 East River Road, Department of Psychology, University of Minnesota, Minneapolis, MN 55455; e-mail: soishi@tc.umn.edu.

PSPB, Vol. 27 No. 12, December 2001 1674-1682

© 2001 by the Society for Personality and Social Psychology, Inc.

goals that reflect the desires, wishes, and needs of ingroup members (see also Schwartz, Sagiv, & Boehnke, 2000, for the link between values and daily concerns). In their seminal *Psychological Review* article, Markus and Kitayama (1991) also emphasized the interconnected nature of goals in the interdependent culture and noted that "the goals of others may become so focal in consciousness that the goals of others may be experienced as personal goals" (p. 229). Consistent with this thesis, Iyengar and Lepper (1999) have recently discovered that Asian American schoolchildren enjoyed and performed anagram and math problems better in an imposed condition (i.e., when they were told that the task was chosen by their mother or classmates) than in a free-choice condition. In contrast, European American schoolchildren enjoyed and performed the same problems better in a free-choice condition than in a chosen condition.

Based on the cultural variation in the type of salient goals, Markus and Kitayama (1994) proposed the culture-specific genesis of emotional well-being. These researchers posited that the attainment of culturally prescribed goals, or engagement in culturally appropriate behavior, should feel "good." To the extent that culturally prescribed goals in an independent culture are to stand out, feelings of separation and pride should lead to good feelings in an independent culture. On the other hand, to the extent that culturally prescribed goals in an interdependent culture are to fit in and have harmonious relationships, feelings of connection should lead to good feelings in an interdependent culture. Consistent with these hypotheses, Kitayama, Markus, and Kurokawa (2000) found that the frequency of good feelings was most closely associated with the frequency of friendly feelings in Japan, whereas it was most highly correlated with the frequency of pride in the United States. Also, consistent with the basic idea of Markus and Kitayama (1994), self-esteem (Diener & Diener, 1995) and freedom (Oishi, Diener, Lucas, & Suh, 1999) were significantly stronger predictors of life satisfaction in individualist cultures than in collectivist cultures. Similarly, relationship harmony had a predictive power of life satisfaction above and beyond self-esteem among Hong Kong students but not among American students (Kwan, Bond, & Singelis, 1997). In addition, the perception of a person's life by important others played a prominent role in predicting Asians' life satisfaction but played only a minor role in predicting European Americans' life satisfaction (Radhakrishnan & Chan, 1997; Suh, 1999). These findings suggest that the well-being of Asians may depend not only on how they view themselves but also on how they are viewed by important others (Heine et al., 1999; Triandis, 1995). Furthermore, the salience of the

external perspective among Asians (Suh, 1999) suggests that the type of goal progress conducive to Asians' well-being might be different in an important way from European Americans'.

The Present Studies

Although the previous cross-cultural studies (Diener & Diener, 1995; Heine & Lehman, 1999; Kwan et al., 1997; Oishi et al., 1999; Suh, 1999; Suh, Diener, Oishi, & Triandis, 1997) found important cultural variations in correlates of well-being, they were limited in two ways. First, because the previous studies relied entirely on global self-reports at one point in time, knowledge of specific processes and causal chains involving subjective well-being (SWB) was notably missing. What predicts changes in well-being? And how do these predictors differ across cultures? Second, despite the fact that goals have been an integral part of the cultural theory of the self (Markus & Kitayama, 1991) and individualism-collectivism (Triandis, 1995), they have not been directly measured and tested in the context of SWB in the previous research. Therefore, the role of goal attainment in SWB has never been examined in the cross-cultural context.

We conducted three studies to address these limitations from the previous research. In these studies, we tested the role of goal attainment and motivation in temporal changes in the well-being of Asians and European Americans. In all studies, participants first evaluated their recent life satisfaction at Time 1. Next, the participants listed the five most important goals for the next month (Study 1) or week (Studies 2 and 3) and rated the degree to which they pursued these goals for independent (Studies 1 and 3) or interdependent (Studies 2 and 3) reasons. Following Sheldon and Kasser (1998), we defined independent goal pursuit as pursuing a goal for the enjoyment and fun that it provides to them. We defined interdependent goal pursuit as pursuing a goal to make parents and friends happy. At Time 2 (i.e., 1 month later in Study 1 and 1 week later in Studies 2 and 3), the participants rated their well-being and their degree of goal attainment. Based on cultural variation in the function of motivation (Heine et al., 1999; Iyengar & Lepper, 1999; Markus & Kitayama, 1991), we hypothesized that progress toward goals pursued for interdependent reasons would lead to positive changes in well-being among Asians, whereas progress toward goals pursued for independent reasons would lead to positive changes in well-being among European Americans. The present studies extend the previous research by (a) providing more direct information on process and causal chains of SWB and (b) examining culture-specific functions of goals and motivation in SWB.

STUDY 1

*Method**PARTICIPANTS*

Participants were 87 European Americans (28 men, 57 women, 2 unknown) and 19 Asian Americans (7 men, 12 women) in a semester-long course on personality and well-being at the University of Illinois. The median age for European Americans was 20 years (range from 18 to 25 and older), whereas the median age for Asian Americans was 21 years (range from 18 to 23 years old). Eight of the 19 Asian American participants were born in the United States, and all but 3 participants have lived in the United States for at least 6 years.

MEASURES AND PROCEDURE

Monthly life satisfaction was measured by a 5-item scale based on the Satisfaction With Life Scale (SWLS) (Diener, Emmons, Larsen, & Griffin, 1985). Sample items include, "In most ways my life during the past month was close to ideal," "The conditions of my life during the past month were excellent," and "During the past month, I was satisfied with my life." Participants indicated their agreement on a 7-point scale (1 = *strongly disagree*, 2 = *disagree*, 3 = *slightly disagree*, 4 = *neither agree nor disagree*, 5 = *slightly agree*, 6 = *agree*, 7 = *strongly agree*). The mean Time 1 monthly satisfaction was 24.24 ($SD = 5.81$) for European Americans and 21.42 ($SD = 6.69$) for Asian Americans, $t = 1.86$, $p = .06$. Cronbach's alpha for the Time 1 monthly satisfaction scale was .89 for European Americans and .90 for Asian Americans. At Time 1, after completing the monthly life satisfaction scale, the participants listed their five most important goals in the coming month on a separate sheet of the paper. We assessed independent goal pursuit by using the scale developed by Sheldon and Kasser (1998); that is, for each goal, participants indicated their agreement on the statement, "I pursue this goal because of the fun and enjoyment that it provides me" using the 7-point scale (1 = *not at all true*, 7 = *absolutely true*). The index of independent goal pursuit was computed by taking the average of the ratings for this statement across the five goals. The mean independent goal pursuit was 3.93 ($SD = 1.22$) for European Americans and 3.67 ($SD = 1.56$) for Asian Americans, $t = .82$, ns . At Time 2 (exactly 1 month after the first assessment), the participants first rated their monthly satisfaction using the scale described above. The mean Time 2 monthly life satisfaction was 24.88 ($SD = 5.23$) for European Americans and 22.79 ($SD = 6.17$) for Asian Americans, $t = 1.52$, $p = .13$. Cronbach's alpha for the Time 2 monthly satisfaction scale was .86 for European Americans and .91 for Asian Americans. Then, the goal list was given back individually and the participants rated the degree of goal progress on each goal (i.e., How much did

you achieve this goal?) on the 7-point scale (1 = 0%, 4 = about 50%, 7 = 100%). The index of goal progress was computed by averaging the ratings for the five goals. The mean goal progress was 4.80 ($SD = 1.02$) for European Americans and 4.51 ($SD = 1.32$) for Asian Americans, $t = .98$, ns . We did not find any gender difference regarding weekly satisfaction. Also, the key interaction between goal progress and goal motives did not differ across gender in all three studies. Thus, we did not include gender in our analyses below.

Results and Discussion

Time 2 monthly life satisfaction was predicted from Time 1 monthly life satisfaction, independent goal pursuit, goal progress, and the interaction between independent goal pursuit and goal progress for each cultural group using a regression analysis with the centering procedure outlined by Aiken and West (1991). This analysis allowed us to examine the degree to which changes in monthly life satisfaction were predicted from independent goal pursuit, goal progress, and the interaction between independent goal pursuit and goal progress (see Cohen & Cohen, 1983, for details). Replicating the findings of Sheldon and Kasser (1998), we found a significant two-way interaction between goal progress and independent goal pursuit among European Americans ($B = 1.27$, $\beta = .22$, $p < .05$). As shown by the dotted lines in Figure 1, the degree of goal progress was, on average, positively associated with an increase in monthly life satisfaction. Furthermore, for European Americans, this tendency was significantly stronger for those who pursued the goals for independent reasons; that is, goal attainment was particularly beneficial to those who pursued their goals for independent reasons among European Americans. On the other hand, the interaction between independent goal pursuit and goal progress was not only nonsignificant but also negative among Asian Americans ($B = -.24$, $\beta = -.07$, ns). In other words, the benefit of goal progress was not greater for those Asian Americans who pursued their goals for independent reasons. In fact, the benefit of goal progress for those who pursued their goals for independent reasons was slightly smaller than those who pursued their goals for interdependent reasons (see solid lines in Figure 1). Thus, Study 1 indicates that whereas independent goal pursuit increases the positive effect of goal attainment on the well-being of European Americans, the positive function of independent goal pursuit does not seem to operate among Asian Americans.

STUDY 2

We conducted Study 2 to extend Study 1 in several ways. First, because of the small sample size of Asians, the estimates in Study 1 might not be as reliable as desired.

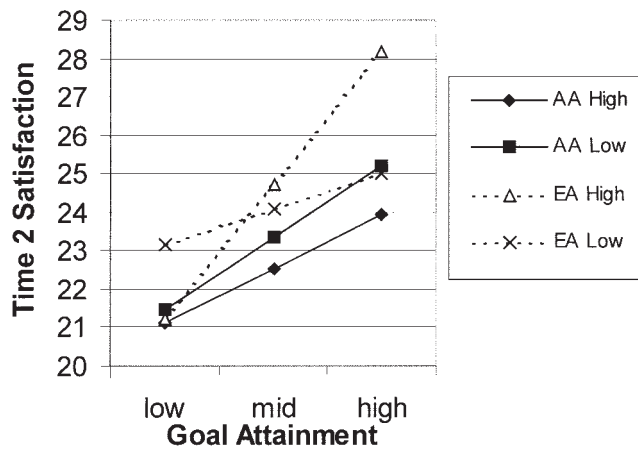


Figure 1 Adjusted Time 2 monthly satisfaction as a function of goal attainment for Asian Americans with high independent goal pursuit (AA High), Asian Americans with low independent goal pursuit (AA Low), European Americans with high independent goal pursuit (EA High), and European Americans with low independent goal pursuit (EA Low).

NOTE: The estimated regression equations for European Americans and Asian Americans are as follows: EA: $LS\ 2 = 20.04 + .18\ LS1 + 2.20\ GA + .30\ IGP + 1.27\ GA * IGP$; AA: $LS\ 2 = 13.73 + .43\ LS1 + 1.63\ GA - .41\ IGP - .24\ GA * IGP$, where $LS2$ = Time 2 monthly satisfaction, $LS1$ = Time 1 monthly satisfaction, GA = standardized goal attainment, and IGP = standardized independent goal pursuit. Following Aiken and West (1991), goal attainment and independent goal pursuit were standardized around the mean before forming the interaction term. The regression lines described above were computed using the mean Time 1 monthly satisfaction and 1 SD above (high) or below (low) the mean independent goal pursuit.

Thus, we obtained more Asian participants in Study 2. Second, Study 1 did not provide any information as to factors that could contribute to positive changes in the well-being of Asians. Finally, retrospective judgment of life satisfaction and goal attainment over 1 month might have led participants to use their general levels of life satisfaction and goal attainment. To reduce such a memory bias in assessment of goal progress and life satisfaction, we shortened the interval from 1 month to 1 week. This time frame should allow for more reality-based judgment of life satisfaction and goal progress in Study 2 than Study 1. Based on Iyengar and Lepper's (1999) findings on Asian Americans, we predicted that interdependent goal pursuit, or goal pursuit to make parents and friends happy, would enhance the positive effect of goal attainment on the well-being of Asian Americans.

Method

PARTICIPANTS

Participants were 67 European Americans (34 men, 27 women, 6 unknown) and 64 Asian Americans (29 men, 30 women, 5 unknown) enrolled in an introductory psychology course at the University of Illinois.

MEASURES AND PROCEDURE

Weekly satisfaction was assessed by a three-item scale based on the SWLS. The items include, "I am satisfied with the past 1 week of my life" and "The conditions of my life during the last week were excellent." Participants indicated their agreement on a 7-point scale (1 = *strongly disagree*, 4 = *neither agree nor disagree*, 7 = *strongly agree*). The index of weekly satisfaction was computed by taking the average of the ratings for three statements. The mean weekly satisfaction was 4.14 ($SD = 1.22$) for European Americans and 4.12 ($SD = 1.38$) for Asian Americans, $t = .09$, *ns*. Cronbach's alpha of the Week 1 satisfaction scale was .88 for European Americans and .87 for Asian Americans. At Time 1, participants listed the five most important goals for the next 7 days. Then, for each goal, they indicated their agreement with the statement, "I pursue this goal because I want to make my parents and friends happy" on a 7-point scale (1 = *not at all true*, 4 = *somewhat true*, 7 = *absolutely true*). The mean interdependent goal pursuit was 3.70 ($SD = 1.57$) for European Americans and 3.71 ($SD = 1.34$) for Asian Americans, $t = .01$, *ns*. At Time 2 (1 week later), the participants returned to the same experimental laboratory and completed the weekly satisfaction scale. The mean Week 2 satisfaction was 4.38 ($SD = 1.35$) for European Americans and 4.57 ($SD = 1.36$) for Asian Americans, $t = -.70$, *ns*. Cronbach's alpha for the Week 2 satisfaction scale was .92 for European Americans and .91 for Asian Americans. Next, participants were provided with their own goal lists from Time 1 and rated their progress on each goal ("How much did you achieve this goal?") on the 7-point scale (1 = 0%, 4 = about 50%, 7 = 100%). The index of goal progress was computed by taking the average of the ratings for these five items. The mean goal progress was 4.83 ($SD = 1.16$) for European Americans and 4.75 ($SD = .95$) for Asian Americans, $t = -.40$, *ns*.

Results and Discussion

As in Study 1, Week 2 satisfaction was predicted from Week 1 satisfaction, interdependent goal pursuit, goal progress, and the interaction between interdependent goal pursuit and goal progress for each group. Consistent with Sheldon and Kasser's (1998) findings, there was a significantly negative interaction between interdependent goal pursuit and goal progress among European Americans ($B = -.32$, $\beta = -.26$, $t = 2.29$, $p < .05$). Among European Americans, the degree to which goal attainment was associated with positive changes in weekly satisfaction was significantly less for those who pursued their goals to make parents and friends happy than for those who did not pursue the goals for interdependent reasons (see dotted lines in Figure 2). On the other hand, the interaction between interdependent goal pursuit and goal progress was positive and nearly

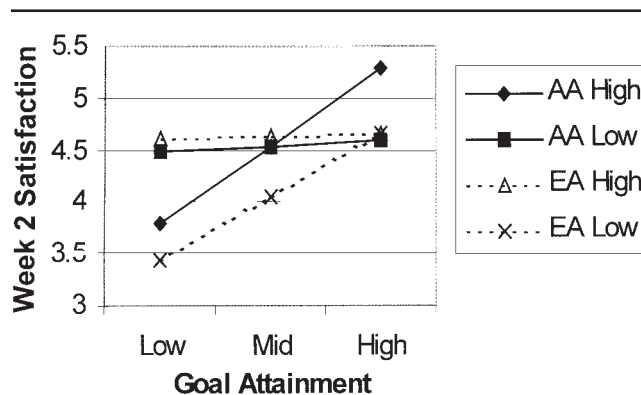


Figure 2 Adjusted Week 2 satisfaction as a function of goal attainment for Asian Americans with high interdependent goal pursuit (AA High), Asian Americans with low interdependent goal pursuit (AA Low), European Americans with high interdependent goal pursuit (EA High), and European Americans with low interdependent goal pursuit (EA Low).

NOTE: The estimated regression equations for European Americans and Asian Americans are as follows: EA: $LS2 = 2.36 + .48 LS1 + .29GA + .29DGP - .32GA \cdot DGP$; AA: $LS2 = 3.63 + .22LS1 + .40GA + .00DGP + .35GA \cdot DGP$, where $LS2$ = Week 2 satisfaction, $LS1$ = Week 1 satisfaction, GA = goal attainment, and DGP = interdependent goal pursuit. The regression lines described above were obtained using the same procedure as in Study 1.

significant among Asian Americans ($B = .35$, $\beta = .20$, $t = 1.46$, $p = .15$) (see solid lines in Figure 2). The 95% confidence interval for the unstandardized regression coefficient for the interaction term obtained in the Asian sample ranged from $-.13$ to $.82$, which excludes the unstandardized regression coefficient for the interaction term obtained in the European American sample ($B = -.32$). Furthermore, a regression analysis including both Asian and European Americans (i.e., predicting Week 2 satisfaction from Week 1 satisfaction, culture, goal progress, interdependent goal pursuit, and all the interaction terms) revealed a significant three-way interaction among culture (European vs. Asian Americans), interdependent goal pursuit, and goal progress ($B = -.33$, $\beta = -.23$, $t = 2.50$, $p = .01$). The three-way interaction indicates the contrasting role of interdependent goal pursuit on the effect of goal attainment on the well-being of Asian and European Americans. As seen in Figure 2, among Asians, goal progress was more conducive to weekly satisfaction for those who pursued their goals for interdependent reasons, whereas among European Americans, goal progress was less conducive to weekly satisfaction for those who pursued their goals for interdependent reasons. In short, although it replicated the previous research (Sheldon & Kasser, 1998) among European Americans, Study 2 revealed that interdependent goal pursuit, which was considered to be detrimental to well-being, could have a beneficial role in the well-being of Asian Americans.

STUDY 3

We conducted Study 3 to address three remaining issues from the first two studies. First, although the first two studies provided support for our hypothesis, we did not examine independent and interdependent goal pursuits in the same study. Second, although we followed the previous studies (e.g., Sheldon & Kasser, 1998) in measuring intrinsic goal pursuit ("because of fun and enjoyment that it provides me"), the item we used in Study 1 might not convey the concept of independent goal pursuit well. Also, the item we used for measuring interdependent goal pursuit in Study 2 ("because I want to make my friends and family happy") might not entirely represent the traditional definition of extrinsic motivation. Third, although we found the expected three-way interaction in Study 2, the two-way interaction between goal attainment and interdependent goal pursuit was not statistically significant among Asian Americans. This could be due to the fact that Asians in Study 2 lived in the United States. Indeed, previous research shows that Asians living in North America tend to show patterns of self-esteem and self-descriptions more individualistic than Asians living in Asia (e.g., Heine et al., 1999; Rhee, Uleman, Lee, & Roman, 1995). To address these issues, in Study 3, we examined both independent and interdependent goal pursuits, included two more items capturing the independent and interdependent nature of goal pursuit, and collected data from Japanese college students living in Japan.

Method

PARTICIPANTS

Participants were 70 Japanese students (20 men, 50 women) at Meisei University in Tokyo, Japan, who were enrolled in a research method course in psychology.

MEASURES AND PROCEDURE

All the materials were prepared in Japanese by the first author and administered in Japanese. Weekly satisfaction was measured by the same three-item scale used in Study 2. The mean weekly satisfaction was 4.44 ($SD = 1.64$) at Time 1. Cronbach's alpha for this scale was .89 at Time 1. As in Study 2, participants listed the five most important goals for the next 7 days at Time 1. Then, for each goal, they indicated their agreement with the following two statements used in Studies 1 and 2 (i.e., "I pursue this goal because of the fun and enjoyment that it provides me," "I pursue this goal because I want to make my parents and friends happy") on a 7-point scale (1 = *not at all true*, 4 = *somewhat true*, 7 = *absolutely true*). In addition, for each goal, they indicated their agreement with two additional statements: "I pursue this goal for myself, not for others" and "I pursue this goal to meet expectations of others," again on a 7-point scale (1 = *not at all true*,

TABLE 1: Descriptive Statistics and Correlations Among Four Goal Motives Among Japanese Participants in Study 3

Goal Motives	1	2	3	4
1. For fun and enjoyment	—	-.16	.29*	.24*
2. For self		—	-.36**	-.11
3. For family and friends			—	.67**
4. For expectations of others				—
<i>M (SD)</i>	3.38 (1.41)	6.49 (.65)	2.86 (1.27)	2.85 (1.39)

NOTE: $N = 70$. Goal motives are reasons why they pursued their goals.

* $p < .05$. ** $p < .01$.

4 = *somewhat true*, 7 = *absolutely true*). The descriptive statistics and correlations among four types of goal pursuit (i.e., the mean goal pursuit score across five goals) are shown in Table 1.

At Time 2 (1 week later), the participants returned to the same experimental laboratory and completed the weekly satisfaction scale. The mean Week 2 satisfaction was 4.28 ($SD = 1.57$) at Time 2. Cronbach's alpha for the Week 2 satisfaction scale was .93. Next, participants were provided with their own goal lists from Time 1 and rated their attainment on each goal ("How much did you achieve this goal?") on the 100-point scale, ranging from 0% to 100%. To make the rating easier, we changed the goal attainment scale from the artificially devised 7-point scale used in Study 2 to the more natural, 100% scale in this study. The index of goal attainment was computed by taking the average of the ratings for the five goals. The mean goal attainment was 54.93% ($SD = 20.96$). As recommended by Judd and McClelland (1989, p. 526), we transformed the percentage ratings provided by the participants using a logit transformation to normalize the distribution and the psychological meaning of intervals in percentages. The logit-transformed goal attainment score was used in the following analyses.

Results and Discussion

Goal motives. As seen in Table 1, Japanese participants pursued their goals for themselves to a greater extent than to make friends and family happy, $t(69) = 18.76$, $p < .01$, to meet the expectations of others, $t(69) = 19.05$, $p < .01$, or for fun and enjoyment, $t(69) = 15.86$, $p < .01$. As expected, goal pursuit for self was negatively correlated with goal pursuit to make friends and family happy. Also, as expected, goal pursuit to make friends and family happy was highly correlated with goal pursuit to meet the expectations of others. Interestingly, intrinsic goal pursuit (i.e., for fun and enjoyment) was positively correlated with goal pursuit to make friends and family happy and to meet the expectations of others. Thus, the

descriptive statistics and patterns of correlations among goal motives reveal an interesting picture of the Japanese participants. On one hand, these Japanese showed that they pursued their goals for independent reasons. On the other hand, the goals they pursued to make friends and family happy and to meet the expectations of others were the goals that were fun and enjoyable. Here, one can see that so-called extrinsic goal motives (e.g., Sheldon & Kasser, 1998) are highly internalized among the Japanese participants.

Hypothesis testing. As in Studies 1 and 2, Week 2 satisfaction was predicted from Week 1 satisfaction, goal pursuit, goal progress, and the interaction between goal pursuit and goal progress. We repeated this multiple regression analysis for each goal pursuit separately. Consistent with Study 1, the interaction between goal progress and intrinsic goal pursuit (i.e., goal pursuit for fun and enjoyment) was nonsignificant among Japanese college students ($B = .02$, $\beta = .01$, $t = .11$, ns); that is, goal progress was no more beneficial for the Japanese who pursued their goals for fun and enjoyment than for those who did not. Similarly, the interaction between goal progress and independent goal pursuit (i.e., goal pursuit for self, not for others) was also nonsignificant ($B = .06$, $\beta = .06$, $t = .41$). Therefore, goal progress was no more beneficial for the Japanese who pursued their goals for themselves than for those who did not. In other words, the previous findings on the positive benefit of intrinsic goal pursuit (e.g., Sheldon & Kasser, 1998) were not replicated with the Japanese.

On the other hand, consistent with Study 2, the interaction between goal progress and goal pursuit to make friends and family happy was marginally positive ($B = .37$, $\beta = .19$, $t = 1.73$, $p = .09$). A simple slope analysis (Aiken & West, 1991) revealed that for the Japanese who pursued their goals to make friends and family happy (1 SD above the mean), goal attainment was associated with a positive change in well-being (e.g., 1 SD increase in goal progress corresponded to .52 increase in Week 2 satisfaction). On the other hand, for the Japanese who did not pursue their goals to make their friends and family happy, 1 SD increase in goal attainment corresponded to .21 decrease in Week 2 satisfaction. Indeed, the obtained regression equation indicates that when goal attainment was average, those low in this goal pursuit reported slightly higher Week 2 satisfaction than those high in this goal pursuit (4.44 vs. 4.10). Nevertheless, when goal attainment was high, those high in family/friends' goal pursuit reported substantially higher Week 2 satisfaction than those low in parental goal pursuit (4.62 vs. 4.23). Consistent with Study 2, therefore, goal progress translated into a positive change in weekly satisfaction for the Japanese who pursued their goals to make their friends and family happy, whereas it did not bring more satisfac-

tion for the Japanese who did not pursue their goals to this end.

Finally, consistent with our hypothesis, the interaction between goal progress and goal pursuit to meet others' expectations was significantly positive ($B = .46$, $\beta = .27$, $t = 2.31$, $p = .02$). A simple slope analysis revealed that, as can be seen in Figure 3, 1 *SD* increase in goal attainment corresponded to .70 increase in Week 2 satisfaction for the Japanese high in this goal pursuit. On the other hand, 1 *SD* increase in goal attainment corresponded to .22 decrease in Week 2 satisfaction for the Japanese low in the goal pursuit for others' expectations. More specifically, when goal attainment was high (1 *SD* above the mean), the Japanese high in this goal pursuit reported much higher satisfaction than those low in this goal pursuit (4.75 vs. 4.25), although when goal attainment was average, the Japanese high in the extrinsic goal pursuit were not as satisfied as those low in the extrinsic goal pursuit (4.08 vs. 4.47). Therefore, goal progress had a more positive benefit for the Japanese who pursue their goals to meet the expectations of others than for those who do not.

GENERAL DISCUSSION

In three studies, we examined the role of independent and interdependent goal pursuit on the well-being of Asians and European Americans. Based on recent cross-cultural findings on motivation (Iyengar & Lepper, 1999) and self-construals (Heine et al., 1999; Markus & Kitayama, 1991), we predicted that the function of independent and interdependent goal pursuit on well-being would differ between Asians and European Americans. Consistent with our predictions, Study 1 showed that independent goal pursuit did not enhance the positive effect of goal attainment on the well-being of Asians while amplifying the benefit of goal attainment on the well-being of European Americans. Furthermore, Study 2 demonstrated that interdependent goal pursuit tended to increase the benefit of goal progress among Asians while diminishing the effect of goal progress among European Americans. Finally, Study 3 showed that goal progress was particularly beneficial for the well-being of the Japanese who pursued their goals to make their friends and family happy and to meet the expectations of others. Altogether, the present findings provide evidence that processes through which Asians and European Americans attain their well-being are different. European Americans appear to gain and maintain their well-being by achieving goals that they pursue for their own enjoyment and fun. On the other hand, Asian Americans seem to attain and maintain their well-being by achieving goals that they pursue to make important others happy and meet the expectations of others.

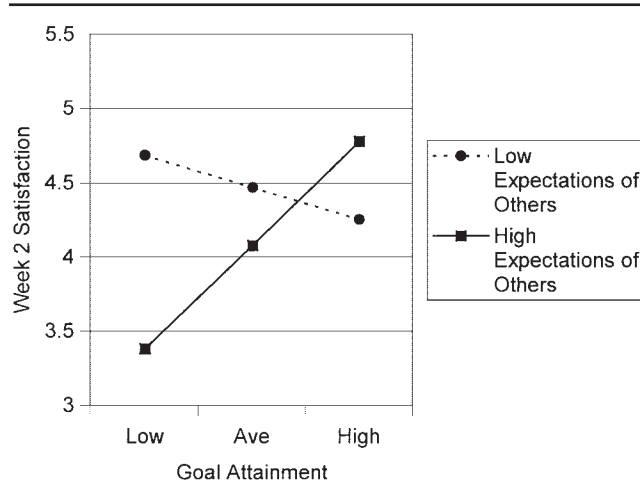


Figure 3 Adjusted Week 2 satisfaction as a function of goal attainment for Japanese high in goal pursuit to meet expectations of others and low in goal pursuit to meet expectations of others in Study 3.

NOTE: The estimated regression equation was as follows: $LS_2 = 2.34 + .436LS_1 + .244GA - .20GPE + .46GA * GPE$, where LS_2 = Week 2 satisfaction, LS_1 = Week 1 satisfaction, GA = standardized goal attainment, and GPE = standardized goal pursuit to meet others' expectations. With the mean Week 1 satisfaction = 4.44 in the equation, simple slopes for high and low in goal pursuit to meet the expectations of others (1 *SD* above or below mean) are as follows: High: $LS_2 = 4.08 + .70 GA$; Low: $LS_2 = 4.48 - .216GA$.

In American psychology, personal choice independent of others has been the sine qua non of spontaneous behavior (Lepper, Greene, & Nisbett, 1973) and mental health (Maslow, 1947; Rogers, 1961). To the extent that individuals pursue and achieve self-chosen goals, and to the extent that individuals feel that they are the driving forces of their lives, they feel good (Sheldon & Kasser, 1998). This theory perfectly captures American icons such as Michael Jordan and Bill Gates. As evidenced by the idealization of the self-made billionaire, the founder of the Softbank Masayoshi Son in Japan, self-determination has recently become a popular ideology in Asia as well. Indeed, Study 3 showed that on average, Japanese college students pursued their goals for themselves (6.49) much more frequently than to make family and friends happy (2.86) or to meet the expectations of others (2.85). Also, on average, the degree to which Asian participants in Studies 1 and 2 pursued their goals for intrinsic reasons or to make friends and family happy was very similar to European American counterparts. In other words, Japanese participants in Study 3 as well as Asian participants in Studies 1 and 2 are not as "collectivist" or "interdependent" as cultural theorists (e.g., Markus & Kitayama, 1991, 1994; Triandis, 1995) might assume in terms of goal motives. What is interesting, however, is that despite the similar levels of independent and interdependent goal motives across cultures, the very function of goal motives differed considerably

across cultures. That is, although both Japanese and European Americans pursue their goals for themselves, such independent goal pursuit does not generate as positive an outcome for Japanese as for European Americans.

Why does independent goal pursuit not work for Asians? One possibility is that because of the traditional value of conformity and deference to authority figures among Asians (e.g., Bond, 1988; Schwartz, 1994), Asians who subscribe to independent goal pursuit are more prone to psychological conflict than European Americans. This conflict, in turn, results in the lack of positive consequence of independent goal pursuit among Asian Americans. Although this explains the cultural difference in the function of independent goal pursuit, this does not fully explain the positive function of interdependent goal pursuit among Asians. Given the ubiquity of the idealization of independence in American culture (Wolfe, 2000), Asians who hold traditional Confucian values may be prone to psychological conflict between conformity and self-determination as much as Asians who prefer the mainstream American values. Thus, the value conflict hypothesis does not seem to fully account for the positive function of interdependent motivation among Asians. Alternatively, a more viable explanation for the current findings can be offered from the cultural theory of the self (Markus & Kitayama, 1991, 1994). According to this theory, Asians' self-concepts are so intertwined with expectations and perceptions by important others that expectations from important others could become their own goals among Asians' interdependent selves. To the extent that their goals overlap with expectations from important others, making their parents and friends happy becomes a key to their own sense of satisfaction. The flip side of this reasoning is that even if Asians achieved the goal they set for themselves, they would not feel satisfied if their parents or friends were not happy about their goals.

Different processes governing the well-being of Asians and European Americans also have an implication for cultural differences in mean levels of SWB. For years, researchers found that people in East Asia were less satisfied with their lives than European and North Americans (e.g., Diener, Diener, & Diener, 1995; Veenhoven, 1993). Whereas the necessary and sufficient condition for happiness for European Americans appears to be to make themselves happy by achieving their self-chosen goals, there seem to be more conditions for Asians. That is, for Asians to be happy they must not only satisfy themselves but also satisfy their parents and friends. To the extent that meeting one condition is easier than multiple ones, European Americans on the average can feel good about their lives more readily than Asians. Although this possibility must be examined more

fully in the future, it seems evident that the processes through which people with different self-construals attain their well-being have an immense implication for the mean level of well-being.

Future Directions

In the past, intrinsic motivation was viewed as fundamental and as innate as biological needs such as thirst and hunger (Maslow, 1947; Rogers, 1961). While evidence for the paramount importance of intrinsic motivation among European Americans continues to accumulate, the current findings suggest that intrinsic motivation may not be as biological or fundamental as once thought. Instead, the present findings suggest that the function of motivation is tailored in an important way by culture. Independent goal pursuit appears to be instilled early in life and positively reinforced by the mainstream American culture, whereas consideration for important others seems to be desirable and sometimes demanded in Asian American communities. As a result, expectations from important others seem to be deeply internalized and become integral parts of the self among Asians, which in turn provide standards for evaluating their own life experiences.

It should be noted, however, that the present explanation from the cultural theory of the self (Markus & Kitayama, 1991) remains incomplete in two respects. First, we did not measure self-concepts of Asians and European Americans. Thus, the above explanation is based on the assumption that the interdependent aspect of the self was salient for Asian Americans, whereas the independent aspect of the self was salient for European Americans in the present studies. Second, it is difficult to pinpoint a crucial factor responsible for cultural differences obtained in the present studies. Is it the salience of "I" versus "We"? Is it the value of conformity versus hedonism? Or is it the familiarity of personal choice and independent decision making? These questions still remain. It is critical, therefore, that future research identify the parsimonious conditions for these cultural differences by examining specific factors, such as accessibility of key concepts (e.g., Gardner, Gabriel, & Lee, 1999; Oishi, Wyer, & Colcombe, 2000) and thinking styles (Peng & Nisbett, 1999) in the context of goal progress, motivation, and SWB.

REFERENCES

- Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage.
- Bond, M. H. (1988). Finding universal dimensions of individual variation in multicultural studies of values: The Rokeach and Chinese value surveys. *Journal of Personality and Social Psychology*, 55, 1009-1015.
- Brunstein, J. (1993). Personal goals and subjective well-being: A longitudinal study. *Journal of Personality and Social Psychology*, 65, 1061-1070.

- Brunstein, J. C., Schultheiss, O. C., & Graessman, R. (1998). Personal goals and emotional well-being: The moderating role of motive dispositions. *Journal of Personality and Social Psychology*, 75, 494-508.
- Cantor, N., & Blanton, H. (1996). Effortful pursuit of personal goals in daily life. In P. M. Gollwitzer & J. A. Bargh (Eds.), *The psychology of action: Linking cognition and motivation to behavior* (pp. 338-364). New York: Guilford.
- Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlational analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.
- Diener, E., & Diener, M. (1995). Cross-cultural correlates of life satisfaction and self-esteem. *Journal of Personality and Social Psychology*, 68, 653-663.
- Diener, E., Diener, M., & Diener, C. (1995). Factors predicting the subjective well-being of nations. *Journal of Personality and Social Psychology*, 69, 851-864.
- Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction With Life Scale. *Journal of Personality Assessment*, 49, 71-75.
- Emmons, R. A. (1986). Personal strivings: An approach to personality and subjective well-being. *Journal of Personality and Social Psychology*, 51, 1058-1068.
- Emmons, R. A. (1991). Personal strivings, daily life events, and psychological and physical well-being. *Journal of Personality*, 59, 453-472.
- Emmons, R. A. (1996). Striving and feelings: Personal goals and subjective well-being. In P. M. Gollwitzer & J. A. Bargh (Eds.), *The psychology of action: Linking cognition and motivation to behavior* (pp. 313-337). New York: Guilford.
- Gardner, W. L., Gabriel, S., & Lee, A. Y. (1999). "I" value freedom, but "we" value relationships: Self-construal priming mirrors cultural differences in judgment. *Psychological Science*, 10, 321-326.
- Heine, S. J., & Lehman, D. R. (1999). Culture, self-discrepancies, and self-satisfaction. *Personality and Social Psychology Bulletin*, 25, 915-925.
- Heine, S. J., Lehman, D. R., Markus, H. R., & Kitayama, S. (1999). Is there a universal need for positive self-regard? *Psychological Review*, 106, 766-794.
- Heine, S. J., Takata, T., & Lehman, D. R. (2000). Beyond self-presentation: Evidence for self-criticism among Japanese. *Personality and Social Psychology Bulletin*, 26, 71-78.
- Iyengar, S. S., & Lepper, M. R. (1999). Rethinking the value of choice: A cultural perspective on intrinsic motivation. *Journal of Personality and Social Psychology*, 76, 349-366.
- Judd, C. M., & McClelland, G. H. (1989). *Data analysis: A model-comparison approach*. Orlando, FL: Harcourt Brace Jovanovich.
- Kitayama, S., Markus, H. R., & Kurokawa, M. (2000). Culture, emotion, and well-being: Good feelings in Japan and the United States. *Cognition and Emotion*, 14, 93-124.
- Kitayama, S., Markus, H. R., Matsumoto, H., & Norasakkunkit, V. (1997). Individual and collective processes in the construction of the self: Self-enhancement in the United States and self-criticism in Japan. *Journal of Personality and Social Psychology*, 72, 1245-1267.
- Kwan, V. S. Y., Bond, M. H., & Singelis, T. M. (1997). Pancultural explanations for life satisfaction: Adding relationship harmony to self-esteem. *Journal of Personality and Social Psychology*, 73, 1038-1051.
- Lepper, M. R., Greene, D., & 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.
- Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, 98, 224-253.
- Markus, H. R., & Kitayama, S. (1994). The cultural construction of self and emotion: Implications for social behavior. In S. Kitayama & H. R. Markus (Eds.), *Emotion and culture* (pp. 89-130). Washington, DC: American Psychological Association.
- Maslow, A. (1947). *Motivation and personality*. New York: Harper & Row.
- Miller, J. (1999). Cultural psychology: Implications for basic psychological theories. *Psychological Science*, 10, 85-91.
- Oishi, S., Diener, E., Lucas, R. E., & Suh, E. M. (1999). Cultural variation in predictors of life satisfaction: A perspective from needs and values. *Personality and Social Psychology Bulletin*, 25, 980-990.
- Oishi, S., Diener, E., Suh, E. M., & Lucas, R. E. (1999). Value as a moderator in subjective well-being. *Journal of Personality*, 67, 157-184.
- Oishi, S., Wyer, R. S., Jr., & Colcombe, S. (2000). Cultural variation in the use of current life satisfaction to predict the future. *Journal of Personality and Social Psychology*, 78, 434-445.
- Peng, K., & Nisbett, R. E. (1999). Culture, dialectics, and reasoning about contradictions. *American Psychologist*, 54, 741-754.
- Radhakrishnan, P., & Chan, D. K.-S. (1997). Cultural differences in the relation between self-discrepancy and life satisfaction. *International Journal of Psychology*, 32, 387-398.
- Rhee, E., Uleman, J. S., Lee, H. K., & Roman, R. J. (1995). Spontaneous self-descriptions and ethnic identities in individualistic and collectivistic cultures. *Journal of Personality and Social Psychology*, 69, 142-152.
- Rogers, C. (1961). *On becoming a person*. Boston: Houghton-Mifflin.
- Sagiv, L., & Schwarz, S. H. (2000). Value priorities and subjective well-being: Direct relations and congruity effects. *European Journal of Social Psychology*, 30, 177-198.
- Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 25, pp. 1-65). Orlando, FL: Academic Press.
- Schwartz, S. H. (1994). Beyond individualism-collectivism: New cultural dimensions of values. In U. Kim, H. C. Triandis, C. Kagitcibasi, S.-C. Choi, & G. Yoon (Eds.), *Individualism and collectivism: Theory, method, and applications* (pp. 85-122). Thousand Oaks, CA: Sage.
- Schwartz, S. H., Sagiv, L., & Boehnke, K. (2000). Worries and values. *Journal of Personality*, 68, 309-346.
- Sheldon, K. M., & Kasser, T. (1998). Pursuing personal goals: Skills enable progress, but not all goal progress is beneficial. *Personality and Social Psychology Bulletin*, 24, 1319-1331.
- Suh, E. M. (1999). *Self and the use of emotion information: Joining culture, personality, and situational influences*. Manuscript submitted for publication, University of California, Irvine.
- Suh, E., Diener, E., Oishi, S., & Triandis, H. C. (1997). The shifting basis of life satisfaction judgments across cultures: Emotions versus norms. *Journal of Personality and Social Psychology*, 74, 482-493.
- Triandis, H. C. (1995). *Individualism and collectivism*. Boulder, CO: Westview.
- Wolfe, A. (2000, May 7). The pursuit of autonomy. *New York Times Magazine*, pp. 53-56.
- Veenhoven, R. (1993). *Happiness in nations: Subjective appreciation of life in 56 nations 1946-1992*. Rotterdam, the Netherlands: Erasmus University Rotterdam.

Received January 23, 2001

Revision accepted February 14, 2001

Assessing Generalization in Perceived Self-Efficacy: Multidomain and Global Assessments of the Effects of Self-Defense Training for Women

Julie C. Weitlauf

Daniel Cervone

University of Illinois at Chicago

Ronald E. Smith

University of Washington

Paul M. Wright

University of Illinois at Chicago

The authors assessed the impact of self-defense training for women on multifaceted aspects of perceived self-efficacy. As compared to a waiting list control condition, training increased self-efficacy perceptions not only for self-defense skills but also across a variety of domains, including self-defense abilities, sports competencies, and coping skills. Trained participants also experienced a significant increase in more global aspects of personality, including perceptions of physical self-efficacy and assertiveness. No changes were detected on a trait measure of global self-efficacy; however, there was a significant change on a composite score of a multidomain self-efficacy questionnaire and on several domain-specific subscales, indicating that trained participants experienced a boost in multiple domains of self-efficacy not directly tapped by the intervention. Implications for constructing more sensitive measures of coping skills generalization effects are discussed.

A major challenge faced by practitioners who seek to foster psychological change is to develop interventions whose effects generalize (Goldfried & Robbins, 1982; Smith, 1989, 1999; Weitlauf, Smith, & Cervone, 2000). Ideally, interventions will benefit participants not only in the particular social and behavioral domains targeted by the treatment but also throughout multiple aspects of individuals' daily lives. This challenge has both practical and theoretical components. Assessing generalization requires a theory that addresses two primary questions: (a) What are the psychological qualities that one should assess to investigate generalization? and (b) What is the proper way of assessing those qualities? The difficulties

involved in answering these questions are evidenced by the fact that the issue of cross-situational generality has been among the most vexing topics in all of personality psychology (Caprara & Cervone, 2000; Mischel, 1968).

We address these theoretical and methodological challenges by drawing on the general framework of social-cognitive theories of personality (reviewed in Cervone & Shoda, 1999) and the more specific framework of self-efficacy theory (Bandura, 1977, 1997). Social-cognitive theory highlights the dynamic cognitive and affective processes underlying personality functioning and coherence. In this perspective, assessment goes beyond the examination of overt psychological tendencies and focuses on dynamic psychological systems that causally contribute to personality functioning. Second, assessment is contextualized; that is, psychological quali-

Authors' Note: This article is based on an ongoing research program, the Women's Self-Defense Project, directed by the first author under the direction of the second author at the University of Illinois at Chicago (UIC). Portions of this research were presented at the 1999 American Psychological Society Convention in Denver, Colorado. The Women's Self-Defense Project was supported by a grant through the UIC Campus Research Board. The authors wish to thank UIC's Residential Life Program and Department of Kinesiology for their cooperation in the facilitation of this project. In addition, we would like to thank Mary Ann Pec for her help in managing the project and to Gloria Balague, David McKirnan, and Donald Hellison for their contributions to the article. Correspondence concerning this article should be directed to Julie Weitlauf, VA Palo Alto Health Care System, Psychology Service (116B), 3801 Miranda Ave., Palo Alto, CA 94304; e-mail: weitlauf@uic.edu.

PSPB, Vol. 27 No. 12, December 2001 1683-1691

© 2001 by the Society for Personality and Social Psychology, Inc.

ties are assessed with respect to the contexts that make up people's daily lives (Cervone, Shadel, & Jencius, 2001). Contextualized assessment showcases patterns of behavioral variability across contexts; such idiosyncratic variations in psychological functioning are a basic fact of social life (Bugenthal, 2000) and can critically distinguish individuals from one another (Mischel, 1999; Mischel & Shoda, 1995).

Perceived Self-Efficacy

In the study of generalization of treatment effects, social cognitive theory places particular emphasis on people's subjective perceptions of their capabilities for performance, or self-efficacy perceptions (Bandura, 1997; Cervone & Scott, 1995). Self-efficacy theory posits that self-efficacy perceptions govern a set of behavioral, cognitive, and affective processes that are critical to performance success (Bandura, 1999). People with a higher sense of self-efficacy on tasks tend to persist in their efforts when faced with challenges (Cervone & Peake, 1986), display less anxiety in response to threats (Bandura, Taylor, Williams, Mefford, & Barchas, 1985), and report fewer anxiety-related cognitions (Sarason, Pierce, & Sarason, 1996).

Research in many important activity domains documents the critical role of self-efficacy beliefs in facilitating adaptive behavior. For example, strong self-efficacy beliefs can facilitate the promotion of health behaviors (Schwarzer, 1995), including recovery from physical setbacks (Ewart, 1992), the control of eating (Glynn & Ruderman, 1986), and safer sex practices (Montoya, 1998). Robust self-efficacy beliefs improve performance in the workplace (Stajkovic & Luthans, 1998) and promote success in athletic pursuits (Escarti & Guzman, 1999).

Assessing Generalization in Perceived Self-Efficacy

Throughout the social cognitive literature, efficacy perceptions generally have been assessed in a contextualized manner. Measures tap people's appraisals of their capabilities to attain specific levels or types of performances in designated contexts (Bandura, 1977; Cervone, 1985). The use of these contextualized measures naturally raises the question of how one should assess cross-context generality. Here there are two options. The first is to employ measures that are decontextualized. Instead of asking individuals to appraise their capabilities to handle designated challenges and circumstances, questionnaire items might tap generalized beliefs, such as whether one sees oneself as a competent, efficacious person. Such procedures yield global measures of the construct "generalized self-efficacy" (Coppel, 1980; Sherer et al., 1982; Schwarzer, Babler, Kwiatek, & Shroder, 1997). The question of whether

intervention produces generalized effects can then be operationalized by examining the influence of the intervention on measures of this generalized construct (e.g. Smith, 1989; Weitlauf et al., 2000).

Despite their popularity, the use of decontextualized measures of global self-efficacy clashes with the recommendations of social cognitive theory (Bandura, 1997). Rather, Bandura (1997) recommends a contextualized approach to assessment. In this approach, self-efficacy assessments tap "beliefs about personal abilities to produce specified levels of performance" (Bandura, 1997, p. 45). "In no case" are these assessments "dissociated from context" (Bandura, 1997, p. 50). Instead, items describe a behavioral setting and activity to be performed and ask people to indicate their certainty that they can perform the designated task in that setting. "Items linked to major activity domains" are seen to be "an improvement over omnibus measures that are dissociated from clearly defined activities and contextual factors" (Bandura, 1997, p. 48). Thus, domain-based assessments are favored because domain-linked measures are stronger predictors of performance than are global indices (Bandura, 1997).

A second possible empirical advantage of domain-based versus global assessments has received little attention. This is the possibility that domain-based assessments will prove to be more sensitive indices of the generalization of treatment effects. People commonly cling to abstract, global beliefs about the self that may be at variance with their day-to-day experiences (Mischel, 1968). As a result, assessments of global beliefs may suggest that a given intervention did not produce effects that generalized. However, a domain-based approach may reveal that the intervention did, in fact, generalize beyond the particular domain in which the intervention occurred. This possibility is a central focus of the present research. To our knowledge, no prior studies have employed both global and multidomain measures of self-efficacy; thus, no studies directly speak to this potential advantage of domain-based self-efficacy assessment.

Global self-efficacy measures, such as the Sherer et al. (1982) scale, assess global feelings of personal competence by asking broad questions such as, "When I make plans, I am certain that I can make them work." By contrast, a domain-based measure of self-efficacy taps specific, self-efficacy beliefs in a particular domain by asking highly contextualized questions (i.e., "How confident are you that you can be on time for an 8:00 class every week throughout a busy semester?"). A number of studies have employed contextualized domain-linked measures and found that treatments can generalize from one domain to another. These include studies of the effects of the clinical treatment of agoraphobia (Williams, Kinney, & Falbo, 1989), self-defense training for women

(Ozer & Bandura, 1990), and the mastery of a high-risk sport (Brody, Hatfield, & Spalding, 1988). These studies, however, provide limited information about generalization in that the self-efficacy measures used tapped relatively narrow ranges of functioning in domains that were highly related to the treatment domain.

In the present study, we developed a contextualized multidomain measure that assesses self-efficacy in a variety of activity domains (i.e., academic achievement, sports, interpersonal coping skills, assertiveness, and conscientiousness). The advantage of this Multidomain Self-Efficacy Scale is that it allows us to assess generalization of self-efficacy beliefs to activity domains that are both highly related to the training as well as to domains that are quite distinct.

Self-Defense Training for Women

Self-defense training for women provided an ideal psychosocial intervention to examine generalization effects. Self-defense training is well suited to the questions asked here in that the intervention itself is highly contextualized. Women are trained in skills that are specifically designed to enable them to cope with physical or sexual assault. The question we ask, then, is whether this training produces changes in self-efficacy domains beyond physical self-protection. We hypothesized that women who participated in self-defense training would experience boosts in self-efficacy across a range of activities that do not directly involve self-defense.

Generalization was assessed using both decontextualized global scales of generalized self-efficacy and a contextualized (multidomain) scale that assessed self-efficacy across a variety of activity domains. We also explored the relationship of participant schematic with several hypotheses. First, we expected to replicate the "empowerment effect" of self-defense training on women demonstrated by Ozer and Bandura (1990) and by Weitlauf et al. (2000). That is, we expected that training would significantly enhance women's confidence in their abilities to defend themselves if attacked but also that such training would have a positive impact on more global beliefs of competence and physical ability. Second, we wanted to explore and compare the sensitivity of contextualized measures of generalized self-efficacy with more global, traitlike constructs. We suspected that the contextualized multidomain self-efficacy measure would be a more sensitive index of generalization than a global trait measure. That is, women who completed training would show significant increases in self-efficacy across a variety of life domains, and this change would be more evident than increases in global perceptions of efficacy as measured by a generalized self-efficacy questionnaire. Finally, we assessed generalization to several relevant

personality domains, including self-esteem, assertiveness, and aggression.

In one experimental condition, we attempted to enhance generalization effects by embellishing training with structured discussions and writing assignments designed to help participants connect self-defense coping skills with other life activities. We reasoned that the generalization of self-efficacy beliefs from one domain to another at least partly reflects constructive cognitive activities. People may reflect on the skills they have gained in one domain, or on the new knowledge about themselves that they have gained while acquiring those skills, and then apply their skills and self-knowledge to other life domains (Smith, 1999).

METHOD

Experimental Design

We used a 3 (groups) \times 2 (measurement times) repeated measures mixed design in which participants were randomly assigned to one of two training conditions or a waiting list control condition. Assessments of the outcome variables were completed at pretraining or baseline (Time 1) and immediately following the conclusion of the intervention at posttraining (Time 2).

Participants

Participants were 125 female undergraduate students enrolled in introductory psychology at a large urban university in the midwestern United States. Participants responded to posted advertisements for free self-defense training and were subsequently screened for eligibility. Women were eligible to participate if they had no previous or concurrent training in the martial arts, self-defense, or weapons training. Of the original 125 women, 115 qualified to participate and 96 completed the training. The total attrition rate of qualified participants was 17% and was not related to condition. Participants ranged in age from 18 to 26 years ($M = 19.02$, $SD = 1.50$). Caucasians constituted 30.20% of the sample, 17.70% were African American, 21.90% were Latino, 24.00% were Asian, 5.20% were multiracial, and 1.0% were Native American.¹

Procedure

Participants assigned to the self-defense conditions were sent schedules of class times and location and began their training immediately. Pre- and posttraining data collection occurred on-site during the first and last class sessions. Participants assigned to the waiting list control condition were informed that they would begin their training immediately following the completion of the first session of classes (a waiting period of 4 weeks) and were instructed to attend two data collection ses-

sions that corresponded to the first and last weeks of the experimental group's training. Waiting list participants received no intervention during the waiting period and were instructed not to pursue self-defense or related training during this time. Control participants also were sent information about the dates, times, and locations of their self-defense course as well as dates and times for data collection.

Self-defense training condition. The basic experimental intervention was a 16-hour, intensive, physical self-defense course specifically designed for college-age women. The training, consisting of four weekly, 4-hour training sessions, was designed to teach verbal and physical resistance to rape. The first and fourth authors, both experienced self-defense instructors, conducted all sessions. Make-up sessions were offered on several occasions so that participants who could not attend their usual group could attend an alternate session. This resulted in full attendance by all experimental group participants.

Emotional resistance to rape was addressed through coping skills training that taught participants emotion management and attentional control techniques that would help them focus their resistance efforts during an attack. Participants were taught techniques to help secure their breathing, remain calm, and "think their way out of an assault." During the verbal resistance phases of training, participants were taught to use their voice as an active tool in resistance.

Participants also were taught basic physical resistance techniques. Techniques and moves were drawn primarily from the martial arts of Aikido, Shotokan Karate, and Hapkido. Striking techniques included basic punches and blocks, groin and knee kicks, and several other upper-body striking techniques such as the hammer fist and knife hand strikes. Participants learned to develop their form and, using a blocking shield, practiced strikes at full force during each subsequent training session. Participants learned a variety of methods for freeing themselves from an assailant when grabbed at the hand, wrist, arm, shoulder, throat, waist, or hair. Special emphasis was placed on breaking out of strangulation holds and resisting choking.

Participants were repeatedly drilled in all physical and verbal skills as a part of training. Continuous corrective feedback was provided to all students by both instructors and, as the course progressed, by participants to one another. In addition, participants used classroom mirrors as an added means of continuous feedback to help with the acquisition and correction of skills. To ensure skill mastery, participants completed 2,000 repetitions of each striking technique and did not take on more advanced techniques until mastering more basic ones.²

Enhanced self-defense training condition. Participants in this condition were given an additional 30 minutes of training per week during which they engaged group discussions and writing assignments designed to help them reflect on how the training might be applicable in their lives. Some of the discussion and journal exercises were free form; however, much of this enhanced training component was guided by specific questions designed to help them think broadly about how this training might affect their lives.

Measures

Self-defense efficacy. Participants completed a 16-item self-efficacy measure that assessed their perceived capability to execute a variety of self-defense skills, recognize the warning signs of a potential attacker, and obtain access to the appropriate medical and legal resources to help them cope with an assault. These items specifically targeted the curriculum offered in the self-defense course. Sample items are as follows: "I have the skill/ability to execute (hurt or disable an attacker via) a variety of close-range techniques, including various forms of a basic hand strike (straight punch, knife hand, hammer fist, etc.)" and "I have the skill/ability to execute (hurt or disable an attacker via) a variety of kicking techniques, including a groin and knee kick." Participants rated their perceived ability to execute each of these skills on 10-point scales, with 1 labeled *not competent at all* and 10 labeled *very competent*. A score was computed by averaging across items. Reliability analyses indicated that this measure has an acceptable level of internal consistency; Cronbach's alpha at Time 1 was .94 and .98 at Time 2.

Global self-efficacy measures. The Physical Self-Efficacy Scale³ (Ryckman, Robbins, Thorton, & Cantrell, 1982) was administered to assess perceived self-efficacy in the physical or athletic domain. This 22-item scale assesses perceptions of one's body, athletic ability, coordination, and motor capabilities. Sample items are as follows: "I take little pride in my ability in sports" and "Because of my agility, I have been able to do things which many others could not do." Each item is scored on a 6-point scale. Internal consistency was adequate; Cronbach's alpha was .82 at Time 1 and .78 at Time 2.

To assess an even more global level of efficacy beliefs, Sherer et al.'s (1982) general self-efficacy subscale was used to assess feelings of general capability, productivity, and self-regulation (see Note 3). Sample items include the following: "When I make plans, I am certain that I can make them work" and "When I have something unpleasant to do, I stick to it until I finish it." Each of the 17 items is scored on a 14-point scale, and a total score was used for analysis. Reliability analyses indicated that Cronbach's alpha was .86 at Time 1 and .81 at Time 2.

Domain-specific self-efficacy. Our domain-linked measure of perceived self-efficacy was a 32-item multidomain scale designed to tap efficacy beliefs in a variety of commonly occurring settings. Specifically, the scale tapped efficacy beliefs in five domains of functioning: sports (“If you’re competing in a 5K/3-mile road race, how confident are you that you would be able to finish the race in the top 25% of the people of your age and sex?”), academic achievement (“If you’re taking a political science class, how confident are you that you would be able to earn at least a B in the class?”), conscientiousness (“If you’re studying for an exam and your favorite show comes on TV, how confident are you that you would be able to get yourself to turn off the TV and stick to your studying?”), assertive communication (“If you are feeling completely lost in a course, how confident are you in your ability to go to the professor or a teaching assistant and get help?”), and coping⁴ (“If you have to give an oral presentation in class, how confident are you in your ability to remain calm and collected when talking in front of the class?”). For each item, participants responded on a 100-point scale, with 1 being *not at all confident*, 50 being *somewhat confident*, and 100 being *very confident*. Items were summed to provide a measure of perceived self-efficacy within each of the five domains. A composite score was computed by aggregating across all items of the scale. Reliability analyses indicated that this composite had an acceptable level of internal consistency, with Cronbach’s alpha being .92 at Time 1 and .94 at Time 2. On the individual domain-linked scales, Time 1 alphas ranged from .60 (sports) to .81 (coping) and Time 2 alphas ranged from .72 (sports) to .83 (coping).

Personality Trait Measures

The Rosenberg Self-Esteem Scale (Rosenberg, 1965) taps global feelings of personal value and self-worth. Sample items include the following: “I feel that I am a person of worth, at least on an equal basis with others” and “I feel I am able to do things as well as most other people.” Each of the 10 items is scored on a 4-point scale, with a higher score indicating a higher level of self-esteem. Reliability analyses indicated that Cronbach’s alpha was .82 at Time 1 and .90 at Time 2.

The Rathus Assertiveness Scale (Rathus, 1973) was used to assess participants’ self-perceived levels of assertiveness. Sample items include the following: “I am open and frank about my feelings” and “I complain about poor service in a restaurant and elsewhere.” This 30-item measure features a possible range of 180 points (–90 to +90). A total score was used for analysis. Reliability analysis indicated that Cronbach’s alpha was .88 at Time 1 and .90 at Time 2.

The Aggression Questionnaire (TAQ) (Buss & Perry, 1992) was used to assess levels of hostility and aggression.

The TAQ contains four subscales: Anger (“I have trouble controlling my temper”), Hostility (“I wonder sometimes why I feel so bitter about things”), Physical Aggression (“Given enough provocation, I may hit another person”), and Verbal Aggression (“When people annoy me, I may tell them what I think of them”). Reliability analyses indicated that this measure had an acceptable level of internal consistency. For the total score (General Aggression), Cronbach’s alpha was .85 at Time 1 and .85 at Time 2. Subscale reliabilities also were acceptable, with Cronbach’s alpha ranging from .62 to .74 at Time 1 and .75 to .78 at Time 2.

Training Goals and Expectancies

At the outset of the course, participants were asked to describe their reasons for taking the course as well as their specific behavioral goals. They also were asked to note if they expected any personal skills or capabilities to change as a result of their participation. Participants then were asked to circle their most important goal and to describe specific situations in which they felt that change in this particular characteristic or skill is highly relevant.

Responses were transcribed verbatim. The first author and a research assistant developed a coding scheme allowing for independent themes to emerge. The coding scheme included emergent themes of coping skills (i.e., improving communication with a spouse), health-related activities (i.e., improving fitness), and achievement (i.e., improving grades). Two independent raters used this scheme to code all responses. Kappa was at or greater than .87 for all responses.

RESULTS

Baseline Analyses

Means and standard deviations for all variables in all conditions are provided in Tables 1 and 2. To assess potential baseline differences among conditions, multivariate analyses of variance (MANOVAs) were conducted for each conceptually related set of measures. Results of MANOVAs on the self-efficacy variables and on the personality variables indicated no differences between participants in the three conditions at baseline, Wilks’s lambda = .90, $F(8, 180) = 1.17$, and Wilks’s lambda = .90, $F(6, 182) = 1.63$, respectively; p s > .05. Follow-up one-way analyses of variance (ANOVAs) on individual variables indicated no significant baseline differences between conditions on self-defense self-efficacy, $F(2, 93) = .17$, Ryckman’s physical self-efficacy, $F(2, 93) = .69$, Sherer et al.’s measure of general self-efficacy, $F(2, 93) = 2.31$, or the composite score on the multidomain self-efficacy questionnaire, $F(2, 93) = 1.53$, all p s > .05. Likewise, no baseline differences were found

TABLE 1: Means and Standard Deviations for the Self-Efficacy Variables in All Conditions

Variable	Condition								
	Self-Defense			Enhanced Self-Defense			Control		
	M	SD	N	M	SD	N	M	SD	N
Self-defense self-efficacy									
Pre	67.00	18.87	26	71.10	30.75	29	69.00	25.72	41
Post	124.35	25.08	26	133.90	16.68	29	73.66	28.51	41
Physical self-efficacy									
Pre	84.15	14.21	26	85.90	16.93	29	88.27	12.13	41
Post	89.62	14.16	26	94.14	14.78	29	90.59	11.95	41
Global self-efficacy									
Pre	160.89	28.75	26	175.17	31.62	29	173.85	23.20	41
Post	163.54	29.90	26	178.31	28.06	29	171.29	23.41	41
Multidomain general self-efficacy									
Pre	1906.50	354.72	26	2103.97	400.21	29	1968.93	497.11	41
Post	2175.96	422.71	26	2260.86	393.15	29	1974.51	473.95	41
Self-esteem									
Pre	21.62	4.07	26	24.31	4.44	29	23.17	4.20	41
Post	22.58	5.47	26	26.17	4.07	29	24.78	4.73	41
Assertiveness									
Pre	.50	19.59	26	17.86	31.25	29	16.88	27.73	41
Post	4.15	22.49	26	19.21	31.91	29	15.44	29.36	41
Aggression									
Pre	44.73	12.78	26	46.07	13.64	29	44.20	14.98	41
Post	43.35	14.03	26	47.93	15.22	29	40.46	14.30	41

TABLE 2: Means and Standard Deviations of Average Item Responses for Self-Efficacy Domains

Variable	Condition								
	Self-Defense			Enhanced Self-Defense			Control		
	M	SD	N	M	SD	N	M	SD	N
Sport self-efficacy									
Pre	39.42	15.13	26	41.56	19.04	29	38.93	16.97	41
Post	54.67	20.23	26	56.88	18.09	29	41.79	19.01	41
Assertiveness self-efficacy									
Pre	63.64	12.08	26	75.00	15.41	29	69.70	19.21	41
Post	71.32	13.76	26	78.52	15.34	29	69.31	16.89	41
Coping self-efficacy									
Pre	62.44	11.19	26	66.56	14.43	29	64.05	16.27	41
Post	71.12	12.62	26	72.16	13.52	29	63.03	15.87	41
Conscientiousness									
Pre	54.63	19.68	26	63.58	17.00	29	57.27	20.29	41
Post	61.06	19.17	26	63.21	18.47	29	58.28	18.94	41
Academic achievement									
Pre	70.21	15.62	26	74.80	15.70	29	67.59	17.76	41
Post	76.73	15.63	26	78.16	16.29	29	67.98	19.31	41

on Rosenberg's Self-Esteem Scale, $F(2, 93) = 2.78$, or for the total score on the Buss-Perry Aggression Scale, $F(2, 93) = .16$; $ps > .05$. However, a significant difference between groups was found on Rathus's trait measure of assertiveness, $F(2, 93) = 3.68$, $p < .05$, $\eta^2 = .07$. Follow-up contrasts indicated that the basic self-defense cohort was, at baseline, significantly lower on assertiveness than both the enhanced self-defense condition, $F(1, 53) = 5.93$, $p < .05$, and the control condition, $F(1, 65) = 6.87$,

$p < .015$. The groups also differed at baseline on the assertiveness subscale of the multidomain self-efficacy questionnaire, $F(2, 93) = 3.28$, $p < .05$, with participants in the enhanced self-defense condition scoring higher on this subscale than their counterparts in the basic self-defense condition, $F(1, 53) = 9.10$, $p < .005$. There were no baseline differences between the basic self-defense and control conditions or between the control and enhanced training conditions, $ps > .05$, *ns*. There were

no additional baseline differences on any of the remaining self-efficacy subscales.

Analyses of participants' training goals and expectancies revealed highly consistent expectations for training. Nearly 100% of the women stated that they were taking self-defense training to learn how to better defend themselves should they be attacked. Some participants indicated a desire to become more physically fit through training, whereas others hoped for an improvement in self-esteem, confidence, and assertiveness.

Treatment Effects

Self-efficacy measures. A repeated-measures MANOVA including all self-efficacy variables revealed a significant Groups \times Time interaction on this variable cluster; Wilks's lambda = .36, $F(8, 180) = 15.24$, $p < .001$, $\eta^2 = .40$. To assess the self-efficacy variables responsible for the multivariate effect, we conducted a series of two-way repeated-measures ANOVAs. These analyses revealed significant Groups \times Time interactions on all indices of self-efficacy except for Sherer et al.'s General Self-Efficacy Scale. Thus, we found a significant Groups \times Time interaction on the Self-Defense Self-Efficacy Scale, $F(2, 93) = 75.31$, $p < .001$, $\eta^2 = .62$. Planned comparisons revealed that both the basic and enhanced training groups exhibited significant increases in self-defense self-efficacy relative to the waiting list control condition, $F(1, 65) = 101.40$, $p < .001$, $\eta^2 = .61$, and $F(1, 68) = 155.47$, $p < .001$, $\eta^2 = .70$, respectively.

A significant Groups \times Time interaction also occurred on Ryckman's measure of physical self-efficacy, $F(2, 93) = 3.72$, $p < .05$, $\eta^2 = .07$. Planned comparisons revealed that participants in the enhanced condition showed a significant increase on this measure relative to untrained controls, $F(1, 68) = 6.56$, $p < .015$, $\eta^2 = .09$. Although participants in the basic self-defense condition also showed significant increases in physical self-efficacy after training, $t(25) = 3.75$, $p < .001$, their improvement was not significantly different from that of the control participants, $F(1, 65) = 3.53$, $p = .07$.

Turning to global self-efficacy, a two-way, between-group, repeated-measures ANOVA revealed no main or interaction effects on the Sherer et al. (1982) measure of global self-efficacy, $F(1, 93) = .30$, $p > .05$, and $F(2, 93) = .98$, $p > .05$, respectively. In contrast, a significant Groups \times Time interaction was found on the composite score of the multidimensional measure of general self-efficacy, $F(2, 93) = 6.39$, $p < .005$, $\eta^2 = .12$. Planned comparisons revealed that relative to their control condition counterparts, participants in the basic self-defense condition showed significant increases on the composite score of the Multidomain Self-Efficacy Scale, $F(1, 65) = 15.72$, $p < .001$, $\eta^2 = .20$. Participants in the enhanced training con-

dition demonstrated a similar relative increase, $F(1, 68) = 3.83$, $p = .05$.

Given the evidence of some degree of change on the composite score, we then conducted a repeated-measures MANOVA involving the individual scales of the Multidomain Self-Efficacy Scale, which revealed a significant Groups \times Time interaction on this variable cluster, Wilks's lambda = .70, $F(10, 178) = 3.51$, $p < .001$, $\eta^2 = .17$. Two-way, between-groups, repeated-measures ANOVAs on the separate domain subscales of the Multidomain Self-Efficacy Scale revealed significant Groups \times Time interactions on domain measures of sport-specific self-efficacy, $F(2, 93) = 11.99$, $p < .001$, $\eta^2 = .21$, coping, $F(2, 93) = 7.00$, $p < .005$, $\eta^2 = .13$, and assertiveness self-efficacy, $F(2, 93) = 3.21$, $p < .05$, $\eta^2 = .06$. For sport self-efficacy, planned comparisons revealed that participants in the basic self-defense condition showed significantly greater gains in sport self-efficacy than did untrained controls, $F(1, 65) = 21.00$, $p < .001$, $\eta^2 = .24$, as did the enhanced condition, $F(1, 68) = 16.90$, $p < .001$, $\eta^2 = .20$. Similarly, participants in the basic and enhanced training conditions exhibited larger increases in coping self-efficacy than did their waiting-list control peers, $F(1, 65) = 12.97$, $p < .001$, $\eta^2 = .17$, and $F(1, 68) = 5.11$, $p < .05$, $\eta^2 = .07$, respectively. Finally, participants in the basic training condition showed significant increases in assertiveness self-efficacy when compared with controls, $F(1, 65) = 8.28$, $p < .015$, $\eta^2 = .11$, whereas participants in the enhanced condition did not. No treatment effects were found on the conscientiousness or academic achievement subscales.

Self-Esteem, Assertiveness, and Aggression

Repeated-measures ANOVAs revealed significant main effects for Time on the Rosenberg Self-Esteem Scale, $F(2, 93) = 24.41$, $p < .001$, $\eta^2 = .21$, but no treatment-related interaction effect. No significant main or interaction effects were found on the Rathus Assertiveness Scale. On the TAQ, results of a two-way, repeated-measures ANOVA on the total score revealed no significant main or treatment-related interaction effects. Results of a two-way MANOVA on the four subscales of this measure (Anger, Hostility, Physical Aggression, Verbal Aggression) revealed a significant main effect of Time, Wilks's lambda = .79, $F(4, 90) = 5.83$, $p < .001$, $\eta^2 = .21$, but a nonsignificant Group \times Time interaction. Follow-up ANOVAs indicated a significant main effect for Time on the Anger subscale. Across conditions, participants reported a significant decrease in anger, $F(1, 93) = 16.29$, $p < .001$, $\eta^2 = .15$.

DISCUSSION

The results of this study indicate that coping skills training can have a generalized impact on people's per-

ceptions of self-efficacy. Analyses of responses to a contextualized, multidomain, self-efficacy questionnaire revealed that self-defense training boosted women's efficacy beliefs in domains beyond those involving physical self-defense and other related skills for coping with physical threats. Compared with women in a waiting-list control condition, those trained in physical self-defense displayed higher levels of self-efficacy across a variety of skill domains, including general coping and self-regulatory skills, sport-specific physical competencies, and interpersonal assertiveness. Likewise, participants in both self-defense conditions exhibited significant increases in physical self-efficacy. Thus, acquisition of self-defense skills affected more general attitudes about their bodies and physical capabilities.

It is noteworthy that the generalization effects detected by our multidomain, contextualized, self-efficacy measure on three of the five subscales and on the composite score were not detected by the global, decontextualized measure of self-efficacy beliefs developed by Sherer et al. (1982) or by Rathus's trait measure of assertiveness. This finding has both theoretical and applied implications. It suggests that measures of global beliefs or personality traits may underestimate the degree to which psychosocial interventions produce generalized self-efficacy effects in life domains beyond those specifically targeted by the intervention. One possible reason is that self-knowledge and global self-referent beliefs are stored separately in memory (Klein & Loftus, 1993). Experiences thus may alter autobiographical knowledge and domain-specific self-appraisals without altering global conceptions of one's dispositional tendencies. The question of whether interventions alter global self-referent beliefs is important and deserving of continued attention (Smith, 1989, 1999; Weitlauf et al., 2000); however, it cannot be equated with the issue of whether interventions generalize at a behavioral level.

We had anticipated that the particular pattern of change in self-efficacy beliefs that women experienced would be predictable from individual differences in the expectancies and personal goals that women brought to self-defense training. Unfortunately, this hypothesis could not be evaluated because there was so little variation in women's pretraining expectations and goals. Virtually all of the women expressed the goal of improving skills needed to cope with a physical assault and—at least initially—saw the intervention as primarily relevant to this domain.

Finally, self-efficacy generalization was not increased by our enhanced training protocol. In part, this may be attributed to the lack of variation in participant expectations for training.

Limitations and Future Directions

Some methodological aspects of the study, most notably the self-selection of participants into training, limit the generalizability of our findings. Practical constraints prevented us from including a follow-up (i.e., Time 3) condition in the present study to assess the durability of generalization effects and "sleeper" effects that might have been detected at follow-up. Fortunately, prior studies have already demonstrated that self-defense training for women has lasting impact (Ozer & Bandura, 1990; Weitlauf et al., 2000).

Our intervention included both physical self-defense skills and cognitive-affective coping skills; thus, we cannot be certain which aspects of training exerted positive effects on specific outcome variables. This issue could be addressed in future dismantling studies. Knowledge of links between specific training components and generalization effects could provide useful information for designing optimally effective interventions.

In conclusion, results of the present study extend previous research findings about the empowering and generalizing benefits of coping skills training. In this case, self-defense training for women not only enhanced women's beliefs about their ability to protect themselves from physical assault but also influenced other domains of self-efficacy unrelated to training as well as global self-referent beliefs about physical efficacy and assertiveness. Most important, our results illustrate the superior sensitivity of a multidomain self-efficacy questionnaire to global trait measures in the detection of changes in beliefs about the self that result from novel social experiences.

NOTES

1. Participant demographics reflect the diversity present at the University of Illinois at Chicago.

2. The authors acknowledge that skills assessment is a critical element of an applied study. However, the primary focus of the present study is that of changes in self-perception and other personality characteristics. Thus, precise measurement of self-defense skills is not central to our findings. Participants engage in 2,000 repetitions of every striking technique, receive constant corrective feedback, and do not progress to more difficult techniques until they are judged by qualified instructors to have mastered the basic ones. Thus, there is relatively little variability in the acquisition of skills in this program.

3. Ryckman, Robbins, Thornton, and Cantrell's (1982) and Sherer et al.'s (1982) scales are decontextualized and contain items that do not explicitly reference personal capabilities but more general perceptions of competence across activity domains. Thus, they do not fully accord with the social cognitive measurement rationale of self-efficacy theory emphasized by Bandura (1997). However, we included these scales to compare their efficacy against a more sensitive and contextualized index of self-efficacy perceptions across activity domains (our multidomain scale) to compare their levels of sensitivity to generalization measurement.

4. It is important to distinguish our conceptualization of coping self-efficacy from other constructs of coping, such as the use of specific

coping strategies (i.e., ways of coping). We are seeking to measure beliefs in one's ability to cope in a particular manner (i.e., controlling anger) in a particular setting (i.e., interpersonal). By assessing individual participants' assessments of their perceived ability to cope in very specific ways with very specific threats, we were able to use this subscale as an index of self-efficacy generalization following treatment in a manner that conforms to the tenets of self-efficacy theory (Bandura, 1997).

REFERENCES

- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84, 191-215.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.
- Bandura, A. (1999). Social cognitive theory of personality. In D. Cervone & Y. Shoda (Eds.), *The coherence of personality: Social-cognitive bases of consistency, variability and organization* (pp. 185-241). New York: Guilford.
- Bandura, A., Taylor, C. B., Williams, S. L., Mefford, I. N., & Barchas, J. (1985). Catecholamine secretion as a function of perceived coping self-efficacy. *Journal of Consulting and Clinical Psychology*, 53, 406-414.
- Brody, E. B., Hatfield, B. D., & Spalding, T. W. (1988). Generalization of self-efficacy to a continuum of stressors upon mastery of a high-risk sport skill. *Journal of Sport and Exercise Psychology*, 10, 32-44.
- Bugenthal, D. B. (2000). Acquisition of the algorithms of social life: A domain-based approach. *Psychological Bulletin*, 126, 187-219.
- Buss, A. H., & Perry, M. (1992). The Aggression Questionnaire. *Journal of Personality and Social Psychology*, 63, 452-459.
- Caprara, G. V., & Cervone, D. (2000). *Personality: Determinants, dynamics, and potentials*. New York: Cambridge University Press.
- Cervone, D. (1985). Randomization tests to determine significance levels for microanalytic congruences between self-efficacy and behavior. *Cognitive Therapy and Research*, 9, 357-365.
- Cervone, D., & 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.
- Cervone, D., & Scott, W. D. (1995). Self-efficacy theory of behavioral change: Foundations, conceptual issues, and therapeutic implications. In W. O'Donohue & L. Krasner (Eds.), *Theories in behavior therapy* (pp. 349-383). Washington, DC: American Psychological Association.
- Cervone, D., Shadel, W. G., & Jencius, S. (2001). Social-cognitive theory of personality assessment. *Personality and Social Psychology Review*, 5, 33-51.
- Cervone, D., & Shoda, Y. (Eds.). (1999). *The coherence of personality: Social cognitive bases of consistency, variability, and organization*. New York: Guilford.
- Coppel, D. B. (1980). *The relationship of perceived social support and self-efficacy to major and minor stressors*. Unpublished doctoral dissertation, University of Washington, Seattle.
- Escarti, A., & Guzman, J. F. (1999). Effects of feedback on self-efficacy, performance and choice in an athletic task. *Journal of Applied Sport Psychology*, 11, 83-96.
- Ewart, C. K. (1992). The role of physical self-efficacy in recovery from heart attack. In R. Schwarzer (Ed.), *Self-efficacy: Thought control of action* (pp. 287-304). Washington, DC: Hemisphere.
- Glynn, S. M., & Ruderman, A. J. (1986). The development and validation of an eating self-efficacy scale. *Cognitive Therapy and Research*, 10, 403-420.
- Goldfried, M. R., & Robbins, C. (1982). On the facilitation of self-efficacy. *Cognitive Therapy and Research*, 6, 361-379.
- Klein, S. B., & Loftus, J. (1993). The mental representation of trait and autobiographical knowledge about the self. In T. K. Srull & R. S. Wyer, Jr. (Eds.), *Advances in social cognition* (Vol. 5, pp. 1-49). Hillsdale, NJ: Lawrence Erlbaum.
- Mischel, W. (1968). *Personality and assessment*. New York: John Wiley.
- Mischel, W. (1999). Personality coherence and dispositions in a cognitive-affective processing system (CAPS) approach. In D. Cervone & Y. Shoda (Eds.), *The coherence of personality: Social-cognitive bases of consistency, variability, and organization*. (pp. 37-60). New York: Guilford.
- Mischel, W., & Shoda, Y. (1995). A cognitive-affective system theory of personality: Reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review*, 102, 246-268.
- Montoya, I. D. (1998). Social network ties, self-efficacy and condom use among women who use crack cocaine: A pilot study. *Substance Use and Misuse*, 33, 2049-2073.
- Ozer, E. M., & Bandura, A. (1990). Mechanisms governing empowerment effects: A self-efficacy analysis. *Journal of Personality and Social Psychology*, 58, 472-486.
- Rathus, S. A. (1973). A 30-item schedule for assessing assertive behavior. *Behavior Therapy*, 4, 398-406.
- Rosenberg, M. (1965). *Society and the adolescent self-concept*. Princeton, NJ: Princeton University Press.
- Ryckman, R. M., Robbins, M. A., Thornton, B., & Cantrell, P. (1982). Development and validation of a physical self-efficacy scale. *Journal of Personality and Social Psychology*, 42, 891-900.
- Sarason, I. G., Pierce, G. P., & Sarason, B. R. (1996). *Cognitive interference: Theories, methods and findings*. Mahwah, NJ: Lawrence Erlbaum.
- Schwarzer, R. (1995). Changing risk behaviors and adopting health behaviors: The role of self-efficacy beliefs. In A. Bandura (Ed.), *Self-efficacy in changing societies* (pp. 259-288). New York: Cambridge University Press.
- Schwarzer, R., Babler, J., Kwiatek, P., & Shroder, K. (1997). The assessment of optimistic self-beliefs: Comparison of the German, Spanish, and Chinese versions of the General Self-Efficacy Scale. *Applied Psychology: An International Review*, 46, 69-88.
- Sherer, M., Maddux, J. E., Mercandante, B., Prentice-Dunn, S., Jacobs, B., & Rogers, R. W. (1982). The self-efficacy scale: Construction and validation. *Psychological Reports*, 51, 663-671.
- Smith, R. E. (1989). Effects of coping skills training on generalized self-efficacy and locus of control. *Journal of Personality and Social Psychology*, 56, 228-233.
- Smith, R. E. (1999). Generalization effects in coping skills training. *Journal of Sport and Exercise Psychology*, 21, 189-204.
- Stajkovic, A. D., & Luthans, F. (1998). Self-efficacy and work-related performance: A meta-analysis. *Psychological Bulletin*, 124, 240-261.
- Weitlauf, J. C., Smith, R. E., & Cervone, D. (2000). Generalization of coping skills training: Influence of self-defense training on women's efficacy beliefs, assertiveness, and aggression. *Journal of Applied Psychology*, 85, 625-633.
- Williams, S. L., Kinney, P. J., & Falbo, J. (1989). Generalization of therapeutic changes in agoraphobia: The role of perceived self-efficacy. *Journal of Consulting and Clinical Psychology*, 57, 436-442.

Received July 19, 2000

Revision accepted January 3, 2001

Depression as a Moderator of Relationships Between Positive Daily Events and Day-to-Day Psychological Adjustment

John B. Nezlek

College of William & Mary

Shelly L. Gable

University of California, Los Angeles

For 21 days, 123 participants provided measures of their daily depressogenic adjustment, including Beck's cognitive triad, causal uncertainty, control over the environment, self-esteem, and anxiety, and they described the positive and negative events that occurred. Daily adjustment negatively covaried with the number of negative events occurring each day and, except as measured by anxiety, positively covaried with positive events. The covariance between negative events and adjustment was stronger than the covariance between positive events and adjustment. Participants also provided measures of depressive symptoms. For the self-esteem and cognitive triad measures, adjustment covaried more strongly with negative and positive events for the depressed than they did for the nondepressed.

For more than two decades, psychologists have studied day-to-day variability in psychological states to further the understanding of individual differences in psychological well-being and adjustment. This research suggests that daily psychological adjustment covaries with daily events and that trait levels of adjustment moderate this covariation. Although informative, this research is limited in important ways. First, research on daily events and adjustment has operationalized daily adjustment primarily in terms of mood. Second, studies examining how day-level relationships are moderated by trait-level measures have focused on the moderating role of neuroticism and related constructs; few studies have focused on depression.

The focus of existing research may limit its utility for understanding more specific phenomena such as self-esteem and depression. Although understanding disturbances in mood is clearly important to understanding both self-esteem and depression, considerable research

suggests that state-level constructs other than mood are worth investigating. For example, research indicates that depression is associated with greater lability in self-esteem and that depression may moderate day-level relationships between events and self-esteem (Butler, Hokanson, & Flynn, 1994).

Accordingly, the present study examined day-level relationships between events and state measures of depressogenic adjustment other than mood and how such relationships varied as a function of trait depressogenic adjustment. Each day for 3 weeks, participants described the positive and negative events that occurred and provided measures of state adjustment. Over a 4 1/2-month period, participants also provided four reports of their depressive symptoms, which collectively were used to measure depression.

Three hypotheses guided the study: (a) daily adjustment would covary negatively with daily negative events and positively with positive events, (b) adjustment would covary more strongly with negative events than with positive events, and (c) the covariation between daily events and adjustment would be stronger for people who were

Authors' Note: Different analyses of some of the data presented in this article appear in Gable and Nezlek (1998). We are grateful to Edward Deci, Michael Kernis, and Gifford Weary for their help in developing the measures used in this study and to Harry Reis and David Almeida for their comments on drafts of this article. Correspondence concerning this article should be addressed to either John B. Nezlek, College of William & Mary, Department of Psychology, P.O. Box 8795, Williamsburg, VA 23187-8795, or to Shelly L. Gable, University of California, Los Angeles, Psychology Department, 4560 Franz Hall, Box 951563, Los Angeles, CA 90095; e-mail: john.nezlek@wm.edu or gable@psych.ucla.edu.

PSPB, Vol. 27 No. 12, December 2001 1692-1704

© 2001 by the Society for Personality and Social Psychology, Inc.

less well adjusted at the trait level than for those who were better adjusted. These hypotheses were tested by a series of multilevel random coefficient modeling analyses.

Studies of day-level relationships between events and psychological adjustment have focused on daily variations in mood, with mood usually operationalized in terms of positive and negative affect (PA and NA), dimensions suggested by Watson and Tellegen (1985). Although studies of PA and NA have been informative, understanding specific types of distress such as depression may require the use of other approaches. When introducing the PANAS, Watson, Clark, and Tellegen (1988) noted, "Consistent with previous findings that depressive symptomatology is affectively complex . . . researchers interested in studying depressed affect might therefore want to use the PANAS scales as a complement to more traditional depression measures" (p. 1068). Moreover, considerable research and theory suggest that depression also involves more cognitively focused components such as optimism about the future (e.g., Beck, 1972), perceptions of control over outcomes (e.g., Alloy, Kelly, Mineka, & Clements, 1990), and the ability to detect cause and effect in one's social world (e.g., Weary & Edwards, 1994).

To measure the breadth of constructs that research has suggested reflect depressogenic adjustment, five constructs that have been found to covary with depression at the trait level served as the basis for the daily measures of depressogenic adjustment in the present study. These were as follows: Beck's Cognitive Triad (Beck, 1972), control over the outcomes of one's behavior (Deci & Ryan, 1985), the ability to detect cause and effect in one's world (Weary, Jordan, & Hill, 1985), self-esteem (Rosenberg, 1965), and anxiety.

Research on daily events has examined day-level relationships between mood and both positive and negative events, although there has been a somewhat greater interest in negative events, as suggested by the number of studies examining only negative events (e.g., Affleck, Tennen, Urrows, & Higgins, 1994; Bolger, DeLongis, Kessler, & Schilling, 1989; Bolger & Schilling, 1991; Larsen & Kasimatis, 1991; Marco & Suls, 1993; Suls, Martin, & David, 1998). In general, these studies have found that people experience greater NA on days when more negative events occur than on days when fewer negative events occur. Positive events have been studied in conjunction with negative events (Clark & Watson, 1988; David, Green, Martin, & Suls, 1997; Gable, Reis, & Elliot, 2000; Stone, 1981, 1987; Watson, 1988) and alone (Lewinsohn & Graf, 1973; Lewinsohn & Libet, 1972). In general, these studies find that people experience greater PA on days when more positive events occur than on days when fewer positive events occur, and those that

also study negative events find that NA covaries with negative events.

In contrast to the specific relationships found in mood event studies (PA and positive events covary and NA and negative events covary), research on daily variations in self-esteem has found that self-esteem covaries with both positive and negative events (Butler et al., 1994). This research suggested the hypothesis that daily adjustment would covary positively with the positive events that occurred during a day and would covary negatively with the negative events that occurred. This hypothesis is also consistent with the assumptions that poorer adjustment may predispose people to experience more negative events and that better adjustment may predispose people to experience more positive events (Bolger & Schilling, 1991; Smith & Rhodewalt, 1986).

We also expected, similar to the results of Butler et al. (1994), that the day-level covariation between adjustment and negative events would be stronger than the covariation between adjustment and positive events. This more specific prediction also was based on research demonstrating that negative events have a greater psychological impact than positive events. In an extensive review, Taylor (1991) concluded that "diverse literatures in psychology provide evidence that, other things being equal, negative events appear to elicit more physiological, affective, cognitive, and behavioral activity, and prompt more cognitive analysis than neutral or positive events" (p. 67). In a discussion of attitude evaluation, Cacioppo, Gardner, and Berntson (1997) reached a similar conclusion, labeling the tendency for negative information to be more salient for a negativity bias.

The second major focus of research on daily events and daily psychological states has been how event-state relationships vary as a function of psychological traits. This research has concerned the moderating effects of neuroticism and other trait-level constructs such as Type-A behavior (Larsen & Kasimatis, 1991), social support (Affleck et al., 1994), and extraversion (David et al., 1997). Results of this research are somewhat inconsistent; for example, some studies of the moderating role of trait neuroticism have found that greater neuroticism is associated with stronger covariation between negative daily events and daily mood (e.g., Bolger & Schilling, 1991; Marco & Suls, 1993), whereas other studies have not (e.g., Affleck et al., 1994, David et al., 1997). Interestingly, none of these studies found that relationships between positive events and mood were moderated by a trait, in part because studies of trait moderators have tended to focus on negative events.

Research examining the moderating effects of depression on daily event-adjustment relationships suggests that depression may moderate event-adjustment

relationships. Butler et al. (1994) found that daily self-esteem covaried more strongly with positive and negative events for remitted depressives than for those who had never been depressed. The currently depressed were between the two.

We expected that day-level relationships between events and adjustment would be stronger for people who reported higher levels of depressive symptoms than for those who reported lower levels. This hypothesized moderating effect was suggested in part by Rogers's theory of the self, particularly his beliefs about conditions of self-worth (Rogers, 1961). Within a Rogerian framework, a person's sense of self is defined partially as a function of how contingent self-worth is on environmental events or conditions. Rogers believed that the self-worth of more poorly adjusted people is more contingent (less unconditional in Rogers's terminology) on environmental feedback such as daily events than the self-worth of better adjusted people. The day-to-day psychological states (measures of the self) of those who are more poorly adjusted at the trait level should vary more as a function of daily events than the psychological states of those who are better adjusted.

Such differential sensitivity is also consistent with research suggesting that trait self-esteem is negatively related to people's reactivity or sensitivity to events. Brockner (1984) suggested that trait self-esteem is negatively related to plasticity, the susceptibility to the effects of self-relevant social cues. In a series of studies, Kernis and colleagues have found negative relationships between trait self-esteem and the stability of state self-esteem (Kernis, 1993). Extending this, Gable and Nezlek (1998) found similar relationships between depression and a general factor consisting of the instability of a variety of measures of state adjustment.

Finally, the expectation that more depressed people will be more reactive to negative events is consistent with various theoretical accounts of neuroticism (e.g., Eysenck, 1967). It is noteworthy that research and theory on neuroticism has focused on reactivity to negative events such as stressors rather than on reactivity to positive events. In contrast, Rogers (1961) and other self-focused theorists tend to be concerned with a more general construct, lability, which includes reactivity to both negative and positive events.

The present study concerned the covariation between daily events and adjustment and was not intended to test hypotheses about causal relationships between events and adjustment. Nevertheless, because covariation has traditionally been considered to be a necessary (although not sufficient) condition to establish causality, the implications of the present results for understanding causal relationships between daily events and adjustment are discussed later.

METHOD

Participants

Participants were 128 introductory psychology students, 85 women and 43 men, attending the College of William & Mary who participated in partial fulfillment of class requirements. To ensure that the sample contained a sufficient number of participants with high levels of depressive symptoms, people were invited to participate on the basis of two measures of depressive symptoms. Seven weeks before the study began, participants completed the Beck Depression Inventory (BDI) (Beck, 1967) and the Center for Epidemiological Studies Depression Scale (CES-D) (Radloff, 1977). Approximately 25% of those invited to participate scored greater than 10 on the BDI and greater than 17 on the CES-D, whereas the remaining 75% scored significantly less than both cutpoints on these measures; that is, people with greater than normal levels of depressive symptoms were oversampled.

Measures

For each well-being construct (except anxiety), trait and state measures were collected. Trait measures were collected at the beginning and end of the study and state measures were collected each day, as were descriptions of daily events.

Depressive symptoms were measured using the BDI and CES-D. Daily depression was measured by three items representing the elements of Beck's cognitive triad (Beck, 1972): (a) negative view of self, "Overall, how positively did you feel about yourself today"; (b) negative view of life in general, "Thinking of your life in general, how well did things go today"; and (c) negative view of the future, "How optimistic are you about how your life (in general) will be tomorrow?" Participants answered these questions using 7-point scales (with higher numbers indicating a more positive outlook).

Trait causal uncertainty was measured using the Causal Uncertainty Scale (CUS) (Weary & Edwards, 1994). Daily causal uncertainty was measured using four questions based on items from the CUS that were chosen on the basis of factor loadings from previous studies (Weary & Edwards, 1994) and appropriateness for daily assessment. Using 6-point scales ranging from *strongly disagree* (1) to *strongly agree* (6), participants indicated their agreement with the item, "I did not understand why things happened the way they did" in reference to four topics: thinking back on my day today in terms of the positive interactions I had with others, I did not understand why things happened the way they did; thinking back on my day today in terms of the positive nonsocial events (e.g., schoolwork, sports, etc.) that occurred, I did not understand why things happened the

way they did; thinking back on my day today in terms of the negative interactions I had with others, I did not understand why things happened the way they did; and thinking back on my day today in terms of the negative nonsocial events (e.g., schoolwork, sports, etc.) that occurred, I did not understand why things happened the way they did.

Trait causality orientation was measured using the General Causality Orientation Scale (GCOS) (Deci & Ryan, 1985). The GCOS measures three orientations, autonomy, impersonal, and control, and daily causality orientation was measured using six items, two for each orientation. Using 7-point scales ranging from *not at all* (1) to *very much so* (7) (with higher scores representing greater perceived control), participants answered the following questions. One question concerned social activities, "Thinking back on your day today in terms of your relationships with others and the social events that occurred . . . ?" and the other concerned achievement, "Thinking back on your day today in terms of nonsocial areas of performance (e.g., schoolwork, sports, fitness, etc.) . . . ?" The two autonomy orientation questions concluded with the following, "To what extent did you feel that you had a choice about what you did and to what extent did things happen the way you wanted them to happen?" The two control orientation questions concluded with "To what extent did you do things because either you felt you should do them or because other people felt you should do them?" The two impersonal orientation questions concluded with "To what extent were you able to control the outcomes of these events?"

Self-esteem was measured using Rosenberg's (1965) Self-Esteem Scale (RSE). Daily self-esteem was measured using the 10 items on the trait scale reworded to refer to how participants felt about themselves that day. Daily anxiety was assessed using three items from the Profile of Mood States (Lorr & McNair, 1971) that were used by Bolger (1990) to assess daily anxiety. Participants used 9-point scales ranging from *strongly disagree* (1) to *strongly agree* (9) to respond to these three statements: I felt on edge today; I felt uneasy today; and I felt nervous today. No trait measure of anxiety was collected.

Daily events were measured using items from the Daily Events Survey (Butler et al., 1994), a 40-item measure of events appropriate for college students. In the present study, 22 of these 40 events were measured, 12 positive and 10 negative, with social and achievement domains equally represented. Events included, "went out to eat with a friend/date," "tried to do homework and couldn't understand it," "did well on a school or work task (e.g., test, assignment, job duty)," and "had plans fall through to spend time with someone special." In addition, four items (combinations of positive-negative and social-achievement) were created to measure other

events that may have occurred. For example, other positive social events were measured using the item, "had other type of pleasant event (not listed above) with friends, family, or date."

Each day, participants rated each event using the following scale: 0 = *did not occur*, 1 = *occurred and not important*, 2 = *occurred and somewhat important*, 3 = *occurred and pretty important*, 4 = *occurred and extremely important*. The number of positive events that occurred each day and the number of negative events that occurred were calculated.¹

Procedure

At the beginning of the study, participants came to a laboratory and received instructions and a computer disk containing the data collection programs. They were told they would be using a computer to answer a series of questions every day for 3 weeks and questionnaires on the first and last days of the study. Data collection programs were written using the Micro-Analytic Experimental Laboratory software package (MEL) (Schneider, 1988), and participants were able to run these programs on any IBM-compatible personal computer.

Standard instructions for the measures (with modifications for those with a daily frame of reference) were included in the programs. Data were collected using three different programs, and participants were given a list of which programs to run each day. The first program was run on the first day of the study and collected responses to the RSE, the BDI, and the GCOS. The second program was run every day of the study and collected the five daily measures of adjustment and reports of daily events. The third program was run on the last day of the study and administered another RSE, the CES-D, and the CUS.

A member of the research team maintained regular contact with participants via phone and e-mail. They were told to contact the experimenters should any problems arise, such as disk failure, computer viruses, and so forth. Such problems were rare, and when they occurred, participants were given replacement disks within 48 hours and continued the study.

At the end of the study, participants answered questions about their participation. Participants did not think that participating in the study had changed their daily routine meaningfully. Half (53%) reported spending 5 minutes or less per day running the program, and 99% reported spending 10 minutes or less per day. Participants reported that it was relatively easy to run the program, a mean of 4.6 using a 1 to 5 scale where 1 = *very difficult* and 5 = *very easy*. Finally, using 1 to 5 scales where 1 = *not at all* and 5 = *very much*, participants reported that participating in the study did not make them feel or think differently about themselves (2.1), their relation-

ships with other people (1.8), or their schoolwork or other areas of performance (1.7).

Of the 128 participants who began the study; 5 had failed disks, lost their disks, or did not follow instructions. The 123 remaining participants completed the daily measures an average of 19.6 days; 48% of the participants provided daily measures for all 21 days, 24% provided data for 20 days, and 24% provided data for 16 to 19 days.²

RESULTS

The present data comprised what is referred to as a multilevel (or hierarchically nested) data structure in that observations at one level of analysis (days) were nested within another level of analysis (people). Accordingly, the data were analyzed with a series of multilevel random coefficient models (MRCM) using the program HLM (Bryk, Raudenbush, & Congdon, 1998; Version 4.04). MRCM was chosen over ordinary-least-squares (OLS) methods such as using within-person correlations to measure within-person relationships because MRCM provides better parameter estimates than OLS methods (Bryk & Raudenbush, 1992; Kenny, Kashy, & Bolger, 1998; Kreft & de Leeuw 1998). Descriptions of the advantages of MRCM over comparable OLS techniques and using MRCM to analyze daily diary data are presented in Nezlek (2001).

The superiority of MRCM over comparable OLS analyses is due to various factors. First and foremost, MRCM models within-person coefficients (such as those that were the subject of this study) as random, not fixed, effects. In the present study, the exact days over which data were collected were not critical. In essence, the days comprising the study were sampled from a population of days and were meant to represent participants' typical lives. Presumably, coefficients based on samples of other days would have been just as valid (although not exactly the same) as those based on the sample collected; therefore, within-person coefficients were random in that they were sampled from each participant's population of possible coefficients. This sampling of coefficients constitutes a *prima facie* case for treating (modeling) coefficients describing such within-person relationships as random, not fixed.

Procedures that do not model such coefficients as random, such as OLS analyses that treat days as a repeated-measures factor in an ANOVA or analyze within-subjects coefficients as dependent measures, may provide misleading parameter estimates because they do not account for this additional source of variance. Within a traditional OLS framework, errors at different levels of analysis are mathematically independent. For example, the reliability of within-person coefficients does not contribute to tests of individual differences in

these coefficients. One of the advantages of MRCM is its ability to model errors at all levels of analysis simultaneously. That is, the reliability of within-person coefficients does contribute to tests of individual differences in these coefficients. This simultaneity has implications for significance tests of fixed effects (Is an effect significantly different from 0?) and for estimates of the variance of effects.

Moreover, the advantages of HLM over comparable OLS techniques are more pronounced when the number of observations per unit of analysis (e.g., days provided by different people) are small or vary considerably across units (e.g., different people provide different numbers of days) and when covariances are being modeled instead of means. HLM uses a combination of precision weighting (units of analysis contribute to parameter estimates as a function of their reliability and the number of observations within the unit) and Bayesian modeling to estimate measures of central tendency and variances.

The analyses had three goals: (a) to determine the validity and reliability of the daily measures of psychological adjustment, (b) to examine relationships between these measures and daily events, and (c) to examine how day-level relationships between adjustment and events varied as a function of depression.

Validity and Reliability of Daily Measures of Adjustment

The validity and reliability of the measures of daily adjustment were examined using three-level models in which items were nested within days, which were nested within participants. By treating the items constituting a scale as a nested factor, HLM provided a latent variable analysis of scale scores (Bryk & Raudenbush, 1992, pp. 191-196). This procedure also provided estimates of the day-level reliability of the daily measures. The validity of the daily measure of a construct was operationalized as the strength of the relationship between the trait measure of a construct and the mean daily level of the same construct, expressed as shared variance. Reliability was estimated for day and person levels, and these estimates were provided directly by HLM. Reliability and validity analyses are described in the appendix, and the results are summarized in Table 1.

The analyses of the three items measuring Beck's cognitive triad indicated that this daily measure was reliable and valid. Similarly, the analyses of the four items measuring causal uncertainty and the 10 items measuring self-esteem indicated that these daily measures were reliable and valid. The analyses also indicated that three items measuring anxiety were reliable. Although no trait measure of anxiety was collected, previous research by

TABLE 1: Descriptive Statistics for Daily Measures

	Mean	Person-Level Variance	Day-Level Variance	Person-Level Reliability	Day-Level Reliability	Validity
Cognitive triad	5.14	.74	.89	.94	.81	.56
Self-esteem	7.22	1.31	.74	.97	.80	.86
Causal uncertainty	2.48	.64	.42	.97	.72	.50
Anxiety	3.73	2.43	2.67	.95	.86	NA
GCOS-control	4.76	.67	.72	.95	.66	.22
Positive events	5.25	4.67	4.46	.95	NA	NA
Negative events	2.13	2.00	2.60	.94	NA	NA

NOTE: All validity coefficients were significant at the .0001 level. GCOS = General Causality Orientation Scale.

Bolger (1990) suggested that these three items were valid.

The GCOS daily measures were intended to measure the autonomy, control, and impersonal orientations of the GCOS. Nevertheless, the analyses indicated that four of the six items measured the impersonal orientation, whereas two did not correspond to a trait-level GCOS construct. In light of this, the items designed to measure autonomy and impersonal orientations were recoded into one four-item daily measure of impersonal orientation, and the two items intended to measure control orientation were dropped from the analysis.³ This new measure is referred to as control, referring to control over outcomes, and it was found to be reliable and valid. On the basis of these analyses, for each measure of adjustment, daily scores were operationalized as the mean score for the items constituting that scale.

Daily Events and Day-to-Day Adjustment

Day-level relationships between daily events and adjustment were examined using a two-level MRCM. In essence, for each person, a regression equation was estimated describing the relationships between daily adjustment and daily events with adjustment as a dependent measure and positive and negative event scores as independent measures. The Level 1 model was as follows:

$$y_{ij} = \beta_{0j} + \beta_{1j}\text{PosEvent} + \beta_{2j}\text{NegEvent} + r_{ij},$$

in which y is an adjustment score for person j on day i , β_{0j} is a random coefficient representing the intercept for person j , β_{1j} is a random coefficient for positive events, β_{2j} is a random coefficient for negative events, and r_{ij} represents error.

Mean event scores varied considerably across persons and days, and the average positive event score was higher than the average negative event score (5.22 vs. 2.15, $p < .01$). To eliminate the influence of these differences on parameter estimates, event scores were group-mean centered, with group defined as the individual participant. Thus, coefficients for daily events described relation-

TABLE 2: Day-Level Relationships Between Adjustment and Events

Measure	Intercept	Positive Events	Negative Events	χ^2
Self-esteem	7.21	0.08*0.18*	35.6	
Anxiety	3.73	0.03 0.31*	45.6	
Control	4.76	0.09*0.14*	6.2	
Cognitive triad	5.14	0.14*0.21*	13.8	
Uncertainty	2.48	-0.02*0.09*	14.6	

NOTE: Results of the χ^2 tests of the equality of the positive and negative coefficients are in the column labeled χ^2 . All of these tests had 1 df and were significant at the .01 level or beyond.

* $p < .005$.

ships between daily deviations from each person's mean event scores and deviations from that person's mean adjustment.

Hypotheses about day-level relationships between adjustment and events were tested by analyzing Level 1 coefficients at Level 2 (the person) using the following model:

$$\beta_{0j} = \gamma_{q0} + u_{qj}.$$

In these models, γ_{q0} represented the average of the Level 1 coefficients describing relationships between measures of adjustment and daily events. Error was represented by u_{qj} . For each measure, there were three coefficients ($q = 0, 1, 2$), the intercept, a coefficient (referred to as a slope to distinguish it from an intercept) for positive events, and a slope for negative events. The mean slope between positive events and adjustment was represented by γ_{10} , and γ_{20} represented the mean slope between negative events and adjustment. The results of these analyses are presented in Table 2.

As hypothesized, there were significant relationships between daily adjustment and both positive and negative events. All five γ_{20} slopes representing relationships between adjustment and negative events were significantly different from 0. Moreover, the γ_{10} slopes representing the relationships between positive events and the self-esteem, control, cognitive triad, and uncertainty

TABLE 3: Day-Level Relationships Between Adjustment and Events as Moderated by Depressive Symptoms

	<i>Self-Esteem</i>			<i>Cognitive Triad</i>			<i>GCOS Control</i>			<i>Anxiety</i>			<i>Causal Uncertainty</i>		
	<i>Coef.</i>	<i>t</i>	<i>p</i>	<i>Coef.</i>	<i>t</i>	<i>p</i>	<i>Coef.</i>	<i>t</i>	<i>p</i>	<i>Coef.</i>	<i>t</i>	<i>p</i>	<i>Coef.</i>	<i>t</i>	<i>p</i>
Intercept															
Nondepressed	7.64	77.9	<.001	5.46	72.6	<.001	4.99	62.9	<.001	3.18	22.7	<.001	2.28	28.9	<.001
Depressed	6.07	8.3	<.001	4.26	8.5	<.001	4.13	5.6	<.001	5.22	7.5	<.001	3.01	4.8	<.001
Positive events															
Nondepressed	0.07	4.9	<.001	0.12	8.9	<.001	0.09	6.9	<.001	-0.01	<1	<i>ns</i>	-0.02	1.7	<.11
Depressed	0.12	2.2	<.05	0.19	2.5	<.015	0.10	<1	<i>ns</i>	-0.06	<1	<i>ns</i>	-0.04	1.5	<.12
Negative events															
Nondepressed	-0.16	8.8	<.001	<0.18	10.2	<.001	<0.14	7.8	<.001	0.30	7.7	<.001	-0.10	5.9	<.001
Depressed	-0.23	2.1	<.05	<0.26	2.3	<.02	<0.14	<1	<i>ns</i>	0.34	<1	<i>ns</i>	-0.06	1.4	<.17

NOTE: For coefficients for the nondepressed, the column labeled *p* contains the probability level of tests that the coefficient was not 0, whereas for coefficients for the depressed, the column labeled *p* contains the probability level of tests that the difference between the nondepressed and depressed coefficients was 0.

measures were also significantly different from 0. The relationship between anxiety and positive events did not approach conventional levels of significance ($p > .20$).

Across all participants, adjustment was lower on days when negative event scores were higher than on days when negative events score were lower, and adjustment (except anxiety) was higher on days when positive event scores were higher than on days when they were lower. For example, for daily self-esteem, the average positive event coefficient was .08, and the average negative event coefficient was -.18. On average, for each positive event above his or her mean number of positive events a person experienced, that person's daily self-esteem increased .08. Correspondingly, for each negative event above his or her mean number of negative events a person experienced, daily self-esteem decreased .18.

The strength of day-level relationships between positive events and adjustment and day-level relationships between negative events and adjustment were compared using tests of fixed effects (Bryk & Raudenbush, 1992, pp. 48-52). To account for the fact that coefficients representing these relationships differed in sign, these tests compared the absolute values of coefficients. As hypothesized, the covariation between daily adjustment and negative events was stronger than the covariation between adjustment and positive events. The absolute magnitudes of all γ_{20} coefficients were significantly larger than the magnitudes of the corresponding γ_{10} coefficients. The results of these analyses are presented in Table 2.

Depression as a Moderator of Day-Level Relationships Between Events and Adjustment

The last step in the analysis examined how trait levels of adjustment moderated relationships between daily events and adjustment. Participants were classified as depressed or not based on four reports of depressive symptoms provided over 4 1/2 months, a CES-D and BDI

completed 7 weeks prior to the study, and a CES-D and BDI completed during the study. Thirty-three participants who scored above cutpoints of 10 on the BDI and 17 on the CES-D on at least three out of these four measures were classified as depressed, and the remaining 90 were classified as nondepressed. This procedure ensured that only participants who reported high levels of depressive symptoms over an extended period of time were classified as depressed.⁴

Differences between these two groups were examined using a variant of the model used in the previous analyses:

$$\beta_{qj} = \gamma_{q0} + \gamma_{q1}(\text{DEP}) + u_{qj}.$$

As in the previous analysis, γ_{qj} represented the coefficients generated in the Level 1 models, the relationships between measures of adjustment and events. Depression was dummy coded (1 = depressed); therefore, γ_{q0} represented the average Level 1 coefficient for the nondepressed and γ_{q1} represented the depression effect, how much the average Level 1 coefficient for the depressed differed from the average coefficient for the nondepressed. Differences between depressed and nondepressed participants in the strength of the relationship between daily events and adjustment were tested by the γ_{q1} coefficients.

Consistent with previous research (e.g., Bolger & Schilling, 1991), the mean negative event scores for the depressed participants were higher than for the nondepressed (2.79 vs. 1.89, $p < .01$), although there were no differences between the two groups in positive event scores. As in the previous analysis, event scores were group-mean-centered, eliminating the influence on parameter estimates of these differences in event scores. The results of these analyses are presented in Table 3. To simplify the presentation of these results, the depression effect for each coefficient (γ_{q1}) was added to the coeffi-

TABLE 4: Lagged Relationships From Daily Events to Daily Adjustment

<i>Present Day</i>	<i>Previous Day</i>	<i>Coef.</i>	<i>t</i>	<i>p</i>
Self-esteem	Self-esteem	.08	2.38	<.02
	Positive events	.00	<1	<i>ns</i>
	Negative events	-.04	2.38	<.02
Cognitive triad	Cognitive triad	.12	4.39	<.01
	Positive events	.00	<1	<i>ns</i>
	Negative events	.00	<1	<i>ns</i>
GCOS-Control	GCOS-Control	.09	3.20	<.01
	Positive events	.01	1.03	<i>ns</i>
	Negative events	.00	<1	<i>ns</i>
Anxiety	Anxiety	.20	7.01	<.01
	Positive events	.00	<1	<i>ns</i>
	Negative events	.07	2.46	<.02
Causal uncertainty	Causal uncertainty	.13	4.46	<.01
	Positive events	.02	2.49	<.02
	Negative events	.02	1.38	<i>ns</i>

NOTE: Mean unstandardized coefficients are in the column labeled Coef. GCOS = General Causality Orientation Scale.

cient describing the nondepressed (γ_{q0}) so that the tabled values represent the average coefficients for the two groups.

As hypothesized, relationships between self-esteem and the cognitive triad measures were significantly larger for depressed than for nondepressed participants. For these measures, the γ_{11} and γ_{21} coefficients were significantly different from 0. For example, a unit increase in daily positive events scores was associated with a .12 increase in self-esteem for the depressed, whereas for the nondepressed it was associated with a .07 increase. For negative events, the corresponding figures for a unit increase were decreases of .23 and .16. Contrary to expectation, relationships between events and the other measures did not vary as a function of depression.⁵

Finally, there were significant differences between nondepressed and depressed participants in the daily means of all measures (intercepts in Table 4). In terms of all measures, nondepressed participants were better adjusted on a daily basis than depressed participants were after controlling for daily events.⁶

Lagged Relationships Between Daily Adjustment and Daily Events

Although the present study was not explicitly designed to study causal relationships, examining lagged relationships between constructs can provide some insight into causal relationships (e.g., West & Hepworth, 1991). Accordingly, a series of analyses was conducted in which adjustment on day *i* was modeled as a function of adjustment on day *i*-1 and events on day *i*-1. Parallel analyses were conducted in which events on day *i* were modeled as a function of adjustment on day *i*-1 and

events on day *i*-1. For example, to determine whether changes in self-esteem lead to or were followed by changes in events, the following models were analyzed:

$$\text{Lag 1: ESTEEM}(\text{day } i)_{ij} = \beta_{0j} + \beta_{1j}(\text{ESTEEM day } i-1) + \beta_{2j}(\text{POS-EVENT day } i-1) + \beta_{3j}(\text{NEG-EVENT day } i-1) + r_{ij}.$$

$$\text{Lag 2: POS-EVENT}(\text{day } i)_{ij} = \beta_{0j} + \beta_{1j}(\text{ESTEEM day } i-1) + \beta_{2j}(\text{POS-EVENT day } i-1) + r_{ij}.$$

$$\text{Lag 3: NEG-EVENT}(\text{day } i)_{ij} = \beta_{0j} + \beta_{1j}(\text{ESTEEM day } i-1) + \beta_{2j}(\text{NEG-EVENT day } i-1) + r_{ij}.$$

The critical coefficients in these models are the lagged coefficients on the Lag 1 equation β_{2j} (POS-EVENT day *i*-1) and β_{3j} (NEG-EVENT day *i*-1); and the lagged coefficient, β_{1j} (ESTEEM day *i*-1), in the Lag 2 and Lag 3 equations. A causal sequence from events to self-esteem is suggested by significant β_{2j} (POS-EVENT day *i*-1) or β_{3j} (NEG-EVENT day *i*-1) coefficients in the Lag 1 equation, whereas a sequence from self-esteem to events is suggested by a significant β_{1j} (ESTEEM day *i*-1) coefficient in the Lag 2 or Lag 3 equations. These analyses required that data were provided on consecutive days. Of the 2,412 days recorded, only 2,289 could have data for a previous day because there were 123 participants and for each of them, the first day could not logically have a previous day. Of these 2,289 days, 2,221 (97%) had data for previous days.

These analyses found no statistically significant lags from adjustment to events, and only one lag (from anxiety to negative events) approached conventional levels of significance ($p = .085$). In light of this pattern, these results are not presented or discussed. The analyses of lagged relationships from events to adjustment found three significant lags, one from negative events to self-esteem, one from negative events to anxiety, and one from positive events to causal uncertainty. The results of these analyses are summarized in Table 4. Consistent with the results of the static (within a single day) analyses, more negative events on day *i*-1 were associated with lower self-esteem and greater anxiety on the following day. Somewhat surprisingly, and inconsistent with the results of the static analyses, more positive events on day *i*-1 were associated with increased causal uncertainty on the following day.

Supplementary Analyses of Mood

The independence of the adjustment-event covariance found in this study from the mood-event covariance found in other studies was examined in a series of analyses in which relationships between adjustment and events were controlled for anxiety and mood. Each day, participants rated their mood on a 1 to 9 scale

with endpoints labeled *happy-sad*. When measures of adjustment (other than anxiety) were modeled as a function of events and anxiety and mood, with one exception (the positive events-causal uncertainty relationship), slopes between adjustment and events remained statistically significant (p 's < .01) and meaningful in magnitude. Admittedly, anxiety is only a component of negative affectivity, and happy-sad is only a component aspect of a much more complex affective construct. Nonetheless, in a structurally similar study, Nezlek (1999) found that daily self-esteem and daily depressogenic adjustment (operationalized as they were in this study) covaried meaningfully with negative and positive events after controlling for the covariance between events and both NA and PA. Taken together, these results suggest that daily adjustment as operationalized in the present study covaries meaningfully with daily events above and beyond the covariation between events and mood.

DISCUSSION

As hypothesized, daily depressogenic adjustment negatively covaried with the negative events that occurred each day, and except for anxiety, daily adjustment positively covaried with positive events. These findings complement and extend previous research on daily events by demonstrating that depressogenic adjustment is sensitive to daily events. Moreover, the fact that adjustment covaried with both positive and negative events suggests that daily adjustment as operationalized in this study was not merely a measure of a general negativity factor such as NA. Research examining both positive and negative events and NA has found that NA covaries with negative but not with positive events (e.g., Gable et al., 2000; Nezlek, 1999).

Although considerable evidence indicates that anxiety and depression are closely associated at the trait level (e.g., Feldman, 1993), the present results suggest that anxiety and depression are somewhat distinct phenomena at the state level. This conclusion is similar to those reached by Stader and Hokanson (1998) in a study of the daily covariability of depressive symptoms and psychosocial processes and by Roberts and Gotlib (1997) in a study of temporal variability in self-evaluation.

Also as expected, the day-level covariation between depressogenic adjustment and negative events was stronger than the covariation between adjustment and positive events. This replicates previous research on daily events and daily variability in self-esteem (Butler et al., 1994) and is consistent with research indicating that negative events and stimuli generally have more impact than positive events and stimuli (Cacioppo et al., 1997; Taylor, 1991). This result extends previous research by explicitly documenting that this tendency also characterizes rela-

tionships between daily events and measures of daily depressogenic adjustment.

The present results that are probably the most relevant to understanding depression are those describing differences between depressed and nondepressed people in day-level relationships between events and adjustment. As expected, depressed people were more labile (reacted more strongly to events) than the nondepressed in terms of their self-esteem and depressive thinking. The daily variability of depressive thinking per se has not been examined; therefore, the present results showing that trait depression moderates relationships between daily events and daily depressive thinking meaningfully extends previous research on lability and trait adjustment. The depression effect in the lability of self-esteem also agrees with the results of a previous study of such relationships (Butler et al., 1994). Butler et al. found that the previously depressed were more labile than the currently nondepressed (in terms of self-esteem-relevant self-evaluations), although the currently depressed were not different from either of these two groups.

Butler et al. (1994) discussed a variety of explanations for the greater lability of the depressed, and most of these explanations emphasized the weaker sense of self the depressed may have. The present results confirm this logic and also suggest another complementary explanation. The stronger reactions of the depressed to positive events may have been due to different expectations held by the depressed and nondepressed. A negative view of the future is an essential component of Beck's cognitive triad. Consistent with this premise, in a study of daily plans and goals, Nezlek and Elia (1999) found that higher depressive symptoms were associated with setting goals that would take more time to accomplish and were more difficult and harder to accomplish. Within such a framework, the positive events that occur (e.g., meeting a goal) to the depressed may have a greater influence on psychological well-being than they do for the nondepressed due to a contrast effect or a violation of expectations. It is important to note that such a contrast effect would not be due to the differences in how many positive events occurred. There was no depression effect in the number of positive events reported per day, and because event scores were group-mean-centered, individual differences in event scores did not contribute to parameter estimates.

The assumption that trait depression makes people more vulnerable to stressful or negative events is central to diathesis-stress models, and the present results supported such models in terms of the lability of self-esteem and depressive thinking. At first glance, the fact that depression did not moderate some adjustment-negative event relationships seems to be inconsistent with such

assumptions; however, it may not be. The daily adjustment of the depressed was varying around a much lower mean than the daily adjustment of the nondepressed. For depressed people, decreases in daily adjustment brought about by negative events may lower adjustment to a point where it has implications for trait adjustment, whereas the decreases the nondepressed experience may not. The nondepressed may vary through a normal or adaptive range, whereas the range of the depressed may include a maladaptive segment into which the depressed fall in response to negative events. The relatively infrequent experience of such poor daily adjustment may not have implications for trait adjustment but more frequent experience of such poor daily adjustment may maintain or increase the risk for depression.

Differences in people's lability across different day-level measures of adjustment are not unique to this study. Moderators of event-outcome relationships have been studied with a wide variety of measures of daily adjustment, trait adjustment, and daily events, and as noted earlier, these relationships vary across studies. This may reflect the fact that event-outcome relationships may be more sensitive to the specific ways in which constructs have been operationalized than has been previously supposed. For example, Kennedy-Moore, Greenberg, Newman, and Stone (1992) found different day-level relationships using different operationalizations of mood. Clearly, resolving such issues will require a larger body of research.

The preceding discussion has tacitly assumed that daily events affect daily adjustment, an assumption made in most research on daily events. Such a causal relationship is consistent with the results of a longitudinal study by Suh, Diener, and Fujita (1996), who found that recent positive life events lead to increases in life satisfaction, whereas recent negative life events lead to decreases. Similarly, based on structural equation modeling analyses, Nezlek and Reis (1999) concluded that quality of daily social interaction was causally related to mental health, not the reverse. Moreover, studies of causal relationships between daily events and psychological states suggest an event-state causal sequence. Bolger and Zuckerman (1995) found that distress experienced on a particular day was related to the conflict experienced on a previous day, and Gable et al. (2000) found that events occurring on a preceding day predicted present-day affect, whereas prior day affect did not predict present-day events.

In the present study, analyses of lagged relationships similar to those presented by Gable et al. (2000) found some evidence for a causal link from events to adjustment. In the analyses of self-esteem and anxiety, prior day's negative events predicted present adjustment controlling for prior day's adjustment. In contrast, no lagged

relationship between prior day's adjustment and present events approached conventional levels of significance. Although suggestive, the results of these lagged analyses need to be considered cautiously. Lagged relationships occurred only for some of the measures and in the expected direction only for negative events. Clearly, more research that is explicitly designed to examine such causal relationships is needed to resolve such questions.

One of the more important questions about the present results concerns the extent to which the moderating effects of trait depression on day-level relationships reflect differences in depression per se or differences in neuroticism or a general predisposition to experience negative affect. We think the present results make a valuable contribution to the understanding of depression for three reasons. First, depression was measured four times over a 4 1/2-month period using two different measures. Although individual differences in self-reports of depressive symptoms may reflect individual differences in a more general factor, such a general factor does not account for the unique variation associated with depression. Multiple assessments with multiple measures of depression should maximize the extent to which individual differences reflect depression per se rather than a general negativity factor. Second, in previous studies, neuroticism (or general negativity) has not moderated day-level relationships between positive events and psychological states, and the present measure of depression did. Third, Nezlek (1999) replicated the present results and found that compared to trait PA, NA, and anxiety, the CESD was the most reliable moderator of relationships between daily positive events and daily self-esteem and between positive events and daily depressogenic adjustment.

What implications do the present results have for understanding the etiology and maintenance of depression? Keeping various caveats in mind, they suggest that cognitive reactions to daily events are part of these processes. Cognitively focused measures of depressogenic adjustment covaried with the events that occurred each day, and although such relationships may be consistent with previous research and theory, they have not been demonstrated before. Most important, the present results suggest that understanding psychological well-being (at least in terms of depression) needs to take into account people's reactivity to positive events. Unfortunately, as discussed by Barnett and Gotlib (1988), the temporally static design of the present study does not provide a basis for determining if the relationships described herein are antecedents, concomitants, or consequences of depression. Such determination must await the results of prospective studies.

Appendix

Reliability and Validity Analyses

For each construct, a series of three-level models was conducted in which the items measuring that construct were nested within days and days were nested within participants. Each analysis included 123 participants (Level 3) and 2,412 days (Level 2). The number of entries at Level 1 was the product of the number of items in a measure and the number of days. Equations describing these models are presented below with the nomenclature used by Bryk and Raudenbush (1992).

Model 1: Totally Unconditional

$$\begin{array}{lll} \text{Item level} & (\text{Level 1}) & y_{ijk} = \pi_{0jk} + e_{ijk} \\ \text{Day level} & (\text{Level 2}) & \pi_{0jk} = \beta_{00k} + r_{0jk} \\ \text{Person level} & (\text{Level 3}) & \beta_{00k} = \gamma_{000} + u_{00k} \end{array}$$

Model 2: Traits Included at Person Level

$$\text{Person level (Level 3)} \quad \beta_{00k} = \gamma_{000} + \gamma_{001}(\text{TRAIT}) + u_{00k}$$

The first model is called a totally unconditional model in multilevel random coefficient models (MRCM) terminology. Coefficients from Level 1 (the item level) were modeled only as intercepts at both Level 2 (days) and Level 3 (persons), providing reliability estimates at the day and person levels and estimates of the error variance of the latent daily mean of each construct, the variance of u_{00k} . In the second model, coefficients from Level 2 (daily means) were modeled at Level 3 as a function of the corresponding trait measure. The validity of a daily measure was operationalized as the reduction in error variance from the first and second models. The analyses of self-esteem and depression included two trait measures at the person level in the second model because self-esteem and depression were measured at the beginning and end of the study.

The analysis of the General Causality Orientation Scale (GCOS) was not as straightforward as the analyses of the other measures. Six items were designed to provide daily measures of the autonomy, control, and impersonal orientations of the GCOS (two items each). The validity and reliability of these measures were examined using a three-level HLM with a zero-intercept Level 1 model:

$$y_{ijk} = \pi_{1jk}(\text{AUT}) + \pi_{2jk}(\text{IMP}) + \pi_{3jk}(\text{CON}) + e_{ijk}.$$

In this model, y was the response, π_{1jk} was a dummy-coded variable representing the two autonomy items (AUT) (autonomy items coded as 1, other items coded as 0), π_{2jk} was a dummy-coded variable representing the two impersonal items (IMP), and π_{3jk} was a dummy-coded variable representing the two control items (CON). The three coefficients, π_{1jk} , π_{2jk} , and π_{3jk} , represented the daily means for each construct for each person.

This analysis provided estimates of the day-level correlations (τ_{π} s) among the three constructs. The day-level correlation between the coefficients representing the autonomy and impersonal constructs was .99, strongly suggesting that the four

items intended to measure these two constructs measured the same construct, control over outcomes of behavior. This analysis also found that daily mean levels of these two constructs were related (both p s < .0001) to the impersonal orientation on the GCOS at the individual level (Level 3). Furthermore, the two items designed to measure control orientation were not significantly related to any of the three GCOS trait measures (all p s > .25). These findings lead us to recode the four items designed to measure the autonomy and impersonal orientations into one four-item measure of impersonal orientation and to drop the two items intended to measure the control orientation. The resulting measure was a reliable and valid measure of the impersonal construct.

NOTES

1. Positive and negative composite scores, the average importance of events, also were created. The results of analyses using composite scores were similar to the results presented in this article. Also, the present study used only a subset of the daily events survey (DES) items because it was felt that some of the items on the DES occurred too infrequently to qualify as a daily event. It appears that no frequent items were eliminated because the mean number of positive and negative events recorded per day in this study was similar to the numbers reported by Butler, Hokanson, and Flynn (1994) using the full scale version.

2. Unfortunately, the date and time participants provided their responses were not recorded. Nevertheless, we are confident that participants complied with instructions. First, they were sent reminders every 3 days to be certain to comply. Second, as instructed, some participants skipped days when they forgot to run the daily program. Third, and most important, the distributions of positive and negative event scores obtained in the present study are similar to the distributions from a similar study in which date and time of response were recorded (Nezlek & Plesko, 2001). In the present study, participants recorded an average of 5.25 positive events per day ($SD = 3.03$) and an average of 2.32 negative events ($SD = 1.41$). The corresponding figures from Nezlek and Plesko were 5.07 (2.06) and 2.39 (1.38). Also, Nezlek and Plesko used similar procedures to collect data and to instruct and maintain contact with participants, and they excluded only 3% of participants and 1% of days from their analyses because responses were not recorded as requested.

3. The results of analyses that included these two items did not differ meaningfully from those reported in this article.

4. For ease of presentation, those who scored above the cutpoints are referred to as *depressed* and those who scored below the cutpoints are referred to as *nondepressed*. These terms have been used only to simplify the discussion, and this use does not imply that participants who scored above the cutpoints had been diagnosed as depressed or were clinically depressed. Also, the results of analyses that operationalized adjustment as a continuous variable (a factor score based on the four reports of symptoms) were similar to the results presented in this article.

5. A variety of different statistical artifacts such as autocorrelations and heteroskedasticity of variances can influence the results of data collected in a multilevel design over time. The results of analyses that controlled for temporal trends and autocorrelations in the data were very similar to those presented in this article. Moreover, examination of predicted scores suggested that the present results were not due to individual or group differences in the variances of coefficients or to floor or ceiling effects for the rating scales. Details of these analyses are available from the first author.

6. Overall, nondepressed participants had higher daily levels of well-being than depressed participants. Differences between depressed and nondepressed participants in mean levels of daily adjustment are discussed in Gable and Nezlek (1998).

REFERENCES

- Affleck, G., Tennen, H., Urrows, S., & Higgins, P. (1994). Person and contextual features of daily stress reactivity: Individual differences of undesirable daily events with mood disturbance and chronic pain intensity. *Journal of Personality and Social Psychology*, 66, 329-340.
- Alloy, L. B., Kelly, K. A., Mineka, S., & Clements, C. M. (1990). Comorbidity of anxiety and depressive disorders: A helplessness-hopelessness perspective. In J. D. Maser & C. R. Cloninger (Eds.), *Comorbidity of mood and anxiety disorders* (pp. 499-544). Washington, DC: American Psychiatric Press.
- Barnett, P. A., & Gotlib, I. H. (1988). Psychological functioning and depression: Distinguishing among antecedents, concomitants, and consequences. *Psychological Bulletin*, 104, 97-126.
- Beck, A. T. (1967). *Depression: Clinical, experimental, and theoretical aspects*. New York: Hoeber.
- Beck, A. T. (1972). *Depression: Causes and treatment*. Philadelphia: University of Pennsylvania Press.
- Bolger, N. (1990). Coping as a personality process: A prospective study. *Journal of Personality and Social Psychology*, 59, 525-537.
- Bolger, N., DeLongis, A., Kessler, R., & Schilling, E. A. (1989). Effects of daily stress on negative mood. *Journal of Personality and Social Psychology*, 57, 808-818.
- Bolger, N., & Schilling, E. A. (1991). Personality and the problems of everyday life: The role of neuroticism in exposure and reactivity to daily stressors. *Journal of Personality*, 59, 355-386.
- Bolger, N., & Zuckerman, A. (1995). A framework for studying personality in the stress process. *Journal of Personality and Social Psychology*, 69, 890-902.
- Brockner, J. (1984). Low self-esteem and behavioral plasticity: Some implications for personality and social psychology. In L. Wheeler (Ed.), *Review of personality and social psychology* (Vol. 4, pp. 181-238). Beverly Hills, CA: Sage.
- Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models*. Newbury Park, CA: Sage.
- Bryk, A. S., Raudenbush, S. W., & Congdon, R. T. (1998). *HLM Version 4.04*. Chicago: Scientific Software International.
- Butler, A. C., Hokanson, J. E., & Flynn, H. A. (1994). A comparison of self-esteem lability and low trait self-esteem as vulnerability factors for depression. *Journal of Personality and Social Psychology*, 66, 166-177.
- Cacioppo, J. T., Gardner, W. L., & Berntson, G. G. (1997). Beyond bipolar conceptualizations and measures: The case of attitudes and evaluative space. *Personality and Social Psychology Review*, 1, 3-25.
- Clark, L. A., & Watson, D. (1988). Mood and the mundane: Relations between daily life events and self-reported mood. *Journal of Personality and Social Psychology*, 54, 296-308.
- David, J. P., Green, P. J., Martin, R., & Suls, J. (1997). Differential roles of neuroticism, extraversion, and event desirability for mood in daily life: An integrative model of top-down and bottom-up influences. *Journal of Personality and Social Psychology*, 73, 150-159.
- Deci, E. L., & Ryan, R. M. (1985). The General Causality Orientations Scale: Self-determination in personality. *Journal of Research in Personality*, 19, 109-134.
- Eysenck, H. J. (1967). *The biological basis of personality*. Springfield, IL: Charles C Thomas.
- Feldman, L. A. (1993). Distinguishing depression and anxiety in self-report: Evidence from confirmatory factor analyses on nonclinical and clinical samples. *Journal of Consulting and Clinical Psychology*, 61, 631-638.
- Gable, S. L., & Nezlek, J. B. (1998). Level and instability of day-to-day psychological well-being and risk for depression. *Journal of Personality and Social Psychology*, 74, 129-138.
- Gable, S. L., Reis, H. T., & Elliot, A. J. (2000). Behavioral activation and inhibition in everyday life. *Journal of Personality and Social Psychology*, 78, 1135-1149.
- Kennedy-Moore, E., Greenberg, M. A., Newman, M. G., & Stone, A. A. (1992). The relationships between daily events and mood: The mood measure may matter. *Motivation and Emotion*, 16, 143-155.
- Kenny, D. A., Kashy, D. A., & Bolger, N. (1998). Data analysis in social psychology. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 1, pp. 233-265). New York: McGraw-Hill.
- Kernis, M. H. (1993). Stability and instability of self-esteem. In R. Baumeister (Ed.), *Self-esteem: The puzzle of low self regard* (pp. 167-182). New York: Plenum.
- Kreft, I., & de Leeuw, J. (1998). *Introducing multilevel modeling*. Thousand Oaks, CA: Sage.
- Larsen, R. J., & Kasimatis, M. (1991). Day-to-day physical symptoms: Individual differences in the occurrence, duration, and emotional concomitants of minor daily illnesses. *Journal of Personality*, 59, 387-423.
- Lewinsohn, P. M., & Graf, M. (1973). Pleasant activities and depression. *Journal of Consulting and Clinical Psychology*, 41, 261-268.
- Lewinsohn, P. M., & Libet, J. (1972). Pleasant events, activity schedules, and depressions. *Journal of Abnormal Psychology*, 79, 291-295.
- Lorr, M., & McNair, D. M. (1971). *The Profile of Mood States manual*. San Diego, CA: Educational and Industrial Testing Service.
- Marco, C. A., & Suls, J. (1993). Daily stress and the trajectory of mood: Spillover, response assimilation, contrast, and chronic negative affectivity. *Journal of Personality and Social Psychology*, 64, 1053-1063.
- Nezlek, J. B. (1999). Trait adjustment as a moderator of the interactive effects of positive and negative daily events on daily psychological adjustment. In J. B. Nezlek (Chair), *Studying daily variability in psychological states*. Symposium conducted at the meeting of the Society of Experimental Social Psychology, St. Louis, Missouri.
- Nezlek, J. B. (2001). Multilevel random coefficient analyses of event and interval contingent data in social and personality psychology research. *Personality and Social Psychology Bulletin*, 27, 771-785.
- Nezlek, J. B., & Elia, L. (1999). *Daily goal fulfillment and psychological well-being*. Manuscript in preparation, College of William & Mary, Williamsburg, Virginia.
- Nezlek, J. B., & Plesko, R. M. (2001). Day-to-day relationships among self-concept clarity, self-esteem, daily events, and mood. *Personality and Social Psychology Bulletin*, 27, 201-211.
- Nezlek, J. B., & Reis, H. T. (1999, June). *Causal relationships between quality and quantity of social interaction and psychological well-being*. International Network on Personal Relationships Conference, Louisville, Kentucky.
- Radloff, L. S. (1977). The CES-D Scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, 1, 385-401.
- Roberts, J. E., & Gotlib, I. H. (1997). Temporal variability in global self-esteem and specific self-evaluation as prospective predictors of emotional distress: Specificity in predictors and outcomes. *Journal of Abnormal Psychology*, 106, 521-529.
- Rogers, C. (1961). *On becoming a person: A therapists view of psychotherapy*. Boston: Houghton Mifflin.
- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.
- Schneider, W. (1988). Micro Experimental Laboratory: An integrated system for IBM-PC compatibles. *Behavior Research Methods, Instrumentation, and Computers*, 20, 206-217.
- Smith, T. W., & Rhodewalt, F. (1986). On states, traits, and processes: A transactional alternative to the individual differences in Type A behavior and physiological reactivity. *Journal of Research in Personality*, 20, 229-251.
- Stader, S. R., & Hokanson, J. E. (1998). Psychosocial antecedents of depressive symptoms: An evaluation using daily experiences methodology. *Journal of Abnormal Psychology*, 107, 17-26.
- Stone, A. A. (1981). The association between perceptions of daily experience and self- and spouse-rated mood. *Journal of Research in Personality*, 15, 510-522.
- Stone, A. A. (1987). Event content in a daily survey is differentially associated with concurrent mood. *Journal of Personality and Social Psychology*, 52, 56-58.

- Suh, E., Diener, E., & Fujita, F. (1996). Events and subjective well-being: Only recent events matter. *Journal of Personality and Social Psychology*, 70, 1091-1102.
- Suls, J., Martin, R., & David, J. P. (1998). Person-environment fit and its limits: Agreeableness, neuroticism, and emotional reactivity to interpersonal conflict. *Personality and Social Psychology Bulletin*, 24, 88-97.
- Taylor, S. E. (1991). Asymmetrical effects of positive and negative events: The mobilization-minimization hypothesis. *Psychological Bulletin*, 110, 67-85.
- Watson, D. (1988). Intraindividual and interindividual analyses of positive and negative affect: Their relation to health complaints, perceived stress, and daily activities. *Journal of Personality and Social Psychology*, 54, 1020-1030.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS Scales. *Journal of Personality and Social Psychology*, 54, 1063-1070.
- Watson, D., & Tellegen, A. (1985). Toward a consensual model of mood. *Psychological Bulletin*, 98, 219-235.
- Weary, G., & Edwards, J. A. (1994). Individual differences in causal uncertainty. *Journal of Personality and Social Psychology*, 67, 308-318.
- Weary, G., Jordan, J. S., & Hill, M. G. (1985). The attributional norm of internality and depressive sensitivity to social information. *Journal of Personality and Social Psychology*, 49, 1283-1293.
- West, S. G., & Hepworth, J. T. (1991). Statistical issues in the study of temporal data: Daily experiences. *Journal of Personality*, 59, 611-662.

Received January 23, 2001

Revision accepted February 14, 2001

Self-Enhancement: Is It Restricted to Individualistic Cultures?

Jenny Kurman
University of Haifa

Two studies investigated how self-enhancement relates to interdependent and independent self-construals typical of collectivist and individualist cultures, respectively. Participants from three cultures were surveyed, two of them collectivist (Singaporean Chinese and Israeli Druze) and one individualist (Israeli Jews). Study 1 shows that the two collectivist cultures differ in their self-enhancement level among university students: Self-enhancement level among Singaporeans was weaker for academic self-enhancement and for other agentic traits. No cultural difference in self-enhancement of communal traits was found (N = 418). Study 2 replicated these results for high school students and evaluated the relations between self-construals, modesty, and self-enhancement. Regression analyses show that self-enhancement of agentic traits is predicted by independent self-construal and modesty (negatively), whereas self-enhancement of communal traits is predicted by interdependent self-construal (N = 362). The role of modesty norms in self-enhancement is discussed.

Differences in self-concept across cultures have been studied extensively in recent years. A comprehensive review by Markus and Kitayama (1991) suggested that individualist and collectivist societies are characterized by different self-systems. Individualist cultures tend to promote an independent self that is autonomous and self-contained. Collectivist cultures, in contrast, tend to foster an interdependent self that is part of a comprehensive social relationship and that is partially defined by others in that relationship. One of the suggested differences between these two self-systems is the strength and centrality of self-enhancement.

There is ample evidence to suggest that self-enhancement is a basic motivating factor for the self. The most direct evidence for the predominance of self-enhancement is the “above-average effect”: More than the expected 50% of many samples evaluate themselves as above average on many desirable traits (Brown, 1998). Among the possible mechanisms for self-enhancement

are applying selective memory to recalling positive events (Kunda, 1987), performing downward social comparisons (Wills, 1981), and using self-serving attributions (Miller & Ross, 1975).

Nevertheless, many findings suggest that this self-enhancement tendency is not shared by all cultures. For example, respondents of Chinese ethnicity reported lower self-evaluations than did American-born respondents (Bond & Cheung, 1983; White & Chan, 1983). A review of 23 Japanese studies (Kitayama, Takagi, & Matsumoto, 1995) concluded that Japanese fail to exhibit self-serving attributional biases. Furthermore, Takata (1987) reported a “modesty bias” in Japanese participants: They tended to accept negative feedback concerning themselves as more valid than positive feedback. These findings seem to differ substantially from responses of American participants (J. M. Schwartz & Smith, 1976). A study that compared the importance assigned to success situations with that assigned to failure situations (Kitayama, Markus, Matsumoto, & Norasakkunkit, 1997) showed that Americans emphasize success (self-enhancement), whereas Japanese emphasize failure (self-criticism). Hence, it was concluded that self-enhancement is a tendency that prevails much more strongly in the United States.

Markus and Kitayama (1991, p. 242) suggested that motivational differences between the independent and the interdependent self can explain these findings. In contrast to the self-enhancing independent self, the interdependent self is motivated to fit in, restrain itself, and maintain social harmony. In their analysis, Markus and Kitayama imply that the interdependent self has less need for positive self-evaluation in self-components that

Author's Note: Correspondence concerning this article should be sent to Jenny Kurman, Department of Psychology, University of Haifa, Mt. Carmel, 31905, Haifa, Israel; e-mail: jennyk@psy.haifa.ac.il.

PSPB, Vol. 27 No. 12, December 2001 1705-1716

© 2001 by the Society for Personality and Social Psychology, Inc.

are not related to others than does the independent self because the link between personal attributes and self-esteem is less pronounced in collectivist societies. In these societies, lesser value is attached to personal attributes and greater value is assigned to the ability to fit in and be part of the group. Having more positive self-attributes, which is the main advantage of self-enhancement, seems to be less central to the interdependent self than to the independent self. Furthermore, the cost of self-enhancement may be higher for the interdependent self because the group may regard enhanced self-evaluation as an attempt to be singled out. Such an attempt may pose a threat to the harmony of the group. Markus and Kitayama (1991), therefore, assumed that for the interdependent self, self-enhancement does not constitute a motivating factor (see also Heine & Lehman, 1995, p. 605; Heine, Lehman, Markus, & Kitayama, 1999, p. 770).

The present study challenges this idea that lack of self-enhancement is intrinsic to the interdependent self. Moreover, it claims that cultural differences reflected in degree of self-enhancement are not necessarily the inevitable result of the collectivist nature of the interdependent self. This claim will be supported by two studies. The first compares two collectivist cultures whose endemic values differ on two types of traits. The purposes of this study are (a) to question the assumption that low self-enhancement is not characteristic of all collectivist cultures and (b) to show that self-enhancement of communal traits, which may facilitate good relationships and do not reflect personal success, can thrive in both individualist and collectivist cultures. The second study focuses on predicting self-enhancement and tests the relative contribution of self-construals and of modesty to self-enhancement of agentic and communal traits.

STUDY 1: A CROSS-CULTURAL COMPARISON OF SELF-ENHANCEMENT LEVEL IN TWO TYPES OF TRAITS

The vast majority of findings on lack of self-enhancement in collectivist societies were obtained from participants of Japanese or Chinese origin. From this data, it is not possible to determine whether this well-established tendency reflects collectivist values or values endemic to these specific cultures. One way to differentiate between these two explanations is to compare the level of self-enhancement in different collectivist cultures. If a general collectivist perspective is the main cause for lower self-enhancement, it would be valid to predict that all collectivist cultures will fail to exhibit self-enhancement regardless of their endemic values. On the other hand, if specific values are responsible for the lack of self-enhancement found in East Asian cultures, collectivist cultures with differing endemic values may exhibit self-enhancement. To determine which of these assumptions

is supported by the data, it is necessary to investigate collectivist cultures whose heritage differs from East-Asian cultures. Findings of lower self-enhancement in the two collectivist cultures tested compared to the individualist culture would support the collectivism assumption. On the other hand, if one collectivist culture more closely resembles an individualist society than it does another collectivist culture on measures of self-enhancement, the hypothesis of specific cultural determinants would be supported.

Study 1 compares self-enhancement among three cultural groups: Singaporean-Chinese, Israeli Druze, and Israeli Jews. The Druze, a unique subgroup of the Arab people, intuitively appear to be a collectivist society. Unfortunately, there is no one consensual ranking of collectivism that can determine the level of collectivism among the Druze. The characteristics of the culture will be analyzed according to the typical characteristics of collectivism. *The Handbook of Cross-Cultural Psychology* presents several independent descriptions of collectivism (Kagitcibasi, 1997). The main characteristics that appeared in most of the descriptions included the following: emphasis on group views, needs, and goals; emphasis on social norms and duty as defined by the group; beliefs shared with the group; readiness to cooperate with the group; seeking security from the group; and high conformity. The Druze society is characterized by the supremacy of their distinctive religion, in which religious leaders provide guidance for everyday life (Dana, 1974). Their lifestyle is still highly traditional. Much importance is attached to the extended family, called the *Hamula* (Ben-Dor, 1979), which is organized around a patriarchal family structure. Explicit regulations regarding relationship between the sexes are customary (Wainryb & Turiel, 1994). It seems that the Druze could be considered as collectivist according to the emphasis they put on the group for support and behavioral guidance. Their traditional lifestyle probably involves a great deal of conformity. The lifestyle of the Druze is very different from that of urban, secular Israeli Jews, who have generally adopted Western ways (Seginer & Halabi, 1991). Jewish influences on the Druze are quite small, because the Druze community is relatively isolated from other communities in Israel. Druze live in their own homogeneous villages, which are preserved by within-community marriages. The assumption that the Druze are collectivist also is supported by a large cross-cultural study of values that encompassed 39 cultures (S. H. Schwartz, 1994). Seven culture-level value dimensions were presented, two of which were found to be highly correlated with individualism-collectivism: conservatism and affective autonomy. On conservatism, Israeli Druze were ranked 1st, Singaporeans were ranked 4th, and Israeli-Jews were ranked 17th (corresponding

means were 4.51, 4.38, and 4.08, respectively). For affective autonomy, Israeli Jews were ranked 13th, Israeli Druze 29th, and Singaporeans 35th (corresponding means were 3.62, 3.16, and 3.04, respectively). It could be concluded, then, that Singaporeans and Druze are both collectivist, with emphasis on different dimensions of collectivism. Israeli Jews showed the most individualist pattern of the three groups.

There is reason to believe that the two collectivist groups differ in some of their endemic values. Such values could include modesty, or the tendency to underpresent one's favorable traits and abilities in public (Cialdini, Wosinka, Dabul, Whestone-Dion, & Heszen, 1998). Modesty is considered one of the primary values in East Asian cultures, together with obedience, self-restraint, and humility (Cheng, 1946; Chiu & Yang, 1987; Hsu, 1949; Hwang, 1982; Tseng, 1973). It was shown empirically in a Japanese sample that the norm of modest self-presentation was internalized between Grades 2 and 5 (Yoshida, Kojo, & Kaku, 1982). The same trend toward lower evaluations of self-enhancers was found by Bond, Leung, and Wan (1982) using Chinese respondents in Hong Kong.

A reasonable assumption is that modesty is an inherent part of the interdependent self-system. If this is the case, it should be a primary value in Druze society as well. The relevant literature is not very definitive about the importance of modesty in this culture. Bravery, truthfulness, and mutual aid are described as the leading values in the Druze community. Other important values include hospitality, respect for the religious teacher, and self-discipline (Abu-Izzedin, 1984). Whereas modesty is not described as a leading value, it is considered important. A more general view of the behavioral code of the Druze can be found in anthropological works describing the mentality of the Orient. In *The Arab Mind*, Patai (1976) describes exaggeration and overassertion as integral to Arab culture. According to Patai, exaggeration is built into the Arabic language, which is spoken by the Druze. Patai provides several examples of Arabic expressions that are much stronger than the equivalent English ones. For example, the Arabic equivalent of "We miss you" is "*Awhashtena*," which means "You made us desolate." This tendency to exaggerate may have implications on self-presentation as well, so that required modesty will have less influence on self-reports.

A different way to investigate the lower self-enhancement tendencies found in some collectivist cultures is to use the important distinction between egoistic and moralistic biases. The egoistic bias has a narcissistic quality and results from agency endorsement. The moralistic bias results from endorsement by communal values and has the quality of "saint-like" attributions (Paulhus & John, 1998). Previous cross-cultural studies did not dis-

tinguish between these two biases and mainly considered the egoistic bias. It may be that the well-documented effect of low egoistic bias in East Asian cultures will not be replicated with the moralistic bias. Two different factors could contribute to a possible high moralistic bias among East-Asian cultures. One is a strong endorsement of communal values, which make self-enhancement of communal traits profitable, and the other is the low sensitivity of the moralistic bias to modesty requirements. By definition, self-enhancement of agentic traits violates modesty norms because such traits deal with personal success and self-promotion and reflect how capable an individual is. Communal traits, on the other hand, can be seen as reflecting the individual's willingness to be a contributing member of the group. Hence, it is expected that the cultures will differ in their self-enhancement of agentic traits (Singaporean will reveal lower self-enhancement than the other groups) but not in self-enhancement of communal traits.

Self-enhancement measures can be divided into two types. One is subjective and is based on a comparison between the way people rate themselves and the way they rate others (e.g., Brown, 1986); the other is based on a discrepancy between self-ratings and valid external criteria (e.g., John & Robins, 1994). It was suggested that the two types of measures can yield different results (Paulhus, 1998). Accordingly, this study will employ two different measures, one of each kind. The first is based on a discrepancy between academic self-evaluations and academic achievement. This content domain was selected for two reasons: It is an important facet of self-evaluation (Marsh & Shavelson, 1985) and valid objective evaluation concerning a participant's success is available in the form of school grades. Academic self-enhancement, which is considered in the present study as an agentic characteristic, was compared among Singaporean Chinese, Israeli Druze, and Israeli Jews. The first hypothesis of the study is that Singaporeans would reveal lower academic self-enhancement than the other two cultural groups.

A second self-enhancement measure is based on the above-average effect. Participants rate whether they are above or below the average. The rationale for using this well-documented effect (Brown, 1998) is that in any random and sufficiently large sample, approximately half of the participants should be below average, whereas the other half should be above average on any given trait. A self-rating of above average from more than the expected 50% indicates that there is self-enhancement in the group. Moreover, as deviation from the expected 50% grows larger, self-enhancement within the group increases. The above-average effect will be measured separately for agentic and for communal traits. Intelligence, sociability, and health would be considered

TABLE 1: Means and Standard Deviations of Academic Self-Enhancement and Self-Enhancement in Agentic and Communal Traits in Three Cultures

	<i>Singaporeans (S)</i>		<i>Israeli Druze (D)</i>		<i>Israeli Jews (J)</i>		<i>Significant Differences^a</i>
	M	SD	M	SD	M	SD	
Study 1							
Reported grades	.62	.14	.61	.15	.65	.24	—
Academic self-evaluation	.63	.13	.76	.14	.70	.14	S < J < D
Academic self-enhancement	−.13	.07	.14	.06	.14	.00	S < J < D
Self-enhancement of agentic traits	.28	.79	.16	.91	.18	.88	S < J, D
Self-enhancement of communal traits	.26	.86	.26	.86	.25	.87	—
Study 2							
Reported grades	.67	.13	.64	.19	.65	.14	—
Academic self-evaluation	.59	.17	.76	.16	.74	.16	S < D, J
Academic self-enhancement	−.17	.11	.15	.06	.15	.04	S < D, J
Self-enhancement of agentic traits	.35	.63	.26	.80	.25	.86	S < D, J
Self-enhancement of communal traits	.29	.85	.27	.84	.29	.80	—

NOTE: For Study 1, participants included 143 Singaporeans, 146 Israeli Druze, and 129 Israeli Jews; for Study 2, participants included 112 Singaporeans, 102 Israeli Druze, and 143 Israeli Jews.

a. Alpha level was set at .05 for two-tailed tests.

agentic traits, whereas honesty, cooperation, and generosity would be considered communal traits. Along with the rationale presented above, the hypothesis of this study is that Singaporeans will reveal lower self-enhancement than Jews and Druze on agentic traits and that the groups will not differ in their self-enhancement of communal traits.

Method

PARTICIPANTS AND PROCEDURE

The sample consisted of 418 students who had just been accepted to institutions of higher education (university or college). The Singaporean group comprises 143 participants (40 men and 103 women), all of them of Chinese origin. The second group comprises 146 Israeli Druze students (59 men and 87 women), all currently living in Druze villages in the northern part of Israel. The Israeli Jewish group comprises 129 participants (40 men and 89 women), all of whom had been accepted to the Academic College of Emek Yezreel. Only students of Western origin living in urban locales were included in the study. All groups were surveyed just prior to or during the first 2 weeks of the school year, and all answered the questionnaire voluntarily.

QUESTIONNAIRE PACKAGE

Academic self-evaluation. Participants were asked to estimate whether they are generally good students on a 7-point scale ranging from 1 (*I'm a very poor student*) to 7 (*I'm a very good student*).

Reported grades. In Singapore, participants reported their A-level grades; in Israel, they reported grades on their matriculation exams. In both countries, these grades are used as criteria for admission to institutions of

higher education. Previous data show that the correlation between reported and actual grades is very high, and sometimes the two are almost identical (Kurman, 2000; Kurman & Sriram, 1997). Reported grades will therefore be considered as a valid criterion for actual academic achievements.

Above-average effect. Participants were asked to indicate whether they consider themselves below average or above average in a population of the same age and gender on six traits (intelligence, health, and sociability as agentic traits and cooperation, honesty, and generosity as communal traits). A possible middle category, average, was omitted to force participants to choose between the extreme categories.

Grades were reported in the last part of the questionnaire, along with demographic details. The questionnaire included other scales that were not used in this study and served as fillers. Singaporeans answered the questionnaire in English, whereas Jews and Druze answered it in Hebrew. The different versions were back-translated.

Results and Discussion

Comparability of the cultural groups. The three samples were compared on their reported grades, which were standardized by range (see below). Results showed no significant differences between the standardized-by-range grades of the different cultural groups, $F(2, 417) = 1.14$, *ns*, which means that they are comparable on their relative academic standing within their own grading systems. The corresponding means and standard deviations are presented in Table 1. The ratio of men to women was different in each of the three cultural groups. The data were therefore analyzed for gender.

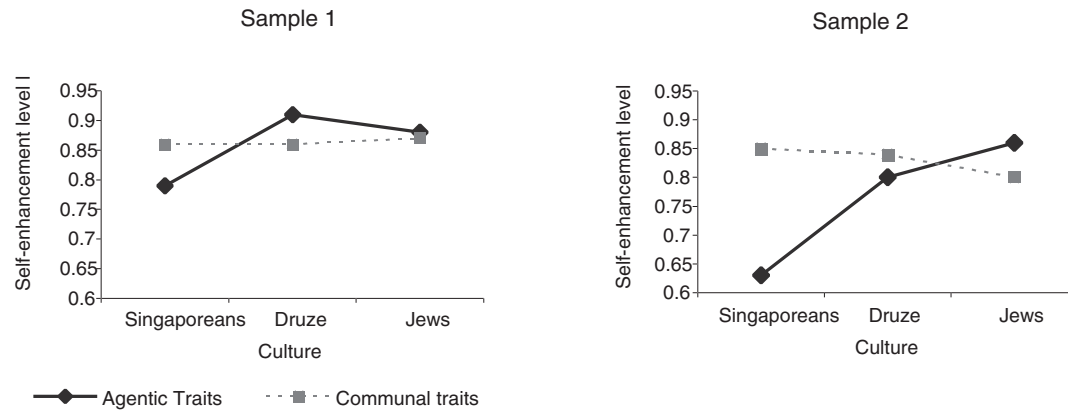


Figure 1 Self-enhancement level in two types of traits in three cultures.

Academic self-enhancement. The first hypothesis was that academic self-enhancement would be less prominent in the Singaporean group than in the Israeli groups (Jewish and Druze). To test this hypothesis, academic self-evaluations were predicted by reported grades, and the residual scores served as the academic self-enhancement measure. A positive residual implies that an individual scored higher on the self-evaluation measure than would be predicted from reported grades (and vice versa in the case of negative residuals). By definition, the correlation between residuals and reported grades is zero. Grades were standardized by range (Milligan & Cooper, 1988) within each culture.¹ An analysis of variance (ANOVA) was conducted to test culture and gender effects. No significant effects for gender or gender by culture were found. A significant culture effect was found, $F(2, 417) = 24.04$, $p < .0001$. Comparisons of the least square means show that all three groups differed significantly. As hypothesized, Singaporeans showed lower academic self-enhancement than did Israelis. Surprisingly, the Druze showed higher academic self-enhancement than Israelis. It may be that this result is somewhat biased because the Druze student sample is more selective than the other student samples. Acceptance to university is less frequent among the Druze than among the other groups. It may be that the relatively high academic self-evaluation exhibited by the Druze students resulted from a downward social comparison, because they compared their academic ability with the whole Druze group and not with the whole student group. Table 1 shows the corresponding means and standard deviations. Note that academic self-evaluation ratings also were standardized by range to facilitate an intuitive impression from the differences between standardized grades and academic self-evaluations.

Above-average effect. The percentage of participants who rated themselves as above average exceeds 50% on

all six traits, in all cultures, and ranged between 74.1% to 94.4%, with an average of 85.68%. Separate self-enhancement scores for communal and agentic traits were computed. As a first step, a factor analysis (with a varimax rotation) was conducted to verify the theoretical grouping into agentic and communal traits. Intelligence, sociability, and health indeed loaded on the first rotated factor (.45, .74, and .59, respectively), which was defined as the agency factor. Honesty and cooperation had high loadings on the second factor (.77, .67), which was defined as the communion factor. Generosity did not load on any of the factors according to a cutting point of .45. Separate analyses conducted for each culture showed that the above traits had loadings higher than .45 on the corresponding factors in all three groups. The self-enhancement scores on agentic and communal traits were computed by averaging the relevant ratings of each participant (a rating of below average was assigned 0 and a rating of above average was assigned 1). An ANOVA was conducted, with culture and gender as independent variables and type of trait as a within-participant variable. Gender and all interactions with gender were insignificant. No significant culture or type of trait effects were found, which can be explained by the significant interaction between culture and type of trait, $F(2, 411) = 5.26$, $p < .006$. Table 1 shows the relevant means and standard deviations, and Figure 1 presents the interaction graphically. Separate ANOVAs conducted for each type of trait showed significant culture effects for self-enhancement of agentic traits, $F(2, 417) = 10.50$, $p < .0001$, which emerged from lower self-enhancement among Singaporeans. The groups did not differ significantly on self-enhancement of communal traits, $F(2, 417) = 1.34$. Comparisons between self-enhancement of agentic and communal traits within each culture also were conducted and showed significant effects for Singaporeans, $F(1, 140) = 5.06$, $p < .026$,

who enhanced communal more than agentic traits, and for Druze, $F(1, 144) = 4.36$, $p < .038$, who enhanced agentic more than communal traits. This trend among the Druze to enhance agentic traits resembles their tendency toward high academic self-enhancement and could be explained by the selectivity of the sample. Note that the Singaporeans revealed less self-enhancement of agentic traits but resembled the other groups in their self-enhancement of communal traits, as was hypothesized.

Results concerning agentic traits were quite similar despite the use of different kinds of self-enhancement measures. The above-average ratings on agency traits, which is a subjective self-enhancement measure, and the discrepancy between academic self-evaluation (an agentic characteristic) and reported grades, which is an objective self-enhancement measure, yielded the same results, showing that Singaporeans reveal less self-enhancement than do Jews and Druze.

It was clearly shown that self-enhancement of agentic traits was evident among the Israeli Druze. It also was shown that a self-enhancement tendency was revealed in Singapore for communal traits. This last finding could be explained in two ways. One explanation is that communal traits are valued much more than agentic traits in Singapore and were accordingly enhanced more. Agentic traits, on the other hand, are not perceived as important and therefore were not enhanced. Another explanation is that self-presentation norms that may prevail in Singapore restrict the manifestation of self-enhancement where agentic traits are concerned but allow for such self-enhancement in communal traits. To distinguish between the two, the direct relations between modesty norms and self-enhancement for both kinds of traits should be tested.

Study 2 investigates the relations among modesty, self-construals, and self-enhancement. Its main purpose is to investigate the relations between modesty and self-enhancement for both types of traits and to show the unique contribution of modesty to self-enhancement of agentic traits. Other purposes of Study 2 are to verify the assumptions concerning group differences in self-construals and in modesty and to replicate the results of Study 1 with two methodological improvements. The first is the use of less selective samples (high school students), and the second is the use of an Arabic version of the questionnaire for the Druze group to eliminate a Western tendency that the Hebrew language may elicit.

STUDY 2: PREDICTION OF SELF-ENHANCEMENT

This study predicts differences in self-enhancement at the individual level by interdependent and independent self-construals and by modesty. The interdependent self-construal reflects an "emphasis on connectedness

and relations often found in non-Western culture," whereas the independent self-construal reflects "the separateness and uniqueness of the individual" (Singelis, 1994, p. 580). Traditionally, these two types of self-construals were considered as two endpoints of the same dimension (Hofstede, 1980). Other theoretical considerations challenge this assumption (Kagitcibasi, 1994) and suggest that the two can be independent of one another. Two different types of self-construals that are independent of each other could have differential relations with self-enhancement. It could be reasoned that the independent self-construal should be related to self-enhancement of agentic traits more than to self-enhancement of communal traits. Uniqueness could be created more easily with enhanced agentic traits, which reflect the capability and competence of the person, than with communal traits. The interdependent self-construal, on the other hand, should be more related to communal traits due to its emphasis on concern for the group. It is therefore hypothesized that the independent self-construal will be more related to self-enhancement of agentic traits than to self-enhancement of communal traits, whereas the opposite is expected for the interdependent self-construal.

The second hypothesis of the study is that modesty would make a unique contribution to the self-enhancement prediction in addition to the contribution of independent and interdependent self-construals. This contribution is expected to appear mainly in predicting self-enhancement of agentic traits. Indeed, it was found that public conditions affected self-descriptions of Japanese on self-profitable traits, which were described less positively in public than in private. Other-profitable traits, on the other hand, were rated more positively in public than in private (Harihara, Yamaguchi, & Niiya, 2000). Accordingly, the hypothesis of the present study is that modesty will make a unique contribution to academic self-enhancement and enhancement of agentic traits as measured by the above-average effect. No such contribution is expected for self-enhancement of communal traits.

Because this study includes explicit measures for the two types of self-construals and for modesty, the basic assumptions of Study 1 will be tested empirically. Jews are expected to be less collectivist than Singaporeans and Druze, and Singaporeans are expected to be more modest than Druze and Jews.

Method

PARTICIPANTS AND PROCEDURE

The sample comprises 362 high school participants in the 11th or 12th grade: 115 Singaporean Chinese (44 men and 71 women), 103 Israeli Druze (43 men and 60 women), and 144 Israeli Jews (48 men and 96 women).

Singaporeans answered the questionnaire in English, Druze filled it out in Arabic, and Israeli Jews answered a Hebrew version. The different versions were identical and were back-translated. The questionnaires were group administered in school classrooms.

QUESTIONNAIRE PACKAGE

All the questionnaires from Study 1 were administered. A factor analysis (with varimax rotation), which was conducted for the above-average traits, yielded solutions equivalent to that of Study 1. Intelligence, sociability, and health loaded on a first factor (.68, .71, and .70, respectively) and honesty and cooperation loaded on a second factor (.74 and .69, respectively). Separate analyses conducted for each culture showed that the above traits loaded on the corresponding factors in all three groups.

Two additional scales were employed: interdependent and independent self-construals scales (Singelis, 1994). The interdependent self-construal scale measures the emphasis on connectedness and relationship (e.g., "It is important for me to maintain harmony within my group"). It also includes items that describe the need to give up self-interest in favor of the group ("I will sacrifice my self-interest for the benefit of the group I am in"). The independent self-construal scale measures the separateness and uniqueness of the individual (e.g., "I enjoy being unique and different from others in many respects"). The original two scales consist of 12 items each. Alpha coefficients were computed for each culture separately. Items that showed low correlations with the total score in any of the cultures were omitted. The final independence scale consisted of 11 items, whereas the final interdependence scale consisted of only 5 items. The massive omission of items for the interdependent self-construal is a result of a very low original alpha among Israeli Jews. The excluded items mostly described the predominance of group needs over self-needs. Cronbach's alphas for the new scales were .56, .56, and .57 for the interdependent scale (which are acceptable, considering the small number of items) and .60, .61, and .64 for the independent scale for Singaporean, Druze, and Jewish groups, respectively.

Modest Responding Scale (MRS). The MRS (Whetstone, Okun, & Cialdini, 1992) consists of 21 items that tap the inclination toward modesty (e.g., "Telling people about my strengths and successes has always been an embarrassing thing for me") and the perceived social desirability of modest responses (e.g., "Bragging on oneself in a group is always socially inappropriate"). It also includes reverse-scored items reflecting the propensity to brag (e.g., "If I've done something well, I like to tell others about it"). Items that had low correlation with the total score of each culture were omitted. Cronbach's alphas of

the final 14-item scale were .83, .76, and .84 for Singaporean, Druze, and Jewish groups, respectively.

Items from the different scales (all of Likert type) were given in a mixed order, all with a 7-point scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*).

Results and Discussion²

Comparability of the cultural groups. The three high school groups were compared on their reported grades, which were standardized by range. Results showed no significant differences between the standardized-by-range grades, $F(2, 355) = 1.33$, *ns*, which means that the samples are comparable on their relative academic standing within their own grading systems. The corresponding means and standard deviations are presented in Table 1. The ratio of men to women was different in each of the three cultures. The data were therefore analyzed for gender.

Self-enhancement level. Results concerning the self-enhancement measures from Study 1 were generally replicated in Study 2. An ANOVA was conducted, with culture and gender as independent variables and academic self-enhancement as a dependent variable. Table 1 shows the relevant means and standard deviations. A strong significant culture effect was found for academic self-enhancement, $F(2, 356) = 37.63$, $p < .0001$. Post hoc comparisons between the least square means showed that Singaporeans scored significantly lower than the other two groups, which did not differ statistically from each other. The finding in Study 1 of higher academic self-enhancement among the Druze was not replicated in the less selective sample of Study 2. No significant effects were found for gender or for the interaction between culture and gender.

All cultures revealed self-enhancement tendencies on the above-average effect. The percentages of above-average responses were somewhat lower in Study 2 compared to Study 1 and ranged between 58.3% and 91.3%, with an average of 79.90%. Self-enhancement was computed for agency and communal traits. An ANOVA with culture and gender as independent variables and type of trait as a within-participant variable was conducted. A significant effect was found for type of trait, $F(1, 356) = 10.45$, $p < .0001$, which showed that self-enhancement was higher for communal than for agentic traits. The interaction between type of trait and culture was also significant, $F(2, 356) = 14.61$, $p < .0001$, and is presented in Figure 1. Separate analyses conducted for the two types of traits showed that the cultural groups differed in their self-enhancement of agentic traits, $F(2, 356) = 16.75$, $p < .0001$, resulting from the lower self-enhancement among Singaporeans. The groups did not differ in self-enhancement of communal traits, $F(2, 356) = .61$. Comparisons between self-enhancement of agentic and com-

TABLE 2: Means and Standard Deviations of Modesty and Interdependent and Independent Self-Construals in Three Cultures

	<i>Singaporeans (S)</i>		<i>Israeli Druze (D)</i>		<i>Israeli Jews (J)</i>		Significant Differences ^a
	M	SD	M	SD	M	SD	
Modesty	4.88	.94	4.52	.98	4.32	1.03	S > D, J
Interdependent self-construal	5.55	.78	6.01	.77	5.40	0.96	S, J ^b > D
Independent self-construal	4.83	.75	5.53	.79	5.53	0.75	S < D, J

a. Alpha level was set at .05 for two-tailed tests.

b. Singaporeans were marginally higher than Jews ($p < .07$).

munal traits within each culture also were conducted and showed a significant strong effect for the Singaporeans, $F(1, 113) = 30.59$, $p < .0001$, who enhanced agentic traits much less than communal ones. A marginal effect was found for Jews, who tended to enhance agentic more than communal traits, $F(1, 142) = 3.53$, $p < .062$. The Druze did not differ in the level they enhanced both kinds of traits. The higher self-enhancement for agentic traits found for this sample in Study 1 was not replicated. No significant effects were found for gender or for the interaction between culture and gender.

As was suggested before, the Druze sample in Study 1 may have been nonrepresentative because it was very selective in terms of academic achievement. It may be that the "ego boost" of acceptance to the university was generalized to other agentic characteristics. Another possibility is that those Druze (especially women) who choose to study in the university are more Westernized in the first place. The results point out that the sometimes problematic generalization of findings obtained with university students is even more problematic when minority groups are concerned. The results of Study 2, which were obtained with a quite representative sample, are a more valid characterization of typical Druze responses.

Interdependent and independent self-construals. Table 2 presents cultural differences for interdependent and independent self-construals. Both had a significant culture effect for the interdependent self-construal, $F(2, 356) = 18.63$, $p < .0001$, and for the independent self-construal, $F(2, 356) = 29.37$, $p < .0001$. Post hoc comparisons of the least square means show that Singaporeans score lower than Druze and Jews for independent self-construal and that Druze score significantly higher than the other two groups for the interdependent self-construal. It seems that Druze emphasize high interdependence, whereas Singaporeans stress low independence of the self-system.

Druze and Jews did not differ on the interdependent self-construal. This trend fits a previous study that showed that Israeli Arab and Jewish students did not dif-

fer on their individualism level (Oyserman, 1993). It also was found that Singaporeans and Jews differed only marginally ($p < .07$) on the interdependent self-construal. It is assumed that this relative resemblance does not reflect a low level of dependency among Singaporeans. It may reflect the fact that connectedness and relationships are valued more among Israeli Jews than they are among other individualist cultures. Indeed, connectedness and loyalty to various peer groups characterize Sabras, native-born Israeli Jews (Almog, 1994).

Modesty. It was assumed that Singaporeans would be characterized by modesty. An ANOVA analysis showed that this was indeed the case: A significant cultural effect, $F(2, 365) = 6.32$, $p < .002$, emerged. Comparisons between the least squares means revealed that Singaporeans exhibited more modesty than did Jews and Druze, who did not differ from each other. The corresponding means and standard deviations are presented in Table 2.

The group comparisons supported the hypothesis that the Druze are collectivist and yet exhibit less modesty and more self-enhancement than do Singaporeans. It could be claimed that these differences are the result of the exposure of the Druze in Israel to the generally Western Jewish Israeli society. This may be partly true for Druze university students but it is not likely for the high school sample. Druze high school students live mostly within their villages, study in Arabic, and maintain a very traditional way of life. Singaporeans, on the other hand, may be highly exposed to Western values. English is an official language in their city-state, many American programs are broadcast on television, and large communities of foreigners live on the island, many of them from Western countries (Australia, New Zealand, and the United States). Thus, the differences between the Druze and the Singaporeans could not be attributed to stronger Western influences on the Druze. The differences found between the groups reflect genuine differences between the two cultures: Druze are more interdependent but also more independent than Singaporeans.

TABLE 3: Correlations Among Interdependent and Independent Self-Construal, Modesty, and Three Self-Enhancement Measures

	1	2	3	4	5
1. Interdependent self-construal					
2. Independent self-construal	.12*				
3. Modesty Response Scale	.13*	-.14*			
4. Academic self-enhancement	.08	.30 [†]	-.30 [†]		
5. Self-enhancement of agentic traits	.02	.17****	-.26 [†]	.32 [†]	
6. Self-enhancement of communal traits	.17****	.05	-.02	.12*	.16****

* $p < .05$. **** $p < .001$. [†] $p < .0001$, for two-tailed tests.

Self-enhancement prediction. To investigate the contribution of modesty to self-enhancement prediction, forced hierarchical regression analyses were computed. First, self-enhancement measures (of academic achievement, agentic traits, and communal traits) were predicted by interdependent and independent self-construals. Modesty was added in a second step. Table 3 shows the correlations between the variables and Table 4 shows the standardized regression coefficients of each step of the regressions as well as the proportion of self-enhancement variance explained in each step.

The regression coefficients of modesty and independent self-construal were significant for academic self-enhancement and for self-enhancement of agentic traits obtained by the above-average ratings. Modesty added 7.3% to the explained variance of academic self-enhancement and 5.1% to the explained variance of self-enhancement of agentic traits, which is quite impressive compared to the proportion of variance explained by self-construals (9.2% and 3.0%, respectively). It seems, then, that the role of modesty in self-enhancement of agentic traits cannot be overlooked.

The fact that individual self-enhancement was assessed by the above-average ratings needs some justification. The above-average effect is clearly a valid self-enhancement measure for comparison between groups. It is problematic for individual differences measurements because above-average ratings are not very extreme and could be too crude to expose meaningful self-enhancement variance. Empirical support for the validity of the above-average ratings as individual self-enhancement measures in the context of the present study lies in the findings that (a) group differences supported the conclusion obtained by the regression analysis and (b) the relations between self-enhancement of agentic traits (as measured by the above-average effect)

TABLE 4: Self-Enhancement Measures Predicted by Interdependent Self-Construal, Independent Self-Construal, and Modesty

	Academic Self-Enhancement	Self-Enhancement in Agentic Traits	Self-Enhancement in Communal Traits
Step 1			
Interdependent self-construal	.04	.00	.17**
Independent self-construal	.30 [†]	.19***	.03
R^2	.10	.03	.03
Step 2			
Interdependent self-construal	.08	.04	.18***
Independent self-construal	.25 [†]	.13***	.04
Modesty Response Scale	-.28 [†]	-.25 [†]	-.04
R^2	.17	.08	.03

** $p < .01$. *** $p < .005$. [†] $p < .0001$, for two-tailed tests.

and other variables is very similar to that obtained by academic self-enhancement. It may be that the use of more than one trait for each trait type (agency and communion) contributed to the validity of the measures. Similar patterns of results were found for the two types of self-enhancement measures of agentic traits (as in Study 1).

GENERAL DISCUSSION

Self-construals and self-enhancement. The relations between self-construals and self-enhancement were found to be very different for agentic and communal traits. Unlike the egoistic bias, the moralistic bias was not related either to the independent self-construal or to modesty. It was affected only by the interdependent self-construal, which did not affect self-enhancement of agentic traits. The moralistic bias was used quite extensively and consistently by all three cultures in the two samples, without significant culture differences. This consistency may imply that the bias helps in maintaining positive self-regard. The relations between the interdependent self and the moralistic bias show that the more that connectedness is valued, the stronger is this bias. Thus, self-enhancement of communal traits reveals the importance assigned to those traits, along with the basic enhancement tendency. Yet, the effect of interdependent self-construal on the moralistic bias was not reflected by cultural differences, and the Druze interdependent group did not reveal a stronger moralistic bias. The relations between interdependent self-construal and the moralistic bias may need further establishment. The lack of a relationship between the independent self-construal and the moralistic bias was supported by group

differences as well as by correlational analyses, as was the lack of relations between modesty and the moralistic bias. The possibility that this bias is a better measure of self-enhancement motives should be tested in the future.

Self-enhancement of agentic traits was positively related to the independent but not to the interdependent self-construal. The data suggest that low independence rather than high interdependence of the self is associated with low self-enhancement of agentic traits. The obtained group differences support this conclusion. The Druze, who revealed both higher interdependent and higher independent self-construal than Singaporeans, also revealed higher self-enhancement in both samples. These findings are in accordance with previous studies, which found that the independent self-construal was more related to self-esteem than was the interdependent self-construal (Heine et al., 1999).

That the interdependent self was found to be unrelated to self-enhancement of agency traits calls for a better understanding of this form of self-construal. Some relevant distinctions were made earlier. A relatedness component, which reflects care for the ingroup and a harmonious way of life, was found to be independent of the dichotomy between self-dominance and group dominance (Kashima et al., 1995). A quite similar distinction emerged in another study, which described a dimension of individualism versus authoritarianism as independent of active collectivism versus withdrawal from the group dimension (Gelfand, Triandis, & Chan, 1996). It may be that strong involvement with the group, which is reflected by interdependent self-construal, can be independent of unique self-identity and emphasis on self-development, which is reflected by independent self-construal. It is quite acceptable that a person could have both a highly developed sense of individual identity and high involvement with the group. It may be that the traditional continuum of individualism-collectivism actually reflects a continuum of individuality, whereas low individuality indicates low emphasis on self-needs. The other dimension (which was called collectivism, relatedness, or interdependence) forms another construct. Such a model can explain some of the existing findings but it still presents some problems. It is difficult to accept the idea that the level of involvement with the group is independent of the necessity to give priority to group needs over self-needs. More studies are needed to test the existence of these separate constructs, consider their explanatory merit, and perhaps further refine them. Another needed line of research should be a search for endemic values, which may contribute to the understanding of cross-cultural differences traditionally explained by self-construals. In the case of self-enhancement, modesty should not be overlooked.

Modesty and self-enhancement. The important role of modesty in self-enhancement of agentic traits was supported by direct relations between modesty and self-enhancement level and by group differences in those traits. It should be noted that modesty was found to be as important to self-enhancement predictions as were the self-construals. These strong relations could be interpreted as a reflection of impression management or of self-deception. It may be that impression management affected participants' responses and they followed the norms to satisfy external observers. Another option is that modesty requirements are internalized and become a dominant belief of the person. Such beliefs or social axioms (Lueng et al., 2000) can generate motivational pressure for their satisfaction and thus affect behavior (Bond, Zagarek, & Spencer-Oatey, 2000). Thus, no external observer is needed to elicit a modest, self-deceptive response. Heine et al. (1999) described findings that show that public conditions do not affect self-presentation of Asians. Such findings could imply that modest self-presentation is a result of personal beliefs concerning self-presentation and not of impression management.³

An important question is whether modesty requirements are endemic to certain cultures or whether, perhaps, they are part of a more general cultural pattern. It is possible that when the development of a unique self is restricted, there is an increase in the demand for modesty. It was suggested that such restrictions could be emphasized in vertical societies (Kurman & Sriram, in press), which try to maintain a hierarchical social order (Triandis, 1995). Various hierarchical cultures with different endemic values should be tested to verify the above suggestion.

Self-enhancement motive in collectivist cultures. The existence or absence of a motive is difficult to study and is usually analyzed indirectly. The current study revealed such indirect findings that support the existence of the self-enhancement motive in the East Asian culture of Singapore. The first supporting finding is the central role that modesty plays in self-enhancement of agentic traits. A possible explanation for this relation is that when requirements for modesty are less internalized and modesty is lower, the self-enhancement motive can function more freely and higher self-enhancement is registered. The second supporting finding is that the moral bias, as was measured by self-enhancement of communal traits, was not lower in Singapore than in the other two groups. As modesty norms were found to be unrelated to self-enhancement of communal traits, a possible conclusion is that an existing self-enhancement motive in Singapore was expressed in such traits that do not contradict the modesty norms.

A different perspective on the question of the existence of a self-enhancement motive in the interdependent self-system is the theoretical explanation for its existence. The interdependent self-system suggested by Markus and Kitayama (1991) implies that the individual's core self is much less developed than those parts of the self defined by others. A recent review (Heine et al., 1999) claimed that the need for positive self-regard may not be shared by Japanese. They supported this claim by many findings and by qualitative analysis of the Japanese mind. Undoubtedly, they present a very illuminating review, which is very convincing in its conclusion that Japanese have low individuality. This study agrees that East Asian cultures have low individuality. The discrepancy between the conclusions of the present study and those of Heine et al.'s lies in the answer to the question of what constitutes the main source of the observed low individuality. It may be a result of social restrictions on self-development, as was suggested in this study and elsewhere (Yamaguchi, 1994), or it could be a result of the fact that self-esteem has sources other than personal success that diminish the effect of personal success on self-esteem. The findings of the present study support the first possibility, which needs further empirical support. Such empirical effort should include a wide spectrum of cultures, which vary independently on the degree of individualism typically exhibited by their members and in the strength of their modesty requirements.

NOTES

1. One of the hypotheses dealt with group differences in self-enhancement. Because Singapore and Israel employ different grading systems, a standardization was required. The more common standardization of Z scores is inappropriate for this purpose because it omits mean differences and defines the mean of each group as zero. A standardization for range was therefore conducted. This standardization transforms all scores to the range [0, 1] by the formula $\frac{x - \text{Minimum}}{\text{Maximum} - \text{Minimum}}$ and allows a comparison between grades from different grading systems. The academic self-enhancement score was the residual score of academic self-evaluations regressed on standardized-by-range grades.

2. Missing data were treated by a pairwise method that resulted in somewhat different numbers of participants in different analyses.

3. It should be noted that some findings that were previously cited (Harihara, Yamaguchi, & Niiya, 2000) did reveal an effect of publicity on self-ratings of Japanese. More studies should try to reconcile these inconsistencies.

REFERENCES

- Abu-Izzedin, N. M. (1984). *The Druze*. Leiden, the Netherlands: E. J. Brill.
- Almog, O. (1994). "Hazabar": *Diokan Sociology* [The "Sabras": A sociological profile]. Unpublished doctoral thesis, University of Haifa.
- Ben-Dor, G. (1979). *The Druzes in Israel: A political study*. Jerusalem: Magnes.
- Bond, M. H., & Cheung, T. S. (1983). The spontaneous self-concept of college students in Hong Kong, Japan, and the United States. *Journal of Cross-Cultural Psychology*, 14, 153-171.
- Bond, M. H., Leung, K., & Wan, K.-C. (1982). The social impact of self-effacing attributions: The Chinese case. *Journal of Social Psychology*, 118, 157-166.
- Bond, M. H., Zagarek, V., & Spencer-Oatey, H. (2000). Culture as an explanatory variable: Problems and possibilities. In H. Spencer-Oatey (Ed.), *Culturally speaking: Managing relations in talk across cultures* (pp. 62-73). London: Cassell.
- Brown, J. D. (1986). Evaluations of self and others: Self-enhancement biases in social judgments. *Social Cognition*, 4, 353-376.
- Brown, J. D. (1998). *The self*. New York: McGraw-Hill.
- Cheng, C. K. (1946). Characteristic traits of the Chinese people. *Social Forces*, 25, 146-155.
- Cialdini, R. B., Wosinska, W., Dabul, A. J., Whetstone-Dion, R., & Heszen, I. (1998). When role salience leads to social role rejection: Modest self-presentation among women and men in two cultures. *Personality and Social Psychology Bulletin*, 24, 473-481.
- Chiu, C. Y., & Yang, C. F. (1987). Chinese subjects' dilemmas: Humility and cognitive laziness as problems in using rating scales. *Bulletin of the Hong Kong Psychological Society*, 18, 39-50.
- Dana, N. (1974). *Hadrusim* [The Druze]. Jerusalem: Misrad Hadatot.
- Gelfand, M. J., Triandis, H. C., & Chan, D. K.-S. (1996). Individualism versus collectivism or versus authoritarianism? *European Journal of Social Psychology*, 26, 397-410.
- Harihara, M., Yamaguchi, S., & Niiya, Y. (2000). *Japanese self-effacement: Low self-regard of self-presentation?* Poster presented at the 15th international congress of the International Association for Cross-Cultural Psychology, Pultusk, Poland.
- Heine, S. J., & Lehman, D. R. (1995). Cultural variation in unrealistic optimism: Does the West feel more invulnerable than the East? *Journal of Personality and Social Psychology*, 68, 595-607.
- Heine, S. J., Lehman, D. R., Markus, H. R., & Kitayama, S. (1999). Is there a universal need for positive self-regard? *Psychological Review*, 106(4), 766-794.
- Hofstede, G. (1980). *Culture's consequences: International differences in work-related values*. London: Sage.
- Hsu, F. L. K. (1949). Suppression versus repression: A limited psychological interpretation of four cultures. *Psychiatry*, 12, 223-242.
- Hwang, C. H. (1982). Studies in Chinese personality: A critical review. *Bulletin of Educational Psychology*, 15, 227-240.
- John, O. P., & Robins, R. W. (1994). Accuracy and bias in self-perception: Individual differences in self-enhancement and the role of narcissism. *Journal of Personality and Social Psychology*, 66, 206-219.
- Kagitcibasi, C. (1994). A critical appraisal of individualism and collectivism: Towards a new formulation. In U. Kim, H. Triandis, C. Kagitcibasi, S. Choi, & G. Yoon (Eds.), *Individualism and collectivism* (pp. 52-65). London: Sage.
- Kagitcibasi, C. (1997). Individualism and collectivism. In J. W. Berry, M. H. Segall, & C. Kagitcibasi (Eds.), *Handbook of cross-cultural psychology* (Vol. 3, pp. 1-50). Boston: Allyn & Bacon.
- Kashima, Y., Yamaguchi, S., Kim, U., Choi, S. C., Gelfand, M. J., & Yuki, M. (1995). Culture, gender and self: A perspective from individualism-collectivism research. *Journal of Personality and Social Psychology*, 69, 925-937.
- Kitayama, S., Markus, H. R., Matsumoto, H., & Norasakkunkit, V. (1997). Individual and collective processes in the construction of the self: Self-enhancement in the United States and self-criticism in Japan. *Journal of Personality and Social Psychology*, 72, 1245-1267.
- Kitayama, S., Takagi, H., & Matsumoto, H. (1995). Seiko to shippai no kiin: Nihonteki jiko no bunka-shinrigaku [Causal attribution of success and failure: Cultural psychology of Japanese selves]. *Japanese Psychological Review*, 38, 247-280.
- Kunda, Z. (1987). Motivated inference: Self-serving generation and evaluation of causal theories. *Journal of Personality and Social Psychology*, 53, 636-647.
- Kurman, J. (2000). *The relations among vertical and horizontal collectivism, modesty, and self-enhancement*. Manuscript submitted for publication.
- Kurman, J., & Sriram, N. (in press). Interrelationships among vertical and horizontal collectivism, modesty, and self-enhancement. *Journal of Cross-Cultural Psychology*.
- Leung, K., Bond, M. H., Reimel de Carrasquel, S., Mudoz, C., Hernandez, M., Murakami, F., Yamaguchi, S., Bierbrauer, G., & Singelis, T. M. (2000). *Social axioms: The search for universal dimensions of general beliefs about how the world functions*. Manuscript submitted for publication.

- Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, 98, 224-253.
- Marsh, H. W., & Shavelson, R. (1985). Self-concept: Its multifaceted, hierarchical structure. *Educational Psychologist*, 20, 107-123.
- Miller, D. T., & Ross, M. (1975). Self-serving biases in the attribution of causality: Fact or fiction? *Psychological Bulletin*, 82, 213-225.
- Milligan, G. W., & Cooper, M. C. (1988). A study of standardization of variables in cluster analysis. *Journal of Classification*, 5, 181-204.
- Oyserman, D. (1993). The lens of personhood: Viewing the self and others in a multicultural society. *Journal of Personality and Social Psychology*, 65, 993-1009.
- Patai, R. (1976). *The Arab mind*. New York: Charles Scribner's Sons.
- Paulhus, D. L. (1998). Interpersonal and intrapsychic adaptiveness of trait self-enhancement: A mixed blessing? *Journal of Personality and Social Psychology*, 74, 1197-1208.
- Paulhus, D. L., & John, O. P. (1998). Egoistic and moralistic biases in self-perception: The interplay of self-deceptive styles with basic traits and motives. *Journal of Personality*, 66, 1025-1060.
- Schwartz, J. M., & Smith, W. P. (1976). Social comparison and the difference of ability. *Journal of Personality and Social Psychology*, 34, 1268-1275.
- Schwartz, S. H. (1994). Beyond individualism/collectivism, new cultural dimensions of values. In U. Kim, H. Triandis, C. Kagitcibasi, S. Choi, & G. Yoon (Eds.), *Individualism and collectivism* (pp. 85-119). London: Sage.
- Seginer, R., & Halabi, H. (1991). Cross-cultural variations of adolescents' future orientation: The case of Israeli Druze versus Israeli Arab and Jewish males. *Journal of Cross-Cultural Psychology*, 22, 224-237.
- Singelis, T. M. (1994). The measurement of independent and interdependent self-construals. *Personality and Social Psychology Bulletin*, 20, 580-591.
- Takata, T. (1987). Self-depreciative tendencies in self-evaluation through social comparison. *Japanese Journal of Experimental Social Psychology*, 27, 27-36.
- Triandis, H. (1995). *Individualism and collectivism*. Boulder, CO: Westview.
- Tseng, W. S. (1973). The concept of personality in Confucian thought. *Psychiatry*, 36, 191-202.
- Wainryb, C., & Turiel, E. (1994). Dominance, subordination and concepts of personal entitlement in cultural contexts. *Child Development*, 65, 1701-1722.
- Whetstone, M. R., Okun, M. A., & Cialdini, R. B. (1992). *The Modest Responding Scale*. Paper presented at the convention of the American Psychological Society, San Diego, California.
- White, W. G., & Chan, E. (1983). A comparison of self-concept scores of Chinese and White graduate students and professionals. *Journal of Non-White Concerns in Personnel and Guidance*, 11, 138-141.
- Wills, T. A. (1981). Downward comparison principles in social psychology. *Psychological Bulletin*, 90, 245-271.
- Yamaguchi, S. (1994). Collectivism among the Japanese: A perspective from the self. In U. Kim, H. Triandis, C. Kagitcibasi, S. Choi, & G. Yoon (Eds.), *Individualism and collectivism* (pp. 175-188). London: Sage.
- Yoshida, T., Kojo, K., & Kaku, H. (1982). A study on the development of self-presentation in children. *Japanese Journal of Educational Psychology*, 30, 30-37.

Received February 6, 2001

Revision accepted February 6, 2001

Reviewers for 2000

Henk Aarts	Ramona Bobocel	Margaret Clark
A. Abele	Galen Bodenhausen	Russ Clark
Dominic Abrams	Laura Bogart	Heather Claypool
Linda Acitelli	Angela Bonacci	Gerald Clore
Virgil Adams	George Bonanno	Jerry Clore
Chris Agnew	Michael Bond	Erik Coats
E. Aidman	Doug Bonett	Dov Cohen
Dolores Albarracin	Steve Booth	Nancy Collins
Sheldon Alexander	Peter Borkenau	Mark Conner
Mark Alicke	Richard Bouchis	Regina Conti
Scott Allison	Valerie Braithwaite	Luke Conway
Naline Ambady	Markus Brauer	Lynne Cooper
Barbara Anderson	Jack Brehm	Don Corassini
Craig Anderson	Kelly Brennan	Diana Cordova
John Archer	G. G. Brensten	Olivier Corneille
Robert Arkin	Robert Bringle	James Coyne
David Armor	Leslie Brody	Susan Crackenberg
Ximena Arriaga	Nadine Bruce	Chris Crandall
Lisa Aspinall	Ross Buck	William Crano
Ruth Ann Atchley	Lee Budesheim	Jennifer Crocker
Martha Augoustinos	Roger Buehler	Benoit Dardenne
Roya Ayman	Heather Bullock	Peter Darke
Assad Azzi	Rod Bund	Nilanjana Dasgupta
Mark Baldwin	Jerry Burger	Joanne Davila
Anita Barbee	Chris Burns	Keith Davis
John Bargh	Brad Bushman	Mark Davis
Robert Baron	Susan Butt	Carsten De Dreu
Manuela Barreto	Bram Buunk	Kay Deaux
Bruce Barry	John Cacioppo	Anita Delongis
Daniel Batson	Donnah Canavan	Debra DeMeis
Roy Baumeister	Dora Capozza	Roberta Deppe
James Beggan	Gustav Carlo	Eric Depret
Denise Beike	Kathy Carnelley	Valerian Derlega
Kathy Bell	Emanuele Castano	Michael Desert
Paul Bell	Don Cervone	Donna Desforges
Veronica Benet	William Chaplin	David DeSteno
Ida Berger	Alex Chapman	Thierry Devos
Diane Berry	Gretchen Chapman	Shari Diamond
Irene Blair	Alison Chasteen	Richard Dienstbier
Hart Blanton	Lelina Chatman	Ap Dijksterhuis
Herbert Bless	Alan Christensen	James Dillard
Jack Block	Robert Cialdini	Ken Dion

- Bertjan Doosje
 Natalie Dove
 Julie Duck
 Muriel Dumont
 David Dunning
 Emmanuelle Dupont
 Ron Duran
 Alice Eagly
 John Edwards
 Rachel Eggins
 J. R. Eisen
 Robert Eisenberger
 Donna Eisenstadt
 Lee Ellis
 Robert Emmons
 Norm Endler
 Robert Enright
 Nicholas Epley
 Ralph Erber
 Amy Eshleman
 Victoria Esses
 Paul Etcheverry
 Julie Exline
 Tom Farsides
 Norm Feather
 Steve Fein
 Robert Feldman
 Klaus Fiedler
 Kelly Fielding
 John Finch
 Catlin Finkenauer
 Ayelet Fishbach
 Julie Fitness
 Monique Fleming
 Robert Folger
 Geoffrey Fong
 Alan Fontana
 Joseph Forgas
 Deborah Frable
 Chris Fraley
 Cindy Franz
 Steve Franzoi
 Pat Frazier
 Arthur Freedman
 Peter Freytag
 Carrie Fried
 Heidi Fritz
 Kathy Fuegan
 Frank Fujita
 Sam Gaertner
 Adam Galinsky
 Steven Gangestad
 Robert Gardner
 Wendi Gardner
 Fredrico Garducci
 John Paolo Garofalo
 Guido Gendolla
 D. George
 Mary Gergen
 William Gerin
 Meg Gerrard
 Frederick Gibbons
 Byran Gibson
 Robert Gifford
 Daniel Gilbert
 Tom Gilovich
 Peter Glick
 George Goethals
 Joel Gold
 Marti Gonzales
 Richard Gonzalez
 Wind Goodfriend
 Ernestine Gordijn
 Anne Gordon
 Sam Gosling
 Richard Gramzow
 William Graziano
 Jeff Greenberg
 Anthony Greenwald
 Dara Greenwood
 Larry Gregory
 Dale Griffin
 Michael Gurtman
 John Haidt
 Judith Hall
 David Hamilton
 M. Hammeral
 Kent Harber
 Curtis Hardin
 Monica Harris
 M. J. Harris
 John Harvey
 Richard Harvey
 Alex Haslam
 Reid Hastie
 Craig Haugtvedt
 Mikki Hebl
 Susan Hendrich
 Leslie Heinberg
 Stephen Heine
 Greg Herek
 Peter Herman
 E. Tory Higgins
 Denis Hilton
 James Hilton
 Steve Hinkle
 Edward Hirt
 Stevan Hobfoll
 Sara Hodges
 Gordon Hodson
 Wim Hofstee
 Mike Hogg
 Ron Holden
 Charles J. Holehan
 Andrea Hollingshead
 Nick Hopkins
 David Houston
 Daniel Hrubes
 Jay Hull
 Mary Lee Hummert
 Yuen Huo
 Ted Huston
 Janet Hyde
 Lauri Hyer
 William Ickes
 Linda Isbell
 Jay Jackson
 Lynne Jackson
 Jill Jacobson
 Leon James
 Ronnie Janoff-Bulman
 Jolanda Jetten
 Ed Johnson
 Roger Johnson
 Lucy Johnston
 Thomas Joiner
 Randall Jorgeson
 Robert Josephs
 John Jost
 Cheryl Kaiser
 Susanne Kaiser
 Martin Kaplan
 Diane Kappen
 Steve Karav
 Ben Karney
 Minora Karsawa
 Jerry Karylowski
 Kerry Kawakami
 Carrie Keating
 Janice Kelly
 David Kenny
 Douglas Kenrick
 Michael Kernis
 Norb Kerr
 Sara Kiesler
 Sung Hee Kim
 Charles Kimble
 Laura King
 Lee Kirkpatrick
 S. Kitayama
 Mary Kite
 Yechiel Klar

Christopher Klauer
 Stanley Klein
 William Klein
 Cliff Knee
 Eric Knowles
 Diane Kobrynowicz
 Richard Koestner
 Audrey Korsgaard
 Dennis Krebs
 Justin Kruger
 Douglas Krull
 Jennifer Krull
 Mike Kuhlman
 Nick Kuiper
 James Kulik
 Lawrence Kurdek
 Richard Lalonde
 Alan Lambert
 Chris Langston
 Randy Larsen
 James Larson
 Daniel Lassiter
 Jean-Philippe Laurenceau
 G. Lautenschlager
 Howard Lavine
 Martin Lea
 Mark Leary
 Len Lecci
 Yeuh-ting Lee
 Michael Leippe
 Deborah Lendziarski
 Stephen Lepore
 Jenn Lerner
 Shana Levin
 Jacques-Philippe Leyens
 Nora Liberman
 Brian Little
 Glenn Little Page
 Penelope Lockwood
 Fabio Lorenzi-Cioldi
 Tim Loving
 Olivier Luminet
 Nina Lutwak
 John Lydon
 Sonja Lyubomirsky
 Jennifer Ma
 Tara MacDonald
 Diane Mackie
 Neil Macrae
 James Maddox
 Greg Maio
 Bertram Malle
 Lucia Manetti
 James E. Marcia

Keith Markman
 Hazel Markus
 Lenny Martin
 Robin Martin
 Rod Martin
 Walter Martin
 Marianne Mast
 Andrew Matthews
 Gerald Matthews
 Robert Mauro
 Jack Mayer
 Daniel McAdams
 Poppy McCleod
 John McClure
 Allen McConnell
 Don McCreary
 Mike McCullough
 Cathy McFarland
 Craig McGarty
 Richard McGlynn
 R. W. Meertens
 Marilyn Mendolia
 Peter Mersch
 Kristin Mickelson
 Mario Mikulincer
 Art Miller
 Carol T. Miller
 Dale Miller
 Todd Miller
 Roger Millsap
 Carolyn Morgan
 Beth Morling
 Michael Morris
 Robert Morris
 Gordon Moskowitz
 William Movellch
 P. Mudrack
 Brian Mullen
 Thomas Mullweiler
 Keith Murningham
 Sandra Murray
 David Myerson
 Scott Nadey
 Todd Nelson
 Steve Neuberg
 Leonard Newman
 Darren Newton
 F. J. Neyer
 Paula Niedenthal
 Yolanda Niemann
 Paul Norris
 Ed O'Neal
 Penelope Oakes
 Yuen Okuo

Thomas Ollendick
 James Olson
 Allen Omoto
 Don Operario
 Judith Ouellette
 Joop Ouwerkerk
 Dario Paez
 Jennifer Pals
 Bernadette Park
 Brian Parkinson
 J. T. Patacek
 Christopher Patrick
 Keith Payne
 Guido Peeters
 Brett Pelham
 Louis Pendry
 K. P. Peng
 Louis Penner
 Randall Peterson
 Tamarha Pierce
 Aaron Pincus
 Elizabeth Pinel
 Michael Platow
 Robert Plomin
 Gregory Pool
 Tom Postmes
 Joseph Priester
 Vincent Prohaska
 John T. Ptacek
 Tom Pyszczynski
 Sarah Queller
 Brian Quigley
 Diana Quinn
 Kimberly Quinn
 J. Rabble
 Brian Rafferty
 Bert Raven
 Stephen J. Read
 R. Reber
 Glenn Reeder
 Dennis Regan
 Alan Reifman
 William Revelle
 Fred Rhodewalt
 Deborah Richardson
 Robert Ridge
 Monica Rodriguez
 Neal Roesse
 Chris Roney
 Stacy Ropp
 Ralph Rosnow
 Michael Ross
 Alexander Rothman
 Laurie Rudman

Karen Ruggiero
 Brents Rushall
 Janet Ruscher
 James Russell
 Carolyn Rusting
 Ann Ruvalo
 Richard Ryan
 John Sabini
 Steve Samuels
 Catherine Sanderson
 Larry Sanna
 Carol Sansone
 David Sattler
 Ken Savitsky
 Mark Schaller
 Michael Schmitt
 Dave Schroeder
 Norbert Schwarz
 Constantine Sedikides
 Denise Sekaquaptewa
 Gun Semin
 John Seta
 William Shadish
 Don Sharpsteen
 Paschal Sheeran
 Ken Sheldon
 Nicole Shelton
 James Shepperd
 James Sherman
 Jeff Sherman
 Carolyn Showers
 James Sidaniuis
 Dit Silvera
 Paul Silvia
 Dean Simonton
 Ramandhar Singh
 Linda Skitka
 John Skowronski
 Christine Smith
 Eliot Smith
 Heather Smith
 Richard R. Smith
 Stephen Smith
 Steve Smith
 Tim Smith
 Mark Snyder

Sheldon Solomon
 Richard Sorrentino
 Steven Spencer
 Susan Sprecher
 Charles Stangor
 Diederik Stapel
 Walter Stephan
 Robert Sternburg
 Abigail Stewart
 Dennis Stolle
 Jeff Stone
 Alan Strathman
 Christina Studebaker
 Peter Suedfeld
 Mark Suh
 William Swann
 Janet Swiezek
 June Tangney
 Shelley Taylor
 James Tedeschi
 Howard Tennen
 Deborah Terry
 Abraham Tesser
 Phil Tetlock
 David Thissen
 Erik Thompson
 Leigh Thompson
 Suzanne Thompson
 Lisa Thompson Ross
 Diane Tice
 Scott Tindale
 Joseph Tomaka
 Andrew Tomarkin
 David Trafimow
 Paul Trapnell
 Yaacov Trope
 Jo Ann Tsang
 Orit Tykocinski
 Robin Vallacher
 Kees Van den Bos
 Eric Van Dijk
 Ad Van Knippenberg
 Daan Van Knippenberg
 Paul Van Lange
 Frank Van Ouerwalle
 Anneloïs Vandengrouke

Eric Vanman
 Bas Verplanken
 Terry Vescio
 Penny Visser
 Jeff Vittengl
 Kathleen Vohs
 William von Hippel
 Roos Vonk
 Jacquie Vorauer
 Kimberly Wade-Benzoni
 Hugh Wagner
 Lawrence Walker
 Eval Walther
 Daniel Wann
 Duane Wegener
 Bernard Weiner
 Gary Wells
 Mark Whatley
 Christian Wheeler
 Paul White
 Daniel Wigboldus
 David Wilder
 Kip Williams
 Michelle Williams
 Teresa Williams
 Thomas Wills
 Paul Windschitl
 Gail Wittenbaum
 Bernd Wittenbrink
 Michael Wohl
 Bogdan Wojciszke
 Wendy Wood
 Steve Worchel
 Ted Wright
 Laura Yamhure
 Stephen Zaccaro
 Mark Zanna
 Michael Zarate
 Leslie Zebrowitz
 Marcel Zeelenburg
 John Zelenski
 Sarah Zemore
 Marvin Zuckerman

INDEX

to

PERSONALITY AND SOCIAL PSYCHOLOGY BULLETIN

Volume 27

- Number 1 (January 2001) pp. 1-136
Number 2 (February 2001) pp. 137-264
Number 3 (March 2001) pp. 265-384
Number 4 (April 2001) pp. 385-512
Number 5 (May 2001) pp. 513-640
Number 6 (June 2001) pp. 641-768
Number 7 (July 2001) pp. 769-912
Number 8 (August 2001) pp. 913-1072
Number 9 (September 2001) pp. 1073-1240
Number 10 (October 2001) pp. 1241-1392
Number 11 (November 2001) pp. 1393-1564
Number 12 (December 2001) pp. 1565-1728

Authors:

- ABOUFADEL, KATHY, see Madon, S.
ABRAHAM, CHARLES, see Conner, M.
AGNEW, CHRISTOPHER R., see Arriaga, X. B.
ALBARRACIN, DOLORES, and ROBERT S. WYER, JR., "Elaborative and Nonelaborative Processing of a Behavior-Related Communication," 691.
ALLEN, JOHN J. B., see Harmon-Jones, E.
ANDREOLETTI, CARRIE, LESLIE A. ZEBROWITZ, and MARGIE E. LACHMAN, "Physical Appearance and Control Beliefs in Young, Middle-Age, and Older Adults," 969.
ARON, ARTHUR, see Hassebrauck, M.
ARRIAGA, XIMENA B., and CHRISTOPHER R. AGNEW, "Being Committed: Affective, Cognitive, and Conative Components of Relationship Commitment," 1190.
ASHER, TERRILEE, see Biesanz, J. C.
ASHER, TERRILEE, see Guadagno, R. E.
AXSOM, DANNY, see Blanton, H.
AYDUK, OZLEM, GERALDINE DOWNEY, and MINJI KIM, "Rejection Sensitivity and Depressive Symptoms in Women," 868.
BAKER, MARK, see Schmitt, D. P.
BALDWIN, MARK W., and KELLEY J. MAIN, "Social Anxiety and the Cued Activation of Relational Knowledge," 1637.
BAUMEISTER, ROY F., see Ciarocco, N. J.
BELLAH, C. GARTH, see McCullough, M. E.
BELLAVIA, GINA, see Murray, S. L.
BERANT, ETY, MARIO MIKULINCER, and VICTOR FLORIAN, "Attachment Style and Mental Health: A 1-Year Follow-Up Study of Mothers of Infants With Congenital Heart Disease," 956.
BEREL, SUSAN R., see Parks, C. D.
BERG, MICHAEL B., RONNIE JANOFF-BULMAN, and JUSTIN COTTER, "Perceiving Value in Obligations and Goals: Wanting to Do What Should Be Done," 982.
BERRY, JACK W., EVERETT L. WORTHINGTON, JR., LES PARROTT, III, LYNN E. O'CONNOR, and NATHANIEL G. WADE, "Dispositional Forgivingness: Development and Construct Validity of the Transgression Narrative Test of Forgivingness (TNTF)," 1277.
BETSCH, TILMANN, HENNING PLESSNER, CHRISTIANE SCHWIEREN, and ROBERT GÜTIG, "I Like It But I Don't Know Why: A Value-Account Approach to Implicit Attitude Formation," 242.
BIESANZ, JEREMY C., STEVEN L. NEUBERG, DYLAN M. SMITH, TERRILEE ASHER, and T. NICOLE JUDICE, "When Accuracy-Motivated Perceivers Fail: Limited Attentional Resources and the Reemerging Self-Fulfilling Prophecy," 621.
BIZMAN, AHARON, YOEL YINON, and SHARON KROTMAN, "Group-Based Emotional Distress: An Extension of Self-Discrepancy Theory," 1291.
BLANTON, HART, DANNY AXSOM, KIMBERLY P. McCLIVE, and SIMANI PRICE, "Pessimistic Bias in Comparative Evaluations: A Case of Perceived Vulnerability to the Effects of Negative Life Events," 1622.
BLANTON, HART, ANNE E. STUART, and REGINA J.J.M. VANDENIJNDEN, "An Introduction to Deviance-Regulation Theory: The Effect of Behavioral Norms on Message Framing," 848.
BLESKE-RECHEK, APRIL L., and DAVID M. BUSS, "Opposite-Sex Friendship: Sex Differences and Similarities in Initiation, Selection, and Dissolution," 1310.
BLESS, HERBERT, see Wänke, M.
BLOCK, JACK, see Onishi, M.
BOGGIANO, ANN K., see Guay, F.
BOHNER, GERD, and THOMAS WEINERTH, "Negative Affect Can Increase or Decrease Message Scrutiny: The Affect Interpretation Hypothesis," 1417.
BOND, CHARLES F., JR., see Richard, F. D.
BOURGEOIS, MARTIN J., see Sommer, K. L.

PSPB, Vol. 27 No. 12, December 2001 1721-1728

© 2001 by the Society for Personality and Social Psychology, Inc.

- BOWMAN, PHILLIP J., see McAdams, D. P.
 BRADBURY, THOMAS N., see Cobb, R. J.
 BRANNON, LAURA A., and TIMOTHY C. BROCK, "Scarcity Claims Elicit Extreme Responding to Persuasive Messages: Role of Cognitive Elaboration," 365.
 BRANSCOMBE, NYLA R., see Jetten, J.
 BROCK, TIMOTHY C., see Brannon, L. A.
 BROWN, JONATHON D., and MARGARET A. MARSHALL, "Self-Esteem and Emotion: Some Thoughts About Feelings," 575.
 BRYANT, BIRGIT, see Mullen, B.
 BUEHLER, ROGER, and CATHY MCFARLAND, "Intensity Bias in Affective Forecasting: The Role of Temporal Focus," 1480.
 BUIDIN, GENEVIÈVE, see Corneille, O.
 BULL, RAY, see Vrij, A.
 BURGER, JERRY M., SHELLEY SOROKA, KATRINA GONZAGO, EMILY MURPHY, and EMILY SOMERVELL, "The Effect of Fleet- ing Attraction on Compliance to Requests," 1578.
 BURRIS, CHRISTOPHER T., see Jackson, L. M.
 BUSHMAN, BRAD J., and GARY L. WELLS, "Narrative Impressions of Literature: The Availability Bias and the Corrective Properties of Meta-Analytic Approaches," 1123.
 BUSS, DAVID M., see Bleske-Rechek, A. L.
 BUUNK, BRAM P., see Ybema, J. F.
 CALLAN, VICTOR J., see Terry, D. J.
 CAMPBELL, LORNE, JEFFREY A. SIMPSON, DEBORAH A. KASHY, and GARTH J. O. FLETCHER, "Ideal Standards, the Self, and Flexibility of Ideals in Close Relationships," 447.
 CANTOR, NANCY, see Sanderson, C. A.
 CAREY, CRAIG J., see Terry, D. J.
 CERVONE, DANIEL, see Weitlauf, J. C.
 CHANG, EDWARD C., see Sanna, L. J.
 CHANTALA, KIM, see Cleveland, H. H.
 CHEMERS, MARTIN M., see Watson, C. B.
 CHIU, CHI-YUE, see Knowles, E. D.
 CIALDINI, R. B., see Guadagno, R. E.
 CIAROCCHI, NATALIE J., KRISTIN L. SOMMER, and ROY F. BAUMEISTER, "Ostracism and Ego Depletion: The Strains of Silence," 1156.
 CLEVELAND, H. HARRINGTON, J. RICHARD UDRY, and KIM CHANTALA, "Environmental and Genetic Influences on Sex-Typed Behaviors and Attitudes of Male and Female Adolescents," 1587.
 CLORE, GERALD L., see Robinson, M. D.
 COBB, REBECCA J., JOANNE DAVILA, and THOMAS N. BRADBURY, "Attachment Security and Marital Satisfaction: The Role of Positive Perceptions and Social Support," 1131.
 CONNER, MARK, and CHARLES ABRAHAM, "Conscientiousness and the Theory of Planned Behavior: Toward a More Complete Model of the Antecedents of Intentions and Behavior," 1547.
 CORNEILLE, OLIVIER, VINCENT Y. YZERBYT, ANOUK ROGIER, and GENEVIÈVE BUIDIN, "Threat and the Group Attribution Error: When Threat Elicits Judgments of Extremity and Homogeneity," 437.
 COTTER, JUSTIN, see Berg, M. B.
 COUDEN, ANDREA, see Schmitt, D. P.
 CRANDALL, CHRISTIAN S., SILVANA D'ANELLO, NURAY SAKALLI, ELEANA LAZARUS, GRAZYNA WIECZORKOWSKA, and N. T. FEATHER, "An Attribution-Value Model of Prejudice: Anti-Fat Attitudes in Six Nations," 30.
 CRANO, WILLIAM D., see Dobbs, M.
 CRISP, RICHARD J., MILES HEWSTONE, and MARK RUBIN, "Does Multiple Categorization Reduce Intergroup Bias?" 76.
 CROSS, SUSAN E., see Kanagawa, C.
 CROSS, SUSAN E., and NIKI V. VICK, "The Interdependent Self- Construal and Social Support: The Case of Persistence," 820.
 DE GROOT, DAPHNE, see Lea, M.
 DE GROOT, DAPHNE, see Postmes, T.
 D'ANELLO, SILVANA, see Crandall, C. S.
 DEJONG, MAJELLA, see Ybema, J. F.
 DE VRIES, NANNE K., see Gordijn, E. H.
 DARLEY, JOHN M., see Garcia, S. M.
 DAVILA, JOANNE, see Cobb, R. J.
 DECI, EDWARD L., RICHARD M. RYAN, MARYLÈNE GAGNÉ, DEAN R. LEONE, JULIAN USUNOV, and BOYANKA P. KORNAZHEVA, "Need Satisfaction, Motivation, and Well-Being in the Work Organizations of a Former Eastern Bloc Country: A Cross-Cultural Study of Self-Determination," 930.
 DEMAINE, L. J., see Guadagno, R. E.
 DIENER, ED, see Oishi, S.
 DOBBS, MICHAEL, and WILLIAM D. CRANO, "Outgroup Accountability in the Minimal Group Paradigm: Implications for Aversive Discrimination and Social Identity Theory," 355.
 DOLINSKI, DARIUSZ, MAGDALENA NAWRAT, and IZABELA RUDAK, "Dialogue Involvement as a Social Influence Technique," 1395.
 DOVIDIO, JOHN F., see Kawakami, K.
 DOWNEY, GERALDINE, see Ayduk, O.
 ECCLES, JACQUELYNNE, see Madon, S.
 EDWARD, KATHERINE, see Vrij, A.
 ESSES, VICTORIA M., see Jackson, L. M.
 ESSES, VICTORIA M., see Quinn, K. A.
 EVANS, SARAH M., see Sanderson, C. A.
 FAZIO, RUSSELL H., and LAURA E. HILDEN, "Emotional Reactions to a Seemingly Prejudiced Response: The Role of Automatically Activated Racial Attitudes and Motivation to Control Prejudiced Reactions," 538.
 FAZIO, RUSSELL H., see Towles-Schwen, T.
 FEATHER, N. T., and CHARLOTTE JOHNSTONE, "Social Norms, Entitlement, and Deservingness: Differential Reactions to Aggressive Behavior of Schizophrenic and Personality Disorder Patients," 755.
 FEATHER, N. T., see Crandall, C. S.
 FIEDLER, KLAUS, and WOLFRAM SCHENCK, "Spontaneous Inferences From Pictorially Presented Behaviors," 1533.
 FINLAY, KRYSTINA A., see Trafimow, D.
 FISKE, SUSAN T., see Operario, D.
 FLANNERY, WM. PETER, STEVEN P. REISE, and JIAJUAN YU, "An Empirical Comparison of Acculturation Models," 1035.
 FLETCHER, GARTH J. O., see Campbell, L.
 FLORIAN, VICTOR, see Berant, E.
 FLORIAN, VICTOR, MARIO MIKULINCER, and GILAD HIRSCHBERGER, "Validation of Personal Identity as a Terror Management Mechanism: Evidence That Sex-Role Identity Moderates Mortality Salience Effects," 1011.
 GABLE, SHELLEY L., see Nezlek, J. B.
 GAGNÉ, FABIY M., and JOHN E. LYDON, "Mindset and Relationship Illusions: The Moderating Effects of Domain Specificity and Relationship Commitment," 1144.
 GAGNÉ, MARYLÈNE, see Deci, E. L.
 GARCIA, STEPHEN MICHAEL, JOHN M. DARLEY, and ROBERT J. ROBINSON, "Morally Questionable Tactics: Negotiations Between District Attorneys and Public Defenders," 731.
 GARVEY, BRIAN, see Sabini, J.
 GIFFORD, ROBERT, see Reynolds, D. J.
 GILBERT, DANIEL T., see Wilson, T. D.
 GINER-SOROLLA, ROGER, "Affective Attitudes Are Not Always Faster: The Moderating Role of Extremity," 666.
 GJERDE, PER F., see Onishi, M.
 GOMEZ, JOHN P., and STEVEN J. TRIERWEILER, "Does Discrimination Terminology Create Response Bias in Questionnaire Studies of Discrimination?" 630.
 GONZAGO, KATRINA, see Burger, J. M.
 GORDIJN, ERNESTINE H., TOM POSTMES, and NANNE K. DE VRIES, "Devil's Advocate or Advocate of Oneself: Effects of Numerical Support on Pro- and Counterattitudinal Self-Persuasion," 395.
 GREEN, LAURA R., DEBORAH R. RICHARDSON, TANIA LAGO, and ELIZABETH C. SCHATTEN-JONES, "Network Correlates of Social and Emotional Loneliness in Young and Older Adults," 281.
 GREENWALD, ANTHONY G., see Rudman, L. A.
 GRIFFIN, DALE W., see Murray, S. L.
 GUADAGNO, ROSANNA E., TERRILEE ASHER, LINDA J. DEMAINE, and ROBERT B. CIALDINI, "When Saying Yes Leads

- to Saying No: Preference for Consistency and the Reverse Foot-in-the-Door Effect," 859.
- GUAY, FRÉDÉRIC, ANN K. BOGGIANO, And ROBERT J. VALLERAND, "Autonomy Support, Intrinsic Motivation, and Perceived Competence: Conceptual and Empirical Linkages," 643.
- GUAY, FRÉDÉRIC, see Senécal, C.
- GURUNG, REGAN A. R., BARBARA R. SARASON, and IRWIN G. SARASON, "Predicting Relationship Quality and Emotional Reactions to Stress From Significant-Other-Concept Clarity," 1267.
- GÜTIG, ROBERT, see Betsch, T.
- GUYLL, MAX, see Madon, S.
- HALL, AMANDA L., see Sabini, J.
- HARMON-JONES, EDDIE, and JOHN J. B. ALLEN, "The Role of Affect in the Mere Exposure Effect: Evidence From Psychophysiological and Individual Differences Approaches," 889.
- HASLAM, S. ALEXANDER, and MICHAEL J. PLATOW, "The Link Between Leadership and Followership: How Affirming Social Identity Translates Vision Into Action," 1469.
- HASSEBRAUCK, MANFRED, and ARTHUR ARON, "Prototype Matching in Close Relationships," 1111.
- HAWKINS, CHRIS, see Petty, R. E.
- HEMENOVER, SCOTT H., "Self-Reported Processing Bias and Naturally Occurring Mood: Mediators Between Personality and Stress Appraisals," 387.
- HENDERSON-KING, DONNA, EAARON HENDERSON-KING, and LISA HOFFMAN, "Media Images and Women's Self-Evaluations: Social Context and Importance of Attractiveness as Moderators," 1407.
- HENDERSON-KING, EAARON, see Henderson-King, D.
- HENDIN, HOLLY M., see Robins, R. W.
- HESSON-MCINNIS, MATHEW, see Reeder, G. D.
- HEWSTONE, MILES, see Crisp, R. J.
- HILDEN, LAURA E., see Fazio, R. H.
- HIRSCHBERGER, GILAD, see Florian, V.
- HIRT, EDWARD R., see McCrea, S. M.
- HODGES, SARA D., see Klein, K. J. K.
- HOFFMAN, LISA, see Henderson-King, D.
- HOFSTEE, WILLEM K. B., see Koole, S. L.
- HOLMES, JOHN G., see Murray, S. L.
- HONG, YING-YI, see Knowles, E. D.
- HOROWITZ, IRWIN A., see Sommer, K. L.
- IGOU, ERIC R., see Wänke, M.
- JACKSON, LYNNE M., VICTORIA M. ESSES, and CHRISTOPHER T. BURRIS, "Contemporary Sexism and Discrimination: The Importance of Respect for Men and Women," 48.
- JAGER, WANDER, see Koole, S. L.
- JANOFF-BULMAN, RONNIE, see Berg, M. B.
- JASCHIK-HERMAN, BRUCE, see Tracey, T. J. G.
- JETTEN, JOLANDA, NYLA R. BRANSCOMBE, MICHAEL T. SCHMITT, and RUSSELL SPEARS, "Rebels With a Cause: Group Identification as a Response to Perceived Discrimination From the Mainstream," 1204.
- JOHAR, GITA VENKATARAMANI, see Sengupta, J.
- JOHNSON, JOEL T., "On Weakening the Strongest Link: Attributions and Intervention Strategies for Behavior Change," 408.
- JOHNSON, JUDITH L., see McCullough, M. E.
- JOHNSTONE, CHARLOTTE, see Feather, N. T.
- JOST, JOHN T., see Stangor, C.
- JUDICE, T. NICOLE, see Biesenz, J. C.
- JUSSIM, LEE, see Madon, S.
- KAISER, CHERYL R., and CAROL T. MILLER, "Reacting to Impending Discrimination: Compensation for Prejudice and Attributions to Discrimination," 1357.
- KAISER, CHERYL R., and CAROL T. MILLER, "Stop Complaining! The Social Costs of Making Attributions to Discrimination," 254.
- KANAGAWA, CHIE, SUSAN E. CROSS, and HAZEL ROSE MARKUS, "Who am I? The Cultural Psychology of the Conceptual Self," 90.
- KARNEY, BENJAMIN R., see McNulty, J. K.
- KASHY, DEBORAH A., see Campbell, L.
- KASSIN, SAUL M., see Sommers, S. R.
- KAWAKAMI, KERRY, and JOHN F. DOVIDIO, "The Reliability of Implicit Stereotyping," 212.
- KELLY, ANITA E., JULIE A. KLUSAS, RENEE T. VON WEISS, and CHRISTINE KENNY, "What Is It About Revealing Secrets That Is Beneficial?" 651.
- KENNY, CHRISTINE, see Kelly, A. E.
- KILPATRICK, SHELLEY DEAN, see McCullough, M. E.
- KIM, MINJI, see Ayduk, O.
- KING, LAURA A., "The Health Benefits of Writing About Life Goals," 798.
- KLEIN, KRISTI J. K., and SARA D. HODGES, "Gender Differences, Motivation, and Empathic Accuracy: When It Pays to Understand," 720.
- KLEIN, WILLIAM M. P., "Post Hoc Construction of Self-Performance and Other Performance in Self-Serving Social Comparison," 744.
- KLUSAS, JULIE A., see Kelly, A. E.
- KNEE, C. RAYMOND, ARUNI NANAYAKKARA, NATHANIEL A. VIETOR, CLAYTON NEIGHBORS, and HEATHER PATRICK, "Implicit Theories of Relationships: Who Cares If Romantic Partners Are Less Than Ideal?" 808.
- KNOWLES, ERIC D., MICHAEL W. MORRIS, CHI-YUE CHIU, and YING-YI HONG, "Culture and the Process of Person Perception: Evidence for Automaticity Among East Asians in Correcting for Situational Influences on Behavior," 1344.
- KOOLE, SANDER L., WANDER JAGER, AGNES E. VAN DEN BERG, CHARLES A. J. VLEK, and WILLEM K. B. HOFSTEE, "On the Social Nature of Personality: Effects of Extraversion, Agreeableness, and Feedback About Collective Resource Use on Cooperation in a Resource Dilemma," 289.
- KOOMEN, WILLEM, see Stapel, D. A.
- KORNAZHEVA, BOYANKA P., see Deci, E. L.
- KROHSE, JOSHUA O., see Reeder, G. D.
- KROTMAN, SHARON, see Bizman, A.
- KRUEGER, JOACHIM, and DAVID STANKE, "The Role of Self-Referent and Other-Referent Knowledge in Perceptions of Group Characteristics," 878.
- KUIJER, ROELINE G., see Ybema, J. F.
- KUMHYR, SANDRA M., see Vorauer, J. D.
- KURMAN, JENNY, "Self-Enhancement: Is It Restricted to Individualistic Cultures?" 1705.
- KURZBAN, ROBERT, KEVIN MCCABE, VERNON L. SMITH, and BART J. WILSON, "Incremental Commitment and Reciprocity in a Real-Time Public Goods Game," 1662.
- LACHMAN, MARGIE E., see Andreoletti, C.
- LAGO, TANIA, see Green, L. R.
- LANE, MARIAN D., see Peterson, B. E.
- LAZARUS, ELEANA, see Crandall, C. S.
- LEA, MARTIN, RUSSELL SPEARS, and DAPHNE DE GROOT, "Knowing Me, Knowing You: Anonymity Effects on Social Identity Processes Within Groups," 526.
- LEE, ANGELAY, "The Mere Exposure Effect: An Uncertainty Reduction Explanation Revisited," 1255.
- LEONE, DEAN R., see Deci, E. L.
- LEWIS, MARTHA, see McAdams, D. P.
- LIND, E. ALLAN, see van den Bos.
- LIVINGSTON, ROBERT W., "What You See Is What You Get: Systematic Variability in Perceptual-Based Social Judgment," 1086.
- LYDON, JOHN E., see Gagné, F. M.
- MADON, STEPHANIE, MAX GUYLL, KATHY ABOUFADDEL, EULICES MONTIEL, ALISON SMITH, POLLY PALUMBO, and LEE JUSSIM, "Ethnic and National Stereotypes: The Princeton Trilogy Revisited and Revised," 996.
- MADON, STEPHANIE, ALISON SMITH, LEE JUSSIM, DANIEL W. RUSSELL, JACQUELYNNE ECCLES, POLLY PALUMBO, and MICHELE WALKIEWICZ, "Am I as You See Me or Do You See Me as I Am: Self-Fulfilling Prophecies and Self-Verification," 1214.
- MAIN, KELLEY J., see Baldwin, M. W.
- MANSTEAD, ANTONY, S. R., see Nienhuis, A. E.
- MARKUS, HAZEL ROSE, see Kanagawa, C.
- MARSHALL, MARGARET A., see Brown, J. D.

- MCADAMS, DAN P., JEFFREY REYNOLDS, MARTHA LEWIS, ALLISON H. PATTEN, AND PHILLIP J. BOWMAN, "When Bad Things Turn Good and Good Things Turn Bad: Sequences of Redemption and Contamination in Life Narrative and Their Relation to Psychosocial Adaptation in Midlife Adults and in Students," 474.
- MCCABE, KEVIN, see Kurzban, R.
- MCCANN, STEWART J. H., "The Precocity-Longevity Hypothesis: Earlier Peaks in Career Achievement Predict Shorter Lives," 1429.
- MCCCLIVE, KIMBERLY P., see Blanton, H.
- MCCREA, SEAN M., and EDWARD R. HIRT, "The Role of Ability Judgments in Self-Handicapping," 1378.
- MCCULLOUGH, MICHAEL E., C. GARTH BELLAH, SHELLEY DEAN KILPATRICK, and JUDITH L. JOHNSON, "Vengefulness: Relationships With Forgiveness, Rumination, Well-Being, and the Big Five," 601.
- McFARLAND, CATHY, see Buehler, R.
- McGHEE, DEBBIE E., see Rudman, L. A.
- McNULTY, JAMES K., and BENJAMIN R. KARNEY, "Attributions in Marriage: Integrating Specific and Global Evaluations of a Relationship," 943.
- MEIER, SUSANNE, see Sanna, L. J.
- MEYERS, JAY, see Wilson, T. D.
- MICKELSON, KRISTIN D., "Perceived Stigma, Social Support, and Depression," 1046.
- MIKULINCER, MARIO, see Berant, E.
- MIKULINCER, MARIO, see Florian, V.
- MILLER, CAROL T., see Kaiser, C. R.
- MILNE, ALAN B., see Tafarodi, R. W.
- MONTIEL, EULICES, see Madon, S.
- MORRIS, MICHAEL W., see Knowles, E. D.
- MUELLERLEILE, PAIGE, see Mullen, B.
- MULLEN, BRIAN, PAIGE MUELLERLEILE, and BIRGIT BRYANT, "Cumulative Meta-Analysis: A Consideration of Indicators of Sufficiency and Stability," 1450.
- MURPHY, EMILY, see Burger, J. M.
- MURRAY, SANDRA L., JOHN G. HOLMES, DALE W. GRIFFIN, GINA BELLAVIA, and PAUL ROSE, "The Mismeasure of Love: How Self-Doubt Contaminates Relationship Beliefs," 423.
- MUSSWEILER, THOMAS, "Focus of Comparison as a Determinant of Assimilation Versus Contrast in Social Comparison," 38.
- NANAYAKKARA, ARUNI, see Knee, C. R.
- NAWRAT, MAGDALENA, see Dolinski, D.
- NEIGHBORS, CLAYTON, see Knee, C. R.
- NEUBERG, STEVEN L., see Biesanz, J. C.
- NEZLEK, JOHN B., "Multilevel Random Coefficient Analyses of Event- and Interval-Contingent Data in Social and Personality Psychology Research," 771.
- NEZLEK, JOHN B., and SHELLEY L. GABLE, "Depression as a Moderator of Relationships Between Positive Daily Events and Day-to-Day Psychological Adjustment," 1692.
- NEZLEK, JOHN B., and REBECCA M. PLESKO, "Day-to-Day Relationships Among Self-Concept Clarity, Self-Esteem, Daily Events, and Mood," 201.
- NIENHUIS, ANNET E., ANTONY S. R. MANSTEAD, and RUSSELL SPEARS, "Multiple Motives and Persuasive Communication: Creative Elaboration as a Result of Impression Motivation and Accuracy Motivation," 118.
- O'CONNOR, LYNN E., see Berry, J. W.
- OISHI, SHIGEHIRO, and ED DIENER, "Goals, Culture, and Subjective Well-Being," 1674.
- ONISHI, MIYOKO, PER F. GJERDE, and JACK BLOCK, "Personality Implications of Romantic Attachment Patterns in Young Adults: A Multi-Method, Multi-Informant Study," 1097.
- OPERARIO, DON, and SUSAN T. FISKE, "Ethnic Identity Moderates Perceptions of Prejudice: Judgments of Personal Versus Group Discrimination and Subtle Versus Blatant Bias," 550.
- PALUMBO, POLLY, see Madon, S.
- PARKS, CRAIG D., and ANN C. RUMBLE, "Elements of Reciprocity and Social Value Orientation," 1301.
- PARKS, CRAIG D., LAWRENCE J. SANNA, and SUSAN R. BEREL, "Actions of Similar Others as Inducements to Cooperate in Social Dilemmas," 345.
- PARROTT, LES, III, see Berry, J. W.
- PATRICK, HEATHER, see Knee, C. R.
- PATTEN, ALLISON H., see McAdams, D. P.
- PETERSON, BILL E. and MARIAN D. LANE, "Implications of Authoritarianism for Young Adulthood: Longitudinal Analysis of College Experiences and Future Goals," 678.
- PETTY, RICHARD E., see Tormala, Z. L.
- PETTY, RICHARD E., ZAKARY L. TORMALA, CHRIS HAWKINS, and DUANE T. WEGENER, "Motivation to Think and Order Effects in Persuasion: The Moderating Role of Chunking," 332.
- PHAM, LIEN B., SHELLEY E. TAYLOR, and TERESA E. SEEMAN, "Effects of Environmental Predictability and Personal Mastery on Self-Regulatory and Physiological Processes," 611.
- PICKETT, CYNTHIA L., "The Effects of Entitativity Beliefs on Implicit Comparisons Between Group Members," 515.
- PLATOW, MICHAEL J., see Haslam, S. A.
- PLATOW, MICHAEL J., and DAAN VAN KNIPPENBERG, "A Social Identity Analysis of Leadership Endorsement: The Effects of Leader Ingroup Prototypicality and Distributive Intergroup Fairness," 1508.
- PLESKO, REBECCA M., see Nezlek, J. B.
- PLESSNER, HENNING, see Betsch, T.
- POSTMES, TOM, see Gordijn, E. H.
- POSTMES, TOM, RUSSELL SPEARS, KHALED SAKHEL, and DAPHNE DE GROOT, "Social Influence in Computer-Mediated Communication: The Effects of Anonymity on Group Behavior," 1243.
- PRATKANIS, ANTHONY R., see Rucker, D. D.
- PREISER, NATALYA, see Watson, C. B.
- PRICE, SIMANI, see Blanton, H.
- PUCA, ROSA MARIA, and HEINZ-DIETER SCHMALT, "The Influence of the Achievement Motive on Spontaneous Thoughts in Pre- and Postdecisional Action Phases," 302.
- QUINN, KIMBERLY A., ERIN M. ROSS, and VICTORIA M. ESSES, "Attributions of Responsibility and Reactions to Affirmative Action: Affirmative Action as Help," 321.
- REEDER, GLENN D., MATHEW HESSON-McINNIS, JOSHUA O. KROHSE, and ELIZABETH A. SCIALABBA, "Inferences About Effort and Ability," 1225.
- REICH, DARCY A., see Weary, G.
- REISE, STEVEN P., see Flannery, W. P.
- REYNOLDS, D'ARCY J., JR., and ROBERT GIFFORD, "The Sounds and Sights of Intelligence: A Lens Model Channel Analysis," 187.
- REYNOLDS, JEFFREY, see McAdams, D. P.
- RICHARD, F. D., CHARLES F. BOND, JR., and JULI J. STOKES-ZOOTA, "That's Completely Obvious . . . and Important': Lay Judgments of Social Psychological Findings," 497.
- RICHARDSON, DEBORAH R., see Green, L. R.
- ROBINS, RICHARD W., HOLLY M. HENDIN, and KALI H. TRZESNIEWSKI, "Measuring Global Self-Esteem: Construct Validation of a Single-Item Measure and the Rosenberg Self-Esteem Scale," 151.
- ROBINSON, MICHAEL D., and GERALD L. CLORE, "Simulation, Scenarios, and Emotional Appraisal: Testing the Convergence of Real and Imagined Reactions to Emotional Stimuli," 1520.
- ROBINSON, ROBERT J., see Garcia, S. M.
- ROGIER, ANOUK, see Corneille, O.
- ROSE, PAUL, see Murray, S. L.
- ROSS, ERIN M., see Quinn, K. A.
- RUBIN, MARK, see Crisp, R. J.
- RUCKER, DEREK D., and ANTHONY R. PRATKANIS, "Projection as an Interpersonal Influence Tactic: The Effects of the Pot Calling the Kettle Black," 1494.
- RUDAK, IZABELA, see Dolinski, D.
- RUDMAN, LAURIE A., ANTHONY G. GREENWALD, and DEBBIE E. McGHEE, "Implicit Self-Concept and Evaluative Implicit Gender Stereotypes: Self and Ingroup Share Desirable Traits," 1164.

- RUMBLE, ANN C., see Parks, C. D.
- RUSSELL, DANIEL W., see Madon, S.
- RYAN, JENNIFER M., see Tracey, T.J.G.
- RYAN, RICHARD M., see Deci, E. L.
- SABINI, JOHN, BRIAN GARVEY, and AMANDA L. HALL, "Shame and Embarrassment Revisited," 104.
- SAKALLI, NURAY, see Crandall, C. S.
- SAKHEL, KHALED, see Postmes, T.
- SANDERMAN, ROBBERT, see Ybema, J. F.
- SANDERSON, CATHERINE A., and NANCY CANTOR, "The Association of Intimacy Goals and Marital Satisfaction: A Test of Four Mediational Hypotheses," 1567.
- SANDERSON, CATHERINE A., and SARAH M. EVANS, "Seeing One's Partner Through Intimacy-Colored Glasses: An Examination of the Processes Underlying the Intimacy Goals-Relationship Satisfaction Link," 463.
- SANNA, LAWRENCE J., see Parks, C. D.
- SANNA, LAWRENCE J., EDWARD C. CHANG, and SUSANNE MEIER, "Counterfactual Thinking and Self-Motives," 1023.
- SARASON, BARBARA R., see Gurung, R.A.R.
- SARASON, IRWIN G., see Gurung, R.A.R.
- SCHATTEN-JONES, ELIZABETH C., see Green, L. R.
- SCHENCK, WOLFRAM, see Fiedler, K.
- SCHMALT, HEINZ-DIETER, see Puca, R. M.
- SCHMITT, DAVID P., ANDREA COUDEN, and MARK BAKER, "The Effects of Sex and Temporal Context on Feelings of Romantic Desire: An Experimental Evaluation of Sexual Strategies Theory," 833.
- SCHMITT, MICHAEL T., see Jetten, J.
- SCHWIEN, CHRISTIANE, see Betsch, T.
- SCIALABBA, ELIZABETH A., see Reeder, G. D.
- SECHRIST, GRETCHEN B., see Stangor, C.
- SEEMAN, TERESA E., see Pham, L. B.
- SEGERSTROM, SUZANNE C., "Optimism and Attentional Bias for Negative and Positive Stimuli," 1334.
- SENÉCAL, CAROLINE, ROBERT J. VALLERAND, and FRÉDÉRIC GUAY, "Antecedents and Outcomes of Work-Family Conflict: Toward a Motivational Model," 176.
- SENGUPTA, JAIDEEP, and GITA VENKATARAMANI JOHAR, "Contingent Effects of Anxiety on Message Elaboration and Persuasion," 139.
- SIMPSON, JEFFRY A., see Campbell, L.
- SMITH, ALISON, see Madon, S.
- SMITH, DYLAN M., see Biesanz, J. C.
- SMITH, RONALD E., see Weitlauf, J. C.
- SMITH, VERNON L., see Kurzban, R.
- SOMERVELL, EMILY, see Burger, J. M.
- SOMMER, KRISTIN L., see Ciarocco, N. J.
- SOMMER, KRISTIN L., IRWIN A. HOROWITZ, and MARTIN J. BOURGEOIS, "When Juries Fail to Comply With the Law: Biased Evidence Processing in Individual and Group Decision Making," 309.
- SOMMERS, SAMUEL R., and SAUL M. KASSIN, "On the Many Impacts of Inadmissible Testimony: Selective Compliance, Need for Cognition, and the Overcorrection Bias," 1368.
- SOROKA, SHELLEY, see Burger, J. M.
- SPEARS, RUSSELL, see Jetten, J.
- SPEARS, RUSSELL, see Lea, M.
- SPEARS, RUSSELL, see Nienhuis, A. E.
- SPEARS, RUSSELL, see Postmes, T.
- STANGOR, CHARLES, GRETCHEN B. SECHRIST, and JOHN T. JOST, "Changing Racial Beliefs by Providing Consensus Information," 486.
- STANKE, DAVID, see Krueger, J.
- STAPEL, DIEDERIK A., and WILLEM KOOMEN, "When We Wonder What It All Means: Interpretation Goals Facilitate Accessibility and Stereotyping Effects," 915.
- STOKES-ZOOTA, JULI J., see Richard, F. D.
- STUART, ANNE E., see Blanton, H.
- TAFARODI, ROMIN W., JANICE TAM, and ALAN B. MILNE, "Selective Memory and the Persistence of Paradoxical Self-Esteem," 1179.
- TAM, JANICE, see Tafarodi, R. W.
- TAYLOR, SHELLEY E., see Pham, L. B.
- TERRY, DEBORAH J., CRAIG J. CAREY, and VICTOR J. CALLAN, "Employee Adjustment to an Organizational Merger: An Inter-group Perspective," 267.
- TOBIN, STEPHANIE J., see Weary, G.
- TORMALA, ZAKARY L., and RICHARD E. PETTY, "On-Line Versus Memory-Based Processing: The Role of 'Need to Evaluate' in Person Perception," 1599.
- TORMALA, ZAKARY L., see Petty, R. E.
- TOWLES-SCHWEN, TAMARA, and RUSSELL H. FAZIO, "On the Origins of Racial Attitudes: Correlates of Childhood Experiences," 162.
- TRACEY, TERENCE J. G., JENNIFER M. RYAN, and BRUCE JASCHIK-HERMAN, "Complementarity of Interpersonal Circumplex Traits," 786.
- TRAFIMOW, DAVID, "The Effects of Trait Type and Situation Type on the Generalization of Trait Expectancies Across Situations," 1463.
- TRAFIMOW, DAVID, and KRYSTINA A. FINLAY, "An Investigation of Three Models of Multitrait Representations," 226.
- TRIERWEILER, STEVEN J., see Gomez, J. P.
- TROPP, LINDA R., and STEPHEN C. WRIGHT, "Ingroup Identification as the Inclusion of Ingroup in the Self," 585.
- TRZESNIEWSKI, KALI H., see Robins, R. W.
- TYKOCINSKI, ORIT E., "I Never Had a Chance: Using Hindsight Tactics to Mitigate Disappointments," 376.
- UDRY, J. RICHARD, see Cleveland, H. H.
- USUNOV, JULIAN, see Deci, E. L.
- VALLERAND, ROBERT J., see Guay, F.
- VALLERAND, ROBERT J., see Senécal, C.
- VANDENBENDEN, REGINA J.J.M., see Blanton, H.
- VAN DEN BERG, AGNES E., see Koole, S. L.
- VAN DEN BOS, KEES, and E. ALLAN LIND, "The Psychology of Own Versus Others' Treatment: Self-Oriented and Other-Oriented Effects on Perceptions of Procedural Justice," 1324.
- VAN KNIPPENBERG, DAAN, see Platow, M. J.
- VAN OVERWALLE, FRANK, and DIRK VAN ROOY, "How One Cause Discounts or Augments Another: A Connectionist Account of Causal Competition," 1613.
- VAN ROOY, DIRK, see Van Overwalle, F.
- VAN VUGT, MARK, "Community Identification Moderating the Impact of Financial Incentives in a Natural Social Dilemma: Water Conservation," 1440.
- VICK, NIKI V., see Cross, S. E.
- VIETOR, NATHANIEL A., see Knee, C.R.
- VLEK, CHARLES A. J., see Koole, S. L.
- VON WEISS, RENEE T., see Kelly, A. E.
- VORAUER, JACQUIE D., and SANDRA M. KUMHYR, "Is This About You or Me? Self-Versus Other-Directed Judgments and Feelings in Response to Intergroup Interaction," 706.
- VRIJ, ALDERT, KATHERINE EDWARD, and RAY BULL, "Stereotypical Verbal and Nonverbal Responses While Deceiving Others," 899.
- WADE, NATHANIEL G., see Berry, J. W.
- WALKIEWICZ, MICHELE, see Madon, S.
- WÄNKE, MICHAELA, HERBERT BLESS, and ERIC R. IGOU, "Next to a Star: Paling, Shining, or Both? Turning Interexemplar Contrast Into Interexemplar Assimilation," 14.
- WATSON, CARL B., MARTIN M. CHEMERS, and NATALYA PREISER, "Collective Efficacy: A Multilevel Analysis," 1057.
- WEARY, GIFFORD, and DARCY A. REICH, "Attributional Effects of Conflicting Chronic and Temporary Outcome Expectancies: A Case of Automatic Comparison and Contrast," 562.
- WEARY, GIFFORD, DARCY A. REICH, and STEPHANIE J. TOBIN, "The Role of Contextual Constraints and Chronic Expectancies on Behavior Categorizations and Dispositional Inferences," 62.

- WEGENER, DUANE T., see Petty, R. E.
 WEINERTH, THOMAS, see Bohner, G.
 WEITLAUF, JULIE C., DANIEL CERVONE, RONALD E. SMITH, and PAUL M. WRIGHT, "Assessing Generalization in Perceived Self-Efficacy: Multidomain and Global Assessments of the Effects of Self-Defense Training for Women," 1683.
 WELBOURNE, JENNIFER L., "Changes in Impression Complexity Over Time and Across Situations," 1071.
 WELLS, GARY L., see Bushman, B. J.
 WIECZORKOWSKA, GRAZYNA, see Crandall, C. S.
 WILSON, BART J., see Kurzban, R.
 WILSON, TIMOTHY D., JAY MEYERS, and DANIEL T. GILBERT, "Lessons From the Past: Do People Learn From Experience That Emotional Reactions Are Short-Lived?" 1648.
 WORTHINGTON, EVERETT L., JR., see Berry, J. W.
 WRIGHT, PAUL M., see Weitlauf, J. C.
 WRIGHT, STEPHEN C., see Tropp, L. R.
 WYER, ROBERT S., JR., see Albarracin, D.
 YBEMA, JAN F., ROELINE G. KUIJER, BRAM P. BUUNK, G. MAJELLA DEJONG, and ROBBERT SANDERMAN, "Depression and Perceptions of Inequity Among Couples Facing Cancer," 3.
 YINON, YOEL, see Bizman, A.
 YU, JIAJUAN, see Flannery, W. P.
 YZERBYT, VINCENT Y., see Corneille, O.
 ZEBROWITZ, LESLIE A., see Andreoletti, C.

Articles:

- "Actions of Similar Others as Inducements to Cooperate in Social Dilemmas," Parks et al., 345.
 "Affective Attitudes Are Not Always Faster: The Moderating Role of Extremity," Giner-Sorolla, 666.
 "Am I as You See Me or Do You See Me as I Am: Self-Fulfilling Prophecies and Self-Verification," Madon et al., 1214.
 "An Attribution-Value Model of Prejudice: Anti-Fat Attitudes in Six Nations," Crandall et al., 30.
 "An Empirical Comparison of Acculturation Models," Flannery et al., 1035.
 "An Introduction to Deviance-Regulation Theory: The Effect of Behavioral Norms on Message Framing," Blanton et al., 848.
 "An Investigation of Three Models of Multitrait Representations," Trafimow and Finlay, 226.
 "Antecedents and Outcomes of Work-Family Conflict: Toward a Motivational Model," Senécal et al., 176.
 "Assessing Generalization in Perceived Self-Efficacy: Multidomain and Global Assessments of the Effects of Self-Defense Training for Women," Weitlauf et al., 1683.
 "The Association of Intimacy Goals and Marital Satisfaction: A Test of Four Mediation Hypotheses," Sanderson and Cantor, 1567.
 "Attachment Security and Marital Satisfaction: The Role of Positive Perceptions and Social Support," Cobb et al., 1131.
 "Attachment Style and Mental Health: A 1-Year Follow-Up Study of Mothers of Infants With Congenital Heart Disease," Berant et al., 956.
 "Attributional Effects of Conflicting Chronic and Temporary Outcome Expectancies: A Case of Automatic Comparison and Contrast," Weary and Reich, 562.
 "Attributions in Marriage: Integrating Specific and Global Evaluations of a Relationship," McNulty and Karney, 943.
 "Attributions of Responsibility and Reactions to Affirmative Action: Affirmative Action as Help," Quinn et al., 321.
 "Autonomy Support, Intrinsic Motivation, and Perceived Competence: Conceptual and Empirical Linkages," Guay et al., 643.
 "Being Committed: Affective, Cognitive, and Conative Components of Relationship Commitment," Arriaga and Agnew, 1190.
 "Changes in Impression Complexity Over Time and Across Situations," Welbourne, 1071.
 "Changing Racial Beliefs by Providing Consensus Information," Stangor et al., 486.
 "Collective Efficacy: A Multilevel Analysis," Watson et al., 1057.

- "Community Identification Moderating the Impact of Financial Incentives in a Natural Social Dilemma: Water Conservation," Van Vugt, 1440.
 "Complementarity of Interpersonal Circumplex Traits," Tracey et al., 786.
 "Conscientiousness and the Theory of Planned Behavior: Toward a More Complete Model of the Antecedents of Intentions and Behavior," Conner and Abraham, 1547.
 "Contemporary Sexism and Discrimination: The Importance of Respect for Men and Women," Jackson et al., 48.
 "Contingent Effects of Anxiety on Message Elaboration and Persuasion," Sengupta and Johar, 139.
 "Counterfactual Thinking and Self-Motives," Sanna et al., 1023.
 "Culture and the Process of Person Perception: Evidence for Automaticity Among East Asians in Correcting for Situational Influences on Behavior," Knowles et al., 1344.
 "Cumulative Meta-Analysis: A Consideration of Indicators of Sufficiency and Stability," Mullen et al., 1450.
 "Day-to-Day Relationships Among Self-Concept Clarity, Self-Esteem, Daily Events, and Mood," Nezlek and Plesko, 201.
 "Depression and Perceptions of Inequity Among Couples Facing Cancer," Ybema et al., 3.
 "Depression as a Moderator of Relationships Between Positive Daily Events and Day-to-Day Psychological Adjustment," Nezlek and Gable, 1692.
 "Devil's Advocate or Advocate of Oneself: Effects of Numerical Support on Pro- and Counterattitudinal Self-Persuasion," Gordijn et al., 395.
 "Dialogue Involvement as a Social Influence Technique," Dolinski et al., 1395.
 "Dispositional Forgiveness: Development and Construct Validity of the Transgression Narrative Test of Forgiveness (TNTF)," Berry et al., 1277.
 "Does Discrimination Terminology Create Response Bias in Questionnaire Studies of Discrimination?" Gomez and Trierweiler, 630.
 "Does Multiple Categorization Reduce Intergroup Bias?" Crisp et al., 76.
 "The Effect of Fleeting Attraction on Compliance to Requests," Burger et al., 1578.
 "The Effects of Entitativity Beliefs on Implicit Comparisons Between Group Members," Pickett, 515.
 "Effects of Environmental Predictability and Personal Mastery on Self-Regulatory and Physiological Processes," Pham et al., 611.
 "The Effects of Sex and Temporal Context on Feelings of Romantic Desire: An Experimental Evaluation of Sexual Strategies Theory," Schmitt et al., 833.
 "The Effects of Trait Type and Situation Type on the Generalization of Trait Expectancies Across Situations," Trafimow, 1463.
 "Elaborative and Nonelaborative Processing of a Behavior-Related Communication," Albarracin and Wyer, 691.
 "Elements of Reciprocity and Social Value Orientation," Parks and Rumble, 1301.
 "Emotional Reactions to a Seemingly Prejudiced Response: The Role of Automatically Activated Racial Attitudes and Motivation to Control Prejudiced Reactions," Fazio and Hilden, 538.
 "Employee Adjustment to an Organizational Merger: An Intergroup Perspective," Terry et al., 267.
 "Environmental and Genetic Influences on Sex-Typed Behaviors and Attitudes of Male and Female Adolescents," Cleveland et al., 1587.
 "Ethnic Identity Moderates Perceptions of Prejudice: Judgments of Personal Versus Group Discrimination and Subtle Versus Blatant Bias," Operario and Fiske, 550.
 "Ethnic and National Stereotypes: The Princeton Trilogy Revisited and Revised," Madon et al., 996.
 "Focus of Comparison as a Determinant of Assimilation Versus Contrast in Social Comparison," Mussweiler, 38.
 "Gender Differences, Motivation, and Empathic Accuracy: When It Pays to Understand," Klein and Hodges, 720.
 "Goals, Culture, and Subjective Well-Being," Oishi and Diener, 1674.

- "Group-Based Emotional Distress: An Extension of Self-Discrepancy Theory," Bizman et al., 1291.
- "The Health Benefits of Writing About Life Goals," King, 798.
- "How One Cause Discounts or Augments Another: A Connectionist Account of Causal Competition," Van Overwalle and Van Rooy, 1613.
- "I Like It But I Don't Know Why: A Value-Account Approach to Implicit Attitude Formation," Betsch et al., 242.
- "I Never Had a Chance: Using Hindsight Tactics to Mitigate Disappointments," Tykocinski, 376.
- "Ideal Standards, the Self, and Flexibility of Ideals in Close Relationships," Campbell et al., 447.
- "Implications of Authoritarianism for Young Adulthood: Longitudinal Analysis of College Experiences and Future Goals," Peterson and Lane, 678.
- "Implicit Self-Concept and Evaluative Implicit Gender Stereotypes: Self and Ingroup Share Desirable Traits," Rudman et al., 1164.
- "Implicit Theories of Relationships: Who Cares If Romantic Partners Are Less Than Ideal?" Knee et al., 808.
- "Incremental Commitment and Reciprocity in a Real-Time Public Goods Game," Kurzban et al., 1662.
- "Inferences About Effort and Ability," Reeder et al., 1225.
- "The Influence of the Achievement Motive on Spontaneous Thoughts in Pre- and Postdecisional Action Phases," Puca and Schmalz, 302.
- "Ingroup Identification as the Inclusion of Ingroup in the Self," Tropp and Wright, 585.
- "Intensity Bias in Affective Forecasting: The Role of Temporal Focus," Buehler and McFarland, 1480.
- "The Interdependent Self-Conceptual and Social Support: The Case of Persistence," Cross and Vick, 820.
- "Is This About You or Me? Self-Versus Other-Directed Judgments and Feelings in Response to Intergroup Interaction," Vorauer and Kumhyr, 706.
- "Knowing Me, Knowing You: Anonymity Effects on Social Identity Processes Within Groups," Lea et al., 526.
- "Lessons From the Past: Do People Learn From Experience That Emotional Reactions Are Short-Lived?" Wilson et al., 1648.
- "The Link Between Leadership and Followership: How Affirming Social Identity Translates Vision Into Action," Haslam and Platow, 1469.
- "Measuring Global Self-Esteem: Construct Validation of a Single-Item Measure and the Rosenberg Self-Esteem Scale," Robins et al., 151.
- "Media Images and Women's Self-Evaluations: Social Context and Importance of Attractiveness as Moderators," Henderson-King et al., 1407.
- "The Mere Exposure Effect: An Uncertainty Reduction Explanation Revisited," Lee, 1255.
- "Mindset and Relationship Illusions: The Moderating Effects of Domain Specificity and Relationship Commitment," Gagné and Lydon, 1144.
- "The Mismeasure of Love: How Self-Doubt Contaminates Relationship Beliefs," Murray et al., 423.
- "Morally Questionable Tactics: Negotiations Between District Attorneys and Public Defenders," Garcia and Darley, 731.
- "Motivation to Think and Order Effects in Persuasion: The Moderating Role of Chunking," Petty et al., 332.
- "Multilevel Random Coefficient Analyses of Event- and Interval-Contingent Data in Social and Personality Psychology Research," Nezlek, 771.
- "Multiple Motives and Persuasive Communication: Creative Elaboration as a Result of Impression Motivation and Accuracy Motivation," Nienhuis et al., 118.
- "Negative Affect Can Increase or Decrease Message Scrutiny: The Affect Interpretation Hypothesis," Bohner and Weinerth, 1417.
- "Narrative Impressions of Literature: The Availability Bias and the Corrective Properties of Meta-Analytic Approaches," Bushman and Wells, 1123.
- "Need Satisfaction, Motivation, and Well-Being in the Work Organizations of a Former Eastern Bloc Country: A Cross-Cultural Study of Self-Determination," Deci et al., 930.
- "Network Correlates of Social and Emotional Loneliness in Young and Older Adults," Green et al., 281.
- "Next to a Star: Paling, Shining, or Both? Turning Interexemplar Contrast Into Interexemplar Assimilation," Wänke et al., 14.
- "On the Many Impacts of Inadmissible Testimony: Selective Compliance, Need for Cognition, and the Overcorrection Bias," Sommers and Kassir, 1368.
- "On the Origins of Racial Attitudes: Correlates of Childhood Experiences," Towles-Schwen and Fazio, 162.
- "On the Social Nature of Personality: Effects of Extraversion, Agreeableness, and Feedback About Collective Resource Use on Cooperation in a Resource Dilemma," Koole et al., 289.
- "On Weakening the Strongest Link: Attributions and Intervention Strategies for Behavior Change," Johnson, 408.
- "On-Line Versus Memory-Based Processing: The Role of 'Need to Evaluate' in Person Perception," Tormala and Petty, 1599.
- "Opposite-Sex Friendship: Sex Differences and Similarities in Initiation, Selection, and Dissolution," Bleske-Rechek and Buss, 1310.
- "Optimism and Attentional Bias for Negative and Positive Stimuli," Segerstrom, 1334.
- "Ostracism and Ego Depletion: The Strains of Silence," Ciarocco et al., 1156.
- "Outgroup Accountability in the Minimal Group Paradigm: Implications for Aversive Discrimination and Social Identity Theory," Dobbs and Crano, 355.
- "Perceived Stigma, Social Support, and Depression," Mickelson, 1046.
- "Perceiving Value in Obligations and Goals: Wanting to Do What Should Be Done," Berg et al., 982.
- "Personality Implications of Romantic Attachment Patterns in Young Adults: A Multi-Method, Multi-Informant Study," Onishi et al., 1097.
- "Pessimistic Bias in Comparative Evaluations: A Case of Perceived Vulnerability to the Effects of Negative Life Events," Blanton et al., 1627.
- "Physical Appearance and Control Beliefs in Young, Middle-Age, and Older Adults," Andreoletti et al., 969.
- "Post Hoc Construction of Self-Performance and Other Performance in Self-Serving Social Comparison," Klein, 744.
- "The Precocity-Longevity Hypothesis: Earlier Peaks in Career Achievement Predict Shorter Lives," McCann, 1429.
- "Predicting Relationship Quality and Emotional Reactions to Stress From Significant-Other-Concept Clarity," Gurung et al., 1267.
- "Projection as an Interpersonal Influence Tactic: The Effects of the Pot Calling the Kettle Black," Rucker and Pratkanis, 1494.
- "Prototype Matching in Close Relationships," Hassebrauck and Aron, 1111.
- "The Psychology of Own Versus Others' Treatment: Self-Oriented and Other-Oriented Effects on Perceptions of Procedural Justice," van den Bos and Lind, 1324.
- "Reacting to Impending Discrimination: Compensation for Prejudice and Attributions to Discrimination," Kaiser and Miller, 1357.
- "Rebels With a Cause: Group Identification as a Response to Perceived Discrimination From the Mainstream," Jetten et al., 1204.
- "Rejection Sensitivity and Depressive Symptoms in Women," Ayduk et al., 868.
- "The Reliability of Implicit Stereotyping," Kawakami and Dovidio, 212.
- "The Role of Ability Judgments in Self-Handicapping," McCrea and Hirt, 1378.
- "The Role of Affect in the Mere Exposure Effect: Evidence From Psychophysiological and Individual Differences Approaches," Harmon-Jones and Allen, 889.
- "The Role of Contextual Constraints and Chronic Expectancies on Behavior Categorizations and Dispositional Inferences," Weary et al., 62.
- "The Role of Self-Referent and Other-Referent Knowledge in Perceptions of Group Characteristics," Krueger and Stanke, 878.
- "Scarcity Claims Elicit Extreme Responding to Persuasive Messages: Role of Cognitive Elaboration," Brannon and Brock, 365.

- "Seeing One's Partner Through Intimacy-Colored Glasses: An Examination of the Processes Underlying the Intimacy Goals-Relationship Satisfaction Link," Sanderson and Evans, 463.
- "Selective Memory and the Persistence of Paradoxical Self-Esteem," Tafarodi et al., 1179.
- "Self-Enhancement: Is It Restricted to Individualistic Cultures?" Kurman, 1705.
- "Self-Esteem and Emotion: Some Thoughts About Feelings," Brown and Marshall, 575.
- "Self-Reported Processing Bias and Naturally Occurring Mood: Mediators Between Personality and Stress Appraisals," Hemenover, 387.
- "Shame and Embarrassment Revisited," Sabini et al., 104.
- "Simulation, Scenarios, and Emotional Appraisal: Testing the Convergence of Real and Imagined Reactions to Emotional Stimuli," Robinson and Clore, 1520.
- "Social Anxiety and the Cued Activation of Relational Knowledge," Baldwin and Main, 1637.
- "A Social Identity Analysis of Leadership Endorsement: The Effects of Leader Ingroup Prototypicality and Distributive Intergroup Fairness," Platow and van Knippenberg, 1508.
- "Social Influence in Computer-Mediated Communication: The Effects of Anonymity on Group Behavior," Postmes et al., 1243.
- "Social Norms, Entitlement, and Deservingness: Differential Reactions to Aggressive Behavior of Schizophrenic and Personality Disorder Patients," Feather and Johnstone, 755.
- "The Sounds and Sights of Intelligence: A Lens Model Channel Analysis," Reynolds and Gifford, 187.
- "Spontaneous Inferences From Pictorially Presented Behaviors," Fiedler and Schenck, 1533.
- "Stereotypical Verbal and Nonverbal Responses While Deceiving Others," Vrij et al., 899.
- "Stop Complaining! The Social Costs of Making Attributions to Discrimination," Kaiser and Miller, 254.
- "'That's Completely Obvious . . . and Important': Lay Judgments of Social Psychological Findings," Richard et al., 497.
- "Threat and the Group Attribution Error: When Threat Elicits Judgments of Extremity and Homogeneity," Corneille et al., 437.
- "Validation of Personal Identity as a Terror Management Mechanism: Evidence That Sex-Role Identity Moderates Mortality Salience Effects," Florian et al., 1011.
- "Vengefulness: Relationships With Forgiveness, Rumination, Well-Being, and the Big Five," McCullough et al., 601.
- "What Is It About Revealing Secrets That Is Beneficial?" Kelly et al., 651.
- "What You See Is What You Get: Systematic Variability in Perceptual-Based Social Judgment," Livingston, 1086.
- "When Accuracy-Motivated Perceivers Fail: Limited Attentional Resources and the Reemerging Self-Fulfilling Prophecy," Biesanz et al., 621.
- "When Bad Things Turn Good and Good Things Turn Bad: Sequences of Redemption and Contamination in Life Narrative and Their Relation to Psychosocial Adaptation in Midlife Adults and in Students," McAdams et al., 474.
- "When Saying Yes Leads to Saying No: Preference for Consistency and the Reverse Foot-in-the-Door Effect," Guadagno et al., 859.
- "When Juries Fail to Comply With the Law: Biased Evidence Processing in Individual and Group Decision Making," Sommer et al., 309.
- "When We Wonder What It All Means: Interpretation Goals Facilitate Accessibility and Stereotyping Effects," Stapel and Koomen, 915.
- "'Who am I?' The Cultural Psychology of the Conceptual Self," Kanagawa et al., 90.