Journal of Personality and Social Psychology
2000, Vol. 79, No. 4, 631–643

0022-3514/00/$5.00 DOI: 10.1037//0022-3514.79.4.631

Stalking the Perfect Measure of Implicit Self-Esteem: The Blind Men and the Elephant Revisited?

Jennifer K. Bosson, William B. Swann, Jr., and James W. Pennebaker
University of Texas at Austin

Recent interest in the implicit self-esteem construct has led to the creation and use of several new assessment tools whose psychometric properties have not been fully explored. In this article, the authors investigated the reliability and validity of seven implicit self-esteem measures. The different implicit measures did not correlate with each other, and they correlated only weakly with measures of explicit self-esteem. Only some of the implicit measures demonstrated good test–retest reliabilities, and overall, the implicit measures were limited in their ability to predict our criterion variables. Finally, there was some evidence that implicit self-esteem measures are sensitive to context. The implications of these findings for the future of implicit self-esteem research are discussed.

According to Indian folklore, there were once six blind men who had heard of the animal called the elephant but did not know what one was like. To satisfy their curiosity, they decided one day to use their sense of touch to determine the creature's appearance. Matters became confusing, however, when each man touched a different part of the elephant and became convinced that he alone understood its true nature. "The elephant is very like a snake!" proclaimed the man who had touched its trunk. The fellow who had touched its side, however, declared the elephant to be "nothing but a wall," whereas the man who touched the creature's tusk claimed that the elephant was "like a spear," and so on. It is no wonder, then, that the six men could not agree on the true appearance of the elephant (Saxe, 1936).

We see two compelling parallels between the tale of *The Blind Men and the Elephant* and the current state of implicit self-esteem research. First, implicit self-esteem researchers, like the six blind men, are involved in a process of giving shape to something that cannot be seen, something whose characteristics must be inferred. Second, like the blind men, researchers tend to use their own idiosyncratic strategies for measuring implicit self-esteem, yielding many different (and perhaps nonoverlapping) pictures of the underlying construct.[1] In this article, we attempt to uncover the "elephant" by examining the different strategies that researchers have devised in their explorations of implicit self-esteem. To this end, we assess the psychometric properties of several different measures of implicit self-esteem, and we ask whether they tap into a coherent, underlying construct. Our hope is that this project will give researchers a more comprehensive picture of the implicit self-esteem construct.

Implicit Self-Esteem

Implicit self-esteem is defined as an automatic, overlearned, and nonconscious evaluation of the self that guides spontaneous reactions to self-relevant stimuli (Greenwald & Banaji, 1995; Pelham & Hetts, 1999). Although the ontological bases of implicit self-esteem are not yet fully understood, Epstein's (1994) cognitive–experiential self-theory provides a useful framework for understanding the operation of implicit beliefs about the self. Epstein assumed that human thought is characterized by two modes of information processing: One is rational, deliberative, and conscious, whereas the other is automatic, affective, and nonconscious. These two modes of information processing correspond to two broad "theories" about the world—one *cognitive* and the other *experiential*—both of which contain schemas about the self. Schemas in the experiential system are "generalizations about what the self and the world are like," based on "synthesis of emotionally significant experiences" (Teglasi & Epstein, 1998, p. 543). Thus, the experiential belief "I am a lovable (or unlovable) person" reflects an automatic, affective evaluation of the self that exists outside of awareness—in other words, implicit self-esteem.

Measurement of Implicit Self-Esteem

Because implicit self-esteem is nonconscious, it must be measured indirectly, in a manner that is relatively free of contamination by self-presentational processes. There are currently several available methods for measuring implicit self-esteem, all of which draw their logic from one or both of the following assumptions: (a) People assign value to objects that are closely associated with the

Jennifer K. Bosson, William B. Swann, Jr., and James W. Pennebaker, Department of Psychology, University of Texas at Austin.

This research was supported by a University of Texas Continuing Education Fellowship and an American Psychological Association Dissertation Research Award to Jennifer K. Bosson and by National Institute of Mental Health Grants MH57455–01A2 and MH52391. We thank Samantha Doversberger, Sarah Hutchins, Susan Liberty, Tanya Mahdi, and Mandy Morrelli for their help with data collection.

Correspondence concerning this article should be addressed to Jennifer K. Bosson, who is now at the Department of Psychology, Box 454, Vassar College, 124 Raymond Avenue, Poughkeepsie, New York 12604-0273. Electronic mail may be sent to jebosson@vassar.edu.

[1] A reviewer pointed out that virtually *all* researchers who engage in psychological testing are in a position similar to that of the blind men, in that very few psychological constructs (implicit or explicit) are directly observable.

self; and (b) attitudes are activated automatically on encounter with attitude-objects.

On average, people evaluate stimuli associated with the self more favorably than stimuli not associated with the self (e.g., Greenwald & Banaji, 1995). Still, there is variability in the extent to which people feel favorably toward self-associated stimuli (e.g., Pelham & Hetts, 1999). Because people are not necessarily aware of their propensity to attach evaluative meaning to self-associated stimuli, their attitudes toward such stimuli can be interpreted as reflecting their implicit attitudes toward the self. This reasoning gives rise to one popular method of tapping implicit self-esteem: assessing how well people like their initials relative to other letters. To the extent that people feel favorably toward their initials, they can be thought to possess high implicit self-esteem (Hoorens & Nuttin, 1993; Kitayama & Karasawa, 1997; Nuttin, 1985, 1987).

Other implicit self-esteem measures are based on the automatic attitude activation effect. Research shows that when an attitude-object is encountered, the evaluation (e.g., *good*) that one associates with that attitude-object is activated spontaneously and without conscious effort (e.g., Bargh, Chaiken, Govender, & Pratto, 1992; Fazio, Sanbonmatsu, Powell, & Kardes, 1986). Thus, for a man whose automatic evaluation of the self is negative, an encounter with himself—such as an unexpected glimpse of his reflection in a store window—should automatically activate negative affect. Once activated, a particular affective state facilitates the processing of evaluatively similar stimuli and impedes the processing of dissimilar stimuli. So, when primed with a certain attitude-object, people can more quickly identify target words that are affectively consistent with the prime than target words that carry an inconsistent evaluative content. For example, if a woman associates vacations with fun and relaxation and she is primed with the word *vacation*, she is likely to subsequently identify the word *happy* more quickly than the word *sad*. Based on the automatic attitude activation effect, one way to capture implicit self-esteem is to prime people with self-relevant stimuli and assess the ease with which they can identify positive versus negative stimuli (e.g., Hetts, Sakuma, & Pelham, 1999; Spalding & Hardin, 2000).

Because the study of implicit self-esteem is relatively new, there is not yet consensus regarding the best strategy for measuring it. To complicate matters, the "implicit"-ness of the available measures is debatable—some researchers are skeptical that the methods outlined above truly access people's nonconscious attitudes. Nevertheless, in this report we are not interested in questioning the implicit nature of the available measures. Rather, our aim is to compare the psychometric properties of several extant measures. The measures that we include in our study were chosen because researchers have designed and used them to tap implicit self-esteem and, in some instances, have uncovered empirical support for the measures' utility.

Empirical Support for the Implicit Self-Esteem Construct

Several researchers have found that implicit and explicit self-esteem are, at most, only moderately correlated with each other (Farnham, Greenwald, & Banaji, 1999; Greenwald & Farnham, 2000; Hetts et al., 1999; Pelham & Hetts, 1999). Moreover, the two types of self-esteem have been found to predict different outcomes, with implicit self-esteem outperforming explicit self-esteem in predicting people's spontaneous and/or affectively driven responses. For example, implicit (relative to explicit) self-esteem was a better predictor of nonverbal anxiety behaviors (Spalding & Hardin, 2000) and negative mood (Greenwald & Farnham, 2000) in response to threatening feedback. Similarly, implicit—but not explicit—self-esteem predicted people's tendencies to make self-serving judgments following a failure experience and to engage in spontaneous, socially undesirable behavior (Pelham & Hetts, 1999).

Another common finding is that discrepancies between implicit and explicit self-esteem are associated with certain personality variables. Bosson and Swann (1998), for example, found that people with high explicit and low implicit self-esteem scored higher in narcissism than people with any other combination of implicit and explicit self-esteem. Also, Smith (2000) found that having higher implicit than explicit self-esteem was associated with unstable explicit self-esteem, whereas having higher explicit than implicit self-esteem was associated with stable explicit self-esteem. These findings are important because they suggest that the study of explicit self-esteem alone is no longer sufficient if we are to fully understand the scope of self-esteem–related traits and behaviors.

Finally, there appears to be a bias toward favorable implicit self-esteem (e.g., Greenwald & Banaji, 1995) that mirrors the well-documented bias toward high explicit self-esteem (e.g., Taylor & Brown, 1988). This implicit positivity bias has even been noted in collectivist cultures, despite these cultures' tendency to deemphasize expressions of individual esteem (Kitayama & Karasawa, 1997). In sum, past studies of implicit self-esteem have yielded promising, but by no means definitive, findings.

A Methodological Caveat

An important consideration in the study of implicit self-esteem is the context surrounding the administration of implicit measures. Specifically, it is plausible that the thoughts that become activated when people focus on their explicit self-esteem could influence the processing of self-relevant information during a subsequent test of implicit self-esteem. If so, then correlations between implicit and explicit measures might become inflated when implicit self-esteem is assessed immediately after explicit self-esteem; for this reason, it is important to vary the order in which one administers implicit and explicit measures in research that explores both types of self-esteem.

Overview and Research Questions

Our primary goal for this project was to investigate the psychometric properties of the available measures of implicit self-esteem. Because of the exploratory nature of this project, our research was not guided by specific hypotheses but by the following questions: Are implicit self-esteem measures internally reliable and stable across time? Do they, like explicit self-esteem measures, capture self-positivity biases in people's responses? Do implicit measures display good convergent validity—that is, do they tap into the same underlying construct? Do they have good discriminant validity, that is, are they distinct from explicit self-esteem measures? Do implicit self-esteem measures have good predictive validity? Do they display psychometric properties comparable to those of

explicit self-esteem measures? And finally, are implicit measures sensitive to the context in which they are administered?

On two occasions separated by 4 weeks, participants completed a series of implicit and explicit self-esteem measures. The implicit self-esteem measures included all currently available measures of which we were aware, as well as one measure that we created for this study. We included three criterion measures that we thought might relate to people's feelings of implicit self-esteem. First, on the basis of evidence that people prefer feedback that matches their explicit self-views (see Swann, 1992, for a review), we assessed people's feedback preferences to explore whether implicit self-esteem would likewise predict interest in receiving positive versus negative evaluations. Second, because people automatically perceive and interpret incoming social cues in a manner that is consistent with the self-system (e.g., Beck, 1985; Crary, 1966; Markus, 1977; Shrauger & Lund, 1975), we assessed the relations between people's implicit self-esteem and their tendency to interpret ambiguously worded phrases in a positive versus negative manner. Third, because people may be able to perceive covert expressions of emotional distress in others' communications (cf. Shedler, Mayman, & Manis, 1993; Weinberger & McClelland, 1990), we asked judges to read essays that participants wrote during the study, and rate the essay-writers' feelings about themselves. We expected participants' essays to reveal their underlying feelings of self-esteem in a manner perceptible to outside judges. Note that these criterion measures were not intended to represent the full spectrum of traits that might conceivably relate to implicit self-esteem; instead, we chose them because, theoretically, each measure should capture some aspect of people's attitudes toward the self. Moreover, although two of the criterion variables involve self-reports, our judges' ratings of participants' essays should provide a measure of people's nonconscious behaviors.

Finally, for exploratory purposes, we related implicit self-esteem to a number of variables that have been shown to relate to explicit self-esteem such as gender (Major, Barr, Zubek, & Babey, 1999), mood (e.g., Dua, 1993; Tarlow & Haaga, 1996), academic achievement (Khalid, 1990; Newbegin & Owens, 1996), and physical health (e.g., Antonucci & Jackson, 1983; Carroll & Buhrow, 1994; O'Connor & Vallerand, 1998; Vingilis, Wade, & Adlaf, 1998). We also related implicit self-esteem to participants' use of the word *I*, which is thought to indicate self-focused attention (e.g., Kernis, Grannemann, Richie, & Hart, 1988; Mullen, Chapman, & Peaugh, 1989; Wegner & Giuliano, 1980).

Method

Participants and Procedure

A total of 40 male and 44 female undergraduates at the University of Texas at Austin participated in two sessions in exchange for course credit. One participant's data were excluded from analyses because she failed to follow instructions.

At Time 1, participants reported to the lab individually and learned that we were studying the effectiveness of several different types of personality assessment techniques. A female experimenter stressed the confidentiality of participants' responses, as well as the importance of their honest and genuine reactions to all tasks. After signing informed consent forms, participants sat in small cubicles equipped with a PC and began the experimental tasks.

Order of implicit and explicit task presentation was a between-subjects variable: Approximately half of the participants completed all of the explicit measures before they completed the implicit measures (*explicit-first* condition), and the remaining participants completed the implicit measures first (*implicit-first* condition). The specific ordering of the measures is shown in Table 1.[2] Because we did not counterbalance the order of task presentation within the block of implicit measures, we attempted to administer these measures in an order that would minimize participants' awareness of what was being assessed. The implicit task series thus proceeded from most indirect (the subliminal attitude-prime task) to most direct (the supraliminal attitude-prime task).

After completing all tasks, participants scheduled a follow-up session with the experimenter. Time 2 sessions occurred at the same time of day as the participant's first session; the mean length of time that elapsed between Times 1 and 2 was 31.23 days (minimum = 22 days, maximum = 38 days). When they returned at Time 2, participants completed the same series of measures as shown in Table 1, with the exception of the Self-Attributes Questionnaire and the Self-Liking and Self-Competence subscales, which were excluded due to time constraints. At the end of the Time 2 session, participants completed two criterion measures—the Feedback-Seeking Questionnaire and the Ambiguous Statements Task—in that order. Finally, participants were debriefed about the purpose of the tasks and were thanked for their help.

Measures

Explicit Self-Esteem Measures

Rosenberg Self-Esteem Scale (RSES). Rosenberg's (1965) RSES is a 10-item scale that measures people's feelings of global self-worth. All responses are made on scales ranging from 1 (*strongly disagree*) to 4 (*strongly agree*). Cronbach's alpha for the RSES was .87.[3]

Self-Liking and Competence Scale (SLC). Tafarodi and Swann's (1995) SLC is a measure of the self-liking and self-competence components of global self-esteem. The SLC contains two 10-item subscales, which we administered separately. A sample self-liking (SL) item is "I feel good about who I am," and a sample self-competence (SC) item is "I perform well at a number of things." Participants rate all items on scales ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). Cronbach's alphas were .87 and .89 for the SL and SC subscales.

Self-Attributes Questionnaire (SAQ). Pelham and Swann's (1989) SAQ measures participants' beliefs about their standing, relative to other college students their age and gender, on five self-concept domains: intellectual competence, social competence, artistic/musical ability, athletic ability, and physical attractiveness. Participants rate themselves on scales ranging from 0 (*bottom 5%*) to 9 (*top 5%*). Cronbach's alpha was .64 for the SAQ.

Writing task. Participants spent 20 min writing about their "very deepest thoughts and feelings" about themselves (e.g., Pennebaker, 1997). We encouraged participants to reflect honestly about themselves, to write in a stream-of-consciousness format, and to refrain from editing their work as they wrote.

Implicit Self-Esteem Measures

Implicit Association Test (IAT). The self-esteem IAT (Greenwald & Farnham, 2000) is a computerized categorization task that measures automatic associations of self-relevant and non–self-relevant words with pleasant and unpleasant words. Prior to the task, respondents generate lists of *me* and *not me* words; an example *me* word is the respondent's city of origin,

[2] All participants completed the Positive and Negative Affect Scale (PANAS) last because we did not feel comfortable classifying mood as either an explicit or an implicit measure.

[3] For measures that we administered twice, alphas are averaged across Times 1 and 2.

Table 1
Order of Administration of the Implicit and Explicit Measures

Implicit measures first (N = 43)	Explicit measures first (N = 40)
Subliminal attitude-prime task	SC
ISES	RSE
Initials- and birthday-preference task	SL
IAT	SAQ
Stroop task	Writing task
Supraliminal attitude-prime task	Subliminal attitude-prime task
SC	ISES
RSE	Initials- and birthday-preference task
SL	IAT
SAQ	Stroop task
Writing task	Supraliminal attitude-prime task
PANAS	PANAS

Note. IAT = Implicit Association Test; ISES = Implicit Self-Evaluation Survey; PANAS = Positive and Negative Affect Scale; RSE = Rosenberg Self-Esteem Scale; SAQ = Self-Attributes Questionnaire; SC and SL = Self-Competence and Self-Liking subscales of the Self-Liking and Competence Scale.

whereas a *not me* word is a city that the respondent does not associate with herself and neither strongly likes nor strongly dislikes.[4] During the task, respondents press one of two keys to categorize target words that appear in the middle of the screen. Each target word is taken from the *me* or *not me* lists, or from preexisting lists of *pleasant* (e.g., *glory, snuggle*) or *unpleasant* (e.g., *vomit, torture*) words. For one block of 40 trials, the *me* and *pleasant* category labels appear on the same side of the computer screen; thus, correct categorization of *me* and *pleasant* target words is accomplished by pressing the same key. For a subsequent block of 40 trials, the *me* and *unpleasant* category labels appear on the same side of the screen, forcing respondents to categorize self-relevant and unpleasant words together.

Following Greenwald, McGhee, and Schwartz (1998), scores on this task are calculated by recoding response latencies that fall below 300 ms as 300 ms, and those that fall above 3000 ms as 3000 ms. In addition, the responses of participants who make more than 20% errors are deleted (we deleted two people's data). Next, log transformations are performed on the raw reaction-time data, and mean response latencies are calculated separately for the two blocks (excluding the first two trials from each block). Finally, the mean response latency for the *me–pleasant* block is subtracted from the mean response latency for the *me–unpleasant* block. Thus, scores reflect the ease with which participants associate pleasant versus unpleasant words with the self. Cronbach's alpha was .88 for the IAT in the current sample.[5]

Supraliminal attitude-prime task. Hetts et al. (1999) adopted a procedure commonly used in research on the automatic attitude activation effect (e.g., Bargh et al., 1992; Fazio et al., 1986) to measure the accessibility of positively and negatively valenced words following the presentation of self-relevant or non–self-relevant attitude primes. During this task, participants see a series of attitude primes flashed on the center of the computer screen for 200 ms; following each attitude prime, respondents must press one of two keys to identify, as quickly and accurately as possible, whether the next word that appears on the screen is *good* or *bad*. Each of the five attitude primes—one self-relevant (*me*) and four non–self-relevant (*it, them, us, that*)—is paired twice with each of the two target words (*good* and *bad*), for a total of 20 trials. The trials are presented in random order and are preceded by eight practice trials in which the primes *at, how, when,* and *with* are each paired twice with the two target words.

Scores are calculated by first recoding response latencies that fall above 1500 ms as 1500 ms, and then performing log transformations on the raw data. Next, the two response latencies for each prime–target pair are averaged (in cases of error, the error trial is discarded and only the remaining latency is used). Final scores are calculated by subtracting peoples' average response latencies during *me–good* trials from their latencies during *me–bad* trials.[6] Thus, scores reflect the accessibility of positive versus negative words following activation of the self-attitude. Cronbach's alpha was −.16 for this measure in the current sample.

Subliminal attitude-prime task. Spalding and Hardin (2000) used an automatic attitude activation task similar to the one described above, but they presented all prime words subliminally. They flashed self-relevant (*me, myself*) and non–self-relevant (*two, manner*) attitude primes onto the center of the computer screen for 17 ms each, making them accessible only on a subliminal level. Spalding and Hardin's measure also differs from Hetts et al.'s (1999) in their use of six pairs of evaluatively loaded antonyms as target words: *good–bad, love–hate, nice–mean, winner–loser, superior–inferior,* and *fine–poor.* During the task, each prime word is paired twice with each target word, creating a total of 96 prime–target trials. Trials are presented in random order and are preceded by 4 practice trials in which randomly chosen prime–target pairs are presented.

Scores on this measure are calculated by recoding latencies that fall below 300 ms as 300 ms and latencies that fall above 2000 ms as 2000 ms. Log transformations are then performed on all of the raw reaction-time data, and four composite scores are created by averaging across all response latencies (excluding error trials) for trials in which: (a) self-relevant primes are followed by positively valenced targets; (b) self-relevant primes are followed by negatively valenced targets; (c) non–self-relevant primes are followed by positively valenced targets; and (d) non–self-relevant primes are followed by negatively valenced targets. Two difference scores are then created by subtracting the non–self-relevant/positive composite from the self-relevant/positive composite, and the non–self-relevant/negative composite from the self-relevant/negative composite. Finally, the positive difference score is subtracted from the negative difference score; thus, scores reflect the extent to which activation of the self-attitude facilitates responses to positive versus negative words. Cronbach's alpha was .49 for this measure in the current sample.

Stroop Color-Naming task. To measure participants' speed in identifying the color of target phrases that reflect positive versus negative self-attitudes, we created a color-naming task based on the Stroop effect (Stroop, 1935). Via computer, seven positive self-statements (e.g., "I AM GREAT"), and seven negative self-statements (e.g., "I'M NO GOOD") are presented to participants. Each self-statement appears alone in the center of the screen; is printed in either red, blue, or green; and has an accompanying control "statement" that consists of a series of *X*s printed in the same color as the target phrase. The task thus consists of 28 color-naming trials, which are presented in random order. Each time a statement appears on the screen, participants respond by pressing, as quickly as possible, one of three keys that correspond to the three colors in which the statements are printed.

[4] An alternate version of the IAT uses pronouns (e.g., *me, them*) instead of *me* and *not me* words. Scores on the two different versions were correlated at .68 (Greenwald & Farnham, 2000).

[5] In calculating Cronbach's alpha for all reaction-time measures, we controlled for individual differences in baseline reaction times (e.g., Fazio, 1990) by first subtracting each latency associated with a self+pleasant trial from the latency associated with the corresponding self+unpleasant trial (this strategy also controls for order effects that may occur from fatigue). We then computed alphas on these difference scores; alphas thus reflect the internal consistency in people's tendency to associate to positive—relative to negative—stimuli with the self.

[6] Hetts et al. (1999) used the response latencies that followed the non–self-relevant primes in analyses but not in their calculation of implicit self-esteem scores.

Scores are calculated by recoding response latencies that fall above 3000 ms as 3000 ms and performing log transformations on all of the raw data. Next, average response latencies are computed (excluding error trials) for the positive self-statement trials, the negative self-statement trials, and the control trials that correspond to the positive and negative self-statements, respectively. The mean latency for the positive control trials is then subtracted from the mean latency for the positive self-statement trials, and the mean latency for the negative control trials is subtracted from the mean latency for the negative self-statement trials; finally, the positive difference score is subtracted from the negative difference score. Thus, scores reflect the speed with which respondents identify the color of positive versus negative self-statements. Longer response latencies on this task reflect greater interference, as people tend to respond more slowly when identifying the color of stimuli that evoke anxiety or impede cognitive processing. We expected higher scores on this task to reflect lower implicit self-esteem because people with a negative implicit self-attitude should take longer to identify the color of negative than of positive self-statements. Cronbach's alpha was −.38 for this measure in the current sample.

Implicit Self-Evaluation Survey (ISES). Pelham and Hetts' (1999) ISES measures the accessibility of pleasant versus unpleasant words after the self-attitude is primed. Using Likert-type rating scales, respondents indicate their level of agreement with 10 self-statements that are designed to prime their attitudes about themselves (e.g., "I am very sensitive to my inner thoughts and feelings"). Following each self-statement, respondents complete three words by providing different letters to the beginning of the same word fragment, which is presented three times in succession (e.g., 1. __ATE, 2. __ATE, 3. __ATE). Word fragments are designed so that it is possible to create four pairs of pleasant–unpleasant antonyms: *love–hate*, *good–bad*, *nice–mean*, and *fair–poor*. The attitude-prime self-statements and their accompanying word completion tasks are arranged in blocks of four, separated by two filler items (e.g., the attitude statement "Dogs make better pets than cats" followed by the word fragment __ONE). Within each block of four priming statements, two of the target word completions are pleasant (e.g., *love*, *good*) and two are unpleasant (e.g., *mean*, *poor*). Both members of an antonym pair (e.g., *love–hate*) never appear as target word completions within the same block.

Accessibility of pleasant versus unpleasant words is measured as a function of the serial position (i.e., 1, 2, or 3) in which the target words are formed (when a target word is not formed, a value of 4 is assigned as the serial position of that word). Scores are calculated by subtracting the value associated with the serial position of each pleasant target word completion from the value associated with the serial position of its antonym and then summing all four difference scores. Thus, higher scores reflect greater accessibility of pleasant than unpleasant words after the self-attitude is primed. Cronbach's alphas were .59 and .53 for the pleasant and unpleasant target word completions.

Initials- and birthday-preference task. Participants rated how well they like each letter of the alphabet and the numbers *0* through *35*, on scales ranging from 1 (*I dislike this letter [number] very much*) to 7 (*I like this letter [number] very much*). To calculate initials-preference scores, the overall liking rating—averaged across all participants—of each respondent's first and last initials is subtracted from his or her mean rating of his or her initials. Birthday-preference scores are calculated by subtracting the overall liking rating—averaged across all participants—of each respondent's birthday number from his or her own rating of his or her birthday number (if a respondent was born on November 19th, *19* is treated as his or her birthday). Initials- and birthday-preference scores reflect the extent to which respondents like their name initials and their birthday number above and beyond the average popularity of those letters and numbers.[7] As a measure of internal consistency for the initials-preference task, we correlated people's rating of their first initial (minus the overall rating of that letter) with their rating of their last initial (minus the overall rating of that letter); the two were correlated at .57.

Criterion Measures

Feedback-Seeking Questionnaire (FSQ). Swann, Wenzlaff, Krull, and Pelham's (1992) FSQ instructs participants to imagine that a close friend is going to answer five pairs of questions about them, pertaining to the five self-concept domains on the SAQ. Within each self-concept domain, respondents choose the two questions (out of a pool of six) that they most want their friend to answer about them; three of the six questions are worded positively (e.g., "What is some evidence you have seen that your friend has good social skills?"), and three are worded negatively (e.g., "What academic subjects would you expect to prove difficult for your friend?"). Each choice of a positively worded question is coded as 1, and each choice of a negatively worded question is coded as −1; responses to all 10 questions are then summed so that high scores indicate a stronger preference for positive than negative feedback. Cronbach's alpha was .53 for the FSQ.

Ambiguous Statements Task (AST). Tafarodi (1998) designed a procedure that determines participants' tendency to interpret ambiguous statements in a positive versus negative manner. Participants are first asked to vividly imagine an acquaintance directing a series of 13 ambiguous, everyday phrases (e.g., "Is this how you want it?") at them. Next, they indicate whether each phrase reflects positive or negative feeling toward them. Finally, participants rate, on scales ranging from 1 (*very slightly intense*) to 7 (*extremely intense*), the intensity of feeling that is expressed by the imagined speaker. The intensity rating for each phrase is assigned a positive sign if the phrase is interpreted in a positive manner and a negative sign if the phrase is interpreted in a negative manner; scores are calculated by computing the average intensity rating across all 13 phrases. Higher scores reflect a tendency to interpret ambiguous phrases in a positive manner. Cronbach's alpha was .63 for this measure.

Independent ratings of self-esteem. Six anonymous raters "played therapist" by evaluating the content of the self-esteem essays that participants wrote during the two experimental sessions. Raters read each essay and then rated, on scales ranging from 1 (*strongly disagree*) to 9 (*strongly agree*), their agreement with two statements that pertained to the essay writer's self-liking (e.g., "I believe that this person feels very lovable and worthy of affection"), two self-competence statements (e.g., "I believe that this person feels competent, skillful, and capable"), three global self-esteem statements (e.g., "Overall, I believe that this person feels good about himself"), and two self-certainty statements (e.g., "I believe that this person is very sure of her feelings about herself"). All items demonstrated good interrater reliability (all intraclass rs $> .70$). We created indices of essay writers' self-liking, self-competence, global self-esteem, and self-certainty by averaging across the items that pertained to each of these factors (all Cronbach's αs $> .88$).

Ancillary Measures

PANAS. Watson, Clark, and Tellegen's (1988) PANAS measures participants' experience of 20 different positive (e.g., *excited*, *proud*) and negative (e.g., *hostile*, *guilty*) emotions. Instructions for this task request participants to rate how strongly they are feeling each emotion "right now" on a scale of 1 (*very slightly or not at all*) to 5 (*very much*). Cronbach's alphas were .87 and .69 for the Positive Affectivity (PA) and Negative Affectivity (NA) subscales.

Use of positive and negative emotion words. As an additional measure of mood, we used the Linguistic Inquiry and Word Count (LIWC) program

[7] An alternative way of creating scores involves subtracting each respondent's mean rating of all numbers (or letters) from the respondent's rating of his or her birthday number (or mean rating of his or her first and last initials). This strategy does not control for the influence of common letter and number preferences, however. The two scoring strategies are correlated at .89 (initials) and .88 (birthday).

(Pennebaker & Francis, 1999) to determine the frequency with which participants used positive (e.g., *happy*) and negative (e.g., *angry*) emotion words in their self-esteem essays.

Use of "I". We used the LIWC program to determine the frequency with which participants used the word *I* in their self-esteem essays.

Doctor visits, grade point averages (GPAs), and Scholastic Aptitude Test (SAT) scores. A portion of our sample granted us access to their academic and University Health Center records. The doctor-visits variable reflects the number of health center appointments that participants kept during the 1998–99 academic year. GPA reflects participants' cumulative GPA based on nontransferred credit hours; SAT scores reflect participants' combined verbal and quantitative scores.

Results

Descriptive Statistics

We report the minimum and maximum scores, means, and standard deviations for all of the self-esteem measures in Table 2. (Here and throughout the Results section, we averaged Time 1 and Time 2 scores for all measures that were administered twice; for measures that were administered only once, we used Time 1 scores.) To explore how well each measure captures the self-positivity bias that typically characterizes distributions of explicit self-esteem scores, we used one-sample *t* tests to compare the mean of each self-esteem measure with the theoretical midpoint of that scale. As shown in column 5 of Table 2, all of the implicit self-esteem measures except the supraliminal and subliminal attitude prime tasks, and all of the explicit self-esteem measures, exhibited statistically significant positivity biases.

Test–Retest Reliabilities

An important purpose of this project was to determine whether implicit self-esteem measures are stable across time. The test–retest reliabilities for the self-esteem measures appear in the last column of Table 2. Of the implicit measures, only the IAT and the initials- and birthday-preference tasks had acceptable (albeit low) test–retest reliabilities. The explicit self-esteem measure—the RSES—demonstrated good test–retest reliability.

Validity

Convergent Validity

Did the different implicit self-esteem measures tap the same underlying construct? Apparently not. As shown in Table 3, intercorrelations among the implicit self-esteem measures were generally small and/or nonsignificant. In contrast, the measures of explicit self-esteem were closely associated with one another.

Discriminant Validity

We next asked whether implicit and explicit measures truly tap distinct constructs. To answer this question, we computed correlations between all of the implicit and explicit self-esteem measures. The results shown in Table 4 indicate that implicit and explicit self-esteem are independent from one another—in general they are nonsignificantly, but positively, correlated (average r = .07). Note that few implicit measures correlated significantly with the explicit measures, and the magnitude of the observed correlations was small (all rs $< .27$).

Predictive Validity

We next explored whether implicit and explicit self-esteem would predict people's interpretations of ambiguous social communications (the AST), people's preferences for positive versus negative feedback (the FSQ), and independent raters' evaluations

Table 2
Descriptive Statistics, Positivity Biases, and Test–Retest Correlations of the Self-Esteem Measures

Self-esteem measure	Min	Max	*M*	*SD*	*t*	*df*	Test–retest
Explicit							
RSE	2.20	4.00	3.29	0.45	15.85*	80	.80
SAQ	2.20	8.80	6.41	1.09	7.48*	80	—
SL	2.80	7.00	5.42	0.96	13.31*	80	—
SC	3.40	7.00	5.84	0.81	20.49*	80	—
Implicit							
IAT	70.39	887.11	433.84	196.81	23.22*	78	.69
Supraliminal	−316.88	338.50	1.79	112.13	0.19	80	.08
Subliminal	−154.80	174.09	−1.58	57.73	−0.18	80	.28
Stroop task	−197.37	260.00	10.76	81.77	2.00*	80	−.05
ISES	−2.00	9.00	3.14	2.25	12.57*	80	.38
Initials-preference task	−3.71	2.85	0.93	1.37	6.13*	80	.63
Birthday-preference task	−2.78	3.38	0.73	1.54	4.28*	80	.53

Note. For measures that use log-transformed scores (the IAT, supraliminal and subliminal attitude-prime tasks, and Stroop), we show the nontransformed scores here. Dashes indicate that test–retest reliabilities could not be computed for the SAQ, SL, and SC because these measures were administered only once. IAT = Implicit Association Test; ISES = Implicit Self-Evaluation Survey; RSE = Rosenberg Self-Esteem Scale; SAQ = Self-Attributes Questionnaire; SC and SL = Self-Competence and Self-Liking subscales of the Self-Liking and Competence Scale; Subliminal = subliminal attitude-prime task; Supraliminal = supraliminal attitude-prime task.

* $p < .05$.

Table 3
Correlations Among the Implicit and Explicit Self-Esteem Measures

Scale	1	2	3	4	5	6	7
			Explicit measure				
1. RSE	—	.45*	.85*	.79*			
2. SAQ		—	.48*	.56*			
3. SL			—	.79*			
4. SC				—			
			Implicit measure				
1. IAT	—	−.14	−.10	.04	−.07	−.06	−.11
2. Supraliminal		—	.08	.15	.12	−.03	.21†
3. Subliminal			—	−.10	.05	.08	−.05
4. Stroop task				—	−.03	−.08	.08
5. ISES					—	−.02	.08
6. Initials-preference task						—	.23*
7. Birthday-preference task							—

Note. IAT = Implicit Association Test; ISES = Implicit Self-Evaluation Survey; RSE = Rosenberg Self-Esteem Scale; SAQ = Self-Attributes Questionnaire; SC and SL = Self-Competence and Self-Liking subscales of the Self-Liking and Competence Scale; Subliminal = subliminal attitude-prime task; Supraliminal = supraliminal attitude-prime task.
† $p < .10$ (marginally significant). * $p < .05$.

of participants' self-views. As can be seen in Table 5, the implicit self-esteem measures were generally uncorrelated with our criterion variables, with a few notable exceptions (the IAT correlated with raters' evaluations of essay-writers' self-competence, self-esteem, and self-certainty; the initials-preference task correlated with the AST and the FSQ; and the birthday-preference task correlated with the FSQ). In contrast, the explicit measures correlated consistently and in the expected directions with the criterion variables.

Relations With the Ancillary Variables

To gain a broader understanding of the associations of implicit self-esteem with personality and behavioral variables, we correlated the self-esteem measures with frequency of doctor visits, GPA, SAT scores, PA, NA, use of the word *I*, use of positive and negative emotion words, and gender. These correlations appear in Table 6.

Scores on a few of the implicit measures correlated with our mood measures: The IAT and the initials-preference task correlated positively with PA, the IAT correlated negatively with the use of negative emotion words, and the birthday-preference task correlated positively with the use of positive emotion words. Aside from this, there were only two significant correlations between implicit self-esteem scores and the ancillary variables: the ISES was negatively correlated with GPA, and the Stroop task was negatively correlated with doctor visits. With regard to the explicit self-esteem measures, there were consistent correlations with the mood measures such that people who scored high in explicit self-esteem tended to score low in NA and high in PA, and they used fewer negative emotion words when writing about themselves.

Table 4
Correlations Between the Implicit and Explicit Self-Esteem Measures

	Explicit measure			
Implicit measure	RSE	SAQ	SL	SC
IAT	.22†	.20†	.20†	.20†
Supraliminal	.06	.26*	.04	.12
Subliminal	−.13	−.09	−.22*	−.17
Stroop task	−.03	.02	.04	.02
ISES	.11	.23*	.08	.14
Initials-preference task	.13	.11	.09	.13
Birthday-preference task	−.10	.17	.00	.00

Note. IAT = Implicit Association Test; ISES = Implicit Self-Evaluation Survey; RSE = Rosenberg Self-Esteem Scale; SAQ = Self-Attributes Questionnaire; SC and SL = Self-Competence and Self-Liking subscales of the Self-Liking and Competence Scale; Subliminal = subliminal attitude-prime task; Supraliminal = supraliminal attitude-prime task.
† $p < .10$ (marginally significant). * $p < .05$.

Order Effects

In all of the preceding analyses, we explored whether the reliability and validity of the implicit measures were affected by the order in which participants completed the explicit and implicit measures.[8] Our most consistent finding was that completing the explicit measures first tended to increase the strength of the associations between the two types of measures. In this section, we briefly summarize all notable order effects.

[8] It is important to distinguish between order effects that result in mean differences in implicit self-esteem and those that influence rank orderings. Scores on the ISES and subliminal attitude-prime task were higher in the explicit-first than in the implicit-first condition ($Fs > 3.24$, $ps < .08$), but order did not affect mean scores on any other implicit measures ($Fs < 1.68$, $ps > .19$).

Table 5
Correlations of Implicit and Explicit Self-Esteem With the Criterion Measures

Self-esteem measure	Ambiguous statements	Feedback-seeking	Raters' evaluations of essay writers' Self-liking	Self-competence	Self-esteem	Self-certainty
Explicit						
RSE	.33*	.25	.51*	.47*	.50*	.32*
SAQ	.01	.34*	.38*	.34*	.37*	.37*
SL	.33*	.30*	.53*	.51*	.55*	.41*
SC	.19	.47*	.49*	.49*	.49*	.41*
Implicit						
IAT	−.04	.11	.18	.25*	.23*	.23*
Supraliminal	.13	.19†	−.06	.07	.05	−.03
Subliminal	−.07	.10	−.17	−.08	−.16	−.19†
Stroop task	−.05	.12	−.03	.00	−.04	.09
ISES	.18	.14	.00	−.05	−.01	−.03
Initials-preference task	.22*	.23*	.05	.08	−.03	−.01
Birthday-preference task	.02	.26*	−.01	.10	.02	.05

Note. IAT = Implicit Association Test; ISES = Implicit Self-Evaluation Survey; RSE = Rosenberg Self-Esteem Scale; SAQ = Self-Attributes Questionnaire; SC and SL = Self-Competence and Self-Liking subscales of the Self-Liking and Competence Scale; Subliminal = subliminal attitude-prime task; Supraliminal = supraliminal attitude-prime task.
† $p < .10$ (marginally significant). * $p < .05$.

Discriminant Validity

Although the implicit and explicit measures clearly tapped different constructs, there was more overlap between the two types of measures when explicit self-esteem was made salient before implicit self-esteem was assessed. For example, the IAT correlated significantly or marginally with all four explicit self-esteem measures in the explicit-first condition (rs = .42 to .28, ps = .04 to .10) but not in the implicit-first condition (all $ps > .30$). The ISES correlated significantly or marginally with the SAQ, the RSES, and the SL in the explicit-first condition, (rs = .36 to .25, ps = .03 to .13), but it was uncorrelated with these measures in the implicit-first condition (all $ps > .40$). Finally, the initials-preference task correlated significantly or marginally with the SAQ and the SC in the explicit-first condition (rs = .32 and .26, ps = .05 and .11) but not in the implicit-first condition (both $ps > .64$).

To more thoroughly explore these order effects, we performed Fisher's r-to-z transformations on the 56 correlations between the

Table 6
Correlations of Implicit and Explicit Self-Esteem With the Ancillary Variables

Self-esteem measure	Doctor visits	GPA	SAT	PA	NA	Gender	Word use in essays *I*	Positive emotion	Negative emotion
Explicit									
RSE	−.13	.05	.05	.31*	−.36*	−.19†	.12	−.04	−.26*
SAQ	.09	−.13	−.06	.12	−.28*	−.16	−.03	−.02	−.21†
SL	−.10	−.06	−.12	.34*	−.20†	−.14	.07	−.02	−.29*
SC	−.004	−.04	−.16	.29*	−.33*	−.09	.08	.04	−.21†
Implicit									
IAT	.04	−.13	−.03	.24*	.03	−.12	−.02	−.03	−.31*
Supraliminal	−.03	.16	.13	.01	−.17	−.18†	.12	.16	.07
Subliminal	.12	.16	.24†	−.06	−.11	.11	−.03	.16	−.14
Stroop task	−.32*	.03	−.25†	.17	.05	−.14	−.13	.09	−.05
ISES	.16	−.29*	−.06	.11	−.19†	.06	.12	−.10	−.02
Initials-preference task	.14	−.02	−.04	.23*	.11	.02	.19†	−.08	−.02
Birthday-preference task	−.02	.20	.05	−.03	−.04	.01	.17	.25*	.11

Note. Correlations with doctor visits and grade point average (GPA) are based on an N of 59; correlations with the Scholastic Aptitude Test (SAT) are based on an N of 53. IAT = Implicit Association Test; ISES = Implicit Self-Evaluation Survey; PA and NA = Positive and Negative Affectivity subscales of the Positive and Negative Affect Scale; RSE = Rosenberg Self-Esteem Scale; SAQ = Self-Attributes Questionnaire; SC and SL = Self-Competence and Self-Liking subscales of the Self-Liking and Competence Scale; Subliminal = subliminal attitude-prime task; Supraliminal = supraliminal attitude-prime task.
† $p < .10$ (marginally significant). * $p < .05$.

implicit and explicit measures (28 correlations in each of the two order conditions). We then subjected these z values to a 2 (order: explicit-first, implicit-first) × 7 (implicit measure: IAT, ISES, initials-preference, birthday-preference, Stroop, supraliminal, subliminal) between-within mixed-model analysis of variance (ANOVA). Results of this analysis confirmed that correlations in the explicit-first condition (average $r = .14$) were significantly larger than those in the implicit-first condition (average $r = .00$), $F(1, 6) = 15.54$, $p < .01$. A significant order by implicit measure interaction, $F(6, 36) = 2.48$, $p < .05$, revealed that this main effect was driven by the IAT, the ISES, and the initials-preference task: These implicit measures were more strongly related to the explicit measures in the explicit-first condition than in the implicit-first condition ($Fs > 11.69$, $p < .05$) whereas the remaining implicit measures' correlations with the explicit measures were unaffected by order ($Fs < 2.11$, $p > .19$).

Predictive Validity

The implicit self-esteem measures appeared to be more strongly related to our criterion variables in the explicit-first than in the implicit-first condition. In the implicit-first condition, only one implicit measure correlated with the criterion variables: The IAT was marginally related to raters' impressions of essay-writers' self-competence and self-certainty ($rs = .26$ and $.28$, $ps = .10$ and $.08$). In contrast, in the explicit-first condition, six of the implicit measures correlated significantly or marginally with at least one criterion variable. The IAT correlated with raters' impressions of essay-writers' self-liking and global self-esteem ($rs = .29$ and $.36$, $ps = .09$ and $.03$); the initials- and birthday-preference tasks correlated with raters' impressions of essay-writers' self-competence ($rs = .35$ and $.28$, $ps = .03$ and $.09$); the ISES and the initials-preference task correlated marginally with the AST ($rs = .26$ and $.30$, $ps = .11$ and $.07$); and the initials- and birthday-preference tasks, the Stroop, and the supraliminal attitude-prime task correlated with the FSQ ($rs = .26$ to $.46$, $ps = .10$ to $.004$).

Again, we performed Fisher's r-to-z transformations on the 84 correlations between the implicit measures and the criterion variables (42 correlations in each order condition). Results of a 2 (order) × 7 (implicit measure) between-within ANOVA on these z values showed that the implicit measures had better predictive validity in the explicit-first condition (average $r = .15$) than in the implicit-first condition (average $r = -.03$), $F(1, 10) = 26.60$, $p < .001$. Although the order by implicit measure interaction was not significant, $F(6, 60) = 1.64$, $p = .15$, we conducted follow-up analyses to determine which implicit measures were responsible for the order effect. All of the implicit measures—except the IAT and the subliminal attitude-prime task—had marginally or significantly better predictive validity in the explicit-first than in the implicit-first condition ($Fs > 3.61$, $ps < .09$). The predictive validity of the IAT and the subliminal task did not differ by order condition ($Fs < 2.11$, $ps > .17$).

Test–Retest Reliabilities

Order of task administration affected the stability of two implicit measures. The ISES was significantly more temporally stable in the implicit-first than in the explicit-first condition ($rs = .53$ and $.15$, $z = 1.92$, $p < .05$). Similarly, the subliminal attitude-prime task was marginally more stable in the implicit-first than in the explicit-first condition ($rs = .39$ and $.14$, $z = 1.17$, $p = .12$). No other implicit measure's test–retest reliability was affected by order, and a between-subjects ANOVA on the seven transformed z values did not reach statistical significance ($Ms = .39$ and $.33$), $F(1, 12) < 1$.

Discussion

In this article, we attempted to gain a clearer understanding of implicit self-esteem by exploring the psychometric properties of the different measures used to assess this construct. In this section, we summarize our findings and then address some of the issues they raise. Finally, each of the authors ends with his or her take on the implications of the findings for future implicit self-esteem research.

Positivity Biases

With the exception of the supraliminal and subliminal attitude-prime tasks, all of the implicit self-esteem measures demonstrated positivity biases such that the average response was significantly higher than the theoretical midpoint. (Note, though, that for the Stroop measure, high scores reflect a *negative* implicit attitude toward the self.) The remaining measures' positivity biases can be interpreted as evidence of their validity, as a tendency toward positive self-evaluation has been noted in both the explicit (e.g., Taylor & Brown, 1988) and implicit (e.g., Greenwald & Banaji, 1995) self-esteem literatures.

Reliability of Implicit Self-Esteem Measures

The implicit measures varied in the extent to which their individual items or trials tapped common response tendencies: Most of them displayed internal reliabilities that ranged from weak to good (αs between .49 and .88), but two implicit measures—the supraliminal attitude-prime task and the Stroop task—appeared internally inconsistent, with alphas of $-.16$ and $-.38$.

With regard to test–retest reliabilities, some implicit measures—the IAT and the initials- and birthday-preference tasks—demonstrated satisfactory stability across time; in contrast, the remaining implicit measures appeared unstable, with test–retest reliabilities that ranged from $-.05$ to .38. Thus, the majority of the available implicit self-esteem measures are not likely to yield similar scores for the same individual from one administration to the next. How much of a problem does the instability of these measures pose for implicit self-esteem research? The answer to this question may depend on one's theoretical stance regarding the measurement of implicit attitudes. For example, some researchers conceptualize implicit self-attitudes as highly malleable from one moment to the next but durable and resistant to change over the long term (Pelham & Hetts, 1999). According to this perspective, low test–retest reliabilities are not problematic because they simply reflect moment-to-moment changes in implicit self-esteem.

Thus, the low test–retest reliabilities of some implicit measures may have a methodological explanation: Perhaps responses to these measures are stable across time, provided that the psychological situation surrounding their administration is kept carefully controlled. That is, if implicit measures capture very specific

self-attitudes that fluctuate widely from situation to situation, then they may appear stable only if administered under identical conditions. This explanation is supported by Pelham and Hetts' (1999) finding that graduating college seniors experienced temporary decreases in implicit self-esteem, presumably because of the stresses and uncertainties associated with their impending entrance into the "real world." In the current study, we took steps to ensure that the experimental situation was identical at Times 1 and 2, but we did not inquire about potentially important events in participants' personal and psychological lives that might temporarily influence their implicit self-regard. It may be that researchers need to make such inquiries when using implicit self-esteem measures so that the psychological situation can be carefully controlled.

From our viewpoint, though, unstable measures are troubling regardless of the reasons behind their instability. If researchers cannot count on their measures to provide consistent readings of implicit self-esteem across time, then their ability to predict research outcomes becomes compromised. Perhaps more importantly, if implicit self-esteem is continually in flux, then an individual's high or low score at any one time becomes virtually impossible to interpret. Therefore, until the issue of the stability of implicit self-esteem is resolved, we recommend using one of the measures that demonstrated good test–retest reliability.

Finally, we consider one more issue related to the reliability of the implicit self-esteem measures: the number of items or trials on which final scores are based. Self-report measures that contain few items (the initials- and birthday-preference tasks and the ISES) and reaction-time tasks that contain relatively few trials (the supraliminal attitude-prime task and the Stroop) may be particularly vulnerable to problems of unreliability. Advocates of these measures, of course, could quite reasonably point out that they have been used with success in the past (Hetts et al., 1999; Kitayama & Karasawa, 1997; Pelham & Hetts, 1999), suggesting that they are at least valid if not reliable.

Convergent Validity of Implicit Self-Esteem Measures

Intercorrelations among the implicit self-esteem measures were inconsistent, weak, negative, and/or nonexistent, suggesting that the different implicit measures do not tap into the same underlying construct. We see four possible explanations for this state of affairs. First, it may be that only some of the implicit measures used in the current study truly tap implicit self-esteem, whereas the remaining measures do not. Second, it is possible that none of the available measures truly taps implicit self-esteem. Third, the implicit measures' low convergent validity may stem from the fact that these measures do not capture systematic variations in people's self-presentation tendencies. That is, high intercorrelations among explicit self-esteem scores may reflect similarities in self-presentation biases rather than similarities in actual self-esteem level, in which case we would not expect such high convergence among implicit measures because people presumably cannot control their responses to these measures. This explanation, however, fails to clarify why the implicit measures are completely unrelated; although we might expect them to be less intercorrelated than are explicit measures, we should still see evidence of some convergence among implicit measures.

Finally, it is possible that implicit self-esteem is a complicated and multifaceted construct, of which the different measures tap separate, unrelated, components. In the spirit of this last interpretation, Banaji (1999) recently encouraged implicit-attitudes researchers to give up the notion that multiple measures of the same construct must correlate. According to this viewpoint, a failure to find evidence of convergent validity among different implicit self-esteem measures is not problematic but simply reflects the complexity of implicit attitudes. Despite such optimism, we find the lack of intercorrelations among implicit self-esteem measures to be worrisome, both theoretically and empirically. If the measures studied in this article truly assess the same construct, then, by definition, they should overlap to a greater degree than they do. Alternatively, if the measures tap separate components of a complex construct, then they should not all purport to assess "implicit self-esteem"—instead, they should be titled to reflect the specific underlying attitudes that they measure. Unfortunately, the current findings do not allow us to answer the question of why implicit measures do not correlate. Clearly, additional research is needed to address the issue of convergent validity and refine the measurement of implicit self-esteem.

Discriminant Validity of Implicit Self-Esteem Measures

The strongest correlation we obtained between an implicit and explicit self-esteem measure was .26, suggesting that these two constructs are independent. This is not surprising, given that others have obtained similar support for the independence of implicit and explicit self-esteem (Greenwald & Farnham, 2000; Farnham et al., 1999; Hetts et al., 1999; Pelham & Hetts, 1999). We also note that (excluding correlations with the subliminal attitude-prime task) the overall pattern of correlations between the implicit and explicit measures was one of nonsignificant, but positive, associations. In this sense, our findings converge with findings in the prejudice literature that show that implicit and explicit measures of prejudice tend to be weakly, but positively, correlated (see Brauer, Wasel, & Niedenthal, 2000).

What accounts for the independence of these two theoretically related constructs? Pelham and Hetts (1999) attribute the independence of implicit and explicit self-esteem to the different learning processes involved in each: Whereas explicit self-views are based on conscious consideration of the self and self-relevant experiences, implicit self-views are learned through more automatic processes such as classical and operant conditioning and implicit learning (Berry & Dienes, 1993; Seger, 1994). Still, if the content of the messages learned through these two separate routes were similar (e.g., "I am a lovable person"), one would still expect a fair amount of overlap between people's responses to implicit and explicit self-esteem measures. The fact that these correlations tend to be weak raises intriguing questions about the cognitive and/or developmental processes that give rise to discrepancies between implicit and explicit self-esteem; we urge researchers to explore these questions, and thus broaden the field's understanding of the operation of implicit self-esteem.

Predictive Validity of Implicit Self-Esteem Measures

Only explicit self-esteem predicted people's feedback preferences and reactions to ambiguous communications; furthermore, with the exception of the IAT, none of the implicit measures related consistently to the impressions that participants conveyed

in their self-esteem essays. The implicit measures' failure to predict responses to our criterion measures is discouraging, given the past success of some of these measures in predicting people's spontaneous, affective responses (e.g., Pelham & Hetts, 1999; Spalding & Hardin, 2000).

Why were we unable to find support for the predictive abilities of implicit self-esteem measures when previous researchers have found such support? Two possible explanations exist. First, in the studies cited above, implicit self-esteem predicted people's responses after receiving negative and/or failure feedback. In contrast, we administered our criterion measures in a context that did not contain any self-relevant threat. Perhaps implicit self-esteem primarily guides people's responses to experiences that threaten the integrity of the self-system and thus evoke great emotion. If so, then we may have been unable to find consistent correlations between implicit self-esteem and the criterion measures because our safe, nonthreatening questionnaire measures did not evoke people's implicit evaluations of the self.

Second, the fact that two of our criterion variables were self-report scales may have interfered with the implicit measures' ability to predict people's responses. As Shedler et al. (1993) pointed out, "the defensive processes that enable people to look healthy on mental health scales may also enable them to look healthy on the criterion measures used to validate these scales" (p. 1118). The possibility that self-reports are inappropriate criteria for implicit self-esteem measures is supported by our finding that the IAT correlated consistently with the variables that were not, strictly speaking, self-report measures—that is, raters' impressions of participants' self-views. Thus, researchers may be well-advised to steer clear of self-report instruments, and instead rely primarily on indirect or nonconscious criterion measures when studying implicit self-esteem.

Finally, with regard to our ancillary variables, neither the explicit nor implicit self-esteem measures related consistently to participants' gender, physical health, academic achievement, or self-focused attention. This is not particularly troubling, though. Until we know more about the ways in which implicit self-views influence the functioning of the self-system, it may be unduly harsh to condemn implicit measures for failing to predict behavioral variables that are most likely influenced by a multitude of factors. To their credit, the IAT and the initials- and birthday-preference tasks correlated in the expected direction with some of our mood measures, a finding which attests to the implicit measures' ability to capture affectively based responses.

Order Effects

When implicit self-esteem was measured after explicit self-esteem, correlations between implicit and explicit self-esteem measures and between the implicit measures and our criterion variables, tended to be higher. On one hand, this finding makes sense given the logic of the automatic attitude activation effect: Priming the self by assessing explicit self-esteem activates affect that is evaluatively consistent with the self-attitude; this affect then guides people's responses to subsequent implicit measures, increasing the correlations between explicit and implicit measures. On the other hand, our order effects may raise concerns about the "implicit"-ness of some of the implicit self-esteem measures. It is possible that preceding implicit measures with explicit ones brings implicit tasks under greater conscious control (i.e., makes them less implicit). Additional research is needed to tease apart these two possible explanations. Finally, we note that other researchers who have worked extensively with implicit self-esteem measures have failed to find that implicit and explicit measures relate more strongly when explicit measures are completed first (A. G. Greenwald, personal communication, April 4, 2000). Therefore, it is difficult to draw firm conclusions regarding the generalizability of our order effects, and we caution researchers to give careful consideration to the order in which they administer implicit and explicit measures.

Coda

We return now to the story with which we opened, and point out yet one more similarity between the plight of implicit self-esteem researchers and that of the blind men who wanted to know what an elephant looked like: The authors, like the fabled blind men, came away from this investigation with very different ideas about the "big picture." We therefore end by presenting our own, individual thoughts about the implications of our findings.

Pennebaker: My primary question is, What, if anything, do these various markers of self-esteem mean? Explicit self-report measures are essentially tapping broad beliefs or schemas about who we think we are. They are essentially self-theories: "I think I am this way"; "I feel certain I'm not this other way." Self theories, as tapped by the RSES, are quite interesting but, in reality, may not reflect aspects of the self of which respondents are unaware. For example, I have a theory or schema about what elephants are like—big, friendly, good memories, probably conservative Republicans. It's a good theory and probably consistent over time, internally consistent, and shared by others in the culture. But, like explicit self-esteem, my elephant theory may not have much to do with any objective real-world measures.

Superficially, implicit measures are far more compelling. They are outside of awareness and hint at psychodynamic processes without our having to invoke Freud's name. But they do not tap general self-esteem processes. It is possible, of course, that the various implicit measures reflect highly specific domains that are probably more method-based than process-based. More worrisome, however, is that the study of implicit self-esteem may be a boondoggle. Right now, the psychometrics simply are not there. Future, more comprehensive validation studies will ultimately reveal if we are holding an elephant's tail or standing alone in an empty barn with a limp rope in our hands.

Swann: Even the most Panglossian advocate of measures of implicit self-esteem could not help but be discouraged by our findings. And although it is certainly too early for the field to declare that we have a boondoggle on our hands, it is time to think long and hard about the viability of measuring implicit self-esteem. To me, one of the most troubling findings was not that implicit measures were only weakly predictive of the criterion variables (because we may have picked the wrong ones), it was that many of them were most predictive when they were collected *after* the explicit measures. If implicit measures are potent only insofar as they can ride on the coattails of explicit measures, then the most efficient strategy may be to measure explicit self-esteem directly and forget about implicit measures. And our findings tell us that being consigned to only using explicit measures would not be

catastrophic. After all, in a field in which self-report measures seldom break the .30 barrier, correlations as high as .55 between our measures of explicit self-esteem and our criterion measures are impressive. Yet explicit measures do have an Achilles' heel, for people may consciously bias their assessments of themselves or may be unaware of a lingering vulnerability that they have repressed or forgotten. This is the extremely important problem that implicit measures are designed to solve. Our findings do not inspire confidence that the current bevy of implicit measures will fill the bill (although the IAT seems more promising than the others). Nevertheless, I still believe that the goal of finding a reliable and valid measure of implicit self-esteem is important enough that it should remain a high priority.

Bosson: Those familiar with the history of projective personality assessment will note some similarities between the issues raised by implicit self-esteem measures and those raised by the Thematic Apperception Test (TAT; Morgan & Murray, 1935/1981). Like the current measures of implicit self-esteem, the TAT is an indirect assessment tool that was devised in an effort to explore "unconscious processes. . .of which the subject [is] not aware" (Murray, 1973; as cited in Anderson, 1999, p. 36). Also like the implicit self-esteem measures, early versions of the TAT had weak psychometric properties such as poor interrater and internal reliabilities, low test–retest correlations, and weak correlations with self-report measures (e.g., Entwisle, 1972; Fineman, 1977; McClelland, 1980; Weinberger & McClelland, 1990; Winter, 1999).

Nevertheless, gradual changes in scoring strategies yielded dramatic improvements in the psychometrics of the TAT and TAT-based measures (e.g., Costantino & Malgady, 1999; Winter, 1999). For example, McClelland and Atkinson (1948) devised a scoring system that increased the predictive validity of the TAT, allowing researchers to predict important outcome variables such as achievement, persistence at challenging tasks, economic success, adaptation to life changes, job performance, and career choice (see McClelland, 1999, and Winter, 1999, for reviews). Perhaps there is a lesson to be learned here. Implicit self-esteem measures—like the TAT—represent a novel approach to the study of personality and the self; if researchers refuse to be discouraged by the implicit measures' weak psychometric beginnings and persist in their efforts to perfect the measurement of implicit self-esteem, they may ultimately find ways to improve implicit measures to the point of greater utility. On the other hand, if we abandon these measures before fully (and creatively) exploring their capabilities, we will never learn what they could have taught us.

References

Anderson, J. W. (1999). Henry A. Murray and the creation of the Thematic Apperception Test. In L. Gieser & M. I. Stein (Eds.), *Evocative images: The Thematic Apperception Test and the art of projection* (pp. 23–38). Washington, DC: American Psychological Association.

Antonucci, T. C., & Jackson, J. S. (1983). Physical health and self-esteem. *Family and Community Health, 6*, 1–9.

Banaji, M. R. (1999, June). *Implicit attitudes can be measured.* Paper presented at the Society for Personality and Social Psychology Preconference of the 11th Annual Convention of the American Psychological Society, Denver, CO.

Bargh, J. A., Chaiken, S., Govender, R., & Pratto, F. (1992). The generality of the automatic attitude activation effect. *Journal of Personality and Social Psychology, 62*, 893–912.

Beck, A. T. (1985). Cognitive therapy, behavior therapy, psychoanalysis, and pharmacotherapy: A cognitive continuum. In M. J. Mahoney & A. Freeman (Eds.), *Cognition and psychotherapy* (pp. 325–347). New York: Plenum Press.

Berry, D. C., & Dienes, Z. (1993). *Implicit learning: Theoretical and empirical issues.* Hove, England: Erlbaum.

Bosson, J. K., & Swann, W. B., Jr. (May, 1998). *Explicit and implicit self-esteem and narcissism.* Poster session presented at the 10th Annual Convention of the American Psychological Society, Washington, DC.

Brauer, M., Wasel, W., & Niedenthal, P. M. (2000). Implicit and explicit components of prejudice. *Review of General Psychology, 4*, 79–101.

Carroll, J. L., & Buhrow, M. (1994). Concurrent validity of the Culture-Free Self-Esteem Inventory and physical health in college students. *Psychological Reports, 74*, 553–554.

Costantino, G., & Malgady, R. G. (1999). The Tell-Me-A-Story Test: A multicultural offspring of the Thematic Apperception Test. In L. Gieser & M. I. Stein (Eds.), *Evocative images: The Thematic Apperception Test and the art of projection* (pp. 191–206). Washington, DC: American Psychological Association.

Crary, W. G. (1966). Reactions to incongruent self-experiences. *Journal of Consulting Psychology, 30*, 246–252.

Dua, J. K. (1993). The role of negative affectivity and positive affectivity in stress, depression, self-esteem, assertiveness, Type A behaviors, psychological health, and physical health. *Genetic, Social, & General Psychology Monographs, 119*, 515–552.

Entwisle, D. R. (1972). To dispel fantasies about fantasy-based measures of achievement motivation. *Psychological Bulletin, 83*, 1131–1153.

Epstein, S. (1994). Integration of the cognitive and the psychodynamic unconscious. *American Psychologist, 49*, 709–724.

Farnham, S. D., Greenwald, A. G., & Banaji, M. R. (1999). Implicit self-esteem. In D. Abrams & M. A. Hogg (Eds.), *Social identity and social cognition* (pp. 230–248). Malden, MA: Blackwell..

Fazio, R. H. (1990). A practical guide to the use of response latency in social psychological research. In C. Hendrick, M. S. Clark, et al. (Eds.), *Research methods in personality and social psychology: Review of personality and social psychology* (Vol. 11, pp. 74–97). Newbury Park, CA: Sage.

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology, 50*, 229–238.

Fineman, S. (1977). The achievement motive construct and its measurement: Where are we now? *British Journal of Psychology, 68*, 1–22.

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

Greenwald, A. G., & Farnham, S. D. (2000). *Using the implicit association test to measure self-esteem and self-concept.* Manuscript submitted for publication.

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

Hetts, J. J., Sakuma, M., & Pelham, B. W. (1999). Two roads to positive regard: Implicit and explicit self-evaluation and culture. *Journal of Experimental Social Psychology, 35*, 512–559.

Hoorens, V., & Nuttin, J. M. (1993). Overevaluation of own attributes: Mere ownership or subjective frequency? *Social Cognition, 11*, 177–200.

Kernis, M. H., Grannemann, B. D., Richie, T., & Hart, J. (1988). The role of contextual factors in the relationship between physical activity and self-awareness. *British Journal of Social Psychology, 27*, 265–273.

Khalid, R. (1990). Self-esteem and academic achievement: An investiga-

tion of ethnic and sex differences. *Journal of Behavioral Sciences, 1,* 3–17.

Kitayama, S., & Karasawa, M. (1997). Implicit self-esteem in Japan: Name letters and birthday numbers. *Personality and Social Psychology Bulletin, 23,* 736–742.

Major, B., Barr, L., Zubek, J., & Babey, S. H. (1999). Gender and self-esteem: A meta-analysis. In W. B. Swann, J. H. Langlois, & L. A. Gilbert (Eds.), *Sexism and stereotypes in modern society: The gender science of Janet Taylor Spence* (pp. 223–253). Washington, DC: American Psychological Association.

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

McClelland, D. C. (1980). Motive dispositions: The merits of operant and respondent measures. In L. Wheeler (Ed.), *Review of personality and social psychology* (Vol. 1, pp. 10–41). Beverly Hills, CA: Sage.

McClelland, D. C. (1999). How the test lives on: Extensions of the Thematic Apperception Test approach. In L. Gieser & M. I. Stein (Eds.), *Evocative images: The Thematic Apperception Test and the art of projection* (pp. 163–175). Washington, DC: American Psychological Association.

McClelland, D. C., & Atkinson, J. W. (1948). The projective expression of needs: I. The effect of different intensities of the hunger drive on perception. *Journal of Psychology, 25,* 205–232.

Morgan, C. D., & Murray, H. A. (1981). A method for investigation of fantasies: The Thematic Apperception Test. In E. S. Shneidman (Ed.), *Endeavors in psychology: Selections from the personology of Henry A. Murray* (pp. 390–401). New York: Harper & Row. (Original work published 1935)

Mullen, B., Chapman, J. G., & Peaugh, S. (1989). Focus of attention in groups: A self-attention perspective. *Journal of Social Psychology, 129,* 807–817.

Newbegin, I., & Owens, A. (1996). Self-esteem and anxiety in secondary school achievement. *Journal of Social Behavior and Personality, 11,* 521–530.

Nuttin, J. M. (1985). Narcissism beyond Gestalt awareness: The name letter effect. *European Journal of Social Psychology, 15,* 353–361.

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.

O'Connor, B. P., & Vallerand, R. J. (1998). Psychological adjustment variables as predictors of mortality among nursing home residents. *Psychology and Aging, 13,* 368–374.

Pelham, B. W., & Hetts, J. J. (1999). *Implicit self-evaluation.* Unpublished manuscript, State University of New York at Buffalo.

Pelham, B. W., & Swann, W. B., Jr. (1989). From self-conceptions to self-worth: On the sources and structure of global self-esteem. *Journal of Personality and Social Psychology, 57,* 672–680.

Pennebaker, J. W. (1997). *Opening up: The healing power of expressing emotion.* New York: Guilford Press.

Pennebaker, J. W. & Francis, M. E. (1999). *Linguistic Inquiry and Word Count (LIWC).* Mahwah, NJ: Erlbaum.

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

Saxe, J. G. (1936). The blind men and the elephant. In H. Felleman (Ed.), *The best loved poems of the American people* (pp. 521–522). New York: Doubleday.

Seger, C. A. (1994). Implicit learning. *Psychological Bulletin, 115,* 163–196.

Shedler, J., Mayman, M., & Manis, M. (1993). The illusion of mental health. *American Psychologist, 48,* 1117–1131.

Shrauger, J. S., & Lund, A. (1975). Self-evaluations and reactions to evaluations from others. *Journal of Personality, 43,* 94–108.

Smith, S. M. (2000). *Self-esteem accessibility: Measurement and correlates.* Unpublished manuscript, North Georgia College and State University.

Spalding, L. R., & Hardin, C. D. (2000). Unconscious unease and self-handicapping: Behavioral consequences of individual differences in implicit and explicit self-esteem. *Psychological Science, 10,* 535–539.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18,* 643–662.

Swann, W. B., Jr. (1992). Seeking truth, finding despair: Some unhappy consequences of a negative self-concept. *Current Directions in Psychological Science, 1,* 15–18.

Swann, W. B., Jr., Wenzlaff, R. M., Krull, D. S., & Pelham, B. W. (1992). The allure of negative feedback: Self-verification strivings among depressed persons. *Journal of Abnormal Psychology, 101,* 293–306.

Tafarodi, R. W. (1998). Paradoxical self-esteem and selectivity in the processing of social information. *Journal of Personality and Social Psychology, 74,* 1181–1196.

Tafarodi, R. W., & Swann, W. B., Jr. (1995). Self-liking and self-competence as dimensions of global self-esteem: Initial validation of a measure. *Journal of Personality Assessment, 65,* 322–342.

Tarlow, E. M., & Haaga, D. A. F. (1996). Negative self-concept: Specificity to depressive symptoms and relation to positive and negative affectivity. *Journal of Research in Personality, 30,* 120–127.

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

Teglasi, H., & Epstein, S. (1998). Temperament and personality theory: The perspective of cognitive-experiential self-theory. *School Psychology Review, 27,* 534–550.

Vingilis, E., Wade, T. J., & Adlaf, E. (1998). What factors predict student self-rated physical health? *Journal of Adolescence, 21,* 83–97.

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

Wegner, D. M., & Giuliano, T. (1980). Arousal-induced attention to self. *Journal of Personality and Social Psychology, 38,* 719–726.

Weinberger, J., & McClelland, D. C. (1990). Cognitive versus traditional motivational models: Irreconcilable or complementary? In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of motivation and cognition* (Vol. 2, pp. 562–597). New York: Guilford Press.

Winter, D. G. (1999). Linking personality and "scientific" psychology: The development of empirically derived Thematic Apperception Test measures. In L. Gieser & M. I. Stein (Eds.), *Evocative images: The Thematic Apperception Test and the art of projection* (pp. 107–124). Washington, DC: American Psychological Association.

Received July 6, 1999
Revision received May 9, 2000
Accepted May 15, 2000 ■