Journal of Personality and Social Psychology
2007, Vol. 93, No. 5, 798–813

0022-3514/07/$12.00 DOI: 10.1037/0022-3514.93.5.798

Implicit Self-Esteem Compensation: Automatic Threat Defense

Laurie A. Rudman, Matthew C. Dohn, and Kimberly Fairchild
Rutgers, The State University of New Jersey

Four experiments demonstrated implicit self-esteem compensation (ISEC) in response to threats involving gender identity (Experiment 1), implicit racism (Experiment 2), and social rejection (Experiments 3–4). Under conditions in which people might be expected to suffer a blow to self-worth, they instead showed high scores on 2 implicit self-esteem measures. There was no comparable effect on explicit self-esteem. However, ISEC was eliminated following self-affirmation (Experiment 3). Furthermore, threat manipulations increased automatic intergroup bias, but ISEC mediated these relationships (Experiments 2–3). Thus, a process that serves as damage control for the self may have negative social consequences. Finally, pretest anxiety mediated the relationship between threat and ISEC (Experiment 3), whereas ISEC negatively predicted anxiety among high-threat participants (Experiment 4), suggesting that ISEC may function to regulate anxiety. The implications of these findings for automatic emotion regulation, intergroup bias, and implicit self-esteem measures are discussed.

Keywords: implicit social cognition, implicit self-esteem, intergroup bias, emotion regulation, Implicit Association Test

On a daily basis, people suffer both large and small events that are likely to threaten their self-esteem. Events that encompass criticism, betrayal by loved ones, or vivid mortality reminders are likely to shake people's faith in themselves, at least temporarily. When the "stern and bitter" law of hedonic asymmetry (Frijda, 1988, p. 354) is included, which suggests that people quickly habituate to pleasure but that pain can be unending, the human mind appears designed for despair, not happiness. However, as Frijda (1988) noted, the law of hedonic asymmetry "predicts a negative balance for the quality of life, unless self-deceit and self-defense intervene, which of course they do" (p. 354).

The present research was concerned with a self-defense process that automatically buffers people from a variety of threats. Because it concerns overcompensating for a blow to self-esteem, we refer to it as *implicit self-esteem compensation* (ISEC). We first encountered ISEC when people who lost a self-relevant contest showed surprisingly high scores on the self-esteem Implicit Association Test (IAT; Greenwald & Farnham, 2000). Specifically, men who lost a football knowledge contest and women who lost a children's developmental skills contest showed higher self-esteem IAT scores compared with reversed conditions (Rudman & Fairchild, 2004, Experiment 1). Because defeat in self-relevant domains might be expected to reduce self-esteem (Tesser, 1988), we reported this outcome as atheoretical. Subsequently, we set out to conceptually replicate ISEC and to examine its function and social consequences.

Laurie A. Rudman, Matthew C. Dohn, and Kimberly Fairchild, Department of Psychology, Rutgers, The State University of New Jersey.

This research was partially supported by National Science Foundation Grants BCS-0109997 and BCS-0417335. We thank David Dunning for his helpful comments on an earlier version of the manuscript.

Correspondence concerning this article should be addressed to Laurie A. Rudman, Department of Psychology, Tillett Hall, Rutgers, The State University of New Jersey, 53 Avenue E, Piscataway, NJ 08854-8040. E-mail: rudman@rci.rutgers.edu

As is shown below, threats to the self can be automatically thwarted in line with theoretical frameworks that support the possibility. Indeed, it would perhaps be more surprising to discover that people were implicitly defenseless. Consistent with Frijda's (1988) proposal that defensive mechanisms routinely counter the law of hedonic asymmetry, optimism prevails over pessimism in the general populace, as research on positive illusions (Taylor & Brown, 1988), self-esteem among the stigmatized (Crocker & Major, 1989), and a host of ego-defensive processes attests (e.g., Dunning, 1999; Greenwald, 1980; Kunda, 1990; Kunda & Sinclair, 1999).

Compensatory and Defensive Self-Esteem

One means of warding off threat concerns compensatory cognitions, including self-affirmation (SA; e.g., Steele, Spencer, & Lynch, 1993). The "fluid compensation" principle (Steele, 1988, p. 267) posits that people can recover from threat in one domain (e.g., a failed exam) by emphasizing their positive qualities in a different domain (e.g., close relationships). Similarly, Tesser (2000) argued that threatened people draw on alternate sources of self-esteem and that this process can proceed effortlessly, without conscious awareness. This reasoning concurs with response latency evidence that people with high self-esteem automatically recruit their positive qualities (and repress their weaknesses) following failure feedback (Dodgson & Wood, 1998).

How is compensatory self-esteem best detected? On the one hand, some evidence suggests that threat can lead to explicit self-esteem compensation (ESEC; Baumeister, 1982; Baumeister & Jones, 1978; Brown & Smart, 1991; Greenberg & Pyszczynski, 1985; Pyszczynski, Greenberg, Solomon, Arndt, & Schimel, 2004). In other words, self-reports have yielded effects that mirror ISEC. On the other hand, there is a vast literature showing that threats negatively affect explicit self-esteem, including social rejection (e.g., Leary & Baumeister, 2000; Rudman & Fairchild, 2004, Experiment 3) and failure feedback (Williams, Cheung, &

Choi, 2000). Thus, at the very least, the picture is unclear using self-reports.

Because the present investigation was concerned with implicit self-esteem, recent research on implicit egotism was pertinent. For example, people under mild self-concept threat (i.e., asked to write about a personal flaw) showed greater liking for strangers whose arbitrary participant numbers (e.g., 12-03) resembled their birthdays (Jones, Pelham, Carvallo, & Mirenberg, 2004). Mild self-concept threat also increased preference for Japanese teas whose brand names resembled participants' own first names (Brendl, Chattopadhyay, Pelham, & Carvallo, 2005). Similarly, self-concept threats have yielded ISEC using the name-letter effect (NLE; Nuttin, 1987). For example, Jones, Pelham, Mirenberg, and Hetts (2002) found that a mild self-concept threat increased the NLE (see also Brendl et al., 2005). According to funneled debriefings, participants were unaware of the basis of their choices, suggesting that threat nonconsciously leads to favoring stimuli associated with the self.

Although implicit measures reflect a combination of automatic and controlled processes (Conrey, Sherman, Gawronski, Hugenberg, & Groom, 2005), the degree to which the NLE (and similar measures relying on choice) is automatic is uncertain. For this reason, Experiments 1–3 used the IAT, which has a known automatic component (Conrey et al., 2005) and on which responses are immune to faking (Banse, Seise, & Zerbes, 2001; Kim, 2003). Furthermore, the self-esteem IAT has been validated in several contexts. For example, people who scored high were subsequently resistant to failure feedback, in support of its use as a trait measure (Greenwald & Farnham, 2000). It has also performed well in balanced identity designs that test cognitive consistency among self- and ingroup-related associations (e.g., Greenwald et al., 2002). In addition, research on fragile self-esteem (defined as high explicit, low implicit) has relied successfully on the self-esteem IAT (for a review, see Jordan, Spencer, & Zanna, 2002). Finally, our own research originally revealed ISEC using the self-esteem IAT (Rudman & Fairchild, 2004). However, because the IAT measures relative attitudes (i.e., preference for self over others), we also tested ISEC using the signature effect, which uses changes in signature size as an indirect, nonrelative index of self-esteem (Stapel & Blanton, 2004).

The Function of Compensatory Self-Esteem

Our discovery that ISEC is evoked by threats to self-esteem led us to question its purpose. Although self-esteem theories differ, they commonly assume that self-esteem defense mechanisms function to regulate negative emotions caused by threats to self-worth (e.g., Fries & Frey, 1980; Leary & Baumeister, 2000; Mehlman & Snyder, 1985; Pyszczynski et al., 2004; Steele, 1988; Stephan & Gollwitzer, 1981; Tesser, 1988). The present research tested our suspicion that ISEC might be a means by which anxiety is automatically controlled. The idea of automatic affect regulation is compelling given the variety of circumstances in which people need to control their negative emotions (Mauss, Evers, Wilhelm, & Gross, 2006). According to Frijda (1988), it is likely that people learn to do so sufficiently well that the process becomes routinized (i.e., automatic) and, therefore, involuntary (Bargh, 1989). Similarly, Wegner and Bargh (1998) posited that "inhibitory processes may occur prior to or during emotional response, and these need not be initiated through conscious control" (p. 481). If threats generally signal a need to avoid or control anxiety, ISEC might serve this general function.

Using similar reasoning, Koole and Jostmann (2004) investigated what they termed "intuitive affect regulation." They found that people high in action orientation down-regulated stress under conditions designed to produce tension (i.e., having to perform well to receive a reward). This effect was found using both explicit and implicit measures. However, they argued that although intuitive affect regulation is efficient (for some people), it is nonetheless under voluntary control and, therefore, is not automatic. By contrast, our objective was to pursue a more purely automatic form of affect regulation in response to self-threat.

To date, research on emotion regulation has been primarily confined to deliberative processes, which vary in their effectiveness and often have detrimental consequences, in part because of ironic processes that can exacerbate the very emotion people are trying to avoid (Wegner, 1994; Wegner & Bargh, 1998). In an exception, researchers found, using the IAT, that people who automatically preferred concepts related to emotion regulation (as opposed to emotion expression) were able to reduce insult-provoked anger on explicit mood and physiological measures (Mauss et al., 2006). Reported preference for emotion regulation was not predictive of this success. Because the outcomes measured showed only positive benefits of automatic anger regulation, the authors concluded that the regulating mechanism was likely benign. However, their research used implicit attitudes toward emotion regulation as a proxy for the mechanism by which affect was controlled. By contrast, our aim was to directly investigate ISEC as a possible means by which unpleasant affect is avoided; moreover, we explored its potentially negative social outcomes.

The Social Consequences of Self-Esteem Defense

Although vigilance against threat likely benefits the self, it may also have unfavorable social consequences. For example, motives to avoid anxiety have been argued to underlie mortality salience effects, which include negative outgroup stereotyping (Greenberg et al., 2003; Schimel et al., 1999). Beyond mortality salience, other types of threat have yielded negative social consequences (e.g., Kunda & Sinclair, 1999; McGregor, Zanna, Holmes, & Spencer, 2001; Twenge, Baumeister, Tice, & Stucke, 2001; van den Bos, Poortvliet, Maas, Miedema, & van den Ham, 2005). For example, negative feedback on an intelligence test provoked outgroup member stereotyping (Fein & Spencer, 1997), suggesting that prejudice functions as self-image maintenance.

More pertinent to the present research, threats can exacerbate implicit intergroup bias. For example, White participants given negative feedback on an achievement test showed greater automatic stereotyping of Asians compared with controls (Spencer, Fein, Wolfe, Fong, & Dunn, 1998). In addition, Australians who were excluded by their peers during a computer game showed stronger implicit bias on an Australian–Aboriginal IAT compared with their included counterparts (Govan, Williams, & Case, 2005). Finally, Whites who believed that the IAT measured racism scored higher on the Black–White IAT than did a comparison group (who were told the IAT measured cultural bias; Frantz, Cuddy, Burnett, Ray, & Hart, 2004). That is, the threat of being racist ironically led to increased automatic bias, particularly for those who most

wished to avoid it (i.e., people motivated to be nonprejudiced; Frantz et al., 2004, Experiment 2). Thus, it appears that a variety of threats can provoke automatic intergroup bias.

Yet how do threats exacerbate implicit bias? In the present research, we hypothesized that the phenomenon could occur as a byproduct of ISEC. We are not positing that at the implicit level, people deliberately downgrade others to boost their self-worth. If that were true, participants motivated not to appear racist would have avoided prejudiced Black–White IAT scores, yet they did not (Frantz et al., 2004). Instead, we suspect that ISEC serves as a keystone to automatic emotion regulation and that implicit intergroup bias is an unintended consequence of that process. The self benefits by controlling anxiety, but outgroup members may pay the price. Presuming that ISEC begins early in life and is well practiced, it likely becomes an automatic self-defense mechanism (i.e., beyond deliberate control; Bargh, 1989). As a consequence, automatic intergroup bias is likely exacerbated because the process is the same whether favoring the self or ingroup members is involved. That is, evaluating the self as superior should carry over to more favorable evaluation of ingroup than outgroup members but only when responses are involuntary. When people control their responses, they are not likely to appear to be biased, as a wealth of prejudice and stereotyping research suggests (see Rudman, 2004a, for a review).

Explicit Versus Implicit Self-Esteem

Above, we noted that direct self-esteem measures have sometimes shown threat-congruent effects, but they have also shown defensive reactions to threat (i.e., ESEC). Thus, we might predict that ESEC will emerge (as well as ISEC) under threat manipulations. However, the types of threats used in the present research were those that have diminished explicit self-esteem in the past. For example, gender identity threat tends to lower reports of self-esteem, particularly for men (Rudman & Fairchild, 2004, Experiment 3). Peer rejection also decreases self-esteem, consistent with sociometer theory (Leary & Baumeister, 2000; Williams, in press). These results bolstered our assumption that such manipulations are indeed threatening, but whether they would result in ESEC or threat-congruency effects on direct measures was tangential to our main prediction that threats to the self will evoke ISEC.

Because we are positing that threat can result in ISEC even when explicit self-esteem is diminished or unaffected, it is important to cite theoretical frameworks that support conceptualizing implicit and explicit self-esteem as distinguishable. Attitudes, including self-esteem (i.e., attitude toward the self), can be defined as associations in long-term memory between objects and valence (positive or negative) that are routinized to the point of being automatically accessed (Fazio, 1990). Because people can control their explicit (but not implicit) responses, response latency measures yield attitude estimates that are comparatively upstream (i.e., closer to the source). That is, even when people are truthful, self-reports can reflect only what they believe about their attitudes, whereas implicit measures bypass this limitation. Implicit attitudes are thought to be automatic not only because they are fast acting but also because they can operate without (a) intention (i.e., are involuntary and uncontrollable) and (b) conscious awareness (Bargh, 1989). For this reason, implicit attitudes have also been described as nonconscious (e.g., Blair, 2001).

More recently, it has been argued that implicit and explicit attitudes stem from different sources, with the former being more influenced by affective experiences (Rudman, 2004b; Rudman & Goodwin, 2004). A parallel view is provided by cognitive-experiential self theory (Epstein & Morling, 1995), in which self-evaluations may reflect the experiential system (automatic and affect driven) or the rational system (deliberative and duty driven). Similarly, it has been argued that implicit attitudes stem from an associative learning system, whereas explicit attitudes stem from a rule-based learning system; the former is more influenced by emotion, whereas the latter is more influenced by accuracy (DeCoster, Banner, Smith, & Semin, 2006; Sloman, 1996; Smith & DeCoster, 2000; Strack & Deutsch, 2004). Some evidence suggests that the memory systems are linked to different neural structures (McClelland, McNaughton, & O'Reilly, 1995). Although the systems are thought to be independent, they interact with each other so that they can provide competing influences on behavior. Taken together, the picture that emerges is not of a unitary self but of one composed of multiple subsystems that can operate in tandem or apart (Greenwald, 1982). As a result, it should not be surprising that correlations between implicit and explicit measures are heterogeneous (Blair, 2002; Fazio & Olson, 2003) and moderated by factors such as self-presentation concerns (Nosek, 2005).

Overview of Research

In Experiment 1, we introduced ISEC in response to gender identity threat by extending our original findings (Rudman & Fairchild, 2004). In Experiment 2, we tested ISEC as a mediator of the ironic implicit racism effect (Frantz et al., 2004) and compared our model with two alternative models. Experiment 3 provided a conceptual replication of our mediational model, using the threat of peer rejection. In addition, we sought to reduce ISEC by affording some participants a chance to self-affirm prior to peer rejection (Steele, 1988). In Experiment 4, we used the signature effect to check on the generalizability of ISEC and to remove the relative comparison of self versus others (as operationalized in the IAT). This afforded a means of testing whether ISEC involves self-esteem compensation (as opposed to downgrading others). Finally, Experiments 3 and 4 tested the role of ISEC vis-à-vis anxiety regulation.

The overarching goal of these experiments was to test the scope of an implicit self-defense mechanism that, if proven to be robust, might account for the negative social effects of various types of threat. In addition, if threat generally signals a need to reduce or preempt anxiety and frequently results in ISEC, it seemed likely that ISEC could provide a means by which automatic emotion regulation occurs.

Experiment 1

As noted above, we first observed ISEC while investigating defeat in gender-relevant contests (Rudman & Fairchild, 2004, Experiment 1). In Experiment 1, we led some men and women to believe they were gender deviant through false positive feedback. After participants had performed a masculine or feminine knowl-

edge test, we told them all that they had scored remarkably high. Thus, deviance was instantiated as success in a cross-gendered domain (Cherry & Deaux, 1978; Yoder & Schleicher, 1996). The masculine knowledge test contained information related to finances and physical dominance (sports and weapons); the feminine knowledge test concerned topics related to child care and physical appearance (e.g., cosmetics and fashion).

Because the feminine competencies we used connote lower status than the masculine competencies, we expected gender deviance to be more threatening to men than to women in this experiment. As a result, we predicted that only gender-deviant men would show evidence of ISEC following success on a cross-gendered test. By contrast, women might show some evidence of defensiveness following failure in the masculine (high-status) task given that male gender is associated with more prestige than female gender (e.g., Rudman & Kilianski, 2000). Finally, we expected gender identity threat to decrease explicit self-esteem, especially for men, as it has in the past (Rudman & Fairchild, 2004, Experiment 3).

Method

Participants

One hundred and eighty-two volunteers (86 men, 96 women) participated in exchange for partial fulfillment of their Introductory Psychology experimental requirement. Of these, 103 (57%) were White, 34 (18%) were Asian, 15 (8%) were Latino, 12 (7%) were Black, and the remaining 10% reported another ethnicity.

Materials

Computerized knowledge tests. The masculine knowledge test included 30 items about finance, cars and motorcycles, sports, and physical violence (e.g., the best way to punch an opponent). The feminine knowledge test included 30 items about beauty, fashion, women's health issues, cooking, and etiquette. The tests purported to assess knowledge that "society expects college-aged men [women] to possess." In reality, they were designed to assess fairly obscure knowledge so that participants would believe the false feedback (i.e., be unable to ascertain their true score). For example, the masculine test required identifying the first people to use flamethrowers in battle (Turks or Greeks) and the stock exchange signifiers of various companies, whereas the feminine test required identifying the first company to invent hair coloring (L'Oréal or Clairol) and where Manolo Blahniks are worn (the head or the feet).

The self-esteem IAT. As in past research (e.g., Greenwald et al., 2002), we used target constructs related to the self (I, me, mine) or others (they, them, theirs) and pleasant versus unpleasant attributes (e.g., smile, vacation, pain, disaster). In addition, block order was counterbalanced such that half of the participants performed the self + pleasant task first, whereas the other half performed the self + unpleasant task first (a procedural variable that did not affect results). The IAT effect was formed by subtracting response latencies for the self + pleasant tasks from the self + unpleasant tasks. Scoring for the IAT followed recent recommendations (Greenwald, Nosek, & Banaji, 2003). Specifically, we used the *D* statistic because it is less influenced by procedural variables (e.g., counterbalancing) and cognitive ability (Huajian, Sriram, Greenwald, & McFarland, 2004).

Explicit self-esteem. Participants responded to the performance and social subscales ($n = 14$ items) of the State Self-Esteem Scale (SSES; Heatherton & Polivy, 1991). Sample items include "I feel confident about my abilities," "I feel inferior to others at this moment," and "I feel displeased with myself." Responses ranged from 1 (*strongly disagree*) to 5 (*strongly agree*). After appropriate recoding, we averaged these items to form the SSES ($\alpha = .89$), on which high scores reflected greater state self-esteem.

Procedure

Participants volunteered to serve as pilot subjects for the gender knowledge tests. Prior to participants' taking the test, a list of five former high scorers appeared on the screen. To heighten perceptions of deviance or normativeness, the names were unambiguously male in the masculine test condition and female in the feminine test condition. After participants completed either the masculine or feminine test (by random assignment), the program administered success feedback following a short interval ("Congratulations! You scored in the 96th percentile!"), accompanied by an explanation of how to interpret the score ("This means you scored higher than 96 out of 100 past test takers"). We then administered the SSES and the self-esteem IAT (in counterbalanced order). Participants were subsequently thanked and debriefed.

Results and Discussion

To demonstrate ISEC, the key analysis concerns a 2 (gender test: masculine, feminine) × 2 (participant gender) × 2 (IAT order: first, second) analysis of variance (ANOVA), which showed only the expected Test × Participant Gender interaction for the self-esteem IAT, $F(1, 174) = 6.01$, $p < .05$. No other effects approached significance, all Fs$(1, 174) < 1.00$, *ns.* As expected, men who succeeded in the feminine test had higher implicit self-esteem ($M = 0.79$, $SD = 0.35$) compared with men who succeeded in the masculine test ($M = 0.65$, $SD = 0.30$), $t(84) = 2.10$, $p < .05$, resulting in a nontrivial effect size ($d = .45$). By contrast, women who succeeded in the masculine test had slightly higher implicit self-esteem ($M = 0.70$, $SD = 0.24$) compared with women who succeeded in the feminine test ($M = 0.61$, $SD = 0.33$), $t(94) = 1.43$, *ns,* resulting in a small effect size ($d = .28$).

A comparable analysis for the SSES yielded only a main effect for test, $F(1, 174) = 6.35$, $p < .05$. Success in the masculine condition led to higher state self-esteem compared with success in the feminine condition for men (Ms = 4.21 vs. 3.80), $t(84) = 2.07$, $p < .05$ ($d = .43$) as well as for women (Ms = 4.08 vs. 3.75), $t(94) = 1.53$, $p = .13$ ($d = .35$), although the effect was only reliable for men. Thus, gender identity threat decreased state self-esteem for deviant men, as it has in the past (Rudman & Fairchild, 2004, Experiment 3). For women, the nontrivial effect size supports our reasoning that success in the masculine test would confer higher status than feminine test success.

Finally, we examined the relationships between the SSES and the IAT. For men, it was negative in the feminine test condition, $r(38) = -.38$, $p < .05$, but weakly positive in the masculine test

condition, $r(44) = .15$, *ns.* For women, it was weakly negative in the feminine test condition, $r(51) = -.24$, *ns,* but marginally positive in the masculine test condition, $r(41) = .27$, $p = .10$. For both genders, these correlations were reliably different (both zs $>$ 2.44, ps $< .05$). The pattern suggests that ISEC is a defensive response because people who reported low self-worth under threat tended to score higher (i.e., compensate) on the IAT.

In sum, men under gender identity threat showed higher self-esteem IAT scores compared with unthreatened men, as predicted. Because the SSES showed the reverse effect, it appears that ISEC is an automatic (not a deliberative) response to threat. Finally, because threatened men who were low on state self-esteem also scored high on the IAT, the compensatory nature of ISEC was supported. The next step was to replicate ISEC using a type of threat that should not show gender differences (specifically, the fear of being racist) and to test our interpretation of the ironic implicit racism effect (Frantz et al., 2004).

Experiment 2

Past research found that Whites threatened by the charge of racism were more (rather than less) likely to show implicit pro-White bias on the IAT compared with low-threat counterparts (Frantz et al., 2004). Moreover, this ironic effect was particularly evident for Whites high on egalitarian values. The fact that the threatened people most eager to be fair were least able to control their pro-White bias under threat strongly suggests an automatic process. Yet is ISEC responsible for it? In Experiment 2, we predicted that the threat of racism would promote both ISEC and automatic bias but that ISEC would account for the latter effect. In other words, favoring the self over generalized others should promote favoring ingroup over outgroup members for people under threat.

However, there were two alternative models to be considered. First, implicit biases might serve to boost self-esteem (i.e., the model might be reversed; Fein & Spencer, 1997). This seemed unlikely because we observed ISEC under gender identity threat in Experiment 1 without evoking intergroup attitudes, but the possibility remains. Second, ISEC might not be necessary for increased automatic bias (or vice versa); instead, threat-induced arousal might promote dominant responses in general (e.g., Zajonc & Sales, 1966). Because IAT research has shown that self- and ingroup-preferences are dominant responses for White respondents (Greenwald et al., 2002; Nosek, Banaji, & Greenwald, 2002), under this model, we expected threat to increase both tendencies, but neither effect should account for the other.

Finally, we predicted that people threatened with being a racist would likely show diminished explicit self-esteem given considerable evidence showing that people feel guilt and compunction when they believe they are being bigoted (see Monteith & Voils, 2001, for a review).

Method

Participants

Eighty-six White volunteers (51 female, 35 male) participated in exchange for partial Introductory Psychology experimental credit.

Materials

IAT measures. In addition to Experiment 1's self-esteem IAT, participants completed the Black–White attitude IAT, which obliged participants to categorize pleasant or unpleasant attributes with photographs of Blacks and Whites (Nosek et al., 2002). Positive IAT effects reflected greater self-esteem or pro-White bias, respectively.

Threat manipulation. Following Frantz et al. (2004), participants received instructions prior to performing the Black–White attitude IAT. Participants in the high-threat condition were told that the IAT "compares your attitudes toward two different groups" and is "a measure of racial bias." Participants in the low-threat condition were told that the IAT is "a measure of cultural stereotypes" and were assured that this knowledge was unrelated to personal stereotypes or interracial attitudes and behavior (see Frantz et al., 2004, p. 1615, for the full instructions).[1]

Explicit measures. To provide better IAT counterparts, we used relative explicit measures. Participants reported how they felt toward self, others, Whites, and Blacks on scales ranging from 1 (*very cold*) to 10 (*very warm*). The difference between self and others formed the explicit self-esteem index. The difference between Whites and Blacks formed the explicit attitude index. As with the IATs, high scores reflected greater self-esteem and pro-White bias.

Procedure

Participants were randomly assigned to threat condition and escorted to a private cubicle. Prior to taking the Black–White attitude IAT, participants read instructions designed to instigate high or low threat (as described above). They then completed the Black–White IAT, the self-esteem IAT, and the explicit indexes of self-esteem and racial attitude.[2] Participants were then thanked and debriefed.

Results and Discussion

Our first aim was to test whether high-threat participants would show greater implicit self-esteem and intergroup bias compared with low-threat controls. Table 1 shows the results. As can be seen, ISEC was observed in the high-threat condition, resulting in a moderate group difference for self-esteem IAT scores. Replicating Frantz et al. (2004), Black–White IAT scores were also higher for threatened participants, resulting in a moderate group difference. Submitting IAT scores to separate 2 (threat: high, low) $\times$ 2 (participant gender) ANOVAs yielded the expected main effects for threat, both Fs(1, 82) > 12.58, ps $< .001$. The remaining effects were weak, all Fs(1, 82) < 2.71, ps $> .11$.

[1] We also included a no-information condition, in which participants were told only that the IAT "is a challenging task but necessary for the aim of this study." Replicating Frantz et al. (2004), this condition resulted in intergroup bias IAT scores that were between those for the high- and low-threat groups and did not reliably differ from either group.

[2] Following Frantz et al. (2004, Experiment 2), we also administered the Motives to Control Prejudiced Reactions Scale (Dunton & Fazio, 1997). However, this measure showed a ceiling effect in both high- and low-threat groups and did not moderate the implicit racism effect.

Table 1
Self-Esteem and Intergroup Bias as a Function of Threat (Experiment 2)

Measure	High threat		Low threat		Group difference	
	M	*SD*	*M*	*SD*	*t*	*d*
Self-esteem IAT	0.36	0.20	0.23	0.18	3.00**	.65
Black–White IAT	0.44	0.30	0.18	0.36	3.61**	.74
Self-esteem index	0.18	2.08	1.14	2.19	2.10*	−.45
Attitude index	0.44	1.80	0.17	2.25	0.62	.13

Note. There were 45 and 41 White participants in the high- and low-threat conditions, respectively. High scores on the self-esteem and attitude IATs reflect associating self or Whites with pleasant attributes (and others or Blacks with unpleasant attributes) more than reversed associations. High scores on the self-esteem and attitude indexes reflect favoring self or Whites over others or Blacks, respectively. The effect size (Cohen's *d*) is based on the pooled standard deviation. By convention, small, moderate, and large effect sizes correspond to .20, .50, and .80, respectively (Cohen, 1988). IAT = Implicit Association Test.
* $p < .05$. ** $p < .01$.

Comparable analyses for the explicit self-esteem and attitude indexes showed unreliable results, all Fs(1, 82) < 2.32, ps $> .12$. Thus, threat resulted in defensive reactions that were detectable only when using uncontrolled responses. In fact, Table 1 reveals that high-threat participants reported lower self-esteem compared with low-threat counterparts, in stark contrast to the self-esteem IAT (but in line with predictions). That is, although the omnibus ANOVA revealed no effects for threat on explicit self-esteem, simple effects targeting our prediction showed the expected negative influence of threat.

Table 2 shows the relationships among Experiment 2's measures as a function of threat condition. As can be seen, there was IAT correspondence for both groups, suggesting that self-esteem may generally inform intergroup bias (and vice versa), using implicit measures (Greenwald et al., 2002). The remaining relationships were nonsignificant, including the link between explicit self-esteem and pro-White preference.

Table 2
Correlations Among Measures as a Function of Threat (Experiment 2)

Measure	Self-esteem IAT	Black–White IAT	Self-esteem index
High threat			
Black–White IAT	.54**		
Self-esteem index	−.02	.10	
Attitude index	.02	.24	−.13
Low threat			
Black–White IAT	.57**		
Self-esteem index	.11	.18	
Attitude index	−.16	−.04	.02

Note. There were 45 and 41 White participants in the high- and low-threat conditions, respectively. High scores on the self-esteem and attitude IATs reflect associating self or Whites with pleasant attributes (and others or Blacks with unpleasant attributes) more than reversed associations. High scores on the self-esteem and attitude indexes reflect favoring self or Whites over others or Blacks, respectively. IAT = Implicit Association Test.
** $p < .01$.

Can ISEC account for ironic implicit racism? A series of regression analyses (Baron & Kenny, 1986) examined whether ISEC could account for increased intergroup bias on the part of Whites threatened with racism. As shown in Figure 1, the relationship between threat and implicit anti-Black bias was reliably reduced after controlling for self-esteem IAT scores. A Sobel (1982) test for the reliability of ISEC as a mediator was significant ($z = 2.67$, $p < .01$). These findings support ISEC as a mechanism that exacerbates automatic intergroup bias for people under threat. They also contradict an alternative model in which threat-related arousal might independently provoke dominant responses (e.g., Zajonc & Sales, 1966).

To test a second alternative model, in which automatic intergroup bias might boost implicit self-esteem, we reversed the analysis by regressing self-esteem IAT scores on threat after accounting for Black–White IAT scores. Although our procedure afforded the strongest test possible of this model (because the self-esteem IAT immediately followed the Black–White IAT), we did not find support for it. Instead, the relationship between threat and self-esteem IAT scores remained stable after accounting for Black–White IAT scores (both βs = .31, $p < .01$).

In sum, Experiment 2 supported our suspicion that ISEC could (reliably but partially) account for the ironic implicit racism observed by Frantz et al. (2004). In so doing, it supported ISEC as an involuntary defense against threat. Because ISEC can increase intergroup bias, the outcome can be ironic, such that Whites threatened with being racist can wind up confirming the threat—not because they are racist but because the threat is sufficiently disconcerting to evoke ISEC. The paradoxical effect of ISEC on people motivated not to be racist suggests that such an involuntary response likely serves an important function. In Experiment 3, we tested whether ISEC is invoked to regulate anxiety.

Experiment 3

In Experiment 3, we threatened some participants with peer rejection, which has been shown to increase automatic biases in the past (Govan et al., 2005). To conceptually replicate Experiment 2,

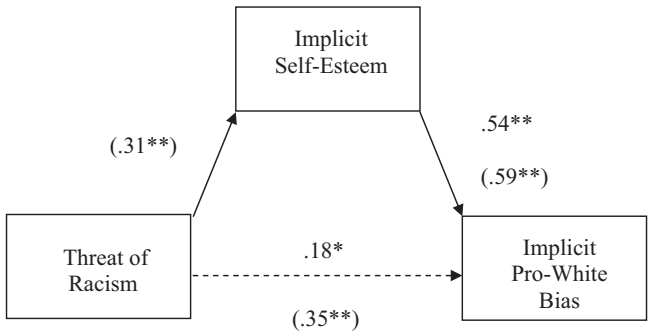

Figure 1. Experiment 2 regression analyses testing implicit self-esteem as a mediator of the relationship between threat and implicit pro-White bias. Coefficients in parentheses reflect a bivariate analysis. A dashed arrow indicates successful mediation. Measures were scored so that positive relationships were expected among them. * $p < .05$. ** $p < .01$.

we expected ISEC to mediate the relationship between threat and implicit bias. New to Experiment 3, we assessed anxiety (prior to ISEC) to determine whether it mediates the relationship between threat and implicit self-esteem. Participants were led to believe they were rejected (high threat) or accepted (low threat) by a phantom peer; a control group did not learn the outcome of the peer's decision. Because social rejection is a powerful threat to self-worth (e.g., Leary & Baumeister, 2000; Rudman & Fairchild, 2004; Williams, in press), we expected it to have a negative effect on participants' reported self-esteem and anxiety.

In addition, we tested SA as a potential moderator of defensive reactions because affirming self-worth promotes security and paves the way to reacting with greater equanimity to threat (Sherman & Cohen, 2002; Sherman & Kim, 2005). Using explicit measures, SA has eliminated threat effects on intergroup bias in the past (e.g., Fein & Spencer, 1997; Schmeichel & Martens, 2005). Therefore, we expected SA to reduce automatic intergroup bias following threat, but for a specific reason—because participants would no longer need ISEC. Thus, we expected only participants high on threat and low on SA to show ISEC (and, as a consequence, increased pro-White bias). To our knowledge, this finding would extend SA as a moderator of implicit threat reactions for the first time.

Method

Participants

Two hundred and forty White volunteers (142 female, 98 male) participated in exchange for partial Introductory Psychology experimental credit.

Materials

Attitude measures. Participants completed the identical self-esteem and Black–White IATs used in Experiment 2. We also adopted the same explicit self-esteem and racial attitude measures.

Anxiety index. Participants completed a state mood measure to indicate how they were feeling "right this moment" on scales ranging from 1 (*not at all*) to 10 (*very much*). The focal items were anxious, nervous, defensive, and calm (reverse scored). These were combined to form the anxiety index ($\alpha = .86$). Filler items (confident, happy, joyful, depressed) were included to defray demand.

Self-affirmation manipulation. Following past SA research (Fein & Spencer, 1997; Schmeichel & Martens, 2005), we asked participants to choose, from among four values (making money, social relationships, art and creativity, or knowledge), the value that was either most or least important to them (in the high- vs. low-SA conditions, respectively). They then wrote an essay about why the value mattered to them (high SA) or why it might matter to others (low SA). On average, high-SA participants chose relationships most often as the value that mattered to them most (59%), followed by knowledge (24%), making money (11%), and creativity (6%). Low-SA participants mirrored this pattern, choosing creativity most often as the value that mattered to them least (56%), followed by making money (28%), knowledge (12%), and relationships (4%).

Personality profile. As part of the cover story, participants responded to 14 open-ended items, including demographic information (age, birth order, hometown) and preferences (e.g., favorite book, color, pet, and ice cream flavor). Other items included "You can be a celebrity for a day—who do you choose?", "How many times in your life have you had a crush on someone?" and "Name one thing you would never do, no matter what."

Threat manipulation. Participants provided their personality profile responses to a phantom confederate (using "networked computers"), who ostensibly would accept or reject them for an interaction task (see *Procedure,* below). Participants were then informed that the confederate had rejected them (high-threat condition) or accepted them (low-threat condition). Control participants were told that no information was available because the program was experiencing technical difficulties.

Procedure

The lab was configured to suggest that other participants had already arrived (e.g., coats and books were prominently displayed). Upon entering, participants were randomly assigned to SA and threat conditions and escorted to a private cubicle. They were told that we were interested in "factors that promote or hinder new friendships" but, to avoid the influence of physical appearance, we were using networked PCs. Participants believed they would exchange personality profiles with an unseen peer, who would decide whether to choose them as a partner for an interaction task. They were informed that teams who worked well together on the task would be eligible for a $50 prize. Participants who were not chosen would be ineligible for the prize; instead, they would serve as pilots for an upcoming project. In fact, all participants completed the dependent variables ostensibly as pilot subjects. Low-threat participants completed them "while waiting for their partner to prepare the interaction task," and controls completed them because technical difficulties prevented the experiment from proceeding normally.

After supposedly uploading their personality profile, participants attempted to download their partner's profile, but they were told it was not yet completed. While waiting, participants completed a short measure to "help us construct a Student Values survey." At this point, high-SA participants chose the value that was most important to them and wrote an essay describing why it mattered to them. Low-SA participants chose the value that was least important to them and wrote an essay describing why it might matter to others. Following this, participants discovered whether they had been chosen for the interaction task (low-threat condition) or rejected (high-threat condition), or whether no information was available (control condition).

Participants then completed the dependent measures in the following order: the anxiety index, the self-esteem IAT, the Black–White IAT, and the explicit self-esteem and attitude measures. They were then thanked and debriefed.

Results and Discussion

Preliminary analyses ruled out reliable effects of procedural variables (counterbalancing block order within the IATs), all Fs < 2.33, *ns*; participant gender, all Fs < 1.00, *ns*; and choice of value to write an essay about, all Fs < 2.13, *ns,* on our dependent variables. We therefore collapsed across these variables for the focal analyses.

Does Self-Affirmation Moderate Threat-Induced ISEC and Implicit Bias?

To test our prediction that SA would reduce implicit ISEC and pro-White bias following threat, we submitted self-esteem and Black–White IAT scores to separate 2 (SA: high, low) × 3 (threat: high, control, low) ANOVAs. Results yielded the expected interaction effect for both IATs, both Fs(2, 232) > 4.53, ps < .01. As Table 3 shows, in the high-threat condition, low-SA participants scored higher than high-SA participants on both IAT measures, resulting in robust effect sizes (seen in the last column). These findings support our expectation that SA would reduce ISEC and automatic racial bias for participants threatened with peer rejection. By contrast, the low-threat group (i.e., participants accepted by their peers) and control group did not differ in their IAT scores as a function of SA.

We also analyzed the interaction effects within SA condition as a function of threat. Within the low-SA group, high-threat participants showed higher self-esteem IAT scores compared with low-threat and control counterparts, both ts > 2.67, ps < .01 (ds = .51 and .52, respectively). Low-SA/high-threat participants also scored higher on the Black–White IAT compared with low-threat and control counterparts, both ts > 2.47, ps < .05 (ds = .45 and .38, respectively). These results are consistent with Experiment 2's findings that threat can automatically increase intergroup bias. However, in support of our prediction that SA would eliminate this effect, high-SA/high-threat participants scored similarly to high-SA participants in the low-threat and control conditions, both ts < 1.10, *ns.*

Table 3
Self-Esteem, Intergroup Bias, and Anxiety as a Function of Self-Affirmation and Threat (Experiment 3)

	Low SA		High SA		Group difference	
Measure	*M*	*SD*	*M*	*SD*	*t*	*d*
High threat[a]						
Self-esteem IAT	0.50_a	0.31	0.12_a	0.45	4.50^{**}	.90
Black–White IAT	0.56_a	0.34	0.32_a	0.61	2.15^{*}	.51
Self-esteem index	0.40_a	2.23	0.58_a	1.52	0.43	−.10
Attitude index	0.26_a	0.88	0.39_a	1.44	0.51	−.08
Anxiety index	4.02_a	1.01	3.92_a	0.91	0.49	.10
Low threat[b]						
Self-esteem IAT	0.28_b	0.38	0.21_a	0.46	0.77	.17
Black–White IAT	0.37_b	0.35	0.22_a	0.64	1.34	.34
Self-esteem index	0.30_a	1.49	0.63_a	1.98	0.83	−.19
Attitude index	0.05_a	1.39	0.50_a	1.78	1.21	−.29
Anxiety index	3.38_b	1.21	3.89_a	0.95	2.06^{*}	−.50
Control group[c]						
Self-esteem IAT	0.27_b	0.46	0.29_a	0.38	0.20	−.05
Black–White IAT	0.35_b	0.39	0.44_a	0.33	1.07	−.19
Self-esteem index	0.37_a	1.39	0.84_a	1.58	1.48	−.27
Attitude index	0.85_b	1.75	0.40_a	1.83	1.17	.29
Anxiety index	3.43_b	1.12	3.67_a	0.88	0.30	.06

Note. All measures are difference scores except the anxiety index. High scores on the self-esteem and attitude IATs reflect associating self or Whites with pleasant attributes (and others or Blacks with unpleasant attributes) more than reversed associations. High scores on the self-esteem and attitude indexes reflect favoring self or Whites over others or Blacks, respectively. Means not sharing a subscript differ among threat conditions at the $p < .05$ level or higher. The effect size (Cohen's d) for SA group differences is based on the pooled standard deviation. By convention, small, moderate, and large effect sizes correspond to .20, .50, and .80, respectively (Cohen, 1988). SA = self-affirmation; IAT = Implicit Association Test.
[a] ns = 43 (low) and 36 (high) SA conditions. [b] ns = 37 (low) and 38 (high) SA conditions. [c] ns = 41 (low) and 45 (high) SA conditions.
$^{*}p < .05$. $^{**}p < .01$.

Does ISEC Mediate the Relationship Between Threat and Implicit Bias?

Another important aim was to conceptually replicate Experiment 2's finding that self-esteem IAT scores can account for the facilitating effect of threat on implicit bias. Because threat increased IAT scores only for low-SA participants, we focused our mediational analyses on this group. As the top half of Figure 2 shows, after accounting for self-esteem, the relationship between threat and implicit bias was reliably reduced (Sobel's $z = 2.07$, $p < .05$). As in Experiment 2, these results are consistent with our hypothesis that threat-induced ISEC can increase automatic ingroup bias. We also conducted the reverse analysis, in which we regressed implicit self-esteem on threat after accounting for Black–White IAT scores. However, as in Experiment 2, there was only a negligible reduction in the effect of threat on self-esteem (βs = .38 and .33, both ps < .001), and thus, implicit pro-White bias was an unreliable mediator (Sobel's $z = 1.70$, *ns*).

The Effect of Threat and Self-Affirmation on Anxiety

Experiment 3's anxiety index afforded a check on whether threat would increase anxiety, but only for low-SA participants. The results of a 2 (SA: high, low) × 3 (threat: high, control, low) ANOVA showed a weak main effect for threat, $F(2, 234) = 2.27$, $p = .11$. Not surprisingly, high-threat participants reported more anxiety compared with low-threat participants (Ms = 3.97 vs. 3.64), $t(152) = 2.00$, $p < .05$. High-threat participants also reported marginally more anxiety than did controls ($M = 3.69$), $t(163) = 1.82$, $p = .07$. In addition, the expected SA × Threat interaction for anxiety was marginally significant, $F(2, 234) = 2.37$, $p < .10$. As can be seen in Table 3, low-SA/high-threat participants reported more anxiety compared with their low-threat and control counterparts, both ts > 2.10, ps < .05. By contrast, high-SA participants reported similar levels of anxiety across threat conditions, all ts < 1.34, *ns.* However, unexpectedly, high-SA/low-threat participants reported more anxiety than their low-SA counterparts (see Table 3). No other reliable differences were found.

Identical 2 (SA: high, low) × 3 (threat: high, control, low) ANOVAs revealed no reliable effects for the explicit self-esteem and attitude indexes, all Fs < 2.21, *ns.* Table 3 shows that explicit self-esteem and attitude scores were unaffected by the manipulations, all ts < 1.21, *ns,* with one surprising exception. Within the low-SA group, control participants reported stronger pro-White bias compared with high-threat participants, $t(82) = 2.00$, $p < .05$, and low-threat participants, $t(76) = 2.21$, $p < .05$. We view these results with caution because the SA × Threat interaction was weak for the attitude index, $F(2, 234) = 1.76$, $p = .17$.

Does Anxiety Mediate the Relationship Between Threat and Implicit Self-Esteem?

Our final goal was to test whether anxiety might account for threat-induced ISEC. Because the relationship between anxiety

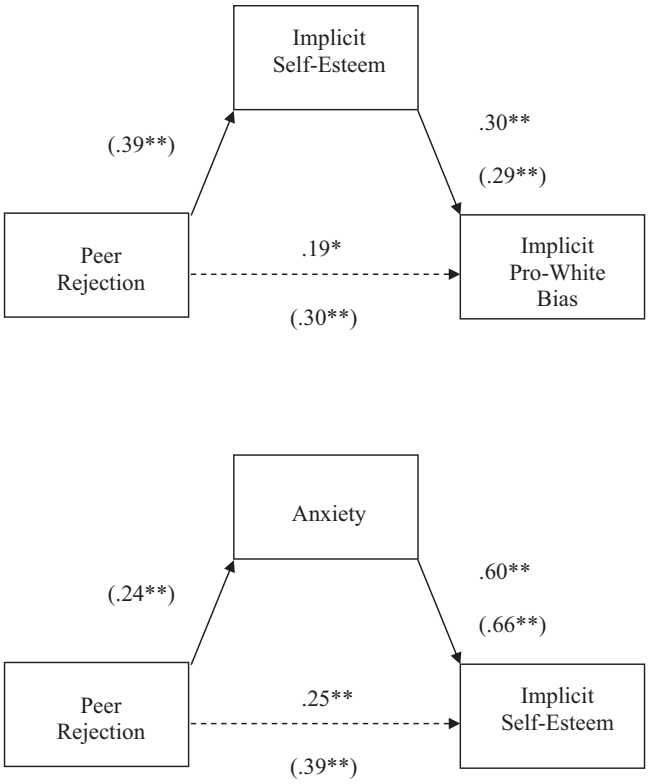

Figure 2. Experiment 3 regression analyses testing (a) implicit self-esteem as a mediator of the relationship between social threat and implicit pro-White bias (top diagram) and (b) anxiety as a mediator of the relationship between social threat and implicit self-esteem (bottom diagram). Coefficients in parentheses reflect a bivariate analysis. A dashed arrow indicates successful mediation. Measures were scored so that positive relationships were expected among them. $^{*}p < .05$. $^{**}p < .01$.

and threat was evident for low-SA participants, $r(119) = .24$, $p < .05$, but not for high-SA participants, $r(117) = .01$, *ns,* we focused our mediational analyses on low-SA participants. As the bottom half of Figure 2 shows, after accounting for anxiety, the relationship between threat and implicit self-esteem was reliably reduced (Sobel's $z = 2.52$, $p < .05$). This finding supports ISEC as a mechanism for anxiety regulation.

If anxiety serves to increase dominant responses in general, we might expect anxiety to also play a role in the positive relationship between threat and implicit bias. However, the correlation between anxiety and Black–White IAT scores was weakly positive for the low-SA group, $r(119) = .12$, *ns,* and weakly negative for the high-SA group, $r(117) = -.07$, *ns.* Thus, anxiety cannot account for threat-induced automatic bias. Instead, Figure 2 suggests that anxiety can lead to ISEC and that ISEC then exacerbates automatic bias in response to threat.

In sum, Experiment 3 replicated ISEC following peer rejection, but only for low-SA (not high-SA) participants. Moreover, peer rejection increased implicit pro-White bias for low-SA (but not high-SA) participants; however, as in Experiment 2, this influence was reliably reduced after accounting for ISEC. Finally, anxiety partially mediated the relationship between threat and implicit self-esteem within the low-SA group. Thus, Experiment 3 provides some evidence that ISEC functions to regulate anxiety in the wake of threat. Because high-SA participants were immune to threat-induced ISEC and showed low levels of anxiety, it appears that SA reduces the need for ISEC as an affect-regulatory device.

Thus far, our predictions for diminished explicit self-esteem were inconsistently supported. In Experiment 1, men's explicit self-esteem decreased under gender identity threat. Yet explicit self-esteem was unreliably reduced following the threat of being a racist (Experiment 2) and not at all affected by peer rejection (Experiment 3), inconsistent with past research (Leary & Baumeister, 2000; Monteith & Voils, 2001). Although these past findings support our manipulations as threatening, why was explicit self-esteem not affected? The most likely explanation is that we administered explicit self-esteem measures following the self-esteem and Black–White IATs, a procedure that could have diluted the predicted effect. In particular, taking two IATs may have been the problem because, in Experiment 1, we counterbalanced the explicit and implicit self-esteem measures and found the predicted decrease in self-esteem following male gender identity threat across conditions. In Experiment 4, using only the signature effect (a brief measure), we expected to find reduced explicit self-esteem following peer rejection compared with peer acceptance or control conditions.

Experiment 4

The IAT is a relative measure, so it is unclear whether ISEC reflects self-enhancement or downgrading others in response to threat. Because implicit egotism is enhanced following mild threats to self-concept (Brendl et al., 2005; Jones et al., 2004), the former seemed likely. Using Experiment 3's threat manipulation, we replaced the IAT with the signature effect (Zweigenhaft, 1977; Zweigenhaft & Marlowe, 1973), which compares baseline to post-threat signature size. If ISEC reflects self-enhancement, high-threat participants should show larger signatures at Time 2 compared with baseline. By contrast, low-threat and control participants should not show the signature effect. Stapel and Blanton (2004) found that subliminally priming people with unintelligent targets yielded the signature effect relative to unprimed controls, suggesting that it is responsive to automatic processes. Moreover, the signature effect was found to be nonreactive (i.e., participants had no idea their self-esteem was being assessed).

In addition, we again assessed anxiety—this time, following ISEC—to examine whether ISEC would reduce anxiety for threatened participants provided they were given sufficient time to benefit. Thus, we expected anxiety to be predicted by a Threat × Delay × Signature Effect interaction. Some participants received the anxiety index immediately following signing their names; others received it following a brief (3-minute) time delay. Thus, Experiment 4 used a 3 (threat: high, low, control) × 2 (time delay: no, yes) between-participants design. Our rationale for manipulating delay was empirically driven; pilot research suggested that immediately assessing anxiety following threat-induced ISEC resulted in inconclusive findings.

Method

Participants

Volunteers ($N = 301$, 183 female, 118 male) participated in exchange for partial Introductory Psychology experimental credit.

Of these, 136 (45%) were White, 99 (33%) were Asian, 22 (7%) were Black, 17 (6%) were Latino, and the remaining 9% reported another ethnicity.

Materials

Threat manipulation. As in Experiment 3, participants shared supposedly uploaded personality profile responses with a phantom peer who ostensibly accepted (low threat) or rejected (high threat) them as a partner for an interaction task. Control participants were again told that no information was available because the computer program was experiencing technical difficulties.

Signature effect. Participants signed two consent forms, one at the beginning of the experiment and one following the threat manipulation (see *Procedure,* below). We then measured the area of each signature and took the difference between them (Stapel & Blanton, 2004). Specifically, signature size was measured by drawing a rectangle around each signature. The topmost point and rightmost point of the signature were identified, and a ruler was used to create a straight lined rectangle. Then, the height and length of this rectangle were measured (in inches). The area covered by the signature was found by multiplying the measured height and width. The difference between the signature's area at Time 1 and at Time 2 was computed. A high score indicated that participants signed their name larger at Time 2 compared with Time 1.

Explicit measures. We administered Experiment 3's anxiety measure (consisting of anxious, nervous, defensive, and calm, reverse scored; $\alpha = .85$). To match the nonrelative nature of the signature effect, participants also rated how they felt toward themselves on a scale ranging from 1 (*very cold*) to 10 (*very warm*).

Procedure

The procedure and cover story were identical to Experiment 3's with the following exceptions. All participants signed a consent form (as a measure of their baseline signature) before beginning the experiment. After supposedly uploading their personality profile and discovering that their partners' profile was not yet completed, participants completed a survey to "assess demographic information" (in fact, to allow sufficient time for the phantom peer to make a decision). The survey consisted of 20 filler questions that asked about their education (e.g., "Do you attend school full time?") and consumer behaviors (e.g., "Do you own a cell phone?").

Following this, participants learned the outcome of their peers' decision (acceptance, rejection, or no information available). They then attempted to access a "pilot survey." However, the program responded with an error message and instructions to elicit the experimenter's aid. Upon entering the booth, the experimenter asked the participant to sign a second consent form, ostensibly because he or she had "forgotten to get a copy for our records." Turning to the computer, the experimenter then pressed a combination of keys to bypass the error message and begin the pilot survey before leaving the cubicle. Participants in the delayed condition then read a passage describing a man having car trouble. They were told to read it carefully because "they would be asked questions about it later." They then completed the anxiety index, followed by the explicit self-esteem measure. Participants in the immediate condition completed the anxiety and explicit self-esteem measures without delay. All participants were then thanked and debriefed.

Results and Discussion

The Signature Effect

One important aim was to test the mechanism of ISEC by using a nonrelative index of self-esteem. We first tested signature effect difference scores against zero. High-threat participants showed larger signatures at Time 2 compared with baseline, $t(97) = 2.19$, $p < .05$ ($M = 0.17$, $SD = 0.43$, $d = .40$). By contrast, low-threat participants showed a weakly reversed signature effect, $t(96) = -1.87$, *ns* ($M = -0.10$, $SD = 0.38$, $d = -.26$), and control participants showed virtually no change, $t(99) = -1.13$, *ns* ($M = -0.05$, $SD = 0.31$, $d = -.16$).

Submitting the signature effect to a 3 (threat: high, low, control) $\times$ 2 (participant gender) ANOVA revealed a main effect for threat, $F(1, 294) = 5.52$, $p < .05$. Simple effects revealed a stronger signature effect for high-threat participants compared with low-threat counterparts, $t(196) = 2.81$, $p < .01$, and controls, $t(199) = 2.19$, $p < .05$. Low-threat participants did not differ from controls, $t(199) = 1.04$, *ns.* These results suggest that the mechanism underlying ISEC is self-esteem enhancement (as opposed to devaluing others). Although social comparisons can also be used to boost self-esteem following threat (Beauregard & Dunning, 1998; Brown & Gallagher, 1992; Dunning & Beauregard, 2000; Dunning & Cohen, 1992), it may not be necessary to involve others in defense of the self.

Does ISEC Regulate Threat-Induced Anxiety?

A primary goal was to test our hypothesis that ISEC functions to down-regulate anxiety. If so, then high-threat participants should show a negative relationship between anxiety and increased signature size following threat. Participants in low-threat or control conditions should not feel the need to regulate their anxiety. To examine whether this would be particularly true for high-threat participants who experienced a delay between signature signing and anxiety assessment, we standardized all variables and then regressed the anxiety index on threat, delay, and the signature effect, including all interaction terms. (Because preliminary analysis ruled out gender as an influence, we excluded this factor from our reported analysis.)

Results showed a marginal main effect for threat on anxiety ($\beta = .12$, $p = .05$). High-threat participants reported more anxiety ($M = 2.36$, $SD = 1.58$) than low-threat participants ($M = 1.90$, $SD = 1.22$), $t(192) = 2.02$, $p < .05$, but not more than controls ($M = 2.08$, $SD = 1.24$), $t(198) = 1.40$, $p = .16$, who did not differ from the low-threat group, $t(96) < 1.00$, *ns.* More important, the expected Threat $\times$ Delay $\times$ Signature Effect interaction was significant ($\beta = -.15$, $p = .01$). Table 4 shows the correlations between the signature effect and anxiety as a function of threat and time delay. As can be seen, for high-threat participants, the signature effect was negatively related to anxiety in the delay condition but not in the immediate condition. This pattern suggests that implicit self-esteem enhancement requires time to down-regulate anxiety. In the control and low-threat conditions, there were no

Table 4
Correlations Among Signature Effect, Anxiety, and Explicit Self-Esteem as a Function of Threat and Time Delay (Experiment 4)

	Signature effect	
Measure	Delay	No delay
High threat[a]		
Anxiety index	−.38**	.10
Self-esteem index	.12	.26
Low threat[b]		
Anxiety index	.27	−.01
Self-esteem index	.12	.06
Control group[c]		
Anxiety index	.06	−.01
Self-esteem index	.13	.20

Note. High scores on the self-esteem index reflect feeling warm toward the self and favoring the self over others.
[a] *n*s = 51 (delay) and 48 (no delay) conditions. [b] *n*s = 49 (delay) and 50 (no delay) conditions. [c] *n*s = 50 (delay) and 52 (no delay) conditions.
** $p < .01$.

reliable correlations between anxiety and the signature effect in either the time delayed or immediate condition (see Table 4).

Explicit Self-Esteem

Because explicit self-esteem was assessed after the signature effect (and the delay manipulation), we regressed this measure on threat, delay, the signature effect, and all of their interaction terms. Results showed a main effect only for threat ($\beta = -.12$, $p < .05$). Planned comparisons revealed that high-threat participants reported lower self-esteem ($M = 6.95$, $SD = 1.82$) than low-threat participants ($M = 7.71$, $SD = 1.62$), $t(93) = 2.12$, $p < .05$, and marginally less self-esteem than control participants ($M = 7.69$, $SD = 1.62$), $t(97) = 1.97$, $p < .06$. The low-threat and control groups did not differ, $t(98) < 1.00$, *ns.* Thus, peer rejection had the expected effect of dampening self-worth when self-reports were used and there were no intervening IATs.

Because explicit self-esteem was also assessed after the anxiety measure, we also regressed this measure on threat, delay, participant gender, anxiety, and all of their interaction terms. Results showed a main effect only for anxiety ($\beta = -.30$, $p < .001$), such that anxious participants were less likely to report high self-esteem. Finally, Table 4 shows the results of correlations between the explicit self-esteem measure and the signature effect as a function of experimental conditions. As can be seen, these were generally positive, ranging from .13 to .26.

In sum, Experiment 4 found that high-threat participants increased their signature following peer rejection, suggesting that ISEC involves self-esteem compensation. Because the signature effect is nonrelative, it was useful for determining whether ISEC relies on increasing self-evaluation, as opposed to devaluing others. Moreover, assessing anxiety following ISEC afforded a test of whether ISEC functions to decrease anxiety. Results supported this hypothesis but only for threatened participants who experienced a brief time delay between signing their name at Time 2 and reporting their anxiety.

General Discussion

In Experiments 1–3, the self-esteem IAT proved to be sensitive to various threats, including gender identity threat, the threat of being racist, and peer rejection. In each case, people who might be expected to suffer a blow to self-regard instead showed ISEC. In addition, in Experiment 4, peer rejection provoked ISEC using the signature effect. Because this measure is nonrelative (i.e., the self is not evaluated relative to others), it supports our claim that ISEC involves maintaining self-esteem. Taken together, our findings suggest that ISEC may be an important self-defense mechanism in response to threat—one that occurs spontaneously and effortlessly.

In retrospect, self-esteem seems an obvious candidate for automatic threat defense given its history as a buffer against ego threats (e.g., Crocker & Major, 1989; Harmon-Jones et al., 1997; McGregor et al., 2001). Moreover, it has been argued that self-esteem is maintained effortlessly through nonconscious processes (Tesser, 2000). Indeed, research using the NLE has found that mild self-concept threats can lead to ISEC (Brendl et al., 2005; Jones et al., 2002). Nonetheless, our initial observation that gender identity threat provoked high (rather than low) implicit self-esteem was unexpected (Rudman & Fairchild, 2004, Experiment 1). As a result, the present research sought to test the scope of ISEC as an automatic defensive response to threat. In so doing, it has provided the most systematic investigation of ISEC to date by examining ISEC across various types of threat, in response to SA, and by relating ISEC to intergroup bias and emotion regulation.

We found a general tendency for threat to evoke ISEC, but we also found moderators of the effect. In Experiment 1, only men (not women) showed ISEC in response to success in a cross-gendered domain, likely because masculine success is not as threatening to women as feminine success is to men (see also Rudman & Fairchild, 2004, Experiment 3). In Experiment 3, SA eliminated ISEC following peer rejection, suggesting that securing self-worth in advance of a threat reduces defensive reactions (Sherman & Cohen, 2002; Sherman & Kim, 2005). Although past research using self-reports has found that SA buffers against threat-induced intergroup biases (e.g., Fein & Spencer, 1997; Schmeichel & Martens, 2005), Experiment 3 provides the first evidence that it does so even on implicit measures. However, in the absence of SA, ISEC was observed in response to peer rejection (Experiments 3–4) and the threat of being a racist (Experiment 2). The common thread among these disparate threats is their ability to instigate anxiety. Moreover, as in past research, they lowered explicit self-esteem in Experiments 1 and 4 (and unreliably in Experiment 2). Thus, our findings implicate ISEC as a means by which people maintain their equanimity following blows to self-worth.

Nonetheless, the value of the present research goes beyond demonstrating ISEC for at least two reasons: first, because our results are consistent with the hypothesis that ISEC is an effective affect-regulatory device—a relatively rare mechanism in the literature to date (Wegner, 1994; see Mauss et al., 2006, and Wegner & Bargh, 1998, for reviews), and second, because we found evidence that it partially mediates the link between threat and automatic intergroup bias (e.g., Spencer et al., 1998), suggesting that a process that benefits the self can have negative social

consequences. These aspects of our research are discussed in greater detail below.

The Role of ISEC in Automatic Emotion Regulation

To date, the evidence for automatic emotion regulation has centered on emotional expression and conditioning (e.g., Ekman, 1984; Lewicki, Hill, & Czyzewska, 1992; Ohman & Mineka, 2001), not on preemption or down-regulation. Although recent research has begun to explore the new territory of "automatic emotion stopping" (Wegner & Bargh, 1998, p. 481), it has used implicit attitudes as a proxy for the mechanism by which preemption occurs (Mauss et al., 2006). Our results support ISEC as a candidate in this regard for two reasons. First, self-reported anxiety partially mediated the relationship between peer rejection and ISEC in Experiment 3, suggesting that ISEC is spurred by threat-induced anxiety. Second, ISEC moderated the effect of threat on anxiety in Experiment 4 such that threatened participants who showed ISEC also reported low anxiety provided they had time to reap ISEC's benefits. In concert, these results suggest that anxiety is an instigator of ISEC but that people who employ this defense can successfully down-regulate their affect. For these reasons, we suspect ISEC deserves a place in the budding pantheon of automatic self-regulatory mechanisms (Amodio et al., 2004; Davis, 1987; Moskowitz, Gollwitzer, Wasel, & Schaal, 1999).

How automatic is ISEC? It is generally agreed that automatic and controlled processes lie on a continuum (e.g., Bargh, 1996; Schneider & Shiffrin, 1977); thus, labeling a process automatic is not meant to imply a strict dichotomy. Nonetheless, we believe ISEC is relatively more automatic than controlled. First, it is revealed by the IAT, on which responses are not likely to be faked (e.g., Banse et al., 2001; Kim, 2003). Second, it appears to be involuntary, much as the ironic implicit racism effect is involuntary (Frantz et al., 2004), in that ISEC differs from reactions to threat observed with direct measures. Moreover, reported self-esteem is more susceptible to many artifacts (Dunning, Heath, & Suls, 2004), including demand, which could be salient for people in high-threat conditions. That is, people might suspect that their self-esteem should be lower if they have just been rejected, suspected of being racists, or having their masculinity threatened during a laboratory experiment. Implicit measures are less problematic in this respect. Finally, if threatened people regulate anxiety by automatically upgrading their self-worth, they would not necessarily show self-esteem decrements or heightened anxiety on self-reports (see also Mauss et al., 2006). That is, ISEC may help people to recover their self-esteem long before they put pen to paper on a direct measure.

The Social Consequences of ISEC

Although ISEC appears to have regulatory benefits for the self, our findings also point to its negative consequences for implicit social cognition. In Experiment 2, we replicated the ironic implicit racism effect (Frantz et al., 2004), but we also found that ISEC partially mediated the link between the threat of being racist and ironically high automatic pro-White bias. Similarly, in Experiment 3, peer rejection resulted in high automatic racial bias (Govan et al., 2005), but ISEC again partially mediated the link between social threat and Black–White IAT scores.

Taken together, the present findings suggest that the automatic tendency to value the self is exacerbated by threat and that this, in turn, can provoke intergroup bias. In other words, ISEC sets the stage for heightened self-favorable preferences that are extended to ingroup members by virtue of their association with the self (Greenwald et al., 2002). As a result, we would expect people under threat to show implicit bias even when this outcome is dreaded. Thus, Experiment 2's ironic implicit racism effect (see also Frantz et al., 2004) can be explained by the fact that ISEC is an involuntary threat response; although it protects the self, it can harm others—even those one wishes to protect.

Implications of the Research for the IAT

Considerable research has shown that the IAT can be used successfully as a trait measure of self-esteem and self-concept (e.g., Asendorpf, Banse, & Mucke, 2002; Egloff & Schmukle, 2002; Greenwald et al., 2002; Greenwald & Farnham, 2000; Pinter & Greenwald, 2005). At the same time, the attitude IAT has been shown to be sensitive to contextual effects, as have other implicit measures (see Blair, 2002, for a review). The fact that the self-esteem and attitude IATs were responsive to threat agrees with recent theorizing about the sources of implicit biases, in which affective experiences are thought to inform implicit responses even when explicit responses are impervious (Rudman, 2004b; Rudman & Goodwin, 2004). As a result, the IAT can be used to detect evidence of threat in cases where other measures have proven to be unpromising.

However, although evidence of ISEC enhances the value of the IAT as a research tool, it also clouds the interpretation of self-esteem IAT scores. In some contexts, high scores might reflect stable self-regard, but in other contexts, they might reflect defensively high state self-esteem. As a result, researchers using the self-esteem IAT as a trait measure may wish to administer it over time to provide an index of stability as well as magnitude, as has been suggested for explicit self-esteem assessment (see Kernis & Paradise, 2002, for a review). Moreover, the self-esteem IAT should not be used to support theories that posit low self-esteem as an outcome of various threats, such as social rejection (e.g., Leary & Baumeister, 2000) or failure in self-relevant domains (Tesser, 1988). Finally, researchers should expect threat to exacerbate implicit intergroup bias, but they should be aware that this effect might be reliably accounted for by ISEC.

Implicit and Explicit Self-Esteem Discrepancies

As noted in the introduction, the fact that implicit and explicit measures sometimes diverge and sometimes cohere suggests they are distinguishable but interdependent. However, the conditions under which they conflict or work together are not well understood. Some authors have argued that implicit and explicit attitudes reside in separate memory systems or stem from different sources of information (DeCoster et al., 2006; Greenwald & Banaji, 1995; Rudman, 2004b; Strack & Deutsch, 2004), but these positions require future research to be fleshed out.

The present experiments suggest that under some types of self-threat, implicit and explicit self-esteem can conflict, but exactly why remains an empirical question. We have suggested that implicit self-esteem may be more defensive in nature than explicit

self-esteem, but it may also be the case that implicit measures are more subjective than self-reports (i.e., less capable of producing a response that is independent of the self; Rudman, 2004b). Research participants may be able to produce a cooler assessment of their self-evaluation when they report their feelings. Although this may sometimes produce a temporarily negative self-opinion ("I was just rejected, so I must have screwed up"), it may also yield no change in self-opinion ("I was just rejected, but I am still a good person"). Implicitly, however, the self may compensate for threat by leaping to an enhanced evaluation ("I was just rejected, but the evaluator knows nothing; I am a magnificent person") to downregulate anxiety. The fact that, as noted in the introduction, explicit self-esteem measures can sometimes also produce compensation adds complexity and intrigue to the picture. This may be particularly likely to occur when self-esteem measures are relatively subtle in nature (e.g., Baumeister & Jones, 1978; Greenberg & Pyszczynski, 1985) or among people with high self-esteem (Baumeister, 1982; Dodgson & Wood, 1998), although whether *high* in this case reflects genuine or stable self-esteem as opposed to defensive self-esteem is not clear (Kernis & Paradise, 2002).

Limitations and Future Directions

The present research used threat manipulations that were known to evoke ISEC (gender identity threat; Rudman & Fairchild, 2004) or implicit bias, including the threat of racism (Frantz et al., 2004) and peer rejection (Govan et al., 2005). As noted above, a likely commonality among these threats is that they provoke anxiety, as seen by threatened participants' reports of high anxiety in Experiment 3 (for the low-SA group, which scored higher than all other groups) and in Experiment 4, in which peer-rejected participants scored higher than accepted counterparts. However, future research should systematically test the nature and magnitude of various threats to determine the generalizability of our results. For example, if people are threatened with criticism from significant others, implicit self-esteem might be diminished, as opposed to enhanced. That is, the relationship between threat and implicit self-esteem could be curvilinear. Consistent with this view, scores on the NLE have increased under mild self-concept threat (e.g., Brendl et al., 2005) but have decreased when researchers have used stronger threats (e.g., failure feedback; for a review, see Koole, Dijksterhuis, & van Knippenberg, 2001) or have measured negative life events (DeHart & Pelham, 2007; Pelham & Hetts, 1999).

Future research should also test the extent to which ISEC regulates various types of emotion. The present findings emphasized threat-induced anxiety, but the types of threat we used have also induced other negative emotions. For example, the threat of being a racist evokes guilt and compunction, as well as anxiety (Monteith & Voils, 2001). Moreover, gender identity threat in men can lead to anger and aggressive behaviors toward women (Maass, Cadinu, Guarnieri, & Grasselli, 2003). Anxiety, anger, and depression have all been linked to peer rejection (e.g., Leary, Twenge, & Quinlivan, 2006; Williams, 2001), and neurological work suggests rejection is experienced as physical pain (see Williams, in press, for a review). Finally, an anger manipulation has been found to exacerbate implicit intergroup bias (DeSteno, Dasgupta, Bartlett, & Cajdric, 2004). Whether this result might also be accounted for by ISEC is an empirical question worthy of pursuit.

The present research cannot speak to the exact reason why threat instigates ISEC. For example, it may be that when participants perform the IAT, anxiety interferes with executive functioning that would normally allow for more control (Croizet et al., 2004; Richeson & Trawalter, 2005). Although this seems somewhat unlikely in Experiment 4, where we used the signature effect, it still remains as a possible explanation for the IAT findings. That is, Experiment 4 suggests that undermining control may not be necessary for ISEC to occur, but it could certainly play a role when people perform the IAT.

More generally, the issue of whether ISEC preempts negative affect or regulates it retrospectively (or both) deserves research attention. For example, people under mortality salience typically do not report anxiety or show physiological evidence of anxiety (Greenberg et al., 2003), perhaps because ISEC is routinely used as a preemptive device. Our own research showed that mortality salience does, in fact, evoke ISEC (Rudman & Dohn, 2005), but its relationship to anxiety when death is salient has yet to be explored. Similarly, peer rejection does not necessarily yield heightened reports of anxiety (Baumeister, Twenge, & Nuss, 2002; Gardner, Pickett, & Brewer, 2000; see Pyszczynski et al., 2004, for a review), a troublesome outcome for sociometer theory (Leary & Baumeister, 2000). Both classic and contemporary theorists have posited that people are capable of preventing negative emotions from reaching conscious awareness (e.g., Freud, 1915/1957; Mauss et al., 2006; Singer, 1990), and ISEC may be a means by which they achieve this result. If so, ISEC may prove useful to self-esteem theorists arguing that the main function of self-esteem is to regulate negative affect (e.g., Fries & Frey, 1980; Leary & Baumeister, 2000; Mehlman & Snyder, 1985; Pyszczynski et al., 2004; Steele, 1988; Stephan & Gollwitzer, 1981; Tesser, 1988).

Conclusion

In concert, our findings suggest that ISEC is a routinized reaction to threat and that it may serve to prevent negative emotions, such as anxiety, from disrupting people's composure in the face of threat. Frijda (1988) argued that without defense mechanisms, people would suffer emotionally from life's vicissitudes. That is, people must have some means of automatically regulating their emotions to avoid chronic anxiety and despair. However, although ISEC is promising as a mechanism in this regard, it also has unfavorable implications for social reality. Protecting the self from threat-induced anxiety may unwittingly exacerbate favoring ingroup over outgroup members. Thus, although ISEC helps to explain automatic biases in response to threat, including ironic implicit racism (Frantz et al., 2004), it offers little comfort to those who suffer the consequences.

References

Amodio, D. M., Harmon-Jones, E., Devine, P. G., Curtin, J. J., Hartley, S. L., & Covert, A. E. (2004). Neural mechanisms for the detection of unintentional race bias. *Psychological Science, 15,* 88–93.

Asendorpf, J. B., Banse, R., & Mucke, D. (2002). Double dissociation between implicit and explicit personality self-concept: The case of shy behavior. *Journal of Personality and Social Psychology, 83,* 380–393.

Banse, R., Seise, J., & Zerbes, N. (2001). Implicit attitudes toward homosexuality: Reliability, validity, and controllability of the IAT. *Zeitschrift fur Experimentelle Psychologie, 48,* 145–160.

Bargh, J. A. (1989). Conditional automaticity: Varieties of automatic influence in social perception and cognition. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 3–51). New York: Guilford Press.

Bargh, J. A. (1996). Automaticity in social psychology. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 169–183). New York: Guilford Press.

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. (1982). Self-esteem, self-presentation, and future interaction: A dilemma of reputation. *Journal of Personality, 50,* 29–45.

Baumeister, R. F., & Jones, E. E. (1978). When self-presentation is constrained by the target's knowledge: Consistency and compensation. *Journal of Personality and Social Psychology, 36,* 608–618.

Baumeister, R. F., Twenge, J. M., & Nuss, C. K. (2002). Effects of social exclusion on cognitive processes: Anticipated aloneness reduces intelligent thought. *Journal of Personality and Social Psychology, 83,* 817–827.

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.

Blair, I. V. (2001). Implicit stereotypes and prejudice. In G. B. Moskowitz (Ed.), *Cognitive social psychology: The Princeton Symposium on the Legacy and Future of Social Cognition* (pp. 359–374). Mahwah, NJ: Erlbaum.

Blair, I. V. (2002). The malleability of automatic stereotypes and prejudice. *Personality and Social Psychology Review, 6,* 242–261.

Brendl, M., Chattopadhyay, A., Pelham, B. W., & Carvallo, M. (2005). Name letter branding: Valence transfers when product specific needs are active. *Journal of Consumer Research, 32,* 405–415.

Brown, J. D., & Gallagher, F. M. (1992). Coming to terms with failure: Private self-enhancement and public self-effacement. *Journal of Experimental Social Psychology, 28,* 3–22.

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.

Cherry, F., & Deaux, K. (1978). Fear of success versus fear of gender-inappropriate behavior. *Sex Roles, 4,* 97–101.

Cohen, J. (1988). *Statistical power for the behavioral sciences.* Hillsdale, NJ: Erlbaum.

Conrey, F. R., Sherman, J. W., Gawronski, B., Hugenberg, K., & Groom, C. (2005). Separating multiple processes in implicit social cognition: The quad-model of implicit task performance. *Journal of Personality and Social Psychology, 89,* 469–487.

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

Croizet, J., Despres, G., Gauzins, M., Huguet, P., Leyens, J., & Meot, A. (2004). Stereotype threat undermines intellectual performance by triggering a disruptive mental load. *Personality and Social Psychology Bulletin, 30,* 721–731.

Davis, P. J. (1987). Repression and the inaccessibility of affective memories. *Journal of Personality and Social Psychology, 53,* 585–593.

DeCoster, J., Banner, M. J., Smith, E. R., & Semin, G. R. (2006). On the inexplicability of the implicit: Differences in the information provided by implicit and explicit tests. *Social Cognition, 24,* 5–21.

DeHart, T., & Pelham, B. W. (2007). Fluctuations in state implicit self-esteem in response to daily negative events. *Journal of Experimental Social Psychology, 43,* 157–165.

DeSteno, D., Dasgupta, N., Bartlett, M. Y., & Cajdric, A. (2004). Prejudice from thin air: The effect of emotion on automatic intergroup attitudes. *Psychological Science, 15,* 319–324.

Dodgson, P. G., & Wood, J. V. (1998). Self-esteem and the cognitive accessibility of strengths and weaknesses after failure. *Journal of Personality and Social Psychology, 75,* 178–197.

Dunning, D. (1999). A newer look: Motivated social cognition and the schematic representation of social concepts. *Psychological Inquiry, 10,* 1–11.

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., Heath, C., & Suls, J. M. (2004). Flawed self-assessment: Implications for health, education, and the workplace. *Psychological Science in the Public Interest, 5,* 71–106.

Dunton, B. C., & Fazio, R. H. (1997). An individual difference measure of motives to control prejudiced reactions. *Personality and Social Psychology Bulletin, 23,* 316–326.

Egloff, B., & Schmukle, S. C. (2002). Predictive validity of an implicit association test for assessing anxiety. *Journal of Personality and Social Psychology, 83,* 1441–1455.

Ekman, P. (1984). Expression and the nature of emotion. In K. R. Scherer & P. Ekman (Eds.), *Approaches of emotion* (pp. 319–344). Hillsdale, NJ: Erlbaum.

Epstein, S., & Morling, S. (1995). Is the self motivated to do more than enhance and/or verify itself? In M. H. Kernis (Ed.), *Efficacy, agency, and self-esteem* (pp. 9–29). New York: Plenum Press.

Fazio, R. H. (1990). Multiple processes by which attitudes guide behavior: The MODE model as an integrative framework. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 23, pp. 75–109). New York: Academic Press.

Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition: Their meaning and use. *Annual Review of Psychology, 54,* 297–327.

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.

Frantz, C. M., Cuddy, A. J. C., Burnett, M., Ray, H., & Hart, A. (2004). A threat in the computer: The race Implicit Association Test as a stereotype threat experience. *Personality and Social Psychology Bulletin, 30,* 1611–1624.

Freud, S. (1957). Repression. In J. Strachey (Ed. & Trans.), *The standard edition of the complete psychological works of Sigmund Freud* (Vol. 14, pp. 143–158). London: Hogarth Press. (Original work published 1915)

Fries, A., & Frey, D. (1980). Misattribution of arousal and the effects of self-threatening information. *Journal of Experimental Social Psychology, 16,* 405–416.

Frijda, N. H. (1988). The laws of emotion. *American Psychologist, 43,* 349–358.

Gardner, W. L., Pickett, C. L., & Brewer, M. B. (2000). Social exclusion and selective memory: How the need to belong influences memory for social events. *Personality and Social Psychology Bulletin, 26,* 486–496.

Govan, C., Williams, K., & Case, T. (2005). *Tolerant on the outside, intolerant on the inside: Dual responses to ostracism.* Manuscript submitted for publication.

Greenberg, J., Martens, A., Jonas, E., Eisenstadt, D., Pyszczynski, T., & Solomon, S. (2003). Psychological defense in anticipation of anxiety: Eliminating the potential for anxiety eliminates the effect of mortality salience on worldview defense. *Psychological Science, 14,* 516–519.

Greenberg, J., & Pyszczynski, T. (1985). Compensatory self-inflation: A response to the threat to self-regard of public failure. *Journal of Personality and Social Psychology, 49,* 273–280.

Greenwald, A. G. (1980). The totalitarian ego. *American Psychologist, 35,* 603–618.

Greenwald, A. G. (1982). Is anyone in charge? Personalysis versus the principle of personal unity. In J. Suls (Ed.), *Psychological perspectives on the self* (pp. 151–181). Hillsdale, NJ: Erlbaum.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102,* 4–27.

Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellott, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review, 109,* 3–25.

Greenwald, A. G., & Farnham, S. D. (2000). Using the Implicit Association Test to measure self-esteem and self-concept. *Journal of Personality and Social Psychology, 79,* 1022–1038.

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology, 85,* 197–216.

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.

Heatherton, T. F., & Polivy, J. (1991). Development and validation of a scale for measuring state self-esteem. *Journal of Personality and Social Psychology, 60,* 895–910.

Huajian, C., Sriram, N., Greenwald, A. G., & McFarland, S. G. (2004). The Implicit Association Test's *D* measure can minimize a cognitive skill confound: Comment on McFarland and Crouch (2002). *Social Cognition, 22,* 673–684.

Jones, J. T., Pelham, B. W., Carvallo, M., & Mirenberg, M. C. (2004). How do I love thee? Let me count the Js: Implicit egotism and interpersonal attraction. *Journal of Personality and Social Psychology, 87,* 665–683.

Jones, J. T., Pelham, B. W., Mirenberg, M. C., & Hetts, J. J. (2002). Name letter preferences are not merely mere exposure: Implicit egotism as self-regulation. *Journal of Experimental Social Psychology, 38,* 170–177.

Jordan, C. H., Spencer, S. J., & Zanna, M. P. (2002). "I love me . . . I love me not": Implicit self-esteem, explicit self-esteem, and defensiveness. In S. J. Spencer, S. Fein, M. P. Zanna, & J. M. Olson (Eds.), *The Ontario Symposium: Vol. 9. Motivated social perception* (pp. 117–145). Mahwah, NJ: Erlbaum.

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

Kim, D. Y. (2003). Voluntary controllability of the Implicit Association Test (IAT). *Social Psychology Quarterly, 66,* 83–96.

Koole, S. L., Dijksterhuis, A., & van Knippenberg, A. (2001). What's in a name: Implicit self-esteem and the automatic self. *Journal of Personality and Social Psychology, 80,* 669–685.

Koole, S. L., & Jostmann, N. (2004). Getting a grip on your feelings: Effects of action orientation and external demands on intuitive affect regulation. *Journal of Personality and Social Psychology, 87,* 974–990.

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

Kunda, Z., & Sinclair, L. (1999). Motivated reasoning with stereotypes: Activation, application, and inhibition. *Psychological Inquiry, 10,* 12–22.

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). New York: Academic Press.

Leary, M. R., Twenge, J. M., & Quinlivan, E. (2006). Interpersonal rejection as a determinant of anger and aggression. *Personality and Social Psychology Review, 10,* 111–132.

Lewicki, P., Hill, T., & Czyzewska, M. (1992). Nonconscious acquisition of information. *American Psychologist, 47,* 796–801.

Maass, A., Cadinu, M., Guarnieri, G., & Grasselli, A. (2003). Sexual harassment under social identity threat: The computer harassment paradigm. *Journal of Personality and Social Psychology, 85,* 853–870.

Mauss, I. B., Evers, C., Wilhelm, F. H., & Gross, J. J. (2006). How to bite your tongue without blowing your top: Implicit evaluation of emotion regulation predicts affective responding to anger provocation. *Personality and Social Psychology Bulletin, 32,* 1–14.

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review, 102,* 419–457.

McGregor, I., Zanna, M. P., Holmes, J. G., & Spencer, S. J. (2001). Compensatory conviction in the face of personal uncertainty: Going to extremes and being oneself. *Journal of Personality and Social Psychology, 80,* 472–488.

Mehlman, R. C., & Snyder, C. R. (1985). Excuse theory: A test of the self-protective role of attributions. *Journal of Personality and Social Psychology, 49,* 994–1001.

Monteith, M. J., & Voils, C. I. (2001). Exerting control over prejudiced responses. In G. B. Moskowitz (Ed.), *Cognitive social psychology: The Princeton Symposium on the Legacy and Future of Social Cognition* (pp. 375–388). Mahwah, NJ: Erlbaum.

Moskowitz, G. B., Gollwitzer, P. M., Wasel, W., & Schaal, B. (1999). Preconscious control of stereotype activation through chronic egalitarian goals. *Journal of Personality and Social Psychology, 77,* 167–184.

Nosek, B. A. (2005). Moderators of the relationship between implicit and explicit evaluation. *Journal of Experimental Psychology: General, 134,* 565–584.

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Harvesting implicit group attitudes and beliefs from a demonstration Web site. *Group Dynamics, 6,* 101–115.

Nuttin, J. M. (1987). Affective consequences of mere ownership: The name letter effect in twelve European languages. *European Journal of Social Psychology, 17,* 381–402.

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

Pelham, B. W., & Hetts, J. (1999). *Implicit self-evaluation.* Unpublished manuscript.

Pinter, B., & Greenwald, A. G. (2005). Clarifying the role of the "other" category in the self-esteem IAT. *Experimental Psychology, 52,* 74–79.

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.

Richeson, J. A., & Trawalter, S. (2005). Why do interracial interactions impair executive function? A resource depletion account. *Journal of Personality and Social Psychology, 88,* 934–947.

Rudman, L. A. (2004a). Social justice in our minds, homes and society: The nature, causes and consequences of implicit bias. *Social Justice Research, 17,* 129–142.

Rudman, L. A. (2004b). Sources of implicit attitudes. *Current Directions in Psychological Science, 13,* 80–83.

Rudman, L. A., & Dohn, M. C. (2005, October). *Implicit self-esteem compensation: Automatic ego-threat defense.* Paper presented at the annual meeting of the Society of Experimental Social Psychology, San Diego, CA.

Rudman, L. A., & Fairchild, K. (2004). Reactions to counterstereotypic behavior: The role of backlash in cultural stereotype maintenance. *Journal of Personality and Social Psychology, 87,* 157–176.

Rudman, L. A., & Goodwin, S. A. (2004). Gender differences in automatic ingroup bias: Why do women like women more than men like men? *Journal of Personality and Social Psychology, 87,* 494–509.

Rudman, L. A., & Kilianski, S. E. (2000). Implicit and explicit attitudes toward female authority. *Personality and Social Psychology Bulletin, 26,* 1315–1328.

Schimel, J., Simon, L., Greenberg, J., Pyszczynski, T., Solomon, S., Waxmonsky, J., & Arndt, J. (1999). Stereotypes and terror management: Evidence that mortality salience enhances stereotypic thinking and preferences. *Journal of Personality and Social Psychology, 77,* 905–926.

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.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review, 84,* 1–66.

Sherman, D., & 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., & Kim, S. H. (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.

Singer, J. L. (Ed.). (1990). *Repression and dissociation: Implications for personality, psychopathology, and health.* Chicago: University of Chicago Press.

Sloman, S. A. (1996). The empirical case of two systems of reasoning. *Psychological Bulletin, 119,* 3–22.

Smith, E. R., & DeCoster, J. (2000). Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. *Personality and Social Psychology Review, 4,* 108–131.

Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equations models. In S. Leinhart (Ed.), *Sociological methodology 1982* (pp. 290–312). San Francisco: Jossey-Bass.

Spencer, S. J., Fein, S., Wolfe, C. T., Fong, C., & Dunn, M. A. (1998). Automatic activation of stereotypes: The role of self-image threat. *Personality and Social Psychology Bulletin, 24,* 1139–1152.

Stapel, D. A., & Blanton, H. (2004). From seeing to being: Subliminal social comparisons affect implicit and explicit self-evaluations. *Journal of Personality and Social Psychology, 87,* 468–481.

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

Stephan, W. G., & Gollwitzer, P. M. (1981). Affect as a mediator of attributional egotism. *Journal of Experimental Social Psychology, 17,* 443–458.

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review, 8,* 220–247.

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

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.

Twenge, J. M., Baumeister, R. F., Tice, D. M., & Stucke, T. S. (2001). If you can't join them, beat them: Effects of social exclusion on aggressive behavior. *Journal of Personality and Social Psychology, 81,* 1058–1069.

van den Bos, K., Poortvliet, P. M., Maas, M., Miedema, J., & van den Ham, E.-J. (2005). An enquiry concerning the principles of cultural norms and values: The impact of uncertainty and mortality salience on reactions to violations and bolstering of cultural worldviews. *Journal of Experimental Social Psychology, 41,* 91–113.

Wegner, D. M. (1994). Ironic processes of mental control. *Psychological Review, 101,* 34–52.

Wegner, D. M., & Bargh, J. A. (1998). Control and automaticity in everyday life. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 1, pp. 446–496). New York: McGraw-Hill.

Williams, K. D. (2001). *Ostracism: The power of silence.* New York: Guilford Press.

Williams, K. D. (in press). Ostracism. *Annual Review of Psychology.*

Williams, K. D., Cheung, C. K. T., & Choi, W. (2000). Cyberostracism: Effects of being ignored over the Internet. *Journal of Personality and Social Psychology, 79,* 748–762.

Yoder, J. D., & Schleicher, T. L. (1996). Undergraduates regard deviation from occupational gender stereotypes as costly for women. *Sex Roles, 34,* 171–188.

Zajonc, R. B., & Sales, S. M. (1966). Social facilitation of dominant and subordinate responses. *Journal of Experimental Social Psychology, 2,* 160–168.

Zweigenhaft, R. L. (1977). The empirical study of signature size. *Social Behavior and Personality, 5,* 177–185.

Zweigenhaft, R. L., & Marlowe, D. (1973). Signature size: Studies in expressive movement. *Journal of Consulting and Clinical Psychology, 40,* 469–473.

Received April 3, 2006
Revision received March 28, 2007
Accepted April 2, 2007 ■