

---

## CURRENT RESEARCH IN SOCIAL PSYCHOLOGY

---

<http://www.uiowa.edu/~grpproc/crisp/crisp.html>

Volume 11, Number 2

Submitted: August 25, 2005

First Revision: October 4, 2005

Accepted: October 31, 2005

Published: October 31, 2005

### **GENDER AND SOCIALLY DESIRABLE RESPONDING AS MODERATORS OF THE CORRELATION BETWEEN IMPLICIT AND EXPLICIT SELF-ESTEEM**

Michael Riketta

University of Tübingen, Germany

#### **ABSTRACT**

*Pelham, Koole, Hardin, Hetts, Seah, and DeHart (2005) found that implicit and explicit self-esteem correlated more strongly among women than men. The goal of this study was to replicate this finding and test whether it was due to gender differences in socially desirable responding (SDR). Ninety-nine German students completed measures of implicit self-esteem (Implicit Association Test and name-letter technique), explicit self-esteem, and SDR. Contrary to the Pelham et al. study, the implicit-explicit correlation was stronger for men than women. SDR (especially its self-deception component) moderated the implicit-explicit correlation independently from gender, albeit in different directions for the two implicit measures.*

## INTRODUCTION

Self-esteem is defined as a global attitude toward the self (Rosenberg, 1965). The common measures of this construct require explicit self-judgments (e.g., responses to items such as "I am satisfied with myself" "I have a positive attitude toward myself"; for an overview of measures, see Blascovich & Tomaka, 1991). Since the 1990s, more and more researchers have used implicit (i.e., unobtrusive) measures of self-esteem instead of, or in combination with, these explicit ones. One motivation for using of implicit measures grows from the assumption that they tap a more spontaneous and intuitive form of self-esteem than do explicit measures (e.g., Bosson, Brown, Zeigler-Hill, & Swann, 2003; Epstein & Morling, 1995; Jordan, Spencer, Zanna, 2003a; Jordan, Spencer, Zanna, Hoshino-Browne, & Correll, 2003b; Koole, Dijksterhuis, & van Knippenberg, 2001; Pelham, Koole, Hardin, Hetts, Seah, & DeHart, 2005; Spalding & Hardin, 1999).

This assumption rests on dual process models of information processing (e.g., Epstein, 1994; Smith & DeCoster, 2001; Wilson, Lindsey, & Schooler, 2000). According to these models, people can process information in both an automatic, intuitive manner and a controlled, more deliberate manner. A common assumption is that these two forms of information processing occur in parallel. Applying these models to self-evaluation, several researchers assumed that people hold two types of global attitudes toward themselves (e.g., Epstein & Morling, 1995; Glen & Banse, 2004; Hetts & Pelham, 2001; Koole et al., 2001; Spalding & Hardin, 1999). These constructs are typically called implicit self-esteem, which is assumed to result from automatic, intuitive processes, and explicit self-esteem, which is assumed to result from controlled, more deliberate processes.

Recently, Pelham et al. (2005) proposed that gender moderates the correlation between implicit and explicit self-esteem (implicit-explicit correlation). They assumed that women generally rely more strongly on their intuition than do men. Pelham et al. concluded that women may express their implicit self-esteem on explicit measures more readily than do men, resulting in a stronger implicit-explicit correlation. To support this argument, Pelham et al. reported results from six samples, in four of which the implicit-explicit correlation was at least marginally significantly stronger among women than among men; in the other two samples, no significant difference emerged. The four samples in which the hypothesis was confirmed consisted of students from the US, Singapore, and the Netherlands (total N = 1264).

However, this pattern of findings does not seem to be robust. Greenwald and Farnham (2000, Exp. 1) found in a sample of American undergraduates that the implicit-explicit correlation was higher among men than women (N = 138).

One reason for this inconsistency may be located at the measurement level. While both studies used the same explicit self-esteem measure (Rosenberg, 1965), they used different implicit measures: the name-letter technique (Nuttin, 1985) and a measure based on word completions (Pelham et al., 1999) in the Pelham et al. study and the self-esteem version of the Implicit Association Test (IAT; see Greenwald et al., 1998, for the original version) in the Greenwald and Farnham (2000) study. All these measures were design to measure implicit self-esteem. However, previous research has found only weak correlations between them at best (e.g., Bosson

et al., 2000; Jordan et al., 2003a). Thus, the implicit measures used in those studies may have tapped different facets of implicit self-esteem. The relations of these facets to explicit self-esteem may have been differentially moderated by gender.

Another reason for the inconsistency between the two studies refers to the mechanism mediating between gender and the implicit-explicit correlation. A candidate mediator is socially desirable responding (SDR). This is one of the few personality variables that has been shown to moderate the correlation between implicit and explicit measures (referring to other attitudes than self-esteem; Nosek, in press; Hofmann, Gschwendner, & Schmitt, 2005). Moreover, men and women have been shown to differ in SDR (e.g., Ones & Visweswaran, 1998; Paulhus, 1991). Thus, any observed gender differences in the implicit-explicit correlation may trace back to gender differences in SDR.

The main goal of this research was to test this mediator hypothesis. Before describing the empirical study, I will define the construct of SDR and discuss more in detail how it may relate to gender and the implicit-explicit correlation.

### **SDR and Its Relation to Implicit and Explicit Self-Esteem**

In an influential model of SDR, Paulhus (e.g., 1991) distinguished between two components of this construct: (a) impression management, which is a "explicit self-presentation toward an audience" (p. 37) with the goal to appear dependable and conventional, and (b) self-deceptive positivity (hereafter called self-deception), which is "the tendency to give self-reports that are honest but positively biased" (p. 37). The present study takes up this approach and looks at these components separately. They may have different relations to implicit and explicit self-esteem and thus may have different effects on the implicit-explicit correlation.

Because impression management is by definition a deliberate process, it is likely to affect responses to implicit (vs. explicit) measures less strongly. Thus, implicit measures may relate less strongly to indicators of impression management than do explicit measures (for this reasoning, see, e.g., Fazio & Olson, 2003; Greenwald & Farnham, 2000; Hetts & Pelham, 2001; Nosek, in press). Greenwald and Farnham (2000) provided some evidence of this, using the IAT as an implicit self-esteem measure.

The relation of self-deception to either explicit or implicit self-esteem is less clear on theoretical grounds. On the one hand, researchers argued that implicit self-esteem is independent of self-deception (e.g., Epstein & Morling, 1995; Jordan et al., 2003b). On the other hand, Paulhus (1991) assumed that self-deception is an automatic process. Thus, it may influence implicit (vs. explicit) self-esteem even more strongly. At the operational level, these two views suggest that implicit (vs. explicit) measures may relate either more weakly or more strongly to self-deception. Only Greenwald and Farnham (2000) have reported relevant data. They found the IAT (vs. several explicit self-esteem measures) to relate more weakly to self-deception. This supports the former theoretical position.

Depending on their relative correlations with implicit and explicit self-esteem, both components of SDR may moderate the implicit-explicit correlation. As argued above, impression

management is likely to influence responses to explicit versus implicit measures more strongly. Thus, responses to explicit versus implicit measures should become increasingly disassociated when impression management tendencies increase (for this reasoning, see Fazio & Olson, 2003; Hofmann et al., 2005; Nosek, in press). This hypothesis has already received support. For example, across 57 attitude objects and over 6,000 participants, Nosek (in press) found a weaker correlation between the IAT and corresponding explicit measures for participants scoring high (vs. low) on a composite measure of self-presentation. Likewise, looking at minorities in Germany (East Germans, Turks) as attitude objects, Hofmann et al. (2005) found a weaker correlation between the IAT and corresponding explicit measures for participants scoring high (vs. low) on two measures of self-presentation (public self-consciousness, motivation to control prejudice).

Because no unequivocal predictions are possible concerning the relation of implicit self-esteem to self-deception, its moderating effect on the implicit-explicit correlation cannot be predicted. Both a positive and a negative moderator effect are conceivable. The implicit-explicit correlation would be weaker under high (vs. low) self-deception if self-deception either affects implicit self-esteem more strongly than explicit self-esteem or vice versa. In this case, strong self-deception should weaken the implicit-explicit correlation (Nosek, in press). However, if self-deception affects implicit and explicit self-esteem equally strongly, the implicit-explicit correlation should be stronger under high (vs. low) self-deception because in this case self-deception increases the variance shared by implicit and explicit self-esteem measures.

### **Gender Differences in SDR and Their Effects on the Implicit-Explicit Correlation**

In a meta-analysis of 66 studies, Ones and Visweswaran (1998) found that men show stronger SDR (as measured with common self-report scales) than do women. Studies that distinguished between the two aforementioned components of SDR, using Paulhus' (1988, 1991) Balanced Inventory of Desirable Responding (BDIR), suggested that this tendency is mainly due to the self-deception component of SDR. For example, Quinn (1988, cited in Paulhus, 1991) found stronger self-deception but weaker self-presentation among male adults compared with female adults ( $N = 884$ ). Paulhus (1988, cited in Paulhus, 1991) ( $N = 433$ ) replicated this pattern among college students. In three German samples (total  $N = 351$ , mostly students), Musch, Brockhaus, and Bröder (2002) found stronger self-deception among men than among women in two samples and no gender differences for impression management. These studies suggest that men show stronger self-deception than women but either weaker or the same level of impression management.

In connection with the above reasoning, this suggests a complex relationship of gender with the implicit-explicit correlation. If impression management is on average weaker among men than women, as suggested by the data of Paulhus (1988) and Quinn (1988), this will work toward a higher implicit-explicit correlation among men. The reason is that impression management may lower the implicit-explicit correlation, as explained above.

Further, if self-deception is on average stronger among men than women, as suggested by the data of Musch et al. (2002), Paulhus (1988), and Quinn (1988), this will work toward either a stronger or a weaker implicit-explicit correlation among men compared with women.

Specifically, the correlation will be weaker among men if self-deception affects implicit self-esteem more strongly than explicit self-esteem or vice versa. In this case, those with the stronger self-deception (prospectively, men) will display a weaker correlation. However, if self-deception affects implicit and explicit self-esteem equally strong, the correlation will be stronger among those with the stronger self-deception (presumably, men).

Thus, both impression management and self-deception may mediate gender differences in the implicit-explicit correlation. Depending on the relative intraindividual strength of these components of SDR and on their relations to gender and implicit and explicit self-esteem, gender may have either positive or negative moderator effects on the implicit-explicit correlation. This reasoning can explain both the gender difference reported by Pelham et al. (2005) and the opposite findings reported by Greenwald and Farnham (2000). Specifically, the differences between these studies may have been due to differences involving impression management or self-deception. Prerequisite is that impression management or self-deception indeed mediates gender differences in the implicit-explicit correlation.

Greenwald and Farnham (2000) already tentatively suggested that impression management may be a mediator of the gender differences they observed. They dismissed this interpretation, however, because impression management was unrelated to most of their explicit measures. Because other studies did find relations of explicit self-esteem measures to SDR (e.g., Astra & Singg, 2000; Riketta, 2004), it still appears promising to test for the mentioned mediator effect.

Drawing on Pelham et al. (2005) and Greenwald and Farnham (2000), this study examines once more whether gender moderated the implicit-explicit correlation. Extending prior research, the present study explores whether SDR (i.e., impression management and/or self-deception) mediates this gender effect. This is the primary goal of this study. Moreover, this is the first study to compare this gender difference on the two most common implicit measures of self-esteem: the name-letter technique and the IAT. Pelham et al. and Greenwald and Farnham used only one of these measures each. Thus, as a secondary goal, this study explores whether the relation of gender to the implicit-explicit correlation depends on the implicit measure used.

## **METHOD**

### **Sample**

Participants were 99 students (30 male, 69 female) of various majors (57 psychology, 39 others, 3 missing) at the University of Tübingen, Germany. They were recruited by announcements in psychology classes and by ads on the campus. All participants took part in an unrelated study after the present one in the same session. They received either course credit or 12 Euros for the two studies together. [1]

### **Materials and Procedure**

Entering the laboratory, participants were greeted by a male experimenter and seated in separate cubicles in front of computers. Participants first worked on the self-esteem IAT, which was closely modelled after Jordan et al.'s (2003a) adaptation of the original procedure by Greenwald

and Farnham (2000). Participants were told that they would see words appearing on the screen. Their task was to classify the words as quickly and accurately as possible by pressing one of two keys. Participants learned that some stimuli had to be classified as self versus non-self while others had to be classified as positive versus negative. Five blocks were conducted. The stimuli appeared in an individual random order in each block. Each stimulus appeared on the screen until the participant pressed a response key. The intertrial interval was 150 ms.

In the first block, evaluative words had to be classified. Ten positive words (Frieden, Sommer, Urlaub, Genuss, Geschenk, Freude, Freiheit, Sonne, Glück, Lächeln; translation: peace, summer, holiday, enjoyment, gift, pleasure, freedom, sun, luck/happiness, smile) and ten negative words (Unfall, Bombe, Hölle, Gewalt, Schmerz, Virus, Trauer, Hass, Verbrechen, Koma; translation: accident, bomb, hell, violence, pain, virus, sadness, hatred, crime, coma) were the stimuli. They were presented once each, resulting in 20 trials. In the second block, self- and non-self-related words had to be classified. Five self-related words (ich, mich, mir, mein, meines; translation: I, me, to me, my, mine) and five non-self-related words (dass, und, wobei, als, wenn; translation: that, and, in/at which, as, when) were the stimuli. They were presented twice each, resulting in 20 trials. In the third block, both tasks were combined. The self- and non-self-related and evaluative words from the preceding blocks appeared in 60 trials. The first 20 trials were practice trials, in which the ten self- and non-self-related stimuli and five positive and five negative stimuli appeared. Participants had to respond self or positive with one key and non-self or negative with the other key. If the response was incorrect, a feedback message appeared after the trial ("You pressed the wrong key") and stayed on the screen until participants pressed the space bar. After the practice trials, participants worked on a block with 40 trials. Each evaluative word appeared once, and each self- and non-self-related word appeared twice. In the fourth block, the same evaluative words as in the first block appeared but the response keys were reversed this time (20 trials). In the fifth block, the same stimuli as in the third block appeared. This time participants had to respond self or negative with one key and non-self or positive with the other key. The first 20 of the 60 trials were practice trials of the same type as in the third block.

The rationale of the IAT is as follows. To the extent that the net (i.e., positive minus negative) valence associated with the self-related stimuli is more positive than the net valence associated with the non-self-related stimuli, interference effects are expected to occur in the fifth block so that reaction times are higher than in the third block. Therefore, the difference in reaction times (for correct responses) between the two blocks is the indicator of self-esteem. The scoring of the IAT in this study followed the recommendations by Greenwald et al. (2003). In essence, the difference in reaction times for correct responses between the fifth and third block (including the practice trials) was divided by the pooled standard deviation. The resulting IAT score is equivalent to Cohen's *d*. In theory, a higher score indicates higher implicit self-esteem.

The present version of the IAT differs from the two versions used by Greenwald and Farnham (2000) in the control words for the self-related words (meaningless non-self words rather than words referring to other people; see Jordan et al., 2003a) and in the scoring procedure. If anything, however, the present version may be more valid than the version used by Greenwald and Farnham. For one, Karpinski (2004) showed that the use of other people as control for the self-related words results in IAT scores that confound self-evaluation with the evaluation of other people. The use of meaningless control words may avoid this confounding. Further, the

scoring algorithm used here (Greenwald et al., 2003) has been shown to reduce the confounding of IAT scores with general task-switching ability (Mierke & Klauer, 2004).

Next, participants received a paper-pencil version of the name letter technique. This version was closely modelled after Koole et al.'s (2001) adaptation of Nuttin's (1985) original procedure. Participants were told that (a) the goal of the questionnaire was to analyze aesthetic judgments of letters and (b) previous studies had shown that such analyses may lead to innovative insights into human emotion. Participants were instructed to indicate their liking of the letters implicitly and intuitively. A list with the 30 letters of the German alphabet in random order followed. Each letter had to be evaluated on a 9-point rating scale anchored with "don't like it at all" and "like it very much". Finally, participants indicated their initials on the questionnaire. When participants had left the laboratory (having completed the measures described below), the experimenter indicated which letters were included in the first name and the surname of the participants. Participants' names were taken from a list on which they had entered their name to confirm the receipt of their compensation for participation.

The rationale of this measure is the mere ownership effect, which denotes the process by which objects associated with the self share the (usually positive) evaluation of the self (Greenwald & Banaji, 1995). Thus, one's self-esteem may color the evaluation of one's name and, as a consequence, of the single letters of one's name. It follows that the evaluation of these letters may be an unobtrusive measure of self-esteem. In the present study, responses on this measure were analyzed with the procedure of Koole et al. (2001). At first, a baseline evaluation for each letter was computed by averaging ratings across all participants who did not have the respective letter in their names. Next, the baseline was subtracted from the evaluation of each name letter of each participant. Two indicators were computed from these difference scores, one by averaging them across all name letters of each participant and the second one by averaging them across the initials of each participant. These two scores (hereafter referred to as name-letter preferences and initials preferences, respectively) indicate the degree to which participants evaluated their name letters and initials overly positively, respectively. In theory, the higher this evaluation, the higher implicit self-esteem. Previous research with the name-letter technique had usually used only one of these two indicators. Thus, this is one of the few studies in which they were compared within the same sample (e.g., like Bosson et al., 2000). Pelham et al. (2005) had used name-letter preferences in one sample and a combination of initials preferences and a related indicator (based on evaluation of participants' own birthday numbers) in another one of the four samples in which they found a gender effect on the implicit-explicit correlation.

Immediately after the name-letter task, participants completed a battery of measures on the computer. Here only those measures are described which are relevant to the present study. Responses were taken on 7-point rating scales ranging from "does not apply at all" to "totally applies". Individual scores were computed by averaging across items.

Impression management and self-deception were measured with a validated German adaptation (Musch et al., 2002) of Paulhus' (1988, 1991) BIDR. The latter is the standard instrument for measuring impression management and self-deception as separate components. The impression management scale of the German version consists of 10 items referring to socially desirable overt behavior (e.g., "I never swear" "I sometimes tell lies if I have to" [reverse coded]), similar

to classic lie scales. The self-deception scale of the German version consists of 10 items referring to socially desirable private thoughts or feelings (e.g., "It is hard for me to shut off a disturbing thought" [reverse-coded], "I am a completely rational person"). All items are translations from the original BIDR. High scores on these scales indicate the tendency to claim that one always behaves in a socially desirable manner and possesses positive cognitive attributes, respectively (i.e., rationality and overconfidence in one's judgments; see also Paulhus, 1991). In three studies, Musch et al. demonstrated sufficient internal consistency (alphas  $> .60$ ) and the convergent and divergent validity of each scale referring to standard measures of personality. One of these studies also showed that only the impression management scale, not the self-deception scale, is sensitive to manipulations of impression management motives (instruction to fake good vs. bad). Thus, these German short forms seem similarly valid as the original version of the BIDR (e.g., see Paulhus, 1991). In the present study, Cronbach alpha was .63 for impression management and .57 for self-deception.

Explicit self-esteem was measured with the Self-Liking Self-Competence Scale (Tafarodi & Swann, 1995) in German translation. It consists of two 10-item subscales, one referring to self-liking (e.g., "I feel comfortable about myself" "I tend to devalue myself" [reverse-coded]) and one referring to self-competence (e.g., "I am a capable person" "I don't succeed at much" [reverse-coded]). Although designed to measure two distinct forms of global self-esteem, the two subscales are usually strongly correlated. This was also true in the present study ( $r = .70$ ). Therefore and because the distinction between self-liking and self-competence is not of interest for the present study, the responses were averaged across all items into the explicit measure of self-esteem (Cronbach alpha = .93). In general, the items of the Self-Liking Self-Competence are similar in wording to the common Rosenberg (1965) Self-Esteem Scale; not surprising, both scales have been shown to correlate very strongly (83% shared variance in a sample of 1648 students, Tafarodi & Milne, 2001). The Rosenberg scale was the explicit self-esteem measure used in the studies by Pelham et al. (2005) and Greenwald and Farnham (2000).

## RESULTS

Appendix A shows the descriptive statistics of the analyzed variables. To explore whether gender moderated the implicit-explicit correlation, I conducted one regression analysis for each implicit measure, with explicit self-esteem as criterion and the implicit measure, gender (dummy-coded and centered, with higher scores for women), and the Implicit Measure x Gender interaction as predictors (see Regression 1 in Appendix B). Here and in the following regression analyses, the criterion and all predictors (except gender and the interaction terms) were z standardized to facilitate the interpretation of the B coefficients and to have centered predictors. The latter is advisable to reduce multicollinearity problems. The interaction terms were computed from the centered predictors.

The interaction was significantly negative for initials preferences ( $B = -0.51$ ,  $SE = 0.22$ ,  $p = .02$ ), indicating that the relation of this implicit measure to explicit self-esteem was more positive for men than for women. The direction of this interaction is contrary to the one that Pelham et al. (2005) observed for this implicit measure. The gender interaction was nonsignificant for name-letter preferences ( $B = -0.28$ ,  $SE = 0.21$ ,  $p = .18$ ) and the IAT ( $B = 0.22$ ,  $SE = 0.22$ ,  $p = .33$ ).



Thus, the negative interaction that Greenwald and Farnham (2000) observed for the IAT did not replicate here.

The following analyses were intended to explore the reason for the significant gender effect on the correlation between initials preferences and the explicit measure. A test for mediated moderation was conducted (Baron & Kenny, 1986). As argued above, SDR (i.e., impression management, self-deception, or both) may be a mediator of the gender effect. This argument would be supported if (a) gender related significantly to SDR, (b) SDR significantly moderated the implicit-explicit correlation when the moderating effect of gender was controlled, and (c) the moderating effect of gender was significantly weaker when the moderating effect of SDR was controlled rather than not controlled (cf. Baron & Kenny, 1986). This was tested here separately for the two components of SDR (impression management and self-deception).

As to (a): Men scored higher than women on self-deception,  $M = 4.20$ ,  $SD = 0.83$  vs.  $M = 3.88$ ,  $SD = 0.70$ ,  $t(97) = 2.01$ ,  $p = .05$ , but did not significantly differ on impression management,  $M = 3.07$ ,  $SD = 0.95$  vs.  $M = 3.22$ ,  $SD = 0.87$ ,  $t(97) = 0.78$ ,  $p = .44$ . This pattern replicates the results of Musch et al. (2002), who used the same measures of SDR. Thus, only self-deception met condition (a) and qualified as a candidate mediator of the moderating effect of gender on the correlation between initials preferences and explicit self-esteem. Steps (b) and (c) were therefore tested only for self-deception, not for impression management.

As to (b) and (c): Self-deception and the Initials Preferences X Self-Deception interaction were entered into the described regression analysis that included initials preferences, gender, and the Initials Preferences X Gender interaction as predictors (see Regression 2 in Appendix B). The Initials Preferences X Self-Deception interaction was marginally significantly positive ( $B = 0.16$ ,  $SE = 0.09$ ,  $p = .08$ ). The Initials Preferences X Gender interaction was reduced to marginal significance ( $B = -0.39$ ,  $SE = 0.20$ ,  $p = .06$ ). However, this reduction failed to reach significance (according to the Sobel test after Baron & Kenny, 1986;  $z = 1.36$ ,  $p = .17$ ). Thus, only condition (b) but not (c) was met. Although self-deception tended to moderate the implicit-explicit correlation, it did not significantly mediate the moderating effect of gender.

In ancillary analyses, the moderating role of SDR in the implicit-explicit correlation was examined more comprehensively, independent from gender and for all three implicit measures as well as for both components of SDR. To test whether impression management moderated the implicit-explicit correlation, I conducted one regression analysis for each implicit measure, with explicit self-esteem as criterion and impression management, implicit self-esteem, and the Impression Management X Implicit Self-Esteem interaction term as predictors (see Regression 3 in Appendix B). As explained above, a negative interaction was expected (e.g., Fazio & Olson, 2003; Nosek, in press), which would indicate a weaker correlation for higher impression management. However, the interaction was marginally significantly positive for name-letter preferences ( $B = 0.15$ ,  $SE = 0.09$ ,  $p = .10$ ), nonsignificantly positive for initials preferences ( $B = 0.15$ ,  $SE = 0.10$ ,  $p = .12$ ), and around zero for the IAT ( $B = 0.01$ ,  $SE = 0.12$ ,  $p = .91$ ).

The moderating role of self-deception was explored in an analogous series of regression analyses (for detailed results, see Regression 4 in the Appendix). For initials preferences, similarly to the regression analyses that included gender, the interaction was significantly positive ( $B = 0.20$ ,  $SE$

= 0.09,  $p = .02$ ), indicating that the relation of initials preferences to explicit self-esteem was more positive under high self-deception. The opposite was true for the IAT ( $B = -0.24$ ,  $SE = 0.10$ ,  $p = .03$ ). The interaction was nonsignificant for name-letter preferences ( $B = -0.07$ ,  $SE = 0.08$ ,  $p = .40$ ).

## DISCUSSION

Pelham et al. (2005) observed stronger implicit-explicit correlations for women than for men in two samples using the name-letter technique. The present study revealed the opposite: The relation between initials preferences and explicit self-esteem was stronger for men than for women. Thus, the gender differences found here resemble those that Greenwald and Farnham (2000) obtained for the IAT. Yet, the present findings do not confirm this study either because the correlation between IAT and explicit self-esteem did not differ significantly between men and women here. In any case, Pelham et al.'s (2005) conclusion that the implicit-explicit correlation is generally stronger among women than among men has to be qualified.

SDR did not significantly mediate the observed moderating effect of gender on the correlation between initials preferences and explicit self-esteem. One reason may be the rather low reliability of measures of the two components of SDR (alphas around .6). Thus, further tests for this mediation effect, using more reliable measures, appear worthwhile.

A remarkable incidental finding was that self-deception, as one component of SDR, moderated the implicit-explicit correlation independently from gender. This moderator effect had the opposite direction for the IAT versus initials preferences. Based on the zero-order correlations observed in this study, one would expect that self-deception weakens the implicit-explicit correlations because self-deception related more strongly to the explicit measure than to the implicit measures. Thus, higher self-deception should reduce the common variance of implicit and explicit measure, resulting in a negative moderator effect (for this reasoning, see Nosek, in press). This was confirmed only for the IAT, whereas the opposite moderator effect was found for initials preferences. In light of this reasoning, the moderator effect for the IAT appears more reasonable than the one for initials preferences. Moreover, the moderator effect for the IAT is consistent with the findings of Nosek (in press) and Hofmann et al. (2005), which refer to non-self-related attitudes. Thus, the IAT might have been more valid as an indicator of implicit self-esteem than the name-letter technique in this study.

Clearly, the latter conclusion is only tentative. At least, the present findings have to be viewed in the broader context of previous validation studies on the self-esteem IAT and the name-letter technique. The results of these studies are inconsistent. While some attested to the validity of these measures (e.g., Baccus, Baldwin, & Parker, 2004; Greenwald & Farnham, 2000; Koole et al., 2001), others failed to find evidence of their validity (e.g., Bosson et al., 2000; Glen & Banse, 2004; Greenwald & Farnham, 2000; Jones et al., 2004; Jordan et al., 2003a; Schimmack & Diener, 2003). The few studies that compared these measures in the same sample also yielded inconsistent findings as to their comparative and absolute validity (e.g., Bosson et al., 2000; Jordan et al., 2003a). In this context, it would be premature to declare the IAT or the name-letter technique the best operationalization of implicit self-esteem or even a sufficiently valid one.

Much remains to be learned about what these measures measure under what circumstances (for similar conclusions, see Bosson et al., 2000; Fazio & Olson, 2003; Hetts & Pelham, 2001).

## **REFERENCES**

Astra, R. L. & Singg, S. (2000). The role of self-esteem in affiliation. *Journal of Social Psychology*, 134, 15-22.

Baccus, J. R., Baldwin, M. W. & Parker, D. J. (2004). Increasing implicit self-esteem through classical conditioning. *Psychological Science*, 15, 498-503.

Baron, R. & Kenny D. (1986). The moderator-mediator variable distinction in social psychological research. *Journal of Personality and Social Psychology*, 51, 1173-1182.

Becker, T. E. (2005). Potential problems in the statistical control of variables in organizational research. *Organizational Research Methods*, 8, 274-289.

Blascovich, J., & Tomaka, J. (1991). Measures of self-esteem. In J. P. Robinson, P. R. Shaver & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 115-160). San Diego, CA: Academic Press.

Bosson, J. K., Brown, R. P., Zeigler-Hill, V. & Swann, W. B. (2003). Self-enhancement tendencies among people with high explicit self-esteem: The moderating role of implicit self-esteem. *Self and Identity*, 2, 169-187.

Bosson, J. K., Swann, W. B., & Pennebaker, J. W. (2000). Stalking the perfect measure of implicit self-esteem. *Journal of Personality and Social Psychology*, 79, 631-643.

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

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

Fazio, R.H., & Olson, M.A. (2003). Implicit measures in social cognition research. *Annual Review of Psychology*, 54, 297-327

Glen, I.S., & Banse, R. (2004). Probing the malleability of implicit and explicit self-esteem: An interview approach. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition*, 22, 133-158.

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.

Higgins, E.T., Shah, J.Y., and Friedman, R. (1997). Emotional responses to goal attainment: Strength of regulatory focus as moderator. *Journal of Personality and Social Psychology*, 72, 515-525.

Hofmann, W., Gschwendner, T., & Schmitt, M. (2005). On implicit-explicit consistency: The moderating role of individual differences in awareness and adjustment. *European Journal of Personality*, 19, 25-49.

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

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

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

Karpinski, A. (2004). Measuring Self-Esteem using the Implicit Association Test: The Role of the Other. *Personality and Social Psychology Bulletin*, 30, 22-34.

Kitayama, S. & Uchida, Y. (2003). Explicit self-criticism and implicit self-regard: Evaluating self and friend in two cultures. *Journal of Experimental Social Psychology*, 29, 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.

Mierke, J., & Klauer, K. C. (2003). Method-Specific Variance in the Implicit Association Test. *Journal of Personality and Social Psychology*, 85(6), 1180-1192.

Musch, J., Brockhaus, R., & Bröder (2002). Ein Inventar zur Erfassung von zwei Faktoren sozialer Erwünschtheit [An inventory for the assessment of two factors of social desirability]. *Diagnostica*, 48, 121-129

Nosek, B. A. (in press). Moderators of the relationship between implicit and explicit evaluation. *Journal of Experimental Psychology: General*.

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

Ones, D.S., & Visweswaran, C. (1998). The effects of social desirability and faking on personality and integrity assessment for personnel selection. *Human Performance*, 11, 245-269.

Paulhus, D. L. (1988). Assessing self-deception and impression management in self-reports: The Balanced Inventory of Socially desirable Responding. Unpublished manual, University of British Columbia, Vancouver, Canada.

Paulhus, D. L. (1991). Measurement and control of response bias. In P. Robinson, P. R. Shaver & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 17-59). San Diego, CA: Academic.

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.

Quinn, B.A. (1989). Religiousness and psychological well-being. Unpublished dissertation, Wayne State University, Detroit.

Riketta, M. (2004). Does social desirability inflate the correlation between self-esteem and anxiety? *Psychological Reports*, 94, 1232-1234.

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

Schimmack, U., & Diener, E. (2003). Predictive validity of explicit and implicit self-esteem for subjective well-being. *Journal of Research in Personality*, 37, 100-106.

Smith, E. R. & DeCoster, J. (2001). Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. *Personality and Social Psychology Review*, 4, 108-131.

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

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

Tafarodi, R. W. & Milne, A.B. (2001). Decomposing global self-esteem. *Journal of Personality*, 70, 443-483.

Wilson, T. D., Lindsey, S. & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review*, 107, 101-126.

**APPENDIX A. DESCRIPTIVE STATISTICS**

Variable	Scale Range	M	SD	Correlations						
				1.	2.	3.	4.	5.	6.	
1. Name-letter preferences	-8: 8	0.44	1.03							
2. Initials preferences	-8: 8	1.27	1.54	.59***						
3. IAT		0.85[a]	0.26	.12	.09					
4. Explicit self-esteem	1; 7	5.25	0.95	.17	.31**	.13				
5. Impression management	1; 7	3.17	0.90	.17	.17	.08	.01			
6. Self-deception	1; 7	3.97	0.75	.07	.15	.15	.45***	.16		
7. Female gender				-.05	.00	.05	.02	.08	-.20*	

[a] Equivalent to Cohen's d, with higher scores denoting higher self-esteem. \* p <= .05. \*\* p < .01. \*\*\* p < .001.

**APPENDIX B. REGRESSION RESULTS (CRITERION: EXPLICIT SELF-ESTEEM)**

Predictors	Measure of ISW											
	IAT				Initials preferences				Name-letter preferences			
	B	SE	Rsq	p	B	SE	Rsq	p	B	SE	Rsq	p
Regression 1			.03	.47			.14	.00			.05	.20
Gender	0.05	0.22		.82	0.05	0.21		.79	0.10	0.22		.65
ISW	0.13	0.10		.22	0.34	0.10		.00	0.14	0.10		.16
ISW X Gender	0.22	0.22		.33	-0.51	0.22		.02	-0.28	0.21		.18
Regression 2							.34	.00				
Gender					0.24	0.19		.20				
ISW					0.28	0.09		.00				
ISW X Gender					-0.39	0.20		.06				
SD					0.43	0.09		.00				
ISW X SD					0.16	0.09		.08				
Regression 3			.01	.67			.12	.01			.06	.15
IM	0.00	0.10		.98	-0.05	0.10		.63	-0.03	0.10		.79
ISW	0.13	0.10		.22	0.34	0.10		.00	0.16	0.10		.12
ISW X IM	0.01	0.12		.91	0.15	0.10		.12	0.15	0.09		.10
Regression 4			.25	.00			.30	.00				
SD	0.42	0.09		.00	0.41	0.09		.00	0.45	0.09		.00
ISW	0.01	0.09		.92	0.26	0.09		.00	0.16	0.10		.09
ISW X SD	-0.24	0.10		.03	0.20	0.09		.02	-0.07	0.08		.40

Notes. Rsq: R-squared. ISW: implicit self-esteem. SD: self-deception. IM: impression management.

## **ENDNOTES**

[1] One might argue that the type of compensation (pay vs. course credit) may have affected the other variables of interest, especially social desirability. (I am grateful to two anonymous reviewers for suggesting this possibility.) To test this, for one, I computed the zero-order correlations of compensation (coded 0=course credit and 1=pay;  $n = 31$  and  $68$ , respectively) with all other variables of interest (self-esteem, IAT, initials preferences, name-letter preferences, self-deception, impression management, and gender). These correlations were nonsignificant, with the highest correlations being those for the IAT ( $r = .12$ ,  $p = .23$ ) and gender ( $r = -.11$ ,  $p = .26$ ; all other  $|r|s < .06$ ,  $ps > .64$ ). Further, I entered compensation (dummy-coded as above) as an additional predictor in all regression analyses that are reported in the text. The results were virtually the same; in particular, by including compensation in the analyses, no B coefficient changed by more than  $0.02$  and no significant result became nonsignificant or vice versa. Thus, as recommended by Becker (2005), only the regression analyses without this control variable are reported in the text.

## **AUTHOR BIOGRAPHY**

Michael Riketta is Assistant Professor of Social and Personality Psychology at the University of Tübingen, Germany. His current research interests cover social identity in organizations, the role of the self-concept in social perception, and multidimensional models of self-esteem. Email: michael.riketta@uni-tuebingen.de