THE PSYCHOLOGY OF SELF-DEFENSE: SELF-AFFIRMATION THEORY

David K. Sherman
Geoffrey L. Cohen

I. Introduction

In major league baseball, a hitter could have a long and productive career by maintaining a .300 average, that is, by getting a base hit 30% of the time. A great deal of money could be earned and fame accrued. Yet the other 70% of the time, this player would have failed. The vast majority of attempts to hit the ball would result in "making an out" and thus pose a potential threat to the player's sense of personal worth and social regard.

Like major league baseball players, people in contemporary society face innumerable failures and self-threats. These include substandard performance on the job or in class, frustrated goals or aspirations, information challenging the validity of long-held beliefs, illness, the defeat of one's political party in an election or of one's favorite sports team in a playoff, scientific evidence suggesting that one is engaging in risky health behavior, negative feedback at work or in school, rejection in a romantic relationship, real and perceived social slights, interpersonal and intergroup conflict, the misbehavior of one's child, the loss of a loved one, and so on. In the course of a given day, the potential number of events that could threaten people's "moral and adaptive adequacy"—their sense of themselves as good, virtuous, successful, and able to control important life outcomes (Steele, 1988)—seems limitless and likely to exceed the small number of events that affirm it. A major undertaking for most people is to sustain self-integrity when faced with the inevitable setbacks and disappointments of daily life—the 70% of the time

ADVANCES IN EXPERIMENTAL
SOCIAL PSYCHOLOGY, VOL. 38
DOI: 10.1016/S0065-2601(06)38004-5

Copyright 2006, Elsevier Inc.
All rights reserved.
0065-2601/06 $35.00

"at bat" when they do not get a base hit. How do individuals adapt to such threats and defend self-integrity?

Much research suggests that people have a "psychological immune system" that initiates protective adaptations when an actual or impending threat is perceived (Gilbert, Pinel, Wilson, Blumberg, & Wheatley, 1998). Psychological adaptations to threats include the various cognitive strategies and even distortions whereby people come to construe a situation in a manner that renders it less threatening to personal worth and well-being. Many of these psychological adaptations can be thought of as defensive in nature, insofar as they alter the meaning of the event in a way that shields people from the conclusion that their beliefs or actions were misguided. Psychologists have documented a wide array of such psychological adaptations that help people to protect their self-integrity in response to threat.

Indeed, defensive adaptations are so stubborn and pervasive that Greenwald (1980) described the ego as "totalitarian" in its ambition to interpret the past and present in a way congenial to its desires and needs. People view themselves as a potent causal agent even over events that they cannot control (Langer, 1975); they view themselves as selectively responsible for producing positive rather than negative outcomes (Greenwald, 1980; Miller & Ross, 1975; Taylor, 1983). They resist change or—if they do change—become more extreme versions of what they were before (Lord, Ross, & Lepper, 1979). People dismiss health information suggesting that they are at risk for disease or should change their risky behavior (Jemmott, Ditto, & Croyle, 1986; Kunda, 1987). Students may disidentify with, or downplay the personal importance of, domains where they fail, thus sustaining self-worth but precluding the opportunity for improvement (Major, Spencer, Schmader, Wolfe, & Crocker, 1998; Steele, 1997). People are overoptimistic in their predictions of future success and estimations of their current knowledge and competence (Dunning, Griffin, Milojkovic, & Ross, 1990; Kruger & Dunning, 1999). Indeed, these defensive adaptations may even benefit psychological and physical health (Taylor & Brown, 1988). Although we suspect that people can be more realistic and more self-critical than this research suggests, and that their optimism and positive illusions may be magnified in certain contexts rather than others (see Armor & Taylor, 2002), the idea that people are ego defensive resonates both with psychological research and lay wisdom. An important question, then, concerns the circumstances under which people are less ego defensive and more open-minded in their relationship with the social world.

We see defensive responses as adaptations aimed at ameliorating threats to self-integrity. The vast research on defensive biases testifies to their robustness and to the frequency with which people use them. Although these defensive responses are adaptive in the sense of protecting or enhancing an individual's sense of self-integrity, they can be maladaptive to the extent they forestall learning from important, though threatening, experiences and

information. Moreover, peoples' efforts to protect self-integrity may threaten the integrity of their relationships with others (Cohen et al., 2005; Murray, Holmes, MacDonald, & Ellsworth, 1998). Yet, these normal adaptations can be ''turned off'' through an altogether different psychological adaptation to threat, an alternative adaptation that does not hinge on distorting the threatening event to render it less significant. One way that these defensive adaptations can be reduced, or even eliminated, is through the process of self-affirmation (Aronson, Cohen, & Nail, 1999; Sherman & Cohen, 2002; Steele, 1988).

Steele (1988) first proposed the theory of self-affirmation. It asserts that the overall goal of the self-system is to protect an image of its self-integrity, of its moral and adaptive adequacy. When this image of self-integrity is threatened, people respond in such a way as to restore self-worth. As noted previously, one way that this is accomplished is through defensive responses that directly reduce the threat. But another way is through the affirmation of alternative sources of self-integrity. Such ''self-affirmations,'' by fulfilling the need to protect self-integrity in the face of threat, can enable people to deal with threatening events and information without resorting to defensive biases.

In this paper, we update the field on research conducted using self-affirmation theory as a framework. This research illuminates both the motivational processes underlying self-integrity maintenance and the implications of such processes for many domains of psychology. We illustrate how self-affirmation affects not only people's cognitive responses to threatening information and events, but also their physiological adaptations and actual behavior. The research presented has implications for psychological and physical health, education, social conflict, closemindedness and resistance to change, prejudice and discrimination, and a variety of other important applied areas. We also examine how self-affirmations reduce threats to the self at the collective level, such as when people confront threatening information about their groups. We then review factors that qualify or limit the effectiveness of self-affirmations, including situations where affirmations backfire, and lead to greater defensiveness and discrimination. We discuss the connection of self-affirmation theory to other motivational theories of self-defense and review relevant theoretical and empirical advances. We conclude with a discussion of the implications of self-affirmation theory for interpersonal relationships and coping.

A. OVERVIEW OF SELF-AFFIRMATION THEORY

Self-affirmation theory (Aronson et al., 1999; Sherman & Cohen, 2002; Steele, 1988) begins with the premise that people are motivated to maintain the integrity of the self. Integrity can be defined as the sense that, on the

whole, one is a good and appropriate person. Cultural anthropologists use the term "appropriate" to refer to behavior that is fitting or suitable given the cultural norms and the salient demands on people within that culture. Thus, the standards for what it means to be a good person vary across cultures, groups, and situations (e.g., Heine, 2005). Such standards of integrity can include the importance of being intelligent, rational, independent, and autonomous, and exerting control over important outcomes. Such standards of integrity can also include the importance of being a good group member and of maintaining close relationships. Threats to self-integrity may thus take many forms but they will always involve real and perceived failures to meet culturally or socially significant standards (Leary & Baumeister, 2000). Consequently, people are vigilant to events and information that call their self-integrity into question, both in their own eyes and in the eyes of others. In such situations, people try to restore or reassert the integrity of the self. Thus, the goal of protecting self-integrity, and the impact of that goal on psychology and behavior, becomes apparent when integrity is threatened.

There are three categories of responses that people deploy to cope with such threats. First, they can respond by accommodating to the threat. That is, they can accept the failure or the threatening information and then use it as a basis for attitudinal and behavioral change. However, to the extent that the threatened domain concerns an important part of one's identity, the need to maintain self-integrity can make it difficult to accept the threatening information and to change one's attitude or behavior accordingly. A second response thus involves ameliorating the threat via *direct* psychological adaptations. While some direct adaptations preserve the fundamental informational value of the event while also changing one's construal of that event (e.g., framing a failure as a learning opportunity; Dweck & Leggett, 1988), other direct psychological adaptations are defensive in nature in that they involve dismissing, denying, or avoiding the threat in some way. We refer to these responses as defensive biases (see Sherman & Cohen, 2002). Although a defensive bias can restore self-integrity, the rejection of the threatening information can lessen the probability that the person will learn from the potentially important information.

Self-affirmation theory proposes a third alternative, a different kind of psychological adaptation—one that, under many circumstances, enables both the restoration of self-integrity and adaptive behavior change. People can respond to threats using the *indirect* psychological adaptation of affirming alternative self-resources unrelated to the provoking threat. Such "self-affirmations" include reflecting on important aspects of one's life irrelevant to the threat, or engaging in an activity that makes salient important values unconnected to the threatening event. Whereas defensive psychological adaptations directly address the threatening information, indirect psychological

adaptations, such as self-affirmation, allow people to focus on domains of self-integrity unrelated to the threat. When self-affirmed in this manner, people realize that their self-worth does not hinge on the evaluative implications of the immediate situation. As a result, they have less need to distort or reconstrue the provoking threat and can respond to the threatening information in a more open and evenhanded manner.

B. BASIC TENETS OF SELF-AFFIRMATION THEORY

Much research within the self-affirmation framework examines whether an affirmation of self-integrity, unrelated to a specific provoking threat, can attenuate or eliminate people's normal response to that threat. If it does, then one can infer that the response was motivated by a desire to protect self-integrity. The self-affirmation framework encompasses four tenets, which are enumerated below:

1. People are Motivated to Protect the Perceived Integrity and Worth of the Self

The most basic tenet of self-affirmation theory (Steele, 1988) is that people are motivated to protect the perceived integrity and worth of the self. As Steele observed, the purpose of the self-system is to "maintain a phenomenal experience of the self ... as adaptively and morally adequate, that is, competent, good, coherent, unitary, stable, capable of free choice, capable of controlling important outcomes ... (p. 262)." These self-conceptions and images making up the self-system can be thought of as the different domains that are important to an individual, or the different contingencies of a person's self-worth (Crocker & Wolfe, 2001). Figure 1 presents a schematic of the self-system. The self is composed of different domains, which include an individual's roles, such as being a student or a parent; values, such as being religious or having a sense of humor; social identities, such as membership in groups or organizations and in racial, cultural, and gender groups; and belief systems, such as political ideologies. The self is also composed of people's goals, such as the value of being healthy or succeeding in school. The self-system is activated when a person experiences a threat to an important self-conception or image. Such threat poses a challenge to a desired self-conception. Thus, failure feedback could threaten a person's identity as a student, negative health information could threaten a person's self-conception as a healthful individual, news about anti-American sentiment could threaten a person's patriotic identity, and evidence of social inequality could challenge a person's belief in a just world (Lerner, 1980). All of these
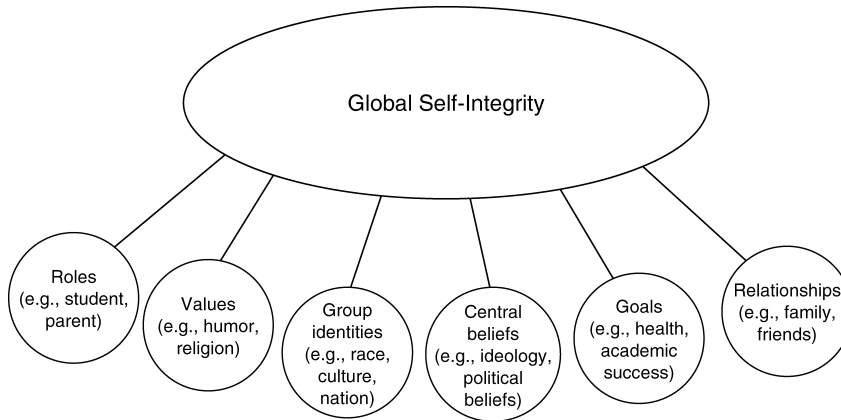
Fig. 1.   Schematic representation of self-system.

events are threatening because they have implications for a person's overall sense of self-integrity.

2. Motivations to Protect Self-Integrity can Result in Defensive Responses

When self-integrity is threatened, people are motivated to repair it, and this motivation can lead to defensive responses. The defensive responses may seem rational and defensible, though they are more "rationalizing" than "rational" (Aronson, 1968; Kunda, 1990; Pyszczynski & Greenberg, 1987). They serve to diminish the threat and consequently, restore the perceived integrity of the self. These defensive responses can be automatic and even unconscious in nature, and indeed, the rapidity with which people respond to threats speaks to the importance of self-integrity maintenance.

3. The Self-System is Flexible

People often compensate for failures in one aspect of their lives by empha-sizing successes in other domains. Personality theorists, such as Allport (1961) and Murphy (1947), have advanced this notion of compensation, and self-affirmation theory is consistent with this claim (see also Brown & Smart, 1991). Because the goal of the self-system focuses on maintaining the overall worth and integrity of the self, people can respond to threats in one domain by affirming the self in another domain. This fungibility in the sources of self-integrity is what can enable smokers, for example, to maintain a perception of worth and integrity despite the potentially threatening conclusion that they are acting in a maladaptive, harmful, and irrational

way (Steele, 1988). Affirmations satisfy the motivation to maintain self-integrity—thus, they reduce the normal psychological adaptations people engage to ameliorate a specific provoking threat.

4. People can be Affirmed by Engaging in Activities that Remind them of "Who They Are" (And Doing so Reduces the Implications for Self-Integrity of Threatening Events)

Those qualities that are central to how people see themselves are potential domains of self-affirmation. Such affirmations can concern friends and family, making art or music, a charity, or the observance of one's religion. In a difficult situation, reminders of these core qualities can provide people with perspective on who they are and anchor their sense of self-integrity in the face of threat. A "self-affirmation" makes salient one of these important core qualities or sources of identity. Operationally, self-affirmations are typically ideographic, in that people first report an important value or life domain, and then they are given the opportunity either to write an essay about it or to complete a scale or exercise that allows them to assert its importance (McQueen & Klein, 2005).

When global perceptions of self-integrity are affirmed, otherwise threatening events or information lose their self-threatening capacity because the individual can view them within a broader, larger view of the self. People can thus focus not on the implications for self-integrity of a given threat or stressor, but on its informational value. When self-affirmed, individuals feel as though the task of proving their worth, both to themselves and to others, is "settled." As a consequence, they can focus on other salient demands in the situation beyond ego protection.

II. Self-Affirmation and Threats to the Individual Self

A great deal of research has used self-affirmation theory to address a wide range of psychological phenomena, including biased information processing, causal attributions, cognitive dissonance, prejudice and stereotyping, stress and rumination. What connects these disparate areas of research is that they all address situations or events where people contend with a threat to a valued self-image. We first review research on the impact of experienced threats to self-identities, such as one's political identity or one's identity as a healthful or intelligent individual. We then review research on the impact of experienced threats to collective identities, such as one's team or racial group.

A. MOTIVATED INFERENCES AND BIASED ASSIMILATION

People often interpret new information in a way that reinforces their beliefs and desires. One striking example occurred during the 2000 presidential controversy in Florida. Evaluations of seemingly arbitrary events (e.g., beliefs about appropriate policies for dealing with hanging chads, whether to count votes from certain counties) were consistently and predictably aligned with people's political identities. The Florida situation presented a series of novel issues where partisanship lines had not, at first, been clearly defined. Yet, in little time, people were able to determine their criteria for establishing voter intent, and in most cases, their criteria were consistent with their partisanship, and the position that they ultimately advocated served the interests of their candidate of choice (cf. Gerber & Green, 1999).

It is interesting to imagine what would have happened during the Florida debacle if the situation were reversed, and if Vice President Albert Gore were leading rather than Governor George W. Bush as the vote was contested. That is, if the situation was "counterbalanced" like a proper psychology experiment, would Republican leaders have argued for the sanctity of the vote and the need to count each chad and would Democratic leaders have argued against them? Research on biased assimilation by Lord et al. (1979) suggests that they would. In this classic study, proponents and opponents of capital punishment evaluated two studies on the efficacy of capital punishment as a deterrent. The two studies featured different designs (a panel design comparing murder rates from states before and after the implementation of capital punishment policy, and a concurrent design comparing murder rates from states that either used capital punishment or did not). Unlike the situation in Florida, the study was completely counterbalanced. That is, two versions of each study were created, one that supported and one that refuted the deterrent value of capital punishment. Thus, the study allows an examination of whether and to what extent prior beliefs (or identities) bias the interpretation of information.

When the study was consistent with participants' prior beliefs, they thought it was better conducted and more convincing regardless of the specific design (concurrent versus panel) of the study. Moreover, reading a mixed bag of evidence—one study supporting and one study contradicting their beliefs—led participants to report becoming *more* confident in the validity of their beliefs about the deterrent efficacy of the death penalty. The study provides an experimental analog to the hypothetical switch in the Florida election scenario. Regardless of what evidence people were presented, they evaluated it in a way consistent with their prior beliefs and, in turn, this biased assimilation appeared to strengthen the very prior beliefs that give rise to the bias in evaluation.

Lord et al. (1979) emphasized the role of cool cognitive inferential processes in producing biased assimilation. Like Bayesian theorists, people use their prior beliefs to evaluate the validity of incoming data. It is thus logically permissible for them to reject belief-incongruent evidence and to accept with little scrutiny belief-congruent evidence (Nisbett & Ross, 1980). In contrast, we argue that such biases issue from the motivation to maintain and protect identity (see also Munro & Ditto, 1997). That is, partisans in the Lord et al. (1979) study scrutinized the evidence in a way that would protect their identity either as a "law and order" conservative or as a "humanitarian" liberal (Cohen, 2003; Ellsworth & Ross, 1983).

According to our analysis, the need to protect a valued identity or self-view is a major source of such biased processing and closed mindedness. Because long-held beliefs are often tied to important identities, they may be given up only with great reluctance, and they may be embraced even when they conflict with the demands of fact, logic, or material self-interest (e.g., Abelson, 1986; Sears & Funk, 1991). Yet, people possess other important identities and values that they can draw on when they encounter belief-threatening information. Providing them with an affirmation of one of these alternative sources of self-integrity should enable them to evaluate the threatening information in a less biased and defensive manner. We tested this logic in a study of biased assimilation in the domain of capital punishment (Cohen, Aronson, & Steele, 2000). In a study patterned after Lord et al. (1979), proponents and opponents of capital punishment read an article from the *Journal of Law and Human Behavior* titled "The Death Penalty: New Evidence Informs an Old Debate." The article was fabricated, but contained facts, statistics, and arguments that waged a persuasive assault on participants' attitudes toward capital punishment. Thus, proponents of capital punishment read an anti-capital punishment report and opponents of capital punishment read a pro-capital punishment report.

Prior to reading the article on capital punishment, all participants completed a writing exercise that constituted our self-affirmation manipulation. Participants in the self-affirmation condition wrote an essay about a personal value that they had rated, during pretest, as personally important (such as their relationships with friends or sense of humor). Specifically, they were asked to describe three to four personal experiences where the value had been important to them and had made them feel good about themselves. The value they wrote about was, in all cases, unrelated to their political views. Participants in the no-affirmation condition wrote about a neutral topic.

In the no-affirmation control condition, our findings mirrored those of Lord et al. (1979). Participants found flaws in the methodology of the studies that contradicted their political beliefs, they suspected bias on the part of the authors of the report, and they persisted in their attitudes about capital

punishment. By contrast, the responses of participants who affirmed a valued self-identity proved more balanced. That is, self-affirmed participants were less critical of the reported research and suspected less bias on the part of the authors of the report. Participants even changed their global attitudes toward capital punishment in the direction of the report they read. That is, proponents of capital punishment supported the death penalty less, and opponents of capital punishment supported it more (Cohen et al., 2000, Study 2). That both partisan groups showed the effect attests to the power of the psychological mechanism. Not only did proponents come to privilege "life" over "law and order" more when self-affirmed than when not, but opponents of capital punishment, when self-affirmed, came to support state-sanctioned execution to a greater extent (see also Jacks & O'Brien, 2004).

A third study examined how people evaluate others who either agree or disagree with their beliefs (Cohen et al., 2000). Pro-choice and pro-life partisans were presented with a debate between two activists on opposite sides of the abortion dispute. Participants who did not receive a self-affirmation judged the activist who shared their convictions more favorably than the activist who did not (see also Hastorf & Cantril, 1954). In contrast, participants who were given a self-affirmation became more balanced in their evaluation of the two activists and asserted that they were relatively less confident of the validity of their abortion attitudes, relative to their nonaffirmed peers.

The motivation to protect identity may prove especially consequential in the context of negotiation. In negotiation—between parties, nations, individuals—there exists a barrier to compromise, a barrier that often leads disputing factions to reject even mutually beneficial settlements and instead persist in mutually destructive conflict (Ross & Ward, 1995; Sherman, Nelson, & Ross, 2003). This barrier issues, in part, from a motivation to defend one's political, national, or regional identity—a motivation that can result in intransigence and stalemate (De Dreu & Carnevale, 2003). To accept compromise entails acting in a way that could exact painful costs to self and social identity. Compromise entails that one must accept policies anathema to one's past ideological commitments, or take courses of action contrary to the interests of one's constituencies and group loyalties. Accordingly, negotiation should be facilitated when people's partisan identities are made less vital to their sense of self-integrity through the affirmation of alternative sources of self-worth.

In one study, pro-choice participants entered into a negotiation with a pro-life advocate about appropriate federal abortion policy (Cohen et al., 2005). To make their partisan identity salient prior to the negotiation, participants were first asked to assert their "true beliefs" on abortion policy. Additionally, prior to the negotiation, half of the participants received a values affirmation and half did not. It was found that affirmation increased

the number of concessions that pro-choice participants made to their pro-life adversary (e.g., they were more likely to agree to the idea of parental notification). More dramatically, affirmation also led participants to evaluate their adversary as more objective and trustworthy (i.e., as less influenced by self-interest and ideology)—a finding with clear implications for real-world negotiation where the cultivation of trust is a critical step in the resolution of conflict.

These studies raise an important question as to *why* people are more open-minded when affirmed than when not. Correll, Spencer, and Zanna (2004) proposed that affirmation leads to greater attitude change via a more careful consideration of the arguments, rather than through more superficial, heuristic processing. That is, affirmation does not lead to change through more peripheral or ''mindless'' routes (e.g., by raising mood and thus promoting agreeableness) but through the more central route of more balanced information processing (Petty & Cacioppo, 1986; Petty & Wegener, 1998). These researchers presented people with pro- and counter-attitudinal arguments that varied in quality. According to their analysis, what should prove most threatening is strong evidence against one's position and weak evidence in favor of one's position. Thus, self-affirmation should increase openness only to *strong* evidence against one's position and increase rejection of *weak* evidence in favor of one's position. This is what Correll et al. (2004) found. Self-affirmation led people to be more responsive to the intrinsic strength of the arguments rather than the concordance of those arguments with their prior beliefs. Also consistent with the claim that self-affirmation helps reduce the bias of personal beliefs in processing information, Cohen et al. (2005) found that self-affirmed participants exhibited more balanced cognitive and affective responses to a counter-attitudinal report than did their non-affirmed peers.

B. THREATENING HEALTH INFORMATION

Defensive processing can be particularly costly when it leads people to reject important health information. Individuals often face information suggesting that they are engaging in behavior that puts their health at risk. Health information can threaten the self by suggesting that people have acted unwisely, for example, by smoking, drinking, or practicing unsafe sex. Although it would be optimal if people responded to personal health information by ceasing risky behavior, people often resist threatening health messages, and subsequently, persist in unhealthful or risky behaviors. For example, one study found that sexually active students who saw an AIDS educational message responded by *lowering* their perceived risk for sexually

transmitted diseases—a defensive response (Morris & Swann, 1996). Several studies have found that when a health message is of high personal relevance, people are more likely to scrutinize the message for fault than when a message is of no special relevance (Ditto & Lopez, 1992; Kunda, 1987; Liberman & Chaiken, 1992). People use different criteria or probative thresholds when evaluating the validity of information that either supports or contradicts their desires (Ditto, Scepansky, Munro, Apanovitch, & Lockhart, 1998). That is, they use a "must I believe this information?" criterion when evaluating evidence that runs contrary to their preexisting health beliefs, whereas they use a "can I believe this information?" criterion when evaluating evidence consistent with those beliefs (Dawson, Gilovich, & Regan, 2002).

People are biased in their assessments of threatening health information, we argue, because being a "healthy person" is an important part of how they see themselves. Making a behavioral change in a positive direction carries the self-evaluative burden of acknowledging that one has engaged in maladaptive behaviors in the past. Thus, out of a desire to protect a self-image of oneself as rational and healthy, people may ultimately persevere in irrational and unhealthful behavior. However, the logic of self-affirmation theory suggests that if individuals can reflect on an alternative source of identity, their overall sense of self-integrity will be secured, and they may be more apt to consider information without resorting to defensive biases.

Sherman, Nelson, and Steele (2000) examined defensive responses to threatening health information in the context of breast-cancer prevention. Participants were women who were either coffee drinkers or noncoffee drinkers, and they reviewed a (fabricated) scientific report linking caffeine consumption to fibrocystic disease, a precursor to breast cancer. The report concluded by suggesting that women can reduce their risk for fibrocystic disease by reducing their caffeine consumption. Those in the no-affirmation control condition exhibited the pattern found in earlier research (Kunda, 1987; Liberman & Chaiken, 1992): the coffee-drinking women were more critical of the scientific article and more resistant to the message than were the noncoffee drinking women (see Fig. 2). By contrast, coffee-drinking women who had completed a scale that enabled them to assert the personal importance of a central value (e.g., their religious or political values) proved more open to the message than any other group and intended to reduce their coffee drinking accordingly. Because the motivation to protect self-worth was satisfied via this self-affirmation, people who would have otherwise felt threatened by the health message proved more open and more willing to engage in adaptive behavior change.

Other researchers have also documented the "de-biasing" effects of self-affirmation. In one study by Reed and Aspinwall (1998), heavy and light
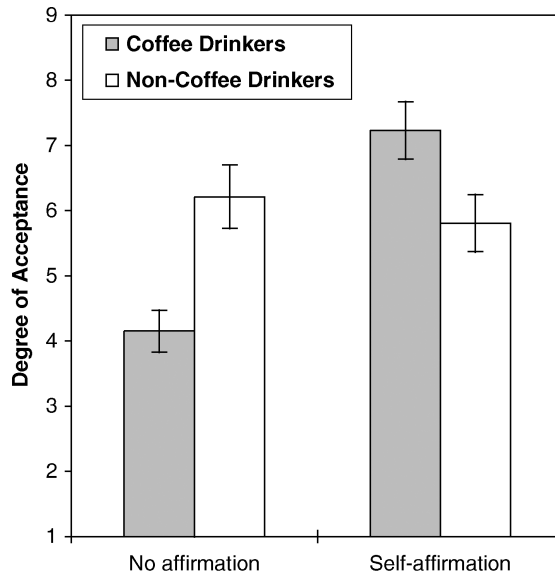
Fig. 2.   Acceptance of article's conclusions as a function of coffee-drinking status and affirmation status. From Sherman, D. A. K., Nelson, L. D., & Steele, C. M. (2000). Do messages about health risks threaten the self? Increasing the acceptance of threatening health messages via self-affirmation. *Personality and Social Psychology Bulletin, 26*, 1046–1058. Adapted with permission.

caffeine consuming women were given information about the link between caffeine consumption and fibrocystic disease. Half of the participants were then given the opportunity to affirm the self by reflecting on acts of kindness that they had performed. The heavy caffeine consumers who were affirmed oriented more rapidly to the risk-confirming information, that is, they looked at this information more quickly and viewed it as more convincing. Furthermore, the affirmed heavy caffeine consumers recalled less risk-disconfirming information at a one-week follow-up.

    If self-affirmation can reduce the biased processing of health information, does it have implications for actual behavior? One study examined whether affirmation could lead to change in health behavior (Sherman et al., 2000). Sexually active undergraduates watched an AIDS-educational video suggesting that their sexual behavior could put them at risk for HIV. Half received a self-affirmation prior to watching the video; the others did not. Although non-affirmed participants tended to resist the presented information (maintaining their perceived risk from pretest levels), affirmed participants responded by increasing their perceived potential risk for contracting AIDS. The affirmation not only affected perceived risk but also affected subsequent health behavior. Whereas 25% of nonaffirmed individuals

purchased condoms after viewing the video, 50% of affirmed participants did so. Additionally, participants who completed the self-affirmation were more likely to take an AIDS-educational brochure (78% did so) than participants in the no-affirmation condition (54% did so). The self-affirmation seemed to buffer individuals from the potential threat of the AIDS-educational message, opened them to the possibility that they were at risk, and motivated them to undertake preventive behaviors.

Do self-affirmation interventions produce long-term effects on increasing the acceptance of threatening health messages? For the effects to have practical significance, they must persist beyond the experimental setting. From a theoretical standpoint, if self-affirmation promotes attitude change via more central route processing (as was found in the Correll et al., 2004 study), then the effects of self-affirmation on attitude change should be durable rather than short-lived. To date, few studies have examined the longer-term effects of self-affirmation.

Harris and Napper (2005) specifically examined the longer-term effects of self-affirmation on health persuasion. Young women (light versus heavy alcohol consumers) were given a pamphlet describing the risks of developing breast cancer from drinking too much alcohol. The pamphlet informed young women that excessive drinking can lead to breast cancer and is thus dangerous. Prior to reading the pamphlet, all participants completed a self-affirmation wherein they either wrote an essay about their most important value or a relatively unimportant value. The researchers assessed perceived risk at the experimental session as well as at one week and one month after the experimental session.

Among the participants most at risk for health problems—that is, the women who drank heavily—those who completed the self-affirmation saw themselves as being at greater risk for breast cancer than those who were not affirmed (as in Sherman et al., 2000), an effect that did not diminish over time. The high-risk self–affirmed participants also reported that they found it easier to imagine themselves with breast cancer—an effect that persisted one month later. Finally, the high-risk participants who completed the self-affirmation had stronger intentions to reduce their alcohol consumption at the experimental session, although this was not associated with more positive health behaviors in the form of drinking reduction during the following month. Thus, the effects of the self-affirmation proved durable (1 month) at affecting risk perceptions, although the long-term effects on behavior have not been established.

An interesting question concerns the factors that facilitate long-term behavioral effects. Long-term effects will be enhanced, we suspect, when situational cues remind individuals of the change in attitude and identity that a persuasive message had engendered in them. For example, Dal Cin, MacDonald, Fong, Zanna, and Elton-Marshall (in press) provided evidence

that an information campaign to reduce unsafe sexual practices could have long-term effects on condom use (as assessed up to 8 weeks later) if participants are given an ''identity cue'' to wear that would remind them both of the ''safe sex'' persuasive video they had previously viewed and to their commitment to the cause of promoting safe sexual practices. Specifically, the impact of the safer sex message on behavior was greatly increased if participants were given a ''friendship bracelet'' that served to affirm their personal concern for people suffering from sexually transmitted diseases (Dal Cin et al., in press). Pairing such identity cues with self-affirmation constitutes an exciting direction for future research.

C. STRESS

If self-affirmation enables people to view otherwise threatening events, such as negative health information, as less threatening, then one intriguing possibility is that it could also reduce evaluative stress. Stress is the process by which environmental events are appraised as threatening, which in turn elicits emotional and physiological responses that can adversely affect health and increase susceptibility to disease (Lazarus, 1993; McEwen, 1998). As Keough, Garcia, and Steele (1998) suggested, stress may arise, in large part, from threats to the perceived worth and integrity of the self (Creswell, et al., 2005; Keough, 1998). Consistent with this notion, common laboratory methods for inducing stress, such as the Trier Social Stress Task, have participants deliver a personally relevant speech and perform difficult mental arithmetic in front of a hostile audience (Kirschbaum, Pirke, & Hellhammer, 1993). These tasks have reliably been shown to threaten the self, induce appraisals of threat, and elicit the stress hormone cortisol (Dickerson & Kemeny, 2004).

The self-affirmation analysis also suggests that stress can be ameliorated through the affirmation of alternative sources of self-integrity. Specifically, self-affirmations can secure self-worth in a domain unrelated to the stressor (Creswell et al., 2005; Keough, 1998), and may, as a result, encourage people to view the stressful event without feeling as though their self integrity is in question. A study examining self-enhancement and physiological responses to stress provides suggestive evidence for this effect. Those who saw themselves as being above average across a number of domains responded to a stressful situation (mental arithmetic challenge in front of a hostile audience) with reduced blood pressure and heart rate, relative to those who were less self-enhancing (Taylor, Lerner, Sherman, Sage, & McDowell, 2003). This finding suggests that those who perceive themselves as having more resources and abilities experience a potentially threatening task in a less physiologically taxing manner because their overall feelings of self-integrity

hinge less on their performance on that particular task (see also Seery, Blascovich, Weisbuch, & Vick, 2004).

An influential model of how to conceptualize psychological resources in stressful situations was advanced by Hobfoll (1989). He defined resources as "those objects, personal characteristics, conditions, or energies that are valued by the individual or that serve as a means for attainment of these objects, personal characteristics, and energies (p. 516)." Psychological resources can thus encompass aspects of self that are unrelated to a given threat or stressor. If this is true, then enhancing the perception of self-resources via self-affirmation could reduce the physiological costs of stress. Consistent with this work, one study found that self-affirmation reduced perceived stress during a difficult serial subtraction task (Keough et al., 1998).

To test the stress-buffering effects of self-affirmation directly, one study had participants complete a self-affirmation procedure—a values scale for their most important value (versus an unimportant value in the control condition)—prior to engaging in the Trier Social Stress Task (Creswell et al., 2005). Thus, those in the self-affirmation condition indicated their agreement with an important value (e.g., religion) in a domain of their life unrelated to the stressful task. Participants were instructed to prepare a speech to be given to two speech evaluators concerning why they would be qualified for a job as an administrative assistant in a psychology department. The speech evaluators (i.e., confederates who were trained to act in a sullen manner) directed the 5-minute speech task and then asked the participant to complete a 5-minute mental arithmetic task by counting aloud back-wards from 2083 by 13's (a task where self-affirmations have been success-ful at improving performance; Schimel, Arndt, Banko, & Cook, 2004). Salivary cortisol measurements were assessed at baseline and at 20, 30, and 45 minutes poststress onset.

As displayed in Fig. 3, during baseline there were no differences in cortisol levels between participants in the self-affirmation and control conditions. Yet, at the 20-minute measurement, and persisting through 45 minutes after the stress task, the participants in the control condition had elevated cortisol levels, whereas the self-affirmation participants did not (Creswell et al., 2005). This finding provides the first experimental evidence that a self-affirmation manipulation can buffer the self not only at a psychological level, but also at a physiological level.

Threatening experiences are not only stressful while they occur, but also after the fact, as people engage in ruminative thinking, and the thoughts of the threatening event reoccur (Martin & Tesser, 1996; Nolen-Hoeksema & Morrow, 1991). One study examined whether this ruminative thinking would be reduced among participants who completed a self-affirmation (Koole, Smeets, van Knippenberg, & Dijksterhuis, 1999). In this research,
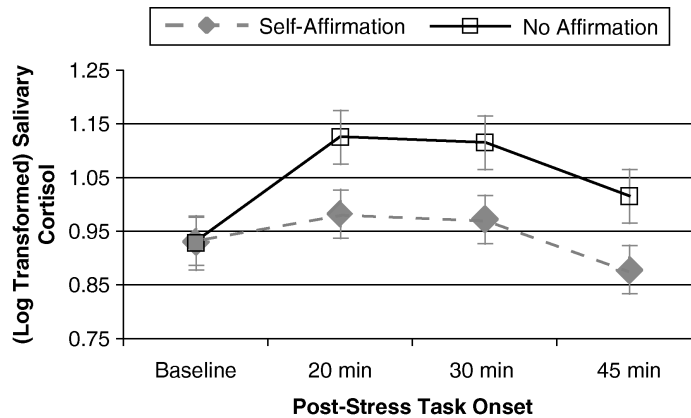
Fig. 3.   Salivary cortisol as a function of time and affirmation status. From Creswell, J. D., Welch, W., Taylor, S. E., Sherman, D. K., Gruenewald, T., & Mann, T. (2005). Affirmation of personal values buffers neuroendocrine and psychological stress responses. *Psychological Science*, *16*, 846–851. Adapted with permission.

participants were given failure feedback on an analogy test that was ostensibly an IQ test. Then, half of them completed a values scale concerning their most important value, whereas the other half completed a values scale of a value they considered to be relatively unimportant to them. The accessibility of the words from the failed IQ test (i.e., words from the analogy items) was assessed, either by recognition accuracy (Study 1) or the use of a lexical decision-making test (Study 2) (Koole et al., 1999). Both measures assessed the extent to which participants were ruminating about the failure. Those participants who affirmed an important aspect of the self by completing the values scale were less likely to ruminate after failure (Koole et al., 1999). Self-affirmation, then, can decrease ruminative thinking, and this may be one mechanism by which it makes stressors less stressful.

If self-affirmations can reduce physiological responses to stress (Creswell et al., 2005), then it is plausible that repeated affirmations might help people cope with daily stressors. As daily stressors have been shown to impair immune functioning (Glaser & Kiecolt-Glaser, 1994), a self-affirmation intervention could lead to beneficial health outcomes. Keough et al. (1998) examined this hypothesis in a sample of undergraduates who wrote essays over winter break on one of four topics. Participants in the affirmation condition wrote about the events of the day in terms of their most important value. Specifically, they integrated the day's events into that value (e.g., by writing about how the happenings of the day were relevant to the importance of their relationships with friends). There were three control conditions: a writing control condition (where people simply recorded what had

happened that day), a positive events control group (where people wrote about the positive things that had happened to them that day), and a nonwriting control group. Compared to all other groups, those who wrote the affirming essays over winter break reported the least amount of stress, and visited the health center significantly less often (when they returned to school). This effect was particularly strong among those who reported experiencing the greatest number of daily hassles and were, presumably, the most stressed.

Thus, self-affirmation interventions were associated with positive physiological and health outcomes with two relatively healthy college student samples (Creswell et al., 2005; Keough et al., 1998). An important question centers on whether similar effects would be obtained in a sample of people confronting disease. A study of women with breast cancer by Creswell et al. (2006) examined the relation between self-affirming writing and recovery from breast cancer. This research was based on a reanalysis of data from a study by Stanton et al. (2002). Stanton and colleagues (Stanton et al., 2002) had initially recruited 60 early-stage breast cancer survivors and randomly assigned them to write four essays on (1) their deepest thoughts and feelings about breast cancer (emotional processing); (2) their positive thoughts and feelings about breast cancer (benefit finding); or (3) the facts of their day (control). Findings from this initial study showed that women in both the emotional-processing and benefit-finding conditions had reduced physical symptoms and doctor visits 3 months after the study.

In the follow-up study (Creswell et al., 2006), all essays were coded for self-affirmation (whether the essay evidenced positive reflection on valued self-domains), cognitive processing (whether the essays evidenced active thinking about the positive aspects of the breast cancer experience), and discovery of meaning statements (whether the essay evidenced efforts to find a larger lesson or significance to their breast cancer), to test for potential mediating mechanisms for the writing effects on breast-cancer recovery. Interestingly, the number of affirming statements made in the essays (e.g., ''We have been married over 31 years and we are very lucky because we still love each other.'') fully mediated the intervention effects found on physical symptoms at 3-month follow-up, whereas (at least in this study) no mediating role was found for cognitive processing or discovery of meaning. Self-affirming cognitions played an important mediating role regardless of whether the intervention itself focused on emotional processing or benefit finding. Additionally, the frequency of self-affirmation statements predicted enhanced coping and reductions in distress at the follow-up session.

Together with the earlier work demonstrating the effects of self-affirmation on increasing the acceptance of threatening health information, the research reviewed in this section suggests that self-affirmation can have both direct

and indirect benefits for health. Directly, it seems to reduce physiological stress reactivity (Creswell et al., 2005), reduce daily stress perceptions (Keough et al., 1998), and potentially, aid in coping with distress in people with chronic disease (Creswell et al., 2006). Indirectly, it may promote more positive health behaviors by making people more open to information that they are engaging in risky health behaviors (Harris & Napper, 2005; Reed & Aspinwall, 1998; Sherman et al., 2000).

D. COGNITIVE DISSONANCE

The notion in self-affirmation theory (Steele, 1988) of a flexible self-system that seeks to maintain its perceived integrity shed new light on cognitive dissonance theory, and in particular, on the assumption that dissonance is aroused by psychological inconsistency (Aronson, 1968; Festinger, 1957). Specifically, the self-affirmation studies suggested that people are primarily motivated to maintain self-integrity rather than psychological consistency. Thus, people should be able to tolerate psychological inconsistency—and to refrain from dissonance-motivated cognitive distortions—if their self-integrity is buttressed through the affirmations of an alternative domain of identity.

For example, Steele, Hopp, and Gonzales (1986; cited in Steele, 1988) had participants rank 10 record albums in order of preference. They then gave the participants a choice of their 5th- or 6th-ranked album—a dissonance provoking choice (Brehm, 1956) that typically leads participants to accentuate the differences in the direction of their choice. Half of the participants then put on a lab coat in anticipation of a second study. For those who indicated that science was their most important value, wearing the lab coat affirmed an important domain. For others who indicated that science was not an important domain, wearing a lab coat was not self-affirming. All students then re-ranked the record albums. As predicted, with one exception, all groups showed the standard "spreading of alternatives" effect. They inflated the value of the option chosen and denigrated the value of the option foregone—a garden-variety rationalization effect. The exception was those science-minded students who wore the lab coat, and who thus had the opportunity to restore their self-integrity through the affirmation of an important self-identity. These participants did not defensively change their attitudes to make them concordant with their choice (see also Steele & Liu, 1983).

A full review of the theoretical debate concerning the circumstances under which self-affirmation is a viable explanation for cognitive dissonance phenomena is beyond the scope of this chapter (see Aronson et al., 1999; Simon, Greenberg, & Brehm, 1995; Stone & Cooper, 2001). However, there have been several important developments that merit review.

In various dissonance-provoking situations, self-affirmations have been found to reduce psychological discomfort and defensive biases. For example, Matz and Wood (2005) found that individuals report experiencing an aversive state of dissonance when others in a social group disagree with them. However, this dissonance discomfort was eliminated among participants who completed a self-affirmation. Another study suggested that dissonance reduction results in overconfident judgments (Blanton, Pelham, Dettart, & Carvallo, 2001). That is, if one cares about the accuracy of one's predictions then any cognitions that call that accuracy into question are likely to arouse dissonance. This dissonance is reduced by finding reasons to bolster one's confidence in the accuracy of one's predictions. Blanton et al. (2001) found that people who cared more about a prediction (involving correctly identifying Pepsi versus Coke based on a taste test) were relatively overconfident in the accuracy of their judgments. However, this overconfidence was eliminated among those who completed a self-affirmation.

Other studies have examined the conditions moderating the effects of self-affirmation on dissonance-reducing justifications (as in Steele & Liu, 1983). When the self-affirmations are in the same domain as the threatening information, they have been found to exacerbate cognitive dissonance (Blanton, Cooper, Skurnik, & Aronson, 1997) because they make salient the personal standards that are violated with the dissonant behavior. Consequently, when given a choice, people tend to choose to affirm the self in a domain unrelated to the perceived threat in order to reduce the dissonance they are experiencing (Aronson, Blanton, & Cooper, 1995). However, even affirmations in alternative domains of self-worth are not impervious to disconfirmation, and when they are disconfirmed (e.g., being told you are not religious shortly after affirming religion as a central value), cognitive dissonance can be reinstated (Galinsky, Stone, & Cooper, 2000). Several studies (e.g., Steele, Spencer, & Lynch, 1993; Stone & Cooper, 2001) have also examined the relationship between self-esteem, self-affirmation, and cognitive dissonance processes—a topic to which we will return in our discussion of individual differences and their moderating influence on self-affirmation processes.

E. MOTIVATED DISTORTIONS IN SOCIAL PERCEPTION

When people make judgments of others—such as whether a person is a good candidate for a job, or whether a person is above or below average—they have the opportunity to put themselves in a positive light through social comparison. Indeed, Dunning (2003) has argued that people are "zealous self-affirmers" in that their social judgments often reflect more about their own self-evaluative needs than about the target of judgment. Whether

students evaluate a person who studies 19 hours a week as "studious" depends upon how many hours a week they study themselves (Dunning & Cohen, 1992)—that is, people egocentrically define what evidence constitutes a given trait or ability. When people define what characteristics are likely to make a person have a successful marriage, they emphasize those characteristics that they themselves possess rather than those they do not (Kunda, 1987). And they do so more after experiencing a threat to self-worth, than after experiencing an affirmation of self-worth (Dunning & Beauregard, 2000; Dunning & Hayes, 1996; Dunning, Leuenberger, & Sherman, 1995).

After experiencing a self-threat, people may engage in any number of strategies to reaffirm self-integrity via social judgment. These strategies include comparing the self with a clearly inferior other (Fein, Hoshino-Browne, Davies, & Spencer, 2003), gossiping negatively about a third party (Wert, 2004; Wert & Salovey, 2004), or harshly judging a political ingroup member who fails to demonstrate as much fervor for the cause as one personally does (Beauregard & Dunning, 1998). At other times, esteem-boosting social judgments take a more disturbing form; people may try to feel better about themselves by putting down members of a marginalized group (Fein & Spencer, 1997). Because stereotypes are "cognitively defensible" justifications for denigrating others, people may be especially apt to use them to restore self-worth. Derogating an outgroup member would not only enhance one's own self-worth via downward comparison processes, but also enhance the integrity of a person's ingroup more generally. In one study, participants who were threatened with negative feedback on an intelligence task showed more stereotyping in their judgments of a gay male than those who received neutral feedback (Fein & Spencer, 1997).

Thus, self-threatening feedback can exacerbate outgroup derogation and the use of stereotypes. This finding suggests that self-image maintenance concerns can motivate prejudicial responses. Is the converse true as well—can an affirmation of an individual's self-integrity reduce the need to stereotype an outgroup member? Fein and Spencer (1997) examined this possibility in one study where participants completed a self-affirmation through writing about an important value (versus a control condition where they wrote about a relatively unimportant value). They did so prior to evaluating a job candidate who was presented either as a member of a negatively stereotyped group (Julie Goldberg, who fit with a "Jewish American Princess" stereotype widely in circulation on campus at the time) or not (Maria D'Agostino, an Italian American). In the no-affirmation condition, participants made more negative evaluations of the candidate's qualification for the job and viewed the personality of the Jewish woman more negatively than that of the Italian woman. In contrast, those participants who completed the self-affirmation rated the Jewish candidate as favorably as the

Italian candidate. Thus, the extent to which a person is threatened or affirmed will affect whether or not they are likely to make prejudicial judgments of an outgroup member (cf. Shrira & Martin, 2005; Zarate & Garza, 2002). Indeed, in this study, self-affirmation did not simply attenuate prejudice and discrimination but eliminated it.

Comparing oneself to another person who is faring worse is another way that people may affirm self-integrity via social perception. When people have a vulnerable or easily threatened self-image, they generally respond with downward social comparisons (Taylor & Lobel, 1989; Wills, 1981). Shelley Taylor and her colleagues (Taylor, Wood, & Lichtman, 1983) found that women with breast cancer responded to their highly threatening situation by making downward social comparisons. They would assert that their illness was, at least, less severe than that of another patient, or that they were coping with their illness better than others. These downward comparisons, Taylor suggested, helped patients to maintain a sense of worth in a situation where they struggled to maintain their sense of control, predictability, and optimism. Such downward comparisons, while helping to sustain self-integrity, may by themselves limit an individual's opportunities to learn from others who are more experienced or performing more optimally (indeed, such ''upward'' comparisons also have motivational benefits; see Lockwood & Kunda, 1997).

If people engage in downward social comparisons to compensate for a threatened self-image, then they should engage in them to a lesser extent when self-affirmed. Indeed, they may even be more prone to engage in *upward* comparisons that they might otherwise view as threatening (Spencer, Fein, & Lomore, 2001). In one study, college students completed a test of intelligence, and they were informed that they performed at the 47th percentile. This mediocre performance presumably threatened participants' self-image as intelligent college students. Half of the students then had the opportunity to affirm the self by writing an essay about an important value, whereas the other half wrote about an unimportant value. Then the participants, in a separate task, were informed that another participant would interview them. For the ostensible purpose of preparing them for their interview, they then listened to excerpts of two previous interviews. In one of the excerpts, the interviewee made a terrible impression (e.g., unintentionally insulting the interviewer, speaking incoherently), whereas in the other the interviewee was smooth and articulate. The participants had a choice of which interview to listen to in full. This choice provided an opportunity to make an upward comparison (if they selected the superior interviewee) or a downward comparison (if they selected the inferior interviewee).

Because all participants had been threatened by the feedback they had received on the intelligence task, the study allowed an examination of social

comparison under threat. In the no-affirmation condition, the participants generally made downward comparisons, as 83% chose to hear the inferior interviewee. By contrast, among those who completed the self-affirmation, 83% chose to hear the superior interviewee, making an upward social comparison. Once again, self-affirmation reduced threat and thereby encouraged people to expose themselves to an informative but potentially threatening learning opportunity.

In summary, predictions derived from self-affirmation theory (Steele, 1988) have been supported in a wide range of situations involving self-threat. When self-integrity is affirmed, people are less biased in their judgments of information related to their political identity (Cohen et al., 2000), their health (Reed & Aspinwall, 1998; Sherman et al., 2000), and their impressions of others (Fein & Spencer, 1997; Spencer et al., 2001). Self-affirmation inoculates people against threat, and thus makes them more open to ideas that would otherwise be too painful to accept. It also reduces the stressfulness of evaluative situations (Creswell et al., 2005) and the cognitive accessibility of threatening cognitions (Koole et al., 1999). Theoretically consistent effects of self-affirmation have been found on self-report measures, physiological responses, and behavior. When self-integrity is secured, people seem less concerned with the self-evaluative implications of social experiences and are more likely to engage their social world in a non-defensive, open manner.

III. Self-Affirmation and Responses to Collective Threats

Originally, self-affirmation theory (Steele, 1988) focused on how people respond to information and events that threaten a valued self-image, such as situations that provoke cognitive dissonance and defensive rationalization of counter-attitudinal behavior (Steele & Liu, 1983), or situations that challenge a sense of personal control (Liu & Steele, 1986). The research detailed in the previous sections extends this theorizing to many other situations where people contend with events that challenge a *personal* identity.

A major advance in self-affirmation theory concerns its relevance to the way people cope with threats to their *social* (i.e., group) identities. This work begins with the premise that social identities—such as affiliation with a sports team, membership in a gender or racial group, citizenship in a country, involvement in an organization—constitute important sources of identity (Abrams & Hogg, 1988; Deaux, 1996; Tajfel & Turner, 1986). Consequently, people will defend against threats to collective aspects of the self much as they defend against threats to individual or personal aspects of self. They may do so even when these events do not directly implicate

oneself (e.g., even when the threat involves the behavior of another group member rather than one's own behavior; Cohen & Garcia, 2005; Norton, Monin, Cooper, & Hogg, 2003; Wiesenfeld, Brockner, & Martin, 1999). We argue that both personal and social (i.e., group) identities are fundamentally confounded (Cohen & Garcia, 2005), in that both types of identities contribute to the same overarching goal of maintaining self-integrity (see also Gaertner, Sedikides, & Graetz, 1999; Sherman & Kim, 2005). When a fellow member of one's racial group confronts the threat of a negative stereotype, it is threatening to self even when one personally faces no risk of being prejudiced against (Cohen & Garcia, 2005). When an athlete's team loses, he or she will resist explanations that implicate the team because doing otherwise would threaten the image of a valued group. When one's country is vilified in an editorial, a patriot will attack the credibility of the source to protect the perceived integrity of his or her national identity.

The self-affirmation analysis of such collective threats, however, asserts that because social identities are only one part of a larger, flexible self-system, people can respond to threats to their group memberships or social identities indirectly. That is, they can maintain an overall self-perception of worth and integrity by affirming some other aspect of the self, unrelated to their group. This insight has applications to a wide range of phenomena related to group identity.

A. GROUP-SERVING JUDGMENTS

A popular aphorism is that "There is no *I* in *team*." It conveys the notion that the goals of a team or a group should be important enough that individuals are willing to sacrifice their personal ambitions and self-interests in the service of team success. Self-serving aspirations are antithetical to the team's progress. However, social groups are a central part of how people see themselves, and people are motivated to defend their social identities (Abrams & Hogg, 1988). Consequently, when explaining the success or failure of their group, people tend to be defensive and group serving. This phenomenon has been observed in several studies examining attributions made for success or failure in real world contexts. Letters to shareholders feature more internal attributions after successful years (e.g., stocks rose because of the work of the management team) than after disappointing years (Bettman & Weitz, 1983), and athletes' explanations for success feature more internal attributions than do their explanations for failure (Lau & Russell, 1980; Winkler & Taylor, 1979). These biased judgments—encompassing both inflated internal attributions for success and attenuated internal attributions for failure—are group serving in the sense that they suggest that

group members are selectively responsible for causing the positive events that befall their group.

In two field studies involving intramural sports team athletes as participants, Sherman and Kim (2005) examined the role of the self-integrity motivations in group-serving judgments. The intramural sports teams, consisting of voluntarily formed groups of friends, played their games where they either won or lost and then participated in the study. They completed a self-affirmation manipulation in which they filled out a values scale concerning either their most important value or a relatively unimportant value. Then, they assessed their attributions for the outcome of the game. They estimated how much their personal performance contributed to the outcome of the game, and how much their team's play and teamwork contributed to the outcome of the game.

Overall, the athletes were both self-serving and group serving in their attributions. That is, winners thought that their personal performance contributed more to their team's victory than losers did to their defeat (a self-serving bias). Additionally, winners thought that their team's performance and teamwork contributed more to their team's victory than losers did to their defeat (a group-serving bias). Both self- and group-serving biases, however, were eliminated among those who completed a self-affirmation. The reduction of group-serving judgments is depicted in Fig. 4. Whereas in the no
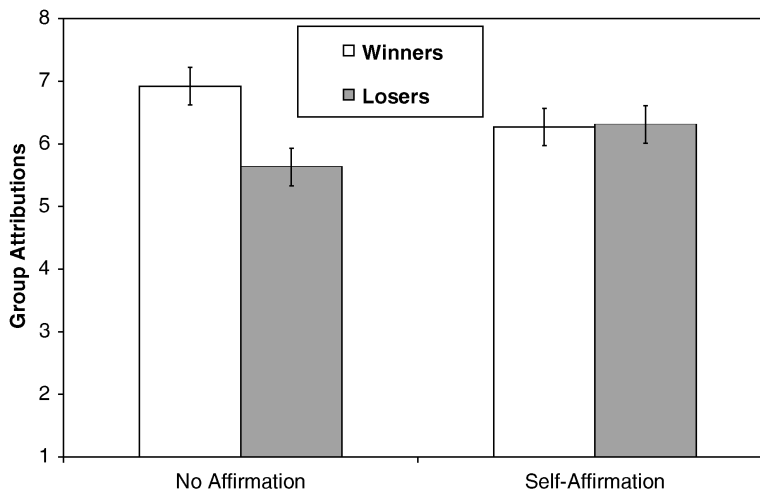

Fig. 4. Group-serving attributions as a function of game outcome and affirmation status. From Sherman, D. K., & Kim, H. S. (2005). Is there an ''I'' in ''team''? The role of the self in group-serving judgments. *Journal of Personality and Social Psychology, 88*, 108–120. Adapted with permission.

affirmation condition, winners made much stronger internal team attributions than did losers, this difference was eliminated in the self-affirmation condition (Sherman & Kim, 2005).

A second goal of this research focused on examining the relation between self- and group-serving judgments. Because one's self-concept and one's social identity are overlapping cognitive constructs (Smith & Henry, 1996; Smith, Coats, & Walling, 1999), people will tend to use the self-concept as an anchor or evaluative base to form judgments of the group (Cadinu & Rothbart, 1996; Gramzow & Gaertner, 2005; Otten, 2002). People generalize their positive evaluation of self to their evaluation of their group. However, if the need to protect self-integrity is satisfied via self-affirmation, then people should be able to evaluate the group independently of how they evaluate the self.

We examined whether such self-group anchoring would be attenuated by a self-affirmation (Sherman & Kim, 2005, Study 2), again with intramural athletes who had just won or lost a game. We replicated the basic findings that self-affirmation reduced both the self- and the group-serving attributional biases. Additionally, we examined the correlations between the attributions to the self and the attributions to the group as a function of affirmation condition. As predicted, in the no-affirmation condition, there was a strong correlation ($r = .60$) between judgments about the self and judgments about the group, indicating that under normal conditions, the athletes generalized their self-evaluations to their evaluations of their group. In contrast, when participants were self-affirmed in an alternative domain, they evaluated the group independently of their self-evaluations, as the correlation was eliminated ($r = -.10$). Self-affirmation, then, allows people to evaluate their groups independently of the way they evaluate themselves.

B. DEFENSE OF COLLECTIVE IDENTITY

Patriotism and pride in one's national citizenship represent manifestations of another collective identity that people try to defend—particularly when their country is under threat. In the wake of September 11th, for example, two responses emerged. On the one hand, some people responded to the national threat by embracing their American identity uncritically, and by finding little if any problem or fault in America's foreign policy in the Middle East. On the other hand, there were those who took a critical eye to American foreign policies and noted how these policies (such as its support of totalitarian regimes in the Middle East) may have sowed the conditions and sense of disenfranchisement among that populace, which contributed, in part, to the terrorist acts against the United States.

We were interested in how people's self-identities as American "patriots" (i.e., people who asserted that they were patriotic, and who felt that the US constituted a force of good in the world) or American "anti-patriots" (i.e., people who asserted that they were unpatriotic, and who felt that the US constituted a force of harm in the world) would affect their evaluation of a persuasive report entitled "Beyond the Rhetoric: Understanding the Recent Terrorist Attacks in Context" (Cohen et al., 2005). The report argued that Islamic terrorism can be understood in terms of the social and economic forces of the Middle East. It further underscored the role that US foreign policy had played in fostering some of the social—economic conditions in the Middle East that later led to the terrorist attacks of 9/11. The arguments were credible (drawn from the writings of several prominent analysts) and buttressed with factual evidence and historical analysis. To amplify the counter-American tone of the report, it was ostensibly written by an author of Arab descent, "Babek Hafezi." To make participants' American identity situationally salient, the experimenter wore a small American flag pin on her lapel.
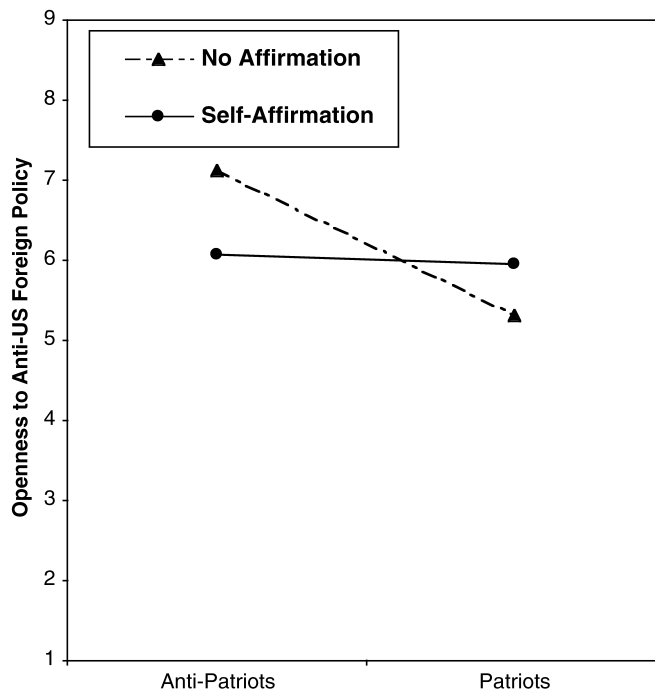

Fig. 5.   Openness to anti-U.S. foreign policy report among patriots and anti-patriots as a function of affirmation status.

Figure 5 presents the relevant results. In the absence of affirmation, there was a partisan divide in perception. Patriots proved far more critical of the report than did antipatriots. Indeed, the correlation between a measure of participants' identification as an American patriot versus antipatriot (administered several weeks prior to the experiment), and their openness to the report was $r = .58$. In other words, participants' preexisting identity accounted for 34% of the variance in their openness to the report. By contrast, in a condition where participants self-affirmed prior to reading the report (by writing about an important value unrelated to their national identity), patriots became more open to the report, and anti-patriots became more skeptical of it. Indeed, the correlation between prior identity and openness was reduced to nil ($r = -.05$). Collective identity ceased to affect the assimilation of new information when participants had affirmed an alternative source of self-integrity.

C. STEREOTYPE THREAT AND PERFORMANCE

Stereotype threat—the potential that one could be judged in light of a negative stereotype about one's group—is a potent type of collective threat (Steele, 1997). Members of negatively stereotyped groups may experience elevated levels of stress when performing on tasks where they risk confirming a stereotype about their group in the eyes of others (Ben-Zeev, Fein, & Inzlicht, 2005; Blascovich, Spencer, Quinn, & Steele, 2001; O'Brien & Crandell, 2003). This stress can in turn undermine performance. Several studies, both in the laboratory and in the field, have now examined whether self-affirmation can reduce stereotype threat and facilitate performance in conditions where people from negatively stereotyped groups have been shown to underperform (Spencer, Steele, & Quinn, 1999; Steele & Aronson, 1995).

One study examined whether stereotype threat could be reduced among women taking a math test (Martens, Johns, Greenberg, & Schimel, 2006). To the extent that women care about performing well in such a situation, they may worry, should they perform poorly, that they could confirm the nega-tive stereotype that women are worse than men at math. If so, then affirma-tion of an alternative aspect of the self could reduce the threat and facilitate performance. Whereas other interventions to reduce stereotype threat directly refute the stereotype or its relevance (e.g., by portraying a test as gender-fair; Cohen, Steele, & Ross, 1999; Spencer et al., 1999), this interven-tion was psychologically farther downstream—aimed at altering students' threat/stress *response* to the perceived relevance of the stereotype.

The participants were female college students who took a difficult math test that was described either as "diagnostic" of their math and reasoning abilities or as a reasoning task unrelated to their ability and still being developed for future research. Replicating the standard stereotype threat effect (Steele, Spencer, & Aronson, 2002), women performed worse in the ability-diagnostic condition (where they were aware that a potentially poor performance on their part could be viewed as evidence of lack of math ability and as ultimately a validation of the negative stereotype), worse than their female counterparts in the ability-nondiagnostic condition, and worse than men overall. However, women in the diagnostic condition who completed a self-affirmation—where they wrote about an important value unrelated to math or to their gender—performed as well as women in the no-threat condition and no different from male college students.

Given the evidence that self-affirmation appears to reduce stereotype threat among students in laboratory situations (Martens et al., 2006; Schimel et al., 2004), perhaps having students engage in affirmation of non-academic aspects of themselves in school settings could produce positive educational benefits by reducing the psychological impact of social identity threat and other sources of psychological distress. In such educational contexts, social identity threat may be exacerbated by the mistrust that minority students sometimes feel when confronted with negative or critical feedback from White evaluators and teachers (Cohen et al., 1999). Because they know that the stereotype could bias members of the outgroup, these students may come to doubt the intentions motivating critical feedback. Is it possible for affirmation to mitigate these detrimental responses?

To examine this question, a field study (Cohen, Garcia, Apfel, & Masters, 2006) was conducted that featured a randomized, double-blind experimental design. Students completed a 20-minute in-class self-affirmation exercise in which they wrote about an important value and why it mattered to them, or, in a control condition, wrote about an unimportant value and why it might be important to someone else. In the control condition, minority students, unlike their majority peers, displayed a decline in trust in their teachers and school administrators over the course of the 7th-grade school year. That is, they judged their grades and treatment in school as less fair and more biased at the end of the year than they had at the beginning. By contrast, affirmed minority students remained constant in their relatively high levels of trust and perceived fairness over the course of the school year (see Fig. 6). Just as affirmation increases trust across partisan lines in the context of negotiation (Cohen et al., 2000, 2005), it cultivates trust and reduces threat across racial lines as well.
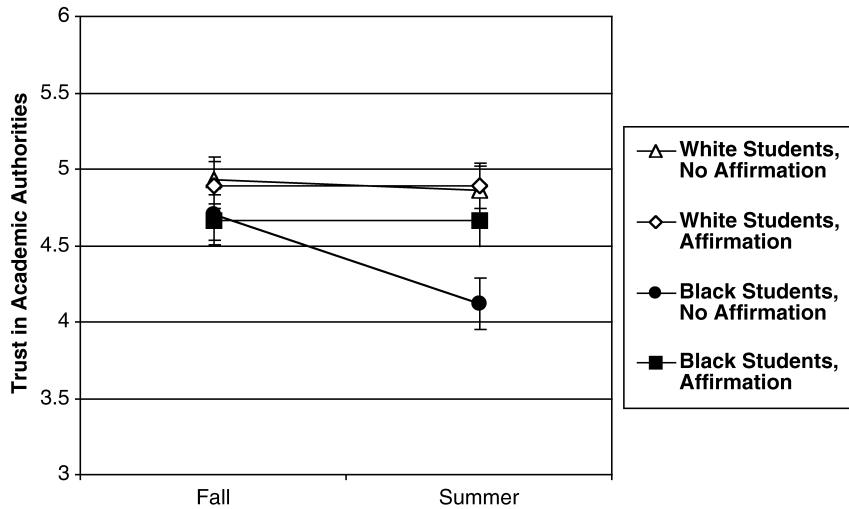
Fig. 6.   Trust in academic authorities as a function of race, time, and affirmation status.

D. PERCEPTIONS OF RACISM

When individuals interact with members of stereotyped groups, they often
want to appear nonprejudiced (e.g., Gaertner & Dovidio, 1986; Monteith &
Voils, 1998). The very idea of prejudice itself could be threatening to an
individual's self-integrity for many reasons. Consider a European American
evaluating whether prejudice against minorities still constitutes a major factor
in the differential academic and economic success of people from different
groups. If prejudice is a major factor, this suggests that one may have been
unfairly privileged, that one may have benefited (indirectly or even directly)
from the oppression of minority groups, and that one's country fails, in some
ways, to live up to the values of equality and egalitarianism. One way people
from majority groups can reduce this threat is by minimizing the perceived
frequency and impact of prejudice against minority groups.

   Building on this logic, Adams, Tormala, and O'Brien (in press) examined
the effect of self-affirmation on perceptions of prejudice against minorities.
The participants were both European Americans and Latinos. Overall,
Latinos perceived more prejudice against minorities in everyday life than
did European American participants. However, this race effect was attenu-
ated by a self-affirmation. There were no group differences in perceived
prejudice among those who affirmed an important value (unrelated to race)

prior to assessing their perceptions of prejudice. Importantly, the effect of self-affirmation was more evident among European Americans than Latino Americans. While Latino Americans showed a slight drop in their perception of prejudice against minority groups when self-affirmed, this drop was not significant. On the other hand, self-affirmation led European Americans to perceive significantly *more* racism against minorities in the United States (Adams et al., in press, Study 2). Affirmed European American participants also agreed, to a far greater extent than their non-affirmed peers, that European Americans in general tend to understate the impact of racism in daily life. Thus, the otherwise threatening idea of racism in America was more acceptable among those who were buffered by a self-affirmation. Indeed, these studies suggest that it is not just minority students who defensively exaggerate the role of prejudice to protect self-worth (cf. Crocker & Major, 1989). It is the majority group members, as well, who defensively *deny* racism in order to protect *their* self-esteem.

IV. Moderator Variables and Qualifying Conditions

Across a variety of potentially threatening situations, self-affirmations reduced perceived threat and the likelihood of engaging in defensive adaptations to threat. Health information suggesting that one has acted in a risky manner, the defeat of one's team, a stereotype directed at oneself or a fellow group member, a report attacking one's political worldview, a failure at an intellectual or athletic task—these events threaten individuals' self-integrity. When the self is under threat, people engage defensive adaptations to ameliorate the threat. Motivated inferences about health information, group-serving biases, biased assimilation of new information, and outgroup derogation are examples of such adaptations. These self-protective strategies can be reduced and even eliminated, however, when people affirm alternative sources of self-integrity unrelated to the provoking threat.

To understand the process of self-affirmation, it is important to examine factors that moderate when and how self-affirmation operates. Cultural differences may moderate the effects of threat and affirmation. Additionally, individual differences in self-esteem and in identification with the domain of threat should moderate whether and to what extent self-affirmation is effective. Situational differences in the salience of a particular identity may prove critical. Moreover, affirmations may have different effects as a function of their form and content. While the majority of the research previously reviewed demonstrates how self-affirmations can lead to greater openness,

some types of affirmations may backfire; they may decrease openness and promote defensive responses. We now review these moderating variables.

A. CULTURE

Given the extensive research and theorizing on cross-cultural differences in the conception and experience of selfhood (e.g., Heine, Lehman, Markus, & Kitayama, 1999; Kim & Markus, 1999; Markus & Kitayama, 1991; Triandis, 1989), an important question concerns the effect of culture on self-affirmation processes. In individualist cultures (such as the US), the self is a more bounded, autonomous entity, whereas in collectivist cultures (such as in East Asia), the self is more interconnected and tied to important relationships. Members of collectivist cultures may thus be less motivated to protect self-integrity because their culture places less emphasis on esteeming the self (Heine et al., 1999). Alternatively, they may be just as motivated to protect self-integrity, but they may differ in terms of the events and issues they find threatening versus affirming.

Heine and Lehman (1997) conducted one of the first studies on culture, self-affirmation, and cognitive dissonance. Among European Canadians, they replicated the Steele and Liu (1983) finding that affirmation of personal values eliminated the spreading of alternatives in the free-choice dissonance paradigm (Brehm, 1956). Yet among Japanese participants, no spreading-of-alternatives effect was found, perhaps because they did not experience the situation as self-threatening (Heine et al., 1999). Consequently, there was no defensive response for the affirmation to ameliorate.

Thus, to identify how self-affirmation operates cross-culturally, it is important to test the effects of a situation that could theoretically threaten self-integrity and produce defensiveness among people in different cultures. Research by Hoshino-Browne and colleagues (Hoshino-Browne et al., 2005) identified such conditions in their examination of the cultural moderation of dissonance and affirmation processes (see also Kitayama, Snibbe, Markus, & Suzuki, 2004). First, they found that when a choice is potentially threatening to their relationships, members of collectivist ethnic groups (i.e., Asian Canadians) experience cognitive dissonance. When Asian Canadians made choices of food preferences, they rationalized their decisions more when they chose for their friends than for themselves (Hoshino-Browne et al., 2005). That is, after making a choice between two comparable food options for a friend, they came to believe that the friend would find the chosen option far superior to the non-chosen option. In contrast, European Canadians rationalized their decisions more when they made choices for themselves rather than for their friends (Study 1 and Study 2 of Hoshino-Browne et al., 2005).

To examine whether this dissonance would be reduced by a self-affirmation, the researchers modified the standard self-affirmation procedure to contrast an independent with an interdependent self-affirmation. For the interdependent affirmation, participants ranked values in terms of how important they were to "themselves and their family" and wrote a paragraph about why this value is shared among them and their families. For the independent self-affirmation, participants ranked values in terms of personal importance and wrote an essay about why that value was personally important. The participants in the study, all of East Asian descent, then made choices for their friends and had the opportunity to rationalize these choices. As predicted, the interdependent affirmation reduced dissonance-motivated rationalization when participants had to make a choice for a friend, whereas the independent self-affirmation did not. Follow-up studies also revealed that, for the European Canadians, the opposite effect occurred: the independent affirmation reduced dissonance-motivated rationalization, but the interdependent affirmation did not. Moreover, both affirmations were effective among bicultural Asian Canadians who identified with both Asian and Canadian culture (see also Hoshino-Browne, Zanna, Spencer, & Zanna, 2004). The latter result suggests that bicultural identity can confer access to different sources of affirmation.

This research sheds light on how culture moderates the effectiveness of self-affirmation. Where the self is structured differently, what constitutes a threat and an affirmation are different as well. However, the general process whereby affirmation reduces defensive responses to threats appears to be culturally invariant.

B. SELF-ESTEEM

If one way that people respond to threats is by affirming the self in an alternative domain, then it stands to reason that those who have more positive views of the self will have more psychological resources with which to self-affirm. Steele et al. (1993) predicted that high self-esteem individuals would have greater affirmational resources and would thus be more resilient to threatening events than would low self-esteem individuals. Consistent with this logic, they found that high self-esteem people were less likely to rationalize a choice they had made (i.e., by inflating the value of a music album they had selected and/or by deflating the value of a desirable music album they had foregone). Interestingly, this effect was found only among those high self-esteem people who had first completed a self-esteem scale that made their self-resources cognitively accessible to them. Under this condition, those with greater affirmational resources were more resilient to

threats to their self-image and less likely to engage in defensive rationaliza-
tion of their behavior (see also Nail, Misak, & Davis, 2004).

There are two schools of thought about the relationship between self-
esteem and the ability to tolerate dissonant acts or beliefs. Whereas the
affirmational resources view posits that people with high self-esteem have
more positive self-concepts, making it easier for them to tolerate their
dissonant actions (Spencer, Josephs, & Steele, 1993; Steele et al., 1993), the
self-consistency view posits that high self-esteem people will view a dissonant
act as more discrepant with their positive self-image (Aronson, 1968;
Thibodeau & Aronson, 1992). According to this latter view, high self-esteem
people are likely to feel more dissonance than low self-esteem people and
are more prone to engage in dissonance reduction strategies such as justifi-
cation. For example, one study examined smokers who were participating
in a smoking cessation clinic (Gibbons, Eggleston, & Benthin, 1997). The
smokers who relapsed and began smoking again reduced the dissonance
associated with this attitude-discrepant behavior (discrepant because they
were committed to quitting smoking) by reducing their self-perceived risk of
acquiring smoking-related illnesses. However, this occurred only among
high self-esteem smokers (Gibbons et al., 1997).

A third theory, the self-standards model of cognitive dissonance (Stone &
Cooper, 2001) proposes that people, at times, use self-esteem as a resource
(as in Steele et al., 1993). However, people may also use their level of self-
esteem as a standard or expectancy that determines whether they view a
particular behavior as dissonant. Once people act, they evaluate their behav-
ior in light of a standard—but what standard comes to mind varies with
situational cues. When the cues focus on normative standards for behavior
(e.g., how most people act), no differences between low– and high–self-
esteem individuals emerge. It is only when the cues focus people on their
own personal standards that self-esteem differences do, in fact, emerge.
Whether low or high self-esteem individuals rationalize more, in turn,
depends on the nature of the self-affirmation. Self-affirmations *related* to
the threatening act (e.g., writing about how compassionate one is after
having behaved uncompassionately) increase the salience of personal stan-
dards of conduct, draw attention to the discrepancy between those standards
and one's previous behavior and thus lead high self-esteem people to experi-
ence more dissonance (and rationalize more) than low self-esteem people. By
contrast, self-affirmations *unrelated* to the threatening act draw attention to
alternative psychological resources, and thus lead high self-esteem people to
experience less dissonance (and rationalize less) than low self-esteem people
(Stone & Cooper, 2003).

One variable that holds important implications for the relation between
self-esteem and cognitive dissonance is whether a person's high self-esteem is

secure and stable versus insecure, defensive, and fragile (Baumeister, Heatherton, & Tice, 1993; Deci & Ryan, 1995; Jordan, Spencer, Zanna, Hoshino-Browne, & Cornell, 2003; see also Kernis & Paradise, 2002). Secure, high self-esteem is operationalized as having high explicit self-esteem (as assessed by a self-report measure of self-esteem such as the Rosenberg Self-Esteem Scale; Rosenberg, 1965) and high implicit self-esteem (i.e., non-conscious, automatic positive associations with the self, as assessed by the Implicit Associations Test; Greenwald, McGhee, & Schwartz, 1998). Such people have high self-esteem both at a conscious level and at a deep, unconscious, and reflexive level. By contrast, defensive high self-esteem is operationalized as having high explicit self-esteem but low implicit self-esteem. Such people have high conscious self-esteem but unconsciously, and reflexively, evaluate the self negatively. Jordan and colleagues find that those with secure high self-esteem show little dissonance reduction (consistent with the affirmational resources view). That is, those with secure high self-esteem are less likely to rationalize their decisions to reduce dissonance, whereas those with defensive high self-esteem are more likely to engage in dissonance reduction (consistent with the self-consistency view) (Jordan, Spencer, & Zanna, 2003). It seems that people with defensive high self-esteem may have fewer self-resources that can be drawn upon when experiencing threat.

C. IDENTITY CENTRALITY AND SALIENCE

Whether a potentially threatening domain is personally important to an individual or constitutes a part of their personal identity affects whether or not they can experience self-threat in that domain (Boninger, Krosnick, & Berent, 1995), and consequently, whether an affirmation will prove effective at ameliorating defensiveness. Ironically, it is the people who view an issue as important rather than unimportant who should prove the most open to affirmation-induced change. For example, a report describing the link between caffeine use and fibrocystic disease concerns a topic of greater importance for heavy caffeine consumers than for light caffeine consumers. Consequently, it is only heavy caffeine consumers who respond defensively to such a report (e.g., Kunda, 1987; Liberman and Chaiken, 1992) and who exhibit greater openness to the report when self-affirmed (Sherman et al., 2000).

Adopting another approach, Correll et al. (2004) examined students' responses to a debate over the merits of a tuition increase. They found that only participants who saw tuition increases as personally important were biased in their assessments of the quality of pro- versus counterattitudinal information on this topic. Further, it was only these participants for whom

affirmation proved an effective debiasing intervention. In a related vein, one study found that sports fans that watched their team win or lose were only biased in their attributions for the outcome of the game when being a fan was viewed by them as an important part of their identity (Sherman, Kinias, Major, Kim, & Prenovost, 2005). Moreover, this group-serving bias could be eliminated when participants affirmed a value of high importance to their group.

Individuals may thus vary in the strength of their allegiance to particular sources of identity and in their resistance to new and challenging ideas that would force them to see the world in other than black and white terms (Correll et al., 2004; Zuwerink & Devine, 1996). Additionally, implicit and explicit situational cues may play a role in heightening these types of identity defense motivations that lead to closed-mindedness and inflexibility. Situational contexts differ in how much they highlight a particular identity and its importance to self-integrity (Cohen et al., 2005). When a given context heightens the salience of an individual's partisan identity and/or commitment to a given position on an identity-relevant issue, that individual may feel relatively more obliged to consider the identity costs of openness and compromise. It is precisely those costs, in turn, that are reduced when an *alternative* source of self-integrity is made salient and affirmed.

In one study, for example, political partisans with varying views regarding US foreign policy evaluated a report critical of that policy. They did so after a self-affirmation or threat was administered, under conditions where either their identities as self-rated "patriots" (versus "anti-patriots") or the goal of rationality was made salient. In the identity salient condition, the experimenter wore an American flag pin on her lapel, whereas in the rationality salient condition, the experimenter wore a white lab coat. A focus on rationality was expected to diminish attention to, and concern with, sources of identity that would otherwise make the prospect of belief change and compromise threatening. The results in the American flag condition were described previously: participants interpreted information critical of the United States in a manner consistent with their national identity—but a self-affirmation eliminated this bias. In contrast, when the experimenter wore the lab coat—minimizing concerns of identity maintenance—affirmation had no effect (Cohen et al., 2005).

D. WHEN AFFIRMATIONS BACKFIRE

An interesting issue concerns classes of affirmations that increase rather than decrease bias and resistance to change. One such class of affirmations is those that are in the *same* domain as the threatening event or information.

Whereas different-domain affirmations decrease bias, closed-mindedness, and inflexibility, same-domain affirmations increase people's sense of self-confidence, certainty, and impunity. They thus produce effects opposite to those observed when the affirmation targets a domain unrelated to the threat (Blanton et al., 1997). Prior to reviewing a counter-attitudinal report, for example, people might affirm a value *related* to the issue that figures in that report. Before reading a report critical of evidence supporting global warming, an environmentalist might reflect on the personal importance of the cause of environmentalism (Sherman et al., 2005). Alternatively, before reading a report arguing against allowing gays in the military, an egalitarian might assert the importance of tolerance (Jacks & O'Brien, 2004). In contrast to domain-irrelevant affirmations, such domain-relevant affirmations consistently increase resistance to change, presumably by highlighting a person's commitment to the issue and the identity at stake. This research suggests that, if attitude and behavior change is the goal, policy-makers and interventionists should resist lay intuitions to provide affirmations targeting the domain of threat. It may do more harm than good to provide academically at-risk students with affirmations of their academic ability, risk-seeking adolescents with affirmations of their rationality, or negotiators with affirmations of their ideological commitments. Instead, the more effective strategy is the counterintuitive one—that of affirming domains *unrelated* to the threat.

An intriguing phenomenon involves the way in which, under some circumstances, affirmations of *moral* worth can lead to a sense of personal impunity, in which people no longer feel obliged to prove themselves in the domain in question and thus feel licensed to act in ways that violate important moral principles. This seems to occur particularly in contexts where the "right" or "moral" course of action is ambiguous. For example, when people affirm their lack of prejudice, they engage in more gender and racial discrimination, at least under some situations that render such discrimination ambiguous in its moral implications (Jacks & O'Brien, 2004; Monin & Miller, 2001). In one line of studies, people made their "moral credentials" clear—in our language, affirmed their self-concepts as "moral" and "egalitarian"—by being led to select a clearly qualified ethnic minority for a job. Relative to those who did not affirm their lack of prejudice, those who did later proved *more* prejudiced; that is, they were more willing to state that a particular job (one that could have posed problems for a minority due to its hostile work environment) was more appropriate for White workers than for Black ones. When people affirm their lack of prejudice, they seem to become more lax in their adherence to egalitarian values.

In an unusually disturbing series of studies, Brown (2000) further demonstrated how moral affirmations not only lead to impunity but also license what many would view as immoral behavior. College students were given

positive feedback concerning the morality of individuals at their school—
they were told that students at their school scored higher on standard moral
development tests than did students attending a rival school. "Morally
affirmed" participants were subsequently more swayed by negative propa-
ganda targeting foreign graduate students. The propaganda in question
asserted that foreign graduate students posed a threat to national security
due to the possibility that they might, for example, steal scientific secrets or
engage in espionage. In response to this propaganda, morally affirmed sub-
jects proved more likely to endorse punitive and "proto-genocidal" policies
for dealing with the "foreign graduate student problem." They advocated to a
greater extent the policy of forcing foreign graduate students to carry identifi-
cation papers with them at all times and even became more favorable to the
proposal to banish all foreign students from studying in American universi-
ties. The moral affirmation seemed to imbue participants with impunity
to discriminate in the name of protecting national security. Clearly, more
research needs to be done to determine the conditions under which such
domain-relevant affirmations lead to open-mindedness, closed-mindedness,
and moral impunity.

   Another class of potentially detrimental affirmations involves those that
activate a person's sense of self-perceived objectivity. When made to feel
objective, people adopt an "I think it, therefore it's true" mindset. That is,
they assume that their own thoughts and beliefs—by virtue of being theirs—
are objective, valid, and therefore worthy of being acted upon (Uhlmann &
Cohen, 2005a). As a consequence, objectivity-affirmed evaluators may be-
come more likely to act on any prejudicial thoughts and beliefs that they
harbor, and whose influence they might have otherwise suppressed. Consis-
tent with this analysis, in a series of studies, Uhlmann and Cohen (2005b)
found that affirming individuals' sense of objectivity made them more likely
to discriminate against women when deciding whom to hire for stereotypi-
cally male, high-status jobs (e.g., corporate representative, police chief).
Objectivity-affirmed subjects also proved more likely to act on stereotypes
that had been nonconsciously primed in their social environment. In one
study, participants were either primed on words related to negative stereo-
types about women (e.g., *pink, Barbie, make-up, emotional*) or not. Those
who affirmed their objectivity (i.e., by answering several easy-to-agree-with
questions concerning their personal objectivity) were more influenced by
this prime. When primed on gender stereotypes, participants proved more
likely to hire a man over a woman when they were objectivity-affirmed than
when they were not (see Fig. 7 for the relevant results). Paradoxically,
making people feel objective causes them to become more subjective. It
confers a sense of impunity about using personal thoughts and beliefs that
happen to be cognitively salient at the particular moment of a decision.
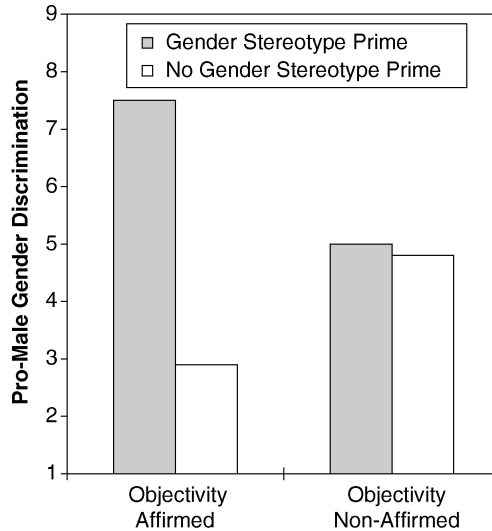
Fig. 7.   Pro-male gender discrimination as a function of gender stereotype prime and affirmation status.

V. Underlying Processes and Superordinate Functions

Given the wide range of studies reported in this review, it is clear that self-affirmations exert effects across many domains of psychological functioning. Yet, important questions remain concerning both the mechanisms underlying these effects and the ultimate function of the self-integrity system. These questions set the stage for exciting and important opportunities for future research.

A. UNDERLYING MECHANISM OF SELF-AFFIRMATION

Through what mechanisms do self-affirmations produce their effects? Affirmations, we suggest, lift people's self-evaluative concerns in the situation at hand and allow other motivations, such as a desire to be even-handed, rational, or healthful, to predominate. But, what are the psychological mechanisms implicated in this process? What do the self-affirmation manipulations do to individuals that result in such effects?

One point seems important to acknowledge. For phenomena as rich and complex as defensive resistance to persuasion, stereotyping, and dissonance reduction, there is unlikely to be one mechanism or mediator through

which affirmation produces the effects it does. It seems likely that there are a number of affective, cognitive, and motivational processes acting in concert to produce self-affirmation effects.

One way to address this question is to probe other theories of self-defense for underlying commonalities. Tesser (2000) has argued that self-esteem maintenance processes are essentially interchangeable with self-affirmation processes. One focus of self-esteem maintenance theory (Tesser, 1988) is how people respond to potentially threatening social comparisons. In performance settings, they behave in ways that maintain their self-esteem, for example, by sabotaging the performance of close others (rather than that of strangers) on important intellectual tasks (Tesser & Smith, 1980). Moreover, supporting the idea of interchangeability or fluidity in self-defensive processes, affirmation of an important value reduces people's tendency to sabotage the performance of friends working on intellectual tasks (Tesser & Cornell, 1991). People are also more likely to spontaneously self-affirm, for example, by writing essays with more self-affirming content, after making a self-threatening upward comparison (Tesser, Crepaz, Collins, Cornell, & Beach, 2000).

Tesser (2000) argues that these studies support a basic interchangeability of self-evaluative processes. Importantly, he asks whether the different self-related processes identified by self-affirmation theory and self-esteem maintenance theory share the common mechanism of maintaining positive affect or mood. From this perspective, the effect of self-affirmation on behavior should be mediated by a reduction in negative affect. Contrary to this expectation, self-affirmation studies generally find that affirmation of important values has no effect on self-reported mood (Fein & Spencer, 1997; Schmeichel & Martens, 2005; Sherman et al., 2000; Spencer et al., 2001) and that manipulations of mood fail to produce the same effects as self-affirmations (e.g., Steele et al., 1993; but see Raghunathan & Trope, 2002). Moreover, if affirmations were operating via positive mood, then one would expect affirmed individuals to be accepting of persuasive arguments regardless of the strength or weakness of those arguments (Mackie & Worth, 1989). The finding that self-affirmation makes people more responsive to the strength of persuasive arguments (Correll et al., 2004) runs contrary to this expectation.

On the other hand, an alternative possibility is that the common mechanism is affect of which a person is unaware (Tesser, Martin, & Cornell, 1996). While it is clear that much cognition lies beneath conscious awareness (Bargh & Chartrand, 1999; Dijksterhuis, 2004; Wilson, 2002), an interesting question concerns whether affect can lie beneath conscious awareness as well. The use of implicit measures of affect (i.e., measures that circumvent conscious control) would help to address this issue. As one example, Koole et al. (1999) found that a measure of nonconscious affect (i.e., a word

fragment completion task) mediated the effect of self-affirmation on the reduction in rumination reported previously. That is, self-affirmation led to increases in implicit positive mood, which in turn was associated with decreases in ruminative thought.

State self-esteem is another mediational candidate. Self-affirmations might boost state self-esteem and thus enable people to accept conclusions that might otherwise lower their state self-esteem below an acceptable threshold. Consistent with this reasoning, Fein and Spencer (1997) gave people positive personality feedback and found that this self-affirmation not only reduced stereotyping but also raised state self-esteem. Moreover, the increase in state self-esteem mediated the reduction in stereotyping. However, in other studies, affirmations of personal values have not been found to boost self-esteem. Schmeichel and Martens (2005) extensively tested the impact of affirmation (i.e., writing about an important personal value) on state self-esteem. While they did find effects on the primary outcomes featured in that study (worldview defense and death-thought accessibility), they found no effects on self-esteem. It appears that self-affirmations can have effects similar to self-esteem effects, but without influencing self-esteem.

One potential mediator that has not been tested, but that may hold promise, is self-certainty. That is, reflecting on a core source of self-integrity, such as one's long-held values, may make people more certain of their identity and their priorities. Such certainty could anchor individuals' self-worth and make them less susceptible to self-integrity threats. Future research examining this issue could adopt measures used to assess self-concept clarity (Campbell, 1990; Markus, 1977), for example by having people indicate "me"/"not me" to a series of trait adjectives. To the extent that affirmation makes people more certain of the self, they may respond more quickly to such questions when they involve self-descriptive traits (see also McGregor & Marigold, 2003).

To close, while different individual studies have found evidence for potential mechanisms underlying self-affirmation effects—including elevations in implicit positive affect (Koole et al., 1999), in state self-esteem (Cohen et al., 2000; Fein & Spencer, 1997), in collective self-esteem (Sherman & Kim, 2005), and in message scrutiny (Correll et al., 2004)—few if any consistent mediators have emerged across multiple studies. Thus, this remains an important topic for future research. In future studies, researchers should also consider the proposal of Spencer, Zanna, and Fong (2005). In questioning whether experiments employing the mediational analyses suggested by Baron and Kenny (1986) are a proper "gold standard" in social psychology, they argue that designs that rely on mediational analyses should only be preferred when measurement of a psychological process is easy and manipulation of it is difficult. In contrast, an alternative strategy in the

present context would be to experimentally manipulate these posited psychological mediators independently of self-affirmation and to assess whether the standard effects of self-affirmation are achieved in their absence.

B. ARE PEOPLE AWARE OF SELF-AFFIRMATION PROCESSES?

One important question concerns whether people are aware of the self-affirmation process or of the effects of affirmations on their behavior. Consider Al Franken's character of Stuart Smalley (1992), who parodied the self-help movement with his so-called daily affirmations, "I'm good enough, I'm smart enough and doggone it, people like me." It appears that his attempts at affirmation fall short, in part, because he is well aware that he is desperately trying to boost his own self-worth. Conversely, the process of self-affirmation, as described in this chapter, tends to occur outside of awareness, in that participants are generally unaware that the goal of the procedure is to modify or enhance their feelings of self-worth. (Moreover, in contrast to Stuart Smalley's self-help attempts, self-affirmation interventions usually entail affirming a *specific* area of self-integrity that is *unrelated* to the domain of threat.) We propose that awareness is a critical moderator of the effects of self-affirmation. When one is made aware of the link between a self-affirmation intervention and one's potential responses to a threatening event, the effectiveness of affirmation may be diminished. In a series of studies, we examined the role of awareness in self-affirmation processes (Sherman et al., 2005).

In one pair of studies, we adopted the methodology of Gilbert et al. (1998) who used two groups of participants—experiencers and forecasters—to compare the actual effects of a stimulus or event with their anticipated effect. In the first group of participants (the experiencers), we demonstrated that self-affirmation of an important value unrelated to identity-threatening news made people more open to counterattitudinal information. In this case, environmentalists who had affirmed an alternative value proved more open to the ostensibly "good" but identity-challenging news that ecological catastrophe in the near future was unlikely. For a second group of environmentalists (the forecasters), we simply *described* the procedure of writing the value-affirming essay and presented them the identity-threatening news report. Then we asked participants whether and how writing the essay would affect their evaluation of the news report. Participants evidenced no awareness that the affirmation exercise would affect openness to the essay. The rated influence of the value-affirming essay was no greater than the rated influence of a neutral essay.

A more direct test of the role of awareness is to make people aware of the potential impact of the self-affirmation exercise and to assess the impact of such awareness on the efficacy of the affirmation. In one study, we examined optimistic illusions about invulnerability to disease; such illusions are notoriously difficult to undo (Weinstein & Klein, 1995). We found that such illusions were reduced via a self-affirmation procedure. Participants who completed a values scale relevant to an important value proved more likely to acknowledge that they, relative to the average student, were at some risk for negative health outcomes. However, affirmation had no impact on participants who had been informed that the value affirmation was linked to their assessments of physical and health risk (i.e., that the researchers were examining the connection between personal values and health assessments). Merely suggesting the potential link between a self-affirmation and one's subsequent response rendered the affirmation impotent.

A third study examined an implication of the argument that affirmation processes occur beneath conscious awareness: that the self-affirmation process can proceed entirely without participants' awareness either of the perceived connection between the affirmation and the outcome or the self-affirming stimulus itself. Adapting the procedure of Bargh, Chen, and Burrows (1996), we primed participants to be in an affirmed state and found that such implicit affirmation can reduce defensiveness (Sherman et al., 2005). Participants failed a difficult math test and then indicated their interest in pursuing various math-related careers (e.g., physicist, mathematician). The typical, defensive response to such failure is to disidentify from domain-relevant careers—that is, to downplay one's personal interest in pursuing math-related professions (e.g., Major et al., 1998). Before failing the math test, participants completed a sentence-unscrambling task. That is, they had to remove one word and unscramble the remaining words to make a sentence (e.g., ''he grape intelligent is''). In one condition, the words featured in some of the items involved self-affirming words (e.g., intelligent, triumphant, proficient), whereas in the other conditions the words involved either threatening words (e.g., unintelligent, defeat) or achievement-related control words (e.g., progress, effort). Participants who had completed the self-affirming, sentence–completion task showed much less defensive disidentification from math-related careers than did participants in either of the two other conditions. Indeed, the effect size of this priming manipulation was equal in magnitude to the effect size of participants' sex in predicting their interest in math-related careers. This was the case in spite of the fact that participants showed little, if any, awareness of the thematic content of the words used in the self-affirmation task and no awareness at all of the potential impact of this task on their interest in math careers.

These findings hold implications for how the process of affirmation may operate in everyday life. Affirmations may operate with subtlety and without a mediating role for awareness or conscious intent. Consider an individual going to the doctor and receiving information that his or her cholesterol is higher than it should be and without significant changes in diet, medication may be necessary. An initial response might be skepticism, the desire to seek a retest or find fault in the information (Ditto & Lopez, 1992). Yet, if people are to profit from this type of diagnostic information, they need to accept it at some level. However, this acceptance need not be immediate. Perhaps the individual returns home and browses the Internet, checking for information on a coming election (affirming a political identity), or examines the scores from last night's game (affirming a valued social identity). In such situations, people may think that they are procrastinating, but this procrastination may serve an important integrity-reparative function. As a consequence, a second appraisal of the health-risk information may cause less psychological pain and avoidance. Contrast this type of "non-conscious" self-affirmation with the type offered in an online self-help webpage for people with bipolar disorder:

> Self affirmation is you telling yourself positive things about yourself, making sure that you impart the same caring for yourself that you would provide for another . . . . Repeat affirmations to yourself each and every day, and each and every time your "dragon" attacks you with a negative thought (bipolarworld.net, 2005).

Such affirmations are intended to help people confront the threats of everyday life. However, we suspect that self-affirmation may prove more effective—as suggested by the results reviewed above—if it occurs without conscious intention. Indeed, for people in chronically stressful situations, affirmations may prove more effective when they are delivered less frequently—the better to camouflage their intended purpose (Cohen et al., 2006). In summary, affirmation appears to operate not in a self-conscious, deliberate manner, but rather in a subtle, nonconscious manner. Future research could address this issue more directly by experimentally manipulating whether participants are informed of the psychological benefits of self-affirmation, and assessing whether such affirmations are less effective when people are unaware, rather than aware, of their beneficial effects.

C. QUESTIONS OF SUPERORDINATE FUNCTIONS

What is the function of self-integrity? And why do people go to such lengths to protect self-integrity, to the point of even misrepresenting reality and misperceiving themselves? Although self-affirmation theory posits a self-system

motivated to protect worth and integrity, it does not speak to the ultimate purpose of such a system. The benefits of positive self-illusions for psychological and even physical health suggest the adaptive value, and perhaps even evolutionary advantage, of maintaining self-integrity (cf. Colvin, Block, & Funder, 1995; Taylor & Brown, 1988; Taylor & Sherman, in press). To return to the baseball analogy noted previously, for a ''good'' hitter who bats .300 but fails nearly 70% of the time, it seems important to maintain a sense of self-worth and efficacy in order to take advantage of those few opportunities where one could get a ''hit.'' Maintaining self-integrity in the face of threat could help sustain optimism and effort in the sometimes long wait for success. Here, we offer some additional, tentative speculations about the ''superordinate'' function of self-integrity.

One possibility is suggested by terror management theory—to wit, that the maintenance of self-integrity serves to stave off the existential terror that accompanies the knowledge of one's own mortality (Solomon, Greenberg, & Pyszczynski, 1991). According to this theory, people cope with the anxiety of impending mortality by identifying with a cultural worldview—a worldview that provides meaning and the possibility of symbolic immortality through participation in culturally sanctioned projects. In the wake of September 11th, for example, American flags appeared in great abundance, and patriotism increased dramatically. After this vivid reminder of death, people found comfort in asserting their American identities, by identifying with their fellow citizens, and by supporting the symbolic leader of their culture, the President (Landau et al., 2004; Pyszczynski, Solomon, & Greenberg, 2003). Another related way that people cope with the knowledge of their mortality is through the cultivation of self-esteem, which is achieved by being a good member of one's culture. According to terror management theorists, maintaining self-esteem (and perhaps self-integrity more generally) serves to buffer people against the terror associated with the knowledge of their own mortality (Greenberg, Solomon, Arndt, & Schimel, 2004; Harmon-Jones et al., 1997).

If cultural worldviews and the maintenance of self-esteem help to protect people from the terror associated with their mortality, then leading individuals to focus on their own mortality should increase both their endorsement of cultural worldviews and engagement in self-esteem protective behaviors. Numerous studies support these hypotheses (see Greenberg, Solomon, & Pyszczynski, 1997 for a review). Research by Schmeichel and Martens (2005) specifically examined the relationship between self-affirmation processes and terror management processes. They examined whether self-affirmation reduces the effects of mortality salience. Writing about an important value, for example, not only has self-affirming effects, but also serves to bolster one's worldview (Schmeichel & Martens, 2005). As

a consequence, people might have less need to stave off thoughts of their own mortality by resorting to various forms of worldview defense, such as derogation of outgroup members who hold views anathema to one's own.

In one study, a mortality salience manipulation—thinking about one's death—led participants to defend their cultural worldview. Participants increased their support of a pro-American essay writer over an anti-American essay writer—a standard terror management effect (e.g., Greenberg et al., 1990). However, this cultural worldview defense was eliminated among participants who had completed a self-affirmation—that is, among participants who had written about an important value unrelated to the threat. Thus, the affirmation of self-integrity seems to buffer people against the existential terror associated with the knowledge of their inevitable death. Such studies, however, do not address the question of motivational primacy. Does the maintenance of self-integrity serve the ultimate purpose of managing existential terror? Such questions of ultimate purpose are, of course, difficult to answer. However, it is clear that both terror management theory and self-affirmation theory provide complementary evidence concerning how the self-integrity maintenance system copes with threats to the self, both existential and otherwise.

From a more evolutionary perspective, an intriguing possibility is that the motivation to protect self-integrity serves to increase social fitness. As Trivers (2000) has argued, people's positive illusions and self-deceptions may be in the service of social deception. Individuals who create a public persona as a person of integrity—as someone whose goodness, strength, and ability to control important outcomes are clear—will acquire material and even reproductive advantages in the social hierarchy. It thus serves the individual well to hide, rationalize, and compensate for personal flaws and mistakes that might otherwise compromise his or her status in the eyes of others. A similar idea is found in Nisbett and Cohen's work (1996) on the Southern culture of honor: by projecting a social image of strength and willingness to retaliate, Southern men maintain their "honor" and consequently deter others from challenging their status or stealing their resources. The critical insight is this: the public posturing, social manipulations, and interpersonal deceptions needed to convince others of one's integrity will prove far more plausible and persuasive if the individual in question actually *believes* in the persona he or she is trying to project (Trivers, 2000). As noted in the sociometer theory of self-esteem (Leary & Baumeister, 2000), the ego-protective strategies people have evolved to maintain self-integrity may thus help to promote their social fitness. Future research might test this account by manipulating whether self-threats and self-affirmations are delivered in private or in public (such as in front of a socially significant audience). If the maintenance of self-integrity serves a social fitness function, then

self-threats may be more threatening and self-affirmations more affirming, when delivered in the presence of others.

VI. Implications for Interpersonal Relationships and Coping

Because people experience many potential threats to their self-integrity, an important question concerns the extent to which people use affirmational strategies in everyday life. To what extent do people spontaneously rely on this mechanism involved in the psychological immune system? Relatedly, what are the real-world analogs to the self-affirmation manipulations that have been featured in the laboratory? Research on close relationships and on coping and resiliency suggests some answers and points the way to exciting directions for future research.

A. INTERPERSONAL RELATIONSHIPS

When people rank their values in terms of their personal importance in self-affirmation studies, they consistently rank relationships with friends and family highly. Additionally, in a writing intervention study designed to elicit thoughts and feelings among patients about their illness—a study that provided an opportunity for self-affirming without an explicit prompt—people most frequently invoked the importance of their close relationships (Creswell et al., 2006). Thus, personal relationships seem to be an important affirmational resource that people draw on in times of stress. Several studies in the close-relationships literature have examined this idea in some depth.

People can use their relationships and their romantic partners as affirmational resources by emphasizing the love they receive from their partners (Kumashiro & Sedikides, 2005; Murray, Bellavia, Feeney, Holmes, & Rose, 2001; Swann, De La Ronde, & Hixon, 1994). People with high self-esteem in particular seem to respond to threats to their self-image by affirming the integrity of their relationships (Murray et al., 1998).

Threats can also exist within relationships, as for example when one partner outperforms the other. Tesser (1988) has found that upward social comparisons within a relationship could constitute a threat, particularly if the domain is of high personal importance. One adaptive way to respond to such a threat is to draw upon the relationship as an affirmational resource (and perhaps also bask in the glory of one's partner). In a series of studies, people in close relationships were able to use the relationship as an affirmation after engaging in a potentially threatening upward social comparison

with their partner (Lockwood, Dolderman, Sadler, & Gerchak, 2004). When dealing with this potential within-relationship threat, those people in close relationships affirmed their relationships by emphasizing the warmth and kindness of the relationship.

Many questions remain for researchers interested in the how personal relationships serve such affirming functions. For example, the affirmation framework may be useful in understanding the relationship between possessing multiple roles and depression. Sociological studies have found that people who have multiple roles, such as family and career, tend to have reduced stress and depression—at least as long as the number of roles remains below a manageable threshold (e.g., Gore & Mangione, 1983). For women in particular, it has been found that although work and family can be great strains on resources, those who participate in multiple roles have reduced levels of depression (Kandel, Davies, & Raveis, 1985). These studies raise questions of causality, as those who have better mental health may be more able to handle multiple roles without becoming depressed (Kandel et al., 1985). This caveat notwithstanding, the self-affirmation framework provides a possible explanation for the observed phenomenon. That is, perhaps multiple roles function as affirmational resources allowing people to use one domain of life to affirm the self when threatened in the other. Adopting the methodological approaches of close relationships researchers, such as use of daily diary reports and experience sampling, is an exciting direction for researchers interested in how people use relationships and other affirmational resources to cope with threats and manage stress on a day-to-day level.

B. COPING AND RESILIENCE

The assumption that people need to work through their grief, and to experience negative emotions after facing loss or trauma, has been challenged by work on resilience—the ability to maintain a stable equilibrium after a loss (Bonanno, 2004; Wortman & Silver, 1989). The failure to experience negative emotions after trauma was thought to be associated with delayed grief and greater psychological problems in the future; however, evidence does not support this intuition that people must experience grief in order to "move on" (Bonanno & Field, 2001). By contrast, those who express positive emotions after a loss are more likely to exhibit resilience in subsequent years (Bonanno & Keltner, 1997).

What are some of the factors that promote resiliency in the face of loss? One factor seems to be the ability to self-enhance (Bonanno, 2004)—to perceive the self as possessing more positive attributes than the average person (Taylor & Brown, 1988). In support of this claim, measured levels

of self-enhancement were associated with better adjustment among Bosnian civilians after the civil war in Sarajevo (Bonanno, Field, Kovacevic, & Kaltman, 2002). In another study involving people near the World Trade Center at the time of the 9/11 terrorist attacks, self-enhancement (as assessed by endorsement of such statements as ''I am fully in control of my own fate,'' and ''I always know why I do things''; see Paulhus, 1984) prospectively predicted a reduction in both depression and in Post Traumatic Stress Disorder symptoms (Bonanno, Rennicke, Dekel, & Rosen, 2005).

It seems plausible that self-affirmation is another mechanism by which people remain resilient after trauma. Just as the ability to self-enhance may help people to muster valuable resources to grapple with life challenges, the ability to affirm alternative domains of self-integrity in the face of threat may facilitate better coping with the trauma or loss (Taylor & Sherman, in press). Affirmation of other aspects of self-integrity—such as religion, relationships, work, or hobbies—may enable people to deal constructively with the threats to perceived control, meaning, and significance that issue from real-world loss and trauma.

Future research is needed to identify whether and how people use self-affirmation to cope with trauma and loss. Just as self-affirmation can lead to greater acceptance and less defensiveness across a wide range of threatening situations featured in laboratory studies, it will be exciting to explore whether and when it can lead to more adaptive coping among those who have experienced trauma.

VII. Conclusions

In her seminal review on motivated cognition, Kunda (1990) observed:

> [M]otivated illusions can be dangerous when they are used to guide behavior and decisions, especially in those cases in which objective reasoning could facilitate more adaptive behavior. For example, people who play down the seriousness of early symptoms of severe diseases such as skin cancer and people who see only weaknesses in research pointing to the dangers of drugs such as caffeine or of behaviors such as drunken driving may literally pay with their lives for their motivated reasoning. Hopefully, once the mechanisms producing such biases are fully understood, it will be possible to help people overcome them.
>
> (Kunda, 1990, p. 495–496)

Self-affirmation theory provides a framework to understand and over-come such biases. At both the individual and collective levels, important domains of functioning—health, political decision-making, conflict, rela-tionships, academic performance—call forth the motivation to defend the

self. People defensively distort, deny, and misrepresent reality in a manner that protects self-integrity. The cost of doing so, of course, is that they miss potential opportunities for learning and growth that, if acted upon, could otherwise increase their adaptiveness in the long term. However, in the face of daily threats, people can also protect self-integrity through the affirmation of alternative sources of self-identity. Doing so helps them to accept experiences and information that, although threatening, hold important lessons for self-change. By illuminating the psychology of self-defense, the research reviewed in this chapter offers practitioners, teachers, clinicians, mediators, and interventionists more generally theory-driven strategies for overcoming self-defense and encouraging self-improvement.

Acknowledgments

References

Abelson, R. (1986). Beliefs are like possessions. *Journal for the Theory of Social Behavior, 16,* 223–250.

Abrams, D., & Hogg, M. (1988). Comments on the motivational status of self-esteem in social identity and intergroup discrimination. *European Journal of Social Psychology, 18,* 317–334.

Adams, G., Tormala, T. T., & O'Brien, L. T. (in press). The effect of self-affirmation on perceptions of racism. *Journal of Experimental Social Psychology.*

Allport, G. W. (1961). *Pattern and growth in personality.* New York: Holt, Rinehart and Wonston.

Armor, D. A., & Taylor, S. E. (2002). When predictions fail: The dilemma of unrealistic optimism. In T. Gilovich, D. Griffin, and D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 334–347). New York: Cambridge University Press.

Aronson, E. (1968). Dissonance theory: Progress and problems. In R. P. Abelson, E. Arson, W. J. McGuire, T. M. Newcomb, M. J. Rosenberg, and P. H. Tannembaum (Eds.), *Theories of cognitive consistency: A sourcebook* (pp. 5–27). Chicago: Rand McNally.

Aronson, J., Blanton, H., & Cooper, J. (1995). From dissonance to disidentification: Selectivity in the self-affirmation process. *Journal of Personality and Social Psychology, 68,* 986–996.

Aronson, J., Cohen, G. L., & Nail, P. R. (1999). Self-affirmation theory: An update and appraisal. In E. Harmon-Jones and J. Mills (Eds.), *Cognitive dissonance theory: Revival with revisions and controversies* (pp. 127–147). Washington, DC: American Psychological Association.

Aronson, J., Fried, C. B., & Good, C. (2002). Reducing the effects of stereotype threat on African American college students by shaping theories of intelligence. *Journal of Experimental Social Psychology, 38,* 113–125.

Bargh, J. A., & Chartrand, T. L. (1999). The unbearable automaticity of being. *American Psychologist, 54,* 462–476.

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

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51,* 1173–1182.

Baumeister, R. F., Heatherton, T. F., & Tice, D. M. (1993). When ego threats lead to self-regulation failure: Negative consequences of high self-esteem. *Journal of Personality and Social Psychology, 64,* 141–156.

Beauregard, K. S., & Dunning, D. (1998). Turning up the contrast: Self-enhancement motives prompt egocentric contrast effects in social judgments. *Journal of Personality and Social Psychology, 74,* 606–621.

Ben-Zeev, T., Fein, S., & Inzlicht, M. (2005). Arousal and stereotype threat. *Journal of Experimental Social Psychology, 41,* 174–181.

Bettman, J. R., & Weitz, B. A. (1983). Attributions in the boardroom: Causal reasoning in corporate annual reports. *Administrative Science Quarterly, 28,* 165–183.

Bipolarworld.net. (2005). *Affirmation.* Retrieved July 1, 2005 from http://www.bipolarworld.net/Treatments/Self%20Help/self_affirmation.htm

Blanton, H., Cooper, J., Skurnik, I., & Aronson, J. (1997). When bad things happen to good feedback: Exacerbating the need for self-justification with self-affirmations. *Personality and Social Psychology Bulletin, 23,* 684–692.

Blanton, H., Pelham, B. W., DeHart, T., & Carvallo, M. (2001). Overconfidence as dissonance reduction. *Journal of Experimental Social Psychology, 37,* 373–385.

Blascovich, J., Spencer, S. J., Quinn, D., & Steele, C. (2001). African Americans and high blood pressure: The role of stereotype threat. *Psychological Science, 12,* 225–229.

Bonanno, G. A., & Field, N. P. (2001). Evaluating the delayed grief hypothesis across 5 years of bereavement. *American Behavioral Scientist, 44,* 798–816.

Bonanno, G. A., & Keltner, D. (1997). Facial expressions of emotion and the course of conjugal bereavement. *Journal of Abnormal Psychology, 106,* 126–137.

Bonanno, G. A., Field, N. P., Kovacevic, A., & Kaltman, S. (2002). Self-enhancement as a buffer against extreme adversity: Civil war in Bosnia and traumatic loss in the United States. *Personality and Social Psychology Bulletin, 28,* 184–196.

Bonanno, G. A., Rennicke, C., Dekel, S., & Rosen, J. (2005). Self-enhancement among high-exposure survivors of the September 11th terrorist attack: Resilience or social maladjustment? *Journal of Personality and Social Psychology, 88,* 984–988.

Bonanno, G. A. (2004). Loss, trauma, and human resilience: Have we underestimated the human capacity to thrive after extremely aversive events? *American Psychologist, 59,* 20–28.

Boninger, D. S., Krosnick, J. A., & Berent, M. K. (1995). Origins of attitude importance: Self-interest, social identification, and value relevance. *Journal of Personality and Social Psychology, 68,* 61–80.

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

Brown, B. R. (2000). When moral entitlement leads to immoral acts. Doctoral dissertation, Stanford University. *Dissertation Abstracts International: Sciences and Engineering, 61,* 1132.

Brown, J. D., & Smart, S. A. (1991). The self and social conduct: Linking self-representations to prosocial behavior. *Journal of Personality and Social Psychology, 60,* 368–375.

Cadinu, M. R., & Rothbart, M. (1996). Self-anchoring and differentiation processes in the minimal group setting. *Journal of Personality and Social Psychology, 70,* 661–677.

Campbell, J. D. (1990). Self-esteem and clarity of the self-concept. *Journal of Personality and Social Psychology, 59,* 538–549.

Cohen, G. L. (2003). Party over policy: The dominating impact of group influence on political beliefs. *Journal of Personality and Social Psychology, 85,* 808–822.

Cohen, G. L., Aronson, J., & Steele, C. M. (2000). When beliefs yield to evidence: Reducing biased evaluation by affirming the self. *Personality and Social Psychology Bulletin, 26,* 1151–1164.

Cohen, G. L., & Garcia, J. (2005). "I am us": Negative stereotypes as collective threats. *Journal of Personality and Social Psychology, 89,* 566–582.

Cohen, G. L., Garcia, J., Apfel, N., & Master, A. (2006). A social-psychological intervention to reduce the racial achievement gap in school. Manuscript under review.

Cohen, G. L., Sherman, D. K., McGoey, M., Hsu, L., Bastardi, A., & Ross, L. D. (2005). Bridging the partisan divide: Self-affirmation and the reduction of ideological bias. Manuscript under review.

Cohen, G. L., Steele, C. M., & Ross, L. D. (1999). The mentor's dilemma: Providing critical feedback across the racial divide. *Personality and Social Psychology Bulletin, 25,* 1302–1318.

Correll, J., Spencer, S. J., & Zanna, M. P. (2004). An affirmed self and an open mind: Self-affirmation and sensitivity to argument strength. *Journal of Experimental Social Psychology, 40,* 350–356.

Colvin, C. R., Block, J., & Funder, D. C. (1995). Overly positive self-evaluations and personality: Negative implications for mental health. *Journal of Personality and Social Psychology, 68,* 1152–1162.

Creswell, J. D., Lam, S., Stanton, A. L., Taylor, S. E., Bower, J. E., & Sherman, D. K. (2006). Does self-affirmation, cognitive processing, or discovery of meaning explain cancer-related health benefits of expressive writing? Manuscript under review.

Creswell, J. D., Welch, W., Taylor, S. E., Sherman, D. K., Gruenewald, T., & Mann, T. (2005). Affirmation of personal values buffers neuroendocrine and psychological stress responses. *Psychological Science, 16,* 846–851.

Crocker, J., & Major, B. (1989). Social stigma and self-esteem: The self-protective properties of stigma. *Psychological Review, 96,* 608–630.

Crocker, J., & Wolfe, C. T. (2001). Contingencies of self-worth. *Psychological Review, 108,* 593–623.

Dal Cin, S., MacDonald, T. K., Fong, G. T., Zanna, M. P., & Elton-Marshall, T. E. (in press). Remembering the message: Using a reminder cue to increase condom use following a safer sex intervention. *Health Psychology.*

Dawson, E., Gilovich, T., & Regan, D. T. (2002). Motivated reasoning and performance on the Wason selection task. *Personality and Social Psychology Bulletin, 28,* 1379–1387.

De Dreu, C. K. W., & Carnevale, P. J. (2003). Motivational bases of information processing and strategy in negotiation and social conflict. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 35, pp. 235–290). San Diego: Academic Press.

Deaux, K. (1996). Social identification. In E. T. Higgins and A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 777–798). New York: Guilford.

Deci, E. L., & Ryan, R. M. (1995). Human autonomy: The basis for true self-esteem. In M. H. Kernis (Ed.), *Efficacy, agency, and self-esteem* (pp. 31–49). New York: Plenum.

Dickerson, S. S., & Kemeny, M. E. (2004). Acute stressors and cortisol responses: A theoretical integration and synthesis of laboratory research. *Psychological Bulletin, 130,* 355–391.

Dijksterhuis, A. (2004). Think different: The merits of unconscious thought in preference development and decision making. *Journal of Personality and Social Psychology, 87,* 586–598.

Ditto, P. H., & Lopez, D. F. (1992). Motivated skepticism: Use of differential decision criteria for preferred and nonpreferred conclusions. *Journal of Personality and Social Psychology, 63,* 569–584.

Ditto, P. H., Scepansky, J. A., Munro, G. D., Apanovitch, A. M., & Lockhart, L. K. (1998). Motivated sensitivity to preference-inconsistent information. *Journal of Personality and Social Psychology, 75,* 53–69.

Dunning, D. (2003). The zealous self-affirmer: How and why the self lurks so pervasively behind social judgment. In S. J. Spencer, S. Fein, M. P. Zanna, and J. M. Olson (Eds.), *Motivated social perception: The Ontario symposium* (Vol. 9, pp. 45–72). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Dunning, D., & Beauregard, K. S. (2000). Regulating impressions of others to affirm images of the self. *Social Cognition, 18,* 198–222.

Dunning, D., & Cohen, G. L. (1992). Egocentric definitions of traits and abilities in social judgment. *Journal of Personality and Social Psychology, 63,* 341–355.

Dunning, D., Griffin, D. W., Milojkovic, J. H., & Ross, L. (1990). The overconfidence effect in social prediction. *Journal of Personality and Social Psychology, 58,* 568–592.

Dunning, D., & Hayes, A. F. (1996). Evidence for egocentric comparison in social judgment. *Journal of Personality and Social Psychology, 71,* 213–229.

Dunning, D., Leuenberger, A., & Sherman, D. A. (1995). A new look at motivated inference: Are self-serving theories of success a product of motivational forces? *Journal of Personality and Social Psychology, 69,* 58–68.

Dweck, C. S., & Leggett, E. (1988). A social cognitive approach to motivation and personality. *Psychological Review, 95,* 256–273.

Ellsworth, P. C., & Ross, L. (1983). Public opinion and capital punishment: A close examination of the views of abolitionists and retentionists. *Crime and Delinquency, 29,* 116–169.

Fein, S., & Spencer, S. J. (1997). Prejudice as self-image maintenance: Affirming the self through derogating others. *Journal of Personality and Social Psychology, 73,* 31–44.

Fein, S., Hoshino-Browne, E., Davies, P. G., & Spencer, S. J. (2003). Self-image maintenance goals and sociocultural norms in motivated social perception. In S. J. Spencer, S. Fein, M. P. Zanna, and J. M. Olson (Eds.), *Motivated social perception: The Ontario symposium* (Vol. 9, pp. 21–44). Mahwah, NJ: Lawrence Erlbaum Associates.

Festinger, L. A. (1957). *A theory of cognitive dissonance.* Evanston, IL: Row, Peterson.

Franken, A. (1992). *I'm good enough, I'm smart enough, and doggone it, people like me! Daily affirmations by Stuart Smalley.* New York: Dell Publishing.

Gaertner, L., Sedikides, C., & Graetz, K. (1999). In search of self-definition: Motivational primacy of the individual self, motivational primacy of the collective self, or contextual primacy? *Journal of Personality and Social Psychology, 76,* 5–18.

Gaertner, S. L., & Dovidio, J. F. (1986). The aversive form of racism. In J. F. Dovidio and S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 61–89). San Diego: Academic Press.

Galinsky, A. D., Stone, J., & Cooper, J. (2000). The reinstatement of dissonance and psychological discomfort following failed affirmations. *European Journal of Social Psychology, 30,* 123–147.

Gerber, A., & Green, D. (1999). Misperceptions about perceptual bias. *Annual Review of Political Science, 2,* 189–210.

Gibbons, F. X., Eggleston, T. J., & Benthin, A. (1997). Cognitive reactions to smoking relapse: The reciprocal relation of dissonance and self-esteem. *Journal of Personality and Social Psychology, 72,* 184–195.

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.

Glaser, R., & Kiecolt-Glaser, J. (1994). *Handbook of human stress and immunity*. New York: Academic Press.

Gore, S., & Mangione, T. W. (1983). Social roles, sex role, and psychological distress: Addditive and interactive models of sex differences. *Journal of Health and Social Behavior, 24,* 300–312.

Gramzow, R. H., & Gaertner, L. (2005). Self-esteem and favoritism toward novel in-groups: The self as an evaluative base. *Journal of Personality and Social Psychology, 88,* 801–815.

Greenberg, J., Pyszczynski, T., Solomon, S., Rosenblatt, A., Veeder, M., Kirkland, S., & Lyon, D. (1990). Evidence for terror management theory II: The effects of mortality salience on reactions to those who threaten or bolster the cultural worldview. *Journal of Personality and Social Psychology, 58,* 308–318.

Greenberg, J., Solomon, S., & Pyszczynski, T. (1997). Terror management theory of self-esteem and cultural worldviews: Empirical assessments and conceptual refinements. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 29, pp. 61–139). San Diego: Academic Press.

Greenwald, A. G. (1980). The totalitarian ego: Fabrication and revision of personal history. *American Psychologist, 35,* 603–618.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology, 74,* 1464–1480.

Harmon-Jones, E., Simon, L., Greenberg, J., Pyszczynski, T., Solomon, S., & McGregor, H. (1997). Terror management theory and self-esteem: Evidence that increased self-esteem reduces mortality salience effects. *Journal of Personality and Social Psychology, 72,* 24–36.

Harris, P. R., & Napper, L. (2005). Self-affirmation and the biased processing of threatening health-risk information. *Personality and Social Psychology Bulletin, 31,* 1250–1263.

Hastorf, A., & Cantril, H. (1954). They saw a game: A case study. *Journal of Abnormal and Social Psychology, 49,* 129–134.

Heine, S. J. (2005). Constructing good selves in Japan and North America. In R. M. Sorrentino, D. Cohen, J. M. Olson, and M. P. Zanna (Eds.), *Cultural and social behavior: The Ontario Symposium* (Vol. 10, pp. 95–116). Mahwah, NJ: Lawrence Erlbaum Associates.

Heine, S. J., & Lehman, D. R. (1997). Culture, dissonance, and self-affirmation. *Personality and Social Psychology Bulletin, 23,* 389–400.

Heine, S. J., Lehman, D. R., Markus, H., & Kitayama, S. (1999). Is there a universal need for positive self-regard? *Psychological Review, 106,* 766–794.

Hobfoll, S. E. (1989). Conservation of resources: A new attempt at conceptualizing stress. *American Psychologist, 44,* 513–524.

Hoshino-Browne, E., Zanna, A. S., Spencer, S. J., & Zanna, M. P. (2004). Investigating attitudes cross-culturally: A case of cognitive dissonance among East Asians and North Americans. In G. Haddock and G. R. Maio (Eds.), *Contemporary perspectives on the psychology of attitudes* (pp. 375–397). New York: Psychology Press.

Hoshino-Browne, E., Zanna, A. S., Spencer, S. J., Zanna, M. P., Kitayama, S., & Lackenbauer, S. (2005). On the cultural guises of cognitive dissonance: The case of Easterners and Westerners. *Journal of Personality and Social Psychology, 89,* 294–310.

Jacks, J. Z., & O'Brien, M. E. (2004). Decreasing resistance by affirming the self. In E. S. Knowles and J. A. Linn (Eds.), *Resistance and persuasion* (pp. 235–257). Mahwah, NJ: Lawrence Erlbaum Associates.

Jemmott, J. B., III, Ditto, P. H., & Croyle, R. T. (1986). Judging health status: Effects of perceived prevalence and personal relevance. *Journal of Personality and Social Psychology, 50,* 899–905.

Jordan, C. H, Spencer, S. J., & Zanna, M. P. (2003). "I love me…I love me not:" Implicit self-esteem, explicit self-esteem, and defensiveness. In S. J. Spencer, S. Fein, M. P. Zanna, and J. M. Olson (Eds.), *Motivated social perception: The Ontario symposium* (Vol. 9, pp. 117–145). Mahwah, NJ: Lawrence Erlbaum Associates.

Jordan, C.,H, Spencer, S. J., Zanna, M. P., Hoshino-Browne, E., & Correll, J. (2003). Secure and defensive high self-esteem. *Journal of Personality and Social Psychology, 85,* 969–978.

Kandel, D. B., Davies, M., & Raveis, V. H. (1985). The stressfulness of daily social roles for women: Marital, occupational, and household roles. *Journal of Health and Social Behavior, 26,* 64–78.

Keough, K. A. (1998). When the self is at stake: Integrating the self into stress and physical health research (Doctoral dissertation, Stanford University, 1998). *Dissertation Abstracts International, 58,* 3959.

Keough, K. A., Garcia, J., & Steele, C. M. (1998). Reducing stress and illness by affirming the self. Unpublished manuscript, Stanford University.

Kernis, M. H., & Paradise, A. W. (2002). Distinguishing between secure and fragile forms of high self-esteem. In E. L. Deci and R. M. Ryan (Eds.), *Handbook of self-determination research* (pp. 339–360). Rochester, NY: University of Rochester Press.

Kim, H., & Markus, H. R. (1999). Deviance or uniqueness, harmony or conformity: A cultural analysis. *Journal of Personality and Social Psychology, 77,* 785–800.

Kirschbaum, C., Pirke, K., & Hellhammer, D. H. (1993). The 'Trier Social Stress Test': A tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology, 28,* 76–81.

Kitayama, S., Snibbe, A. C., Markus, H. R., & Suzuki, T. (2004). Is there any "free" choice? Self and dissonance in two cultures. *Psychological Science, 15,* 527–533.

Koole, S. L., Smeets, K., van Knippenberg, A., & Dijksterhuis, A. (1999). The cessation of rumination through self-affirmation. *Journal of Personality and Social Psychology, 77,* 111–125.

Kruger, J. M., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology, 77,* 1121–1134.

Kumashiro, M., & Sedikides, C. (2005). Taking on board liability-focused information: Close positive relationships as a self-bolstering resource. *Psychological Science, 16,* 732–739.

Kunda, Z. (1987). Motivated inference: Self-serving generation and evaluation of causal theories. *Journal of Personality and Social Psychology, 53,* 636–647.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin, 108,* 480–498.

Landau, M. J., Solomon, S., Greenberg, J., Cohen, F., Pyszczynski, T., Arndt, J., Miller, C. H., Ogilvie, D. M., & Cook, A. (2004). Deliver us from evil: The effects of mortality salience and reminders of 9/11 on support for president George W. Bush. *Personality and Social Psychology Bulletin, 30,* 1136–1150.

Langer, E. J. (1975). The illusion of control. *Journal of Personality and Social Psychology, 32,* 311–328.

Lau, R. R., & Russell, D. R. (1980). Attributions in the sports pages. *Journal of Personality and Social Psychology, 39,* 29–38.

Lazarus, R. S. (1993). From psychological stress to the emotions: A history of changing outlooks. *Annual Review of Psychology, 44,* 1–21.

Leary, M. R., & Baumeister, R. F. (2000). The nature and function of self-esteem: Sociometer theory. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 32, pp. 1–62). San Diego: Academic Press.

Lerner, M. J. (1980). *The belief in a just world: A fundamental delusion*. New York: Plenum Press.

Liberman, A., & Chaiken, S. (1992). Defensive processing of personally relevant health messages. *Personality and Social Psychology Bulletin, 18,* 669–679.

Liu, T. J., & Steele, C. M. (1986). Attributional analysis as self-affirmation. *Journal of Personality and Social Psychology, 51,* 531–540.

Lockwood, P., Dolderman, D., Sadler, P., & Gerchak, E. (2004). Feeling better about doing worse: Social comparisons within romantic relationships. *Journal of Personality and Social Psychology, 87,* 91–103.

Lockwood, P., & Kunda, Z. (1997). Superstars and me: Predicting the impact of role models on the self. *Journal of Personality and Social Psychology, 73,* 91–103.

Lord, C., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology, 37,* 2098–2109.

Mackie, D. M., & Worth, L. T. (1989). Processing deficits and the mediation of positive affect in persuasion. *Journal of Personality and Social Psychology, 57,* 27–40.

Major, B., Spencer, S. J., Schmader, T., Wolfe, C., & Crocker, J. (1998). Coping with negative stereotypes about intellectual performance: The role of psychological disengagement. *Personality and Social Psychology Bulletin, 24,* 34–50.

Markus, H. (1977). Self-schemata and processing information about the self. *Journal of Personality and Social Psychology, 35,* 63–78.

Markus, H., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review, 98,* 224–253.

Martens, A., Johns, M., Greenberg, J., & Schimel, J. (2006). Combating stereotype threat: The effect of self-affirmation on women's intellectual performance. *Journal of Experimental Social Psychology, 42,* 236–243.

Martin, L. L., & Tesser, A. (1996). Some ruminative thoughts. In R. S. Wyer (Ed.), *Advances in social cognition* (Vol. 9, pp. 1–47). Hillsdale, NJ: Erlbaum.

Matz, D. C., & Wood, W. (2005). Cognitive dissonance in groups: The consequences of disagreement. *Journal of Personality and Social Psychology, 88,* 22–37.

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

McGregor, I., & Marigold, D. C. (2003). Defensive zeal and the uncertain self: What makes you so sure? *Journal of Personality and Social Psychology, 85,* 838–852.

McQueen, A., & Klein, W. (2005). Experimental manipulations of self-affirmation: A systematic review. Manuscript under review.

Miller, D. T., & Ross, M. (1975). Self-serving biases in the attribution of causality: Fact or fiction? *Psychological Bulletin, 82,* 213–225.

Monin, B., & Miller, D. T. (2001). Moral credentials and the expression of prejudice. *Journal of Personality and Social Psychology, 81,* 33–43.

Monteith, M. J., & Voils, C. I. (1998). Proneness to prejudiced responses: Toward understanding the authenticity of self-reported discrepancies. *Journal of Personality and Social Psychology, 75,* 901–916.

Morris, K. A., & Swann, W. B. (1996). Denial and the AIDS crisis: On wishing away the threat of AIDS. In S. Oskamp and S. Thompson (Eds.), *Safer sex in the '90s: Understanding and preventing HIV risk behavior* (pp. 57–79). New York: Sage.

Munro, G. D., & Ditto, P. H. (1997). Biased assimilation, attitude polarization, and affect in reactions to stereotype-relevant scientific information. *Personality and Social Psychology Bulletin, 23,* 636–653.

Murphy, G. (1947). *Personality*. New York: Harper and Brothers.

Murray, S. L., Holmes, J. G., MacDonald, G., & Ellsworth, P. (1998). Through the looking glass darkly? When self-doubts turn into relationship insecurities. *Journal of Personality and Social Psychology, 75,* 1459–1480.

Murray, S. L., Bellavia, G., Feeney, B., Holmes, J. G., & Rose, P. (2001). The contingencies of interpersonal acceptance: When romantic relationships function as a self-affirmational resource. *Motivation and Emotion, 25,* 163–189.

Nail, P. R., Misak, J. E., & Davis, R. M. (2004). Self-affirmation versus self-consistency: A comparison of two competing self-theories of dissonance phenomena. *Personality and Individual Differences, 36,* 1893–1905.

Nisbett, R. E., & Ross, L. D. (1980). *Human inference: Strategies and shortcomings of social judgment.* Englewood Cliffs, NJ: Prentice Hall.

Nisbett, R. E., & Cohen, D. (1996). *Culture of honor: The psychology of violence in the South.* Boulder, CO: Westview Press.

Nolen-Hoeksema, S., & Morrow, J. (1991). A prospective study of depression and posttraumatic stress symptoms after a natural disaster: The 1989 Loma Prieta earthquake. *Journal of Personality and Social Psychology, 61,* 115–121.

Norton, M. I., Monin, B., Cooper, J., & Hogg, M. A. (2003). Vicarious dissonance: Attitude change from the inconsistency of others. *Journal of Personality and Social Psychology, 85,* 47–62.

O'Brien, L. T., & Crandall, C. S. (2003). Stereotype threat and arousal: Effects on women's math performance. *Personality and Social Psychology Bulletin, 29,* 782–789.

Otten, S. (2002). "Me and us" or "us and them"? The self as heuristic for defining minimal ingroups. *European Review of Social Psychology, 13,* 1–33.

Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology, 46,* 598–609.

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood method of persuasion. In L. Berkowitz (Ed.), *Advances in experimental social psychology.* (Vol. 19, pp. 123–205). New York: Academic Press.

Petty, R. E., & Wegener, D. T. (1998). Attitude change: Multiple roles for persuasion variables. In D. T. Gilbert, S. T. Fiske, and G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., pp. 323–390). New York: McGraw-Hill.

Pyszczynski, T., & Greenberg, J. (1987). Toward and integration of cognitive and motivational perspectives on social inference: A biased hypothesis-testing model. In L. Berkowitz (Ed.), *Advances in experimental social psycology* (Vol. 20, pp. 297–340). New York: Academic Press.

Pyszczynski, T., Greenberg, J., Solomon, S., Arndt, J., & Schimel, J. (2004). Why do people need self-esteem? A theoretical and empirical review. *Psychological Bulletin, 130,* 435–468.

Pyszczynski, T., Solomon, S., & Greenberg, J. (2003). *In the wake of 9/11: The psychology of terror.* Washington, DC: American Psychological Association.

Raghunathan, R., & Trope, Y. (2002). Walking the tightrope between feeling good and being accurate: Mood as a resource in processing persuasive messages. *Journal of Personality and Social Psychology, 83,* 510–525.

Reed, M. B., & Aspinwall, L. G. (1998). Self-affirmation reduces biased processing of health-risk information. *Motivation and Emotion, 22,* 99–132.

Rosenberg, M. (1965). *Society and the adolescent self-image.* Princeton, NJ: Princeton University Press.

Ross, L., & Ward, A. (1995). Psychological barriers to dispute resolution. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 27, pp. 255–304). San Diego: Academic Press.

Schimel, J., Arndt, J., Pyszczynski, T., & Greenberg, J. (2001). Being accepted for who we are: Evidence that social validation of the intrinsic self reduces general defensiveness. *Journal of Personality and Social Psychology, 80,* 35–52.

Schimel, J., Arndt, J., Banko, K. M., & Cook, A. (2004). Not all self-affirmations were created equal: The cognitive and social benefits of affirming the intrinsic (vs. extrinsic) self. *Social Cognition, 22,* 75–99.

Schmeichel, B. J., & Martens, A. (2005). Self-affirmation and mortality salience: Affirming values reduces worldview defense and death-thought accessibility. *Personality and Social Psychology Bulletin, 31,* 658–667.

Sears, C. L., & Funk, C. L. (1991). Are we reaching undergraduates? A survey of course offering in political psychology. *Political Psychology, 12,* 559–572.

Seery, M., Blascovich, J., Weisbuch, M., & Vick, S. B. (2004). The relationship between self-esteem level, self-esteem stability, and cardiovascular reactions to performance feedback. *Journal of Personality and Social Psychology, 87,* 133–145.

Sherman, D. A. K., Nelson, L. D., & Steele, C. M. (2000). Do messages about health risks threaten the self? Increasing the acceptance of threatening health messages via self- affirmation. *Personality and Social Psychology Bulletin, 26,* 1046–1058.

Sherman, D. K., & Cohen, G. L. (2002). Accepting threatening information: Self-affirmation and the reduction of defensive biases. *Current Directions in Psychological Science, 11,* 119–123.

Sherman, D. K., Cohen, G. L., Nelson, L. D., Bastardi, A., Nussbaum, A. D., & Ross, L. (2005). Affirmed yet unaware: The role of awareness in the process of self-affirmation. Manuscript under review.

Sherman, D. K., & Kim, H. S. (2005). Is there an ''I'' in ''team''? The role of the self in group-serving judgments. *Journal of Personality and Social Psychology, 88,* 108–120.

Sherman, D. K., Kinias, Z., Major, B., Kim, H. S., & Prenovost, M. (2005). Group identification, group affirmation, and group-serving judgments. Manuscript in preparation.

Sherman, D. K., Nelson, L. D., & Ross, L. D. (2003). Naïve realism and affirmative action: Adversaries are more similar than they think. *Basic and Applied Social Psychology, 25,* 275–290.

Shrira, I., & Martin, L. L. (2005). Stereotyping, self-affirmation, and the cerebral hemispheres. *Personality and Social Psychology Bulletin, 31,* 846–856.

Simon, L., Greenberg, J., & Brehm, J. (1995). Trivialization: The forgotten mode of dissonance reduction. *Journal of Personality and Social Psychology, 68,* 247–260.

Smith, E. R., Coats, S., & Walling, D. (1999). Overlapping mental representations of self, in-group, and partner: Further response time evidence and a connectionist model. *Personality and Social Psychology Bulletin, 25,* 873–882.

Smith, E. R., & Henry, S. (1996). An in-group becomes part of the self: Response time evidence. *Personality and Social Psychology Bulletin, 22,* 635–642.

Solomon, S., Greenberg, J., & Pyszczynski, T. (1991). A terror management theory of social behavior: The psychological functions of self-esteem and cultural worldviews. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 24, pp. 93–159). New York: Academic Press.

Spencer, S. J., Fein, S., & Lomore, C. D. (2001). Maintaining one's self-image vis-a-vis others: The role of self-affirmation in the social evaluation of the self. *Motivation and Emotion, 25,* 41–65.

Spencer, S. J., Josephs, R. A., & Steele, C. M. (1993). Low self-esteem: The uphill struggle for self-integrity. In R. Baumeister (Ed.), *Self-esteem: The puzzle of low self-regard* (pp. 21–36). New York: Plenum Press.

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology, 35,* 4–28.

Spencer, S. J., Zanna, M. P., & Fong, G. T. (2005). Establishing a causal chain: Why experiments are often more effective than mediational analyses in examining psychological processes. *Journal of Personality and Social Psychology, 89,* 845–851.

Stanton, A. L., Danoff-Burg, S., Sworowski, L. A., Collins, C. A., Branstetter, A., Rodriguez-Hanley, A., Kirk, S. B., & Austenfeld, J. L. (2002). Randomized, controlled trial of written

emotional expression and benefit finding in breast cancer patients. *Journal of Clinical Oncology, 20,* 4160–4168.

Steele, C. M. (1988). The psychology of self-affirmation: Sustaining the integrity of the self. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 21, pp. 261–302). New York: Academic Press.

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist, 52,* 613–629.

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 69,* 797–811.

Steele, C. M., Hopp, H., & Gonzales, J. (1986). Dissonance and the lab coat: Self-affirmation and the free-choice paradigm. Unpublished manuscript, University of Washington.

Steele, C. M., & Liu, T. J. (1983). Dissonance processes as self-affirmation. *Journal of Personality and Social Psychology, 45,* 5–19.

Steele, C. M., Spencer, S. J., & Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 34, pp. 379–440). New York: Academic Press.

Steele, C. M., Spencer, S. J., & Lynch, M. (1993). Self-image resilience and dissonance: The role of affirmational resources. *Journal of Personality and Social Psychology, 64,* 885–896.

Stone, J., & Cooper, J. (2001). A self-standards model of cognitive dissonance. *Journal of Experimental Social Psychology, 37,* 228–243.

Stone, J., & Cooper, J. (2003). The effect of self-attribute relevance on how self-esteem moderates dissonance processes. *Journal of Experimental Social Psychology, 39,* 508–515.

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.

Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behavior. In S. Worchel and W. G. Austin (Eds.), *Psychology of intergroup relations* (pp. 33–48). Chicago: Nelson-Hall.

Taylor, S. E. (1983). Adjustment to threatening events: A theory of cognitive adaptation. *American Psychologist, 38,* 1161–1173.

Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin, 103,* 193–210.

Taylor, S. E., Lerner, J., Sherman, D. K., Sage, R., & McDowell, N. (2003). Are self-enhancing cognitions associated with healthy or unhealthy biological profiles? *Journal of Personality and Social Psychology, 85,* 605–615.

Taylor, S. E., & Lobel, M. (1989). Social comparison activity under threat: Downward evaluation and upward contacts. *Psychological Review, 96,* 569–575.

Taylor, S. E., & Sherman, D. K. (*in press*). Self-enhancement and self-affirmation: The effects of positive self-thoughts on motivation and behavior. Chapter to appear in J. Shah & W. Gardner (Eds.) *Handbook of motivation science*. New York: Guilford Publications.

Taylor, S. E., Wood, J. V., & Lichtman, R. R. (1983). It could be worse: Selective evaluation as a response to victimization. *Journal of Social Issues, 39,* 19–40.

Tesser, A. (1988). Toward a self-evaluation maintenance model of social behavior. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (pp. 181–227). San Diego, CA: Academic Press.

Tesser, A. (2000). On the confluence of self-esteem maintenance mechanisms. *Personality and Social Psychology Review, 4,* 290–299.

Tesser, A., & Cornell, D. P. (1991). On the confluence of self-processes. *Journal of Experimental Social Psychology, 27,* 501–526.

Tesser, A., Crepaz, N., Collins, J. C., Cornell, D., & Beach, S. R. H. (2000). Confluence of self defense mechanisms: On integrating the self-zoo. *Personality and Social Psychology Bulletin, 26,* 1476–1489.

Tesser, A., Martin, L. L., & Cornell, D. P. (1996). On the substitutability of self-protective mechanisms. In J. A. Bargh (Ed.), *The psychology of action: Linking cognition and motivation to behavior* (pp. 48–68). New York: Guilford Press.

Tesser, A., & Smith, J. (1980). Some effects of task relevance and friendship on helping: You don't always help the one you like. *Journal of Experimental Social Psychology, 16,* 582–590.

Thibodeau, R., & Aronson, E. (1992). Taking a closer look: Reasserting the role of the self-concept in dissonance theory. *Personality and Social Psychology Bulletin, 18,* 591–602.

Triandis, H. C. (1989). The self and social behavior in differing cultural contexts. *Psychological Review, 96,* 506–520.

Trivers, R. (2000). The elements of a scientific theory of self-deception. *Annals of the New York Academy of Sciences, 907,* 114–131.

Uhlmann, E., & Cohen, G. L. (2005a). Constructed criteria: Redefining merit to justify discrimination. *Psychological Science, 16,* 474–480.

Uhlmann, E., & Cohen, G. L. (2005b). I think it, therefore it's true: Self-perceived objectivity licenses the expression of intergroup bias. Manuscript under review.

Weinstein, N. D., & Klein, W. M. (1995). Resistance of personal risk perceptions to debiasing interventions. *Health Psychology, 14,* 132–140.

Wert, S. R. (2004). Negative gossip as a response to threatened social self-esteem. Doctoral dissertation, Yale University. *Dissertation Abstracts International: Section B: The Sciences & Engineering, 65(3-B),* 1601.

Wert, S. R., & Salovey, P. (2004). A social comparison account of gossip. *Review of General Psychology, 8,* 122–137.

Wiesenfeld, B. M., Brockner, J., & Martin, C. (1999). A self-affirmation analysis of survivors' reactions to unfair organizational downsizings. *Journal of Experimental Social Psychology, 35,* 441–460.

Wills, T. A. (1981). Downward comparison principles in social psychology. *Psychological Bulletin, 90,* 245–271.

Wilson, T. D. (2002). *Strangers to ourselves: Discovering the adaptive unconscious.* Cambridge, MA: Harvard University Press.

Winkler, J. D., & Taylor, S. E. (1979). Preference, expectations, and attributional bias: Two field studies. *Journal of Applied Social Psychology, 9,* 183–197.

Wortman, C. B., & Silver, R. C. (1989). The myths of coping with loss. *Journal of Consulting and Clinical Psychology, 57,* 349–357.

Zarate, M. A., & Garza, A. A. (2002). In-group distinctiveness and self-affirmation as dual components of prejudice reduction. *Self and Identity, 1,* 235–249.

Zuwerink, J. R., & Devine, P. G. (1996). Attitude importance and resistance to persuasion: It's not just the thought that counts. *Journal of Personality and Social Psychology, 70,* 931–944.