Research Report

Reporting Tendencies Underlie Discrepancies Between Implicit and Explicit Measures of Self-Esteem

Michael A. Olson,[1] Russell H. Fazio,[2] and Anthony D. Hermann[3]

[1]University of Tennessee, [2]The Ohio State University, and [3]Willamette University

ABSTRACT—*The assumption that implicit measures assess associations that are not accessible to consciousness abounds in current social cognition research. In the present report, we question this assumption, focusing on the construct of implicit self-esteem as a case in point. Although researchers often argue that implicitly measured self-esteem is unconscious, we provide evidence that it is not, and that discrepancies between implicit and explicit measures of self-esteem are the result of reporting tendencies. Study 1 demonstrated that individuals for whom explicitly measured self-esteem is relatively high and implicitly measured self-esteem is relatively low admit to overpresenting themselves. In Study 2, implicit and explicit measures of self-esteem were related when subjects were urged to avoid over- or underpresenting themselves when responding to the explicit measures. We discuss the critical distinction between awareness of one's attitudes and awareness of their influence.*

Freud (1915/1957) argued that self-criticism sometimes lurks beneath conscious awareness, occasionally manifesting itself in personality disorders like narcissism (Kris, 1994). Decades later, researchers armed with modern implicit measurement tools are making similar claims: that one's attitude toward the self can be hidden from consciousness, resulting in the very sorts of problems posited by Freud (e.g., Koole & Pelham, 2003). Can one's unconscious keep such secrets from the prying eye of

consciousness? We question the assumptions of both Freud and current implicit approaches to measuring self-esteem, and argue that people may know themselves better than theories that relegate some aspects of the self to the unconscious would imply.

Evidence suggests that implicitly measured self-esteem relates to psychological functioning (Koole, Dijksterhuis, & van Knippenberg, 2001; Pelham, Mirenberg, & Jones, 2002). For example, Greenwald and Farnham (2000) demonstrated that higher scores on a self-esteem Implicit Association Test (IAT) were associated with less negative reactions to failure feedback, a buttressing effect often found for explicit self-esteem (see also Hetts & Pelham, 2001; Jones, Pelham, Mirenberg, & Hetts, 2002; Spalding & Hardin, 1999). Implicitly and explicitly measured self-esteem can be unrelated (e.g., Bosson, Swann, & Pennebaker, 2000; Koole & Pelham, 2003; Spalding & Hardin, 1999), and such discrepancies have been associated with important consequences. For example, Jordan, Spencer, Zanna, Hoshino-Browne, and Correll (2003) demonstrated that individuals with low implicitly measured self-esteem and high explicitly measured self-esteem show defensiveness, narcissism, increased in-group bias, and greater spreading of alternatives in a dissonance situation.

Many researchers tend to assume that implicit self-esteem is unconscious and that implicit measures tap such unconscious self-views (Kernis, 2003; Koole & Pelham, 2003). For example, Bosson et al. (2000) stated, "Because implicit self-esteem is nonconscious, it must be measured indirectly" (p. 631). This assumption most often comes into play in work with the IAT, probably because of its prevalent use. For example, Rudman and Kilianski (2000) argued that "previous research has supported the IAT as a powerful and flexible measure of unconscious attitudes and beliefs" (p. 1316), and Richeson and Shelton

Address correspondence to Michael Olson, Department of Psychology, University of Tennessee, Knoxville, TN 37996, e-mail: olson@utk.edu.

(2003) stated, "The IAT is a measure of automatic associations, often employed to assess unconscious bias" (p. 288). Other implicit measures are often assumed to tap the unconscious as well (e.g., Baccus, Baldwin, & Packer, 2004; Hetts, Sakuma, & Pelham, 1999).

CONSCIOUS AND UNCONSCIOUS PROCESSES AND THE IAT

We question the assumption that implicit measures tap attitudes of which respondents are unaware (Fazio & Olson, 2003). Instead, we argue that some people are simply reluctant to report such evaluations on explicit measures, and thus create dissociations between the two measurement types. Our perspective matches Fazio's MODE model (Fazio & Towles-Schwen, 1999), which states that attitude measures involving verbal behavior are subject to downstream forces that can counter the influence of automatically activated attitudes. For example, scores on the Modern Racism Scale can be affected not only by automatically activated racial attitudes, but also by motivation to control prejudiced reactions (Fazio, Jackson, Dunton, & Williams, 1995). Individuals with low motivation to control their prejudice are comfortable expressing their automatically activated negativity, but others are more motivated to correct for this negativity. Hence, the relation between implicitly and explicitly measured racial attitudes is moderated by motivation. We argue that similar processes operate in implicit and explicit measures of self-esteem.

Because communicating one's self-evaluation on an explicit measure is necessarily verbal behavior, it can be influenced by forces in addition to one's automatically activated attitude toward the self. Indeed, the very act of responding to an explicit measure of self-esteem can be viewed as an exercise in self-presentation: "Do I risk appearing boastful, or do I risk under-selling myself?" Thus, an explicit query of one's attitude toward the self may automatically activate the self-evaluation, and many individuals may report this attitude on the explicit measure. However, other individuals may be motivated to modify the verbal report in the interest of a self-presentational goal. Responses on the explicit measure may then reflect a desire to present an appearance somewhat different from that implied by one's attitude toward the self.

We are not the first to note the relevance of self-presentation to measures of self-esteem. Indeed, Baumeister, Tice, and Hutton (1989) argued that responses to self-esteem scales are strongly influenced by self-presentational orientation. We argue that self-presentation style might explain cases in which people appear either more or less positive on an explicit measure than would be indicated by the evaluation that is automatically activated in their response to the self. Modest individuals may appear more negative on an explicit measure, and boastful individuals may appear more positive. However, given the prevailing norm that people should feel good about themselves,

we suspect that more people inflate than deflate their self-esteem on explicit measures. Specifically, we argue that the boastful, overpresenting style is associated with more positive scores on an explicit measure of self-esteem than would be implied by implicitly measured self-esteem. Study 1 tested this hypothesis.

STUDY 1

Method

Seventy-one undergraduates participated to fulfill psychology course requirements. Subjects completed one implicit (IAT) measure and two explicit measures of self-esteem on computers in individual cubicles. The IAT's parameters matched the seven-block procedure (Greenwald, Nosek, & Banaji, 2003), with some variations. Self-related words consisted of "Me," "Myself," and "I" and the subject's first, middle, and last names; e-mail address; birthday; hometown; and phone number. Other-related words consisted of "Them," "Him," and "Her" and analogous personal information about someone other than the participant (e.g., first, middle, and last names). Twenty pleasant and 20 unpleasant words were selected from earlier IAT studies. As in our previous work (Olson & Fazio, 2004), we used the "personalized" IAT category labels "I like" and "I don't like" instead of "pleasant" and "unpleasant." In critical blocks, the tasks of categorizing "Self" versus "Other" words and "I like" versus "I don't like" words were interleaved.

After the IAT, subjects completed the Rosenberg (1965) Self-Esteem Questionnaire and a feeling-thermometer measure of self-esteem, on which they used a 101-point scale to record how warmly they felt toward themselves. Finally, subjects rated their self-presentational styles by responding to the question, "How much does the following trait apply to you?" Responses for each of six traits—*modest, humble, meek, proud, boastful,* and *arrogant*—were made on a scale from 1 (*not at all*) to 5 (*very much*).

Results and Discussion

IAT scores reflecting a pro-self bias were computed according to the ratio strategy advocated by Greenwald et al. (2003; with variations as noted in Olson & Fazio, 2004; see Blanton & Jaccard, 2006, for a critique of the scoring and interpretation of the IAT). The mean IAT index was 0.36 ($SD = 0.21$), which differed from zero, $t(70) = 14.90$, $p_{\text{rep}} > .999$, $d = 3.56$. The means on the Rosenberg inventory and feeling thermometer were 3.98 ($SD = 0.71$) and 86.0 ($SD = 9.60$), respectively. After ratings for *modest, humble,* and *meek* were reverse-scored, the trait ratings were subjected to a reliability analysis, which indicated adequate reliability ($\alpha = .61$). They were combined to form an overall trait index on which higher numbers indicate a more overpresenting style ($M = 2.74$, $SD = 0.59$).

Correlations among the IAT, explicit measures, and trait index were computed (Table 1). As in many past studies, the IAT and

TABLE 1

Correlations Between Measures in Study 1

Measure	IAT	Rosenberg	Feeling thermometer	Trait ratings
IAT	—	−.12	−.09	−.06
Rosenberg		—	.62*	.09
Feeling thermometer			—	.29*
Trait ratings				—

Note. IAT = Implicit Association Test; Rosenberg = Rosenberg (1965) Self-Esteem Questionnaire.
*$p < .01$.

explicit measures of self-esteem were uncorrelated. The trait index exhibited no correlation with the IAT or Rosenberg scale, but showed modest correspondence with the feeling thermometer.

Given their strong correlation, the explicit measures were standardized and averaged to form an index of explicitly measured self-esteem. To test the prediction that individuals with relatively low implicitly measured and relatively high explicitly measured self-esteem are more likely than other individuals to have a self-aggrandizing presentational style, we entered IAT scores, the explicit index, and their interaction term into a multiple regression equation predicting the trait index. A nearly significant IAT × Explicit Index interaction was found, $t(70) = 1.88$, $p_{rep} = .864$, $d = 0.45$. As expected, subjects with low implicit and high explicit scores were more likely than others to have high trait-index scores, meaning that they rated themselves relatively high in boastfulness, arrogance, and pride, and low on modesty, humbleness, and meekness. Tests of simple slopes indicated that the correspondence between the explicit index and the trait index was significantly positive among subjects with lower IAT scores ($−1 SD$ from the mean), $b = 0.31$, $t(70) = 1.96$, $p_{rep} = .878$, $d = 0.47$, but was nonsignificant among subjects with higher IAT scores ($+1 SD$ from the mean), $b = −0.18$, $t < 1$.

In sum, Study 1 confirmed our prediction that individuals high in explicitly measured self-esteem but low in implicitly measured self-esteem report a more boastful self-presentation style than other individuals do. This finding is consistent with the notion that discrepancies between implicit and explicit measure of self-esteem are associated with motivation to engage in self-promotion. Moreover, these data also suggest that people are aware of their typical self-presentational style, and perhaps could correct for it if properly motivated to do so. However, these findings suggest only that relatively boastful individuals over-present themselves; they do not support the idea that more modest individuals underpresent themselves. It could be that underpresenting is relatively rare in Western cultures (Kitayama & Uchida, 2003), but it is also likely that self-reported trait ratings are an imperfect means of addressing discrepancies between implicit and explicit measures of self-esteem.

Study 1 also did not directly address the degree to which discrepancies between implicit and explicit measures actually stem from motivated self-presentation. In Study 2, we addressed this issue directly by imploring some subjects to avoid either under- or overpresenting themselves when completing the explicit measures. The experiment examined the relation between implicit and explicit measures with and without such instructions. If people are unaware of their underlying self-sentiments, then such instructions should be ineffective in prompting people to report their automatically activated self-evaluations on an explicit measure. If, however, people are aware of the self-evaluations assessed by the IAT and of their typical self-presentational style, and if they can be induced to abandon that self-presentational orientation, then a positive correlation between the implicit and explicit measures should emerge.

STUDY 2

Method

As part of course requirements, 106 introductory psychology students completed the experiment. Three subjects were excluded from analyses because they made an excessive number of errors (> 30%) on the IAT.

The IAT was identical to that used in Study 1, except that 4 self-related items ("Me," "Myself," "I," and "Self") and 4 other-related terms ("Him," "Her," "Them," and "Other") were used, and the lists of pleasant and unpleasant items were reduced to 10 each. After the IAT, subjects completed the same explicit measures of self-esteem as in Study 1, but those randomly assigned to the honest condition received these instructions prior to completing those measures:

> Please, be honest and accurate when answering the questionnaires. In general, when people answer self-related questions such as the ones you will be presented with, they sometimes tend to overpresent themselves by being proud and boastful, or underpresent themselves by being modest or humble. Please, keep in mind that the answers of your questions will be used only for statistical purposes and they will in no way be linked to you, so please consider each question carefully and answer it as truthfully as possible.

Also, the following message appeared on the computer screen while subjects in the honest condition completed the explicit measures: "Please remember to be completely honest and accurate. Don't be proud or boastful, but don't be modest or humble either." Control subjects received no such instructions.

Results and Discussion

The mean IAT score was 0.58 ($SD = 0.30$), which differed from zero, $t(105) = 19.75$, $p_{rep} > .999$, $d = 3.85$. Means on the Rosenberg inventory and feeling thermometer were 3.88 ($SD = 0.62$) and 78.4 ($SD = 13.12$), respectively. None of the
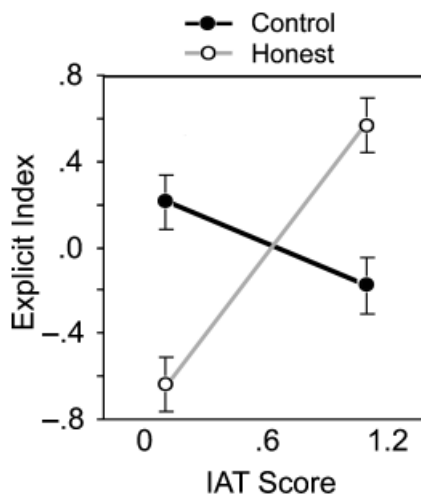
Fig. 1. Composite index of explicitly measured self-esteem as a function of implicitly measured self-esteem (Implicit Association Test, IAT) and honesty condition in Study 2. Higher scores on the explicit index and the IAT reflect higher self-esteem. Error bars represent 1 *SEM*.

measures differed between the two conditions (all ts < 1.15, all p_{rep}s $< .683$). The two explicit measures were highly correlated ($r = .74$), so as in Study 1, they were combined to form a single index.

To determine whether the instructions had an impact on the relation between the implicit and explicit measures, we used a multiple regression equation with IAT scores, a dummy variable coded for condition, and their interaction term to predict the explicit index. The predicted interaction was revealed, $t(102) = 2.23$, $p_{rep} = .940$, $d = 0.44$. IAT scores predicted explicit measures of self-esteem in the honest condition, $b = 0.30$, $t(52) = 2.26$, $p_{rep} = .94$, $d = 0.63$, but not in the control condition, $b = -0.12$, $t < 1$ (see Fig. 1). Thus, explicit and implicit measures of self-esteem corresponded when subjects were implored to avoid over- or underpresenting themselves on the explicit measures.

GENERAL DISCUSSION

The present research investigated an assumption that pervades the social psychology literature: that one's implicitly measured self-esteem is closed to conscious awareness. We have demonstrated that people do have conscious access to their automatically activated attitude toward the self, their implicit self-esteem, and that certain individuals claim to have higher self-esteem on explicit measures than their implicitly measured attitudes would imply. We believe this represents strong evidence that (a) people are aware of their automatic, or implicit, attitudes toward the self, and (b) responses on explicit self-esteem measures are colored by self-presentational strategies (see Dijksterhuis, Albers, & Bongers, in press, for a related conceptual analysis). Future work should test these assertions using other implicit measures.

Although this may be the first demonstration that people are aware of their implicitly measured self-esteem, some previous research is consistent with our arguments. For example, Koole et al. (2001) found that implicit and explicit measures converged when people made their explicit judgments quickly. It appears that when deprived of the opportunity to evoke any self-presentational strategies on the explicit measure, subjects reported what the implicit measure revealed. Pelham et al. (2005) demonstrated that women show greater correspondence than men between implicit and explicit indices of self-esteem, perhaps because women are more likely than men to trust their intuitive feelings about themselves. The present results are also consistent with work by Kitayama and Uchida (2003) demonstrating that cultural norms can influence whether one's explicit reports about the self reflect one's implicitly measured self-views.

We are not implying that people can spontaneously reach total self-insight with a simple prompt. Certainly there are limits to one's conscious ability to retrieve information from the fathoms of memory. Nor are we implying that people are always aware of the influence of their automatic responses. Researchers often say that people are "unaware" of implicitly measured attitudes, and we have shown that it may be wise to consider more carefully whether people are truly unaware of this information. However, there is ample evidence (Nisbett & Wilson, 1977) that people can be unaware of the influence of their attitudes on judgments and other behaviors. We caution future researchers to abide by this distinction and not assume that people lack conscious access to their implicitly measured self-esteem (or other implicitly assessed attitudes). What they are more likely to lack is awareness of the influence that those automatically activated attitudes are exerting in any given judgmental situation.

REFERENCES

Baccus, J.R., Baldwin, M.W., & Packer, D.J. (2004). Increasing implicit self-esteem through classical conditioning. *Psychological Science*, 15, 498–502.

Baumeister, R.F., Tice, D.M., & Hutton, D.G. (1989). Self-presentational motivations and personality differences in self-esteem. *Journal of Personality*, 57, 547–579.

Blanton, H., & Jaccard, J. (2006). Tests of multiplicative models in psychology: A case study using the unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review*, 113, 155–169.

Bosson, J.K., Swann, W.B., Jr., & Pennebaker, J. (2000). Stalking the perfect measure of implicit self-esteem: The blind men and the elephant revisited? *Journal of Personality and Social Psychology*, 79, 631–643.

Dijksterhuis, A., Albers, L.W., & Bongers, K.C. (in press). Digging for the real attitude: Lessons from research on implicit and explicit

self-esteem. In R.E. Petty, R.H. Fazio, & P. Briñol (Eds.), *Attitudes: Insights from the new implicit measures*. Mahwah, NJ: Erlbaum.

Fazio, R.H., Jackson, J.R., Dunton, B.C., & Williams, C.J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology*, *69*, 1013–1027.

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

Fazio, R.H., & Towles-Schwen, T. (1999). The MODE model of attitude-behavior processes. In S. Chaiken & Y. Trope (Eds.), *Dual process theories in social psychology* (pp. 97–116). New York: Guilford.

Freud, S. (1957). The unconscious. In J. Strachey (Ed. & Trans.), *The standard edition of the complete works of Sigmund Freud* (Vol. 14, pp. 159–215). London: Hogarth Press. (Original work published 1915)

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.

Hetts, J.J., & Pelham, B.W. (2001). A case for the nonconscious self-concept. In G.W. Moskowitz (Ed.), *Cognitive social psychology* (pp. 105–123). Mahwah, NJ: Erlbaum.

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.

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., Hoshino-Browne, E., & Correll, J. (2003). Secure and defensive self-esteem. *Journal of Personality and Social Psychology*, *85*, 969–978.

Kernis, M.H. (2003). Toward a conceptualization of optimal self-esteem. *Psychological Inquiry*, *14*, 1–26.

Kitayama, S., & Uchida, Y. (2003). Explicit self-criticism and implicit self-regard: Evaluating self and friend in two cultures. *Journal of Experimental Social Psychology*, *39*, 476–482.

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., & Pelham, B.W. (2003). On the nature of implicit self-esteem: The case of the name letter effect. In S. Spencer, S. Fein, & M.P. Zanna (Eds.), *Motivated social perception: The Ontario Symposium* (pp. 93–116). Hillsdale, NJ: Erlbaum.

Kris, A.O. (1994). Freud's treatment of a narcissistic patient. *International Journal of Psychoanalysis*, *75*, 649–664.

Nisbett, R., & Wilson, T. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*, 231–259.

Olson, M.A., & Fazio, R.H. (2004). Reducing the influence of extrapersonal associations on the Implicit Association Test: Personalizing the IAT. *Journal of Personality and Social Psychology*, *86*, 653–667.

Pelham, B.W., Koole, S.L., Hardin, C.D., Hetts, J.J., Seah, E., & DeHart, T. (2005). Gender moderates the relation between implicit and explicit self-esteem. *Journal of Experimental Social Psychology*, *41*, 84–89.

Pelham, B.W., Mirenberg, M.C., & Jones, J.K. (2002). Why Susie sells seashells by the seashore: Implicit egotism and major life decisions. *Journal of Personality and Social Psychology*, *82*, 469–487.

Richeson, J.A., & Shelton, J.N. (2003). When prejudice does not pay: Effects of interracial contact on executive function. *Psychological Science*, *14*, 287–290.

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

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

Spalding, L.R., & Hardin, C.D. (1999). Unconscious unease and self-handicapping: Behavioral consequences of individual differences in implicit and explicit self-esteem. *Psychological Science*, *10*, 535–539.