

Running head: MASKED AFFECTIVE PRIMING BY NAME LETTERS

Masked affective priming by name letters:
Evidence for a correspondence of explicit and implicit self-esteem

Dirk Wentura¹, Michael Kulfanek², and Werner Greve³

¹ Saarland University, Saarbrücken, Germany

² University of Münster, Germany

³ University of Hildesheim, Germany

Abstract

In an affective priming experiment (N=60), participants' initials were used as masked primes. We found a significant priming effect (denoting positivity of the personal initials) if personal initials (compared to yoked participants' initials) were followed by other-relevant positive or negative targets (e.g., honest, cruel). For possessor-relevant targets (e.g., healthy, lonely), there was no main effect of priming. However, this priming index positively correlated with explicit self-esteem. For participants with high self esteem, a positive effect was found, whereas for participants with low self-esteem, a negative effect was found.

Masked affective priming by name letters: Evidence for a correspondence of explicit and implicit self-esteem

In recent years, social cognition research has been dominated by a strong interest in the indirect measurement of attitudes (e.g., Greenwald & Banaji, 1995). To assess a pure and unbiased evaluation of attitude objects, experimental techniques such as affective priming (Fazio, Sanbonmatsu, Powell, & Kardes, 1986; Fazio, Jackson, Dunton & Williams, 1995; Otten & Wentura, 1999), the Implicit Association Test (Greenwald, McGhee, & Schwartz, 1998; Greenwald & Nosek, 2001), or the affective Simon-task (De Houwer & Eelen, 1998, De Houwer, in press) have been used (for a recent review see Fazio & Olson, 2003). There are three reasons. There are notable reasons for this. First, given the long-standing doubts on the validity of self-report measures, these newly developed techniques seem to open up a new avenue to the non-reactive measurement of attitudes. Second and more importantly, theorizing in social cognition (for an overview see, e.g., Kunda, 1999) is largely based on models of mental representation developed in cognitive psychology (e.g., semantic networks, parallel distributed models of memory etc.) and associated processes (e.g., spread of activation). From this point of view, it is more obvious to predict, e.g., response time differences in priming tasks because they are directly derivable from the basic assumptions of these models compared to the complex action of uttering a judgment. For instance, according to Fazio (e.g., 1990), a strong attitude is equivalent to an association between the attitude object and an evaluation. Presentation of an object-related stimulus (e.g., the name of a political party) will automatically activate the corresponding evaluation. This process can be directly tapped by the evaluation task developed by Fazio and colleagues (1986) if the object-related stimulus is presented briefly and very shortly before a clearly positive or negative target word whose valence has to be determined by a key-press. If the attitude towards the object is, e.g., positive, the automatically activated evaluation decreases the response time to

positive targets and increases it to negative targets, supposedly by a process of response competition (for recent reviews on affective priming see Fazio, 2001; Klauer, 1998; Klauer & Musch, 2003; see also Wentura & Rothermund, 2003).

Whereas the first reasons for the use of indirect measures seem to be widely shared by social cognition researchers, a third reason is a point of more debate. Sometimes the distinction between direct (i.e., self-report) measures and indirect measures of attitudes is coupled with the theoretical distinction of explicit attitudes and implicit attitudes, thereby following suggestions in the field of memory research (e.g., Greenwald & Banaji, 1995). According to this view, explicit attitudes are consciously held and are consciously reportable; implicit attitudes are defined at the level of memory representation (e.g., as an associative link between the representation of the attitude object and an evaluation) and need not be consciously identifiable. Therefore only a modest overlap of variance between indirect and direct measures of attitudes will be expected.

One area of application for indirect measures of attitudes has been self-related affect, i.e., self-esteem (e.g., Greenwald & Farnham, 2000; Farnham, Greenwald, & Banaji, 1999; Hetts, Sakuma, & Pelham, 1999; Spalding & Hardin, 1999). Bosson, Swann, and Pennebaker (2000) recently compared different indirect measures for assessing implicit self-esteem. By and large, the findings were discouraging: Low or virtually zero internal consistencies were found, convergent validity between indirect measures was low, correlations with explicit measures were mostly non-significant¹, they had no predictive validity.

Among the indirect measures of Bosson and colleagues (2000) were two versions of the affective priming task, one with supraliminally presented primes, one with subliminally presented primes. The results for these measures were particularly discouraging. Some correlations were even counter-intuitive. The authors concluded “if researchers ... 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” (p. 642). The present study using the affective priming task for measuring implicit self-esteem might potentially be seen as a step in the direction of improvement. However, because our approach was developed independently of the study of Bosson et al., we will first give our report on the methods and results without reference to those of Bosson et al. Then, we will compare both approaches in the discussion.

Overview

In an affective priming experiment (see Fazio et al., 1986), participants’ initials were used as primes. There were two hypotheses associated with this study. First, following research on the name-letter effect (Nuttin, 1985; see also Jones, Pelham, Mirenberg, & Hetts, 2002; Koole, & Pelham, 2003; Pelham, Mirenberg, & Jones, 2002), for the average participant his/her personal initials should have a positive valence. Thus, priming with one’s own initials should facilitate responses to positive targets and inhibit responses to negative targets compared to an adequate control. Second, despite several failures to find substantial correlations between indirect measures and self-reports (e.g., Bosson et al., 2000; Farnham et al., 1999, Jordan, Spencer, Zanna, Hoshino-Browne & Corell, 2003), we hold on to the hypothesis that this effect should be dependent on interindividual differences in (explicitly measured) self-esteem. With the second hypothesis we do not want to say that indirect and direct (explicit) measures are necessarily based on the same latent variable (i.e., that random error variance is the only difference). However, it is our belief that the automatic self-evaluation that is presumably tapped by the priming task is at minimum a biasing process component that is at work while answering items on a traditional self-esteem measure. We will elaborate on this in the discussion.

For an adequate control, a yoked design was used. Each experimental session was run with two participants such that priming with personal initials can be compared to priming

with the yoked partner's initials. To prevent suspicion and strategic behavior, primes were masked, i.e., they were presented very briefly (i.e., 43 ms), and were overwritten by a row of @. Additionally, we used the response window technique introduced by Draine and Greenwald (1998). Briefly, with this procedure participants are forced to give categorization responses to target stimuli within a time span that is too short to achieve high levels of accuracy. Thus, priming processes show up predominantly in the error rates. Error rates are lower in affectively congruent than in affectively incongruent conditions. Draine and Greenwald (1998) found larger effect sizes with this technique compared to standard procedures. This technique has been successfully used for the implicit measurement of self-related attitudes (Otten & Wentura, 1999) and consumer attitudes (Frings & Wentura, 2003).

Target words in our experiment were clearly positive and negative adjectives. However, Wentura, Rothermund, and Bak (2000) have shown that automatic evaluation might be dependent on a second factor, which is orthogonal to valence. That is, the trait adjectives can be subtyped according to the *kind* of positivity or negativity, which can be termed possessor- vs. other-relevance (Peeters, 1983; Peeters & Czapinski, 1990)². The evaluation of a given trait depends on the perspective of the evaluators — whether they evaluate the trait from the perspective of someone who has to interact with the trait-holder or from the perspective of the trait-holder him/herself. For example, being brutal is primarily bad for the social environment of the brutal person, but not necessarily for the one who is brutal (he might be a sadist). To the contrary, being lonely is primarily bad for those who are depressive, but not necessarily for their social environment (the lonely person might be thankful for contact). The same applies to positive adjectives: To be honest is primarily good for those who interact with the honest person but not necessarily for themselves (the honesty might be abused), whereas intelligence is primarily good for the intelligent persons themselves but not necessarily for the social environment (she or he might be a “bad guy” who abuses his intelligence). To sum up,

possessor-relevant characteristics can be described as primarily having positive or negative implications for the self whereas other-relevant traits can be described as primarily having positive or negative implications for those who have to interact with the trait holder.

Adjectives such as *brutal* or *honest* are called other-relevant, whereas words such as *lonely* or *intelligent* are called possessor-relevant.

From a self-esteem perspective, it seems evident that interindividual differences in explicit self-esteem are more closely linked to possessor-relevant valence than to other-relevant traits. Whereas the evaluation of the self for *high* self-esteem participants is rather unambiguously positive in the light of this distinction, the valence of self for *low* self-esteem participants is clearly ambiguous: On the other-relevant valence dimension, even persons with low self-esteem will see themselves as positive, because they pose no threat to others and they do no harm to others. But on the possessor-relevant valence dimension, persons with low self-esteem have a negative self, because they see themselves as unsuccessful and worthless. Thus, it can be hypothesized that especially priming effects with initials as primes and possessor-relevant adjectives as targets will hinge on explicit self-esteem.

To account for interindividual differences, we measured explicit self-esteem by the Rosenberg-scale (Rosenberg, 1965). Because this measure has variance overlap with momentary mood and non-self-related negative feelings and cognitions, an adjective mood scale (Dalbert, 1992) and the hopelessness scale (Beck, Weissman, Lester, & Trexler, 1974) were added.

Method

Participants

Sixty students (45 women; 15 men) participated in the experiment in partial fulfillment of course requirements or were paid five €. The median age was 23 years (range: 19 to 47 years). All participants were native speakers of German.

Design

The priming manipulation comprised the factorial combination of target valence (positive vs. negative), target type (other-relevant vs. possessor-relevant), and the initials' prime condition (personal initials vs. yoked participants' initials). This 2x2x2 within-subjects design allows the calculation of two priming indices as the dominant dependent variables of the study: Mean error rate for positive targets following the personal initials and mean error rate for negative targets following the yoked participants' initials will be subtracted from mean error rate for negative targets following personal initials and mean error rate for positive targets following yoked participants' initials, separately for other-relevant and possessor-relevant targets. A third prime condition was added as a possible alternative control (instead of yoked partner initials; see *Materials*). Two more priming conditions were added to establish reference priming effects, that is, a fourth prime was the word *Liebe* (*love*; hereafter called *a priori* positive prime) and a fifth prime was the word *Krieg* (*war*; here after called *a priori* negative prime).

There were three blocks. In each block, each target stimulus (see *Materials*) was presented five times to a participant, once in each priming condition. The sequence of priming was counterbalanced in a Latin-square design with five groups of participants and five sets of stimuli. For the assessment of interindividual differences, we used the self-esteem scale (Rosenberg, 1965; German version by Ferring & Filipp, 1996), an adjective mood scale (modified version of Dalbert, 1992), and the hopelessness scale (Beck et al. 1974, German version Krampen, 1994).

Materials

The target set comprised 10 positively and 10 negatively valenced German adjectives with a length of five to eight letters. Adjectives were selected on the basis of their pleasantness values (absolute values of 49 and more on a scale ranging from -100 to +100)

according to a norm list of 908 adjectives (Hager, Mecklenbräuker, Möller, & Westermann, 1985, and Möller & Hager, 1991). Within each valence set, five adjectives were other-relevant and five adjectives possessor-relevant according to norm data (Wentura, Rothermund, & Bak, 1998; see appendix for a list of the materials used). There are a few slight differences between other-relevant and self-relevant negative words with regard to pleasantness values (see Appendix). Other-relevant words have somewhat higher values than possessor-relevant ones.³ For the a priori positive and negative priming condition the words *Liebe* (love) and *Krieg* (war) were used, respectively.

The initials' primes were two-letter strings (in upper cases), consisting of the first letter of the first name and the first letter of the surname.⁴ The alternative control (see *Design*) was a two-letter string with letters randomly selected while disregarding all letters that appeared in either the participant's (full) name or the yoked partner's (full) name.

Procedure

All participants were tested pairwise. They were seated in front of IBM-compatible personal computers, separated by partition walls. To obtain their full names, they acknowledged receipt of money (or receipt of the course requirement form) in advance. Then, they filled out the adjective mood scale. Unobserved by the participants, the experimenter typed the full names of both participants into a third computer that was linked to the two experimental computers. Three letter strings (personal initials, yoked partner's initials, alternative control string) were generated and transferred to the experimental computers.

Next, instructions for the priming task were given on the CRT screen. Participants were told that they had to classify common words with regard to their valence. The response-window priming task roughly followed Draine and Greenwald (1998; see also Frings & Wentura, 2003; Otten & Wentura, 1999). The sequence of events was as follows: A string of seven @ marked the beginning of the trial and also served as a forward mask, which remained

on the screen for 300 ms and was immediately overwritten by the prime on position 4 and 5. Thus, as the prime was always shorter than the forward mask, it was embedded into the string of @s (e.g., @@@DW@@). The prime remained on the screen for 43 ms (three refresh cycles of the screen) and was immediately replaced by a string of seven @ that served as a backward mask that was displayed for 14 ms (one cycle). The subjective impression of the presentation sequence was a brief flicker of a senseless string that was completely replaced by the target stimulus. Finally, the target word was replaced after 300 ms by an exclamation mark that remained on the screen for 450 ms. Onset of the exclamation point defined the beginning of the response window. The window width was 150 ms. If the response fell within the range defined by the response window (i.e., from 300 ms to 450 ms after onset of the target), the exclamation point changed colors (from white to red), giving feedback to the participant that the response was in time. If it did not change the color, the participant knew that the response had been too slow. If the response was given too quickly, the target was erased immediately and the exclamation point never appeared. Participants were instructed to press the correct key (i.e., the right key for positive targets and the left key for negative targets) within the response window. The inter-trial interval was 2500 ms following the response.

At the beginning of the priming task, participants worked through 20 practice trials (i.e., each target word was presented once) in order to become familiarized with the task, and to receive additional feedback if their responses were either too fast or too slow. After the practice trials, a summarized feedback was given, indicating the percentage of correct trials, median RT, percentage of trials with an RT falling within the response window, and percentage of trials with a response that was too fast. The following message was added: "Your goal should be to maximize the rate of responses within the 'time window.' Simultaneously, the rate of correct responses should be about 70 % to 80 %."

The main part of the experiment consisted of three blocks of 100 trials each. Within a block, each target word was presented once in each of the five priming conditions, with the sequence determined by a Latin-square design (see *Design*). Each block was further subdivided into five 20-trials sequences. Each 20-trials sequence comprised each target word once, and was followed by the summarized feedback (see above). The response window was adjusted contingent on the participant's accuracy. That is, if the error rate of a 20-trial block was above 45 % and the RT median was 100 ms or more above the center of the response window, the onset of the response window was increased by 33 ms. If, however, the error rate was below 20 % and the RT median was below the center of the response window plus 100 ms, the beginning of the response window was decreased by 33 ms. The onset of the response window varied between 168 ms and 333 ms across participants and blocks.

Following the three blocks, a direct measure of prime valence recognition was administered. Forty more trials (eight trials per prime) were presented with a row of nine question marks instead of a target word. Before these trials, participants were informed that the sequence of flickers contained a word and they were instructed to categorize these words with regard to valence. Then, the participants filled out the self-esteem scale and the hopelessness scale.

Results

Priming effects

Mean error rates were $M = 28.65\%$ ($SD = 10.56\%$), $M = 29.56\%$ ($SD = 11.52\%$), and $M = 31.08\%$ ($SD = 11.15\%$) for the first, second, and third block, respectively. Discarding trials with response times below 100 ms and above 1000 ms (0.6 % of all trials), priming differences were computed separately (a) for the priming conditions with initials and the a priori priming conditions, (b) for other-relevant targets and possessor-relevant targets, and (c) blocks. A priming difference is defined by subtracting the mean error rate (in percent) for

congruent prime-target pairs (i.e., positive targets following either *love* or own initials or negative targets following either *war* or yoked participants' initials) from the mean error rate for incongruent pairs (i.e., negative targets following either *love* or own initials and positive targets following either *war* or yoked participants' initials). A criterion of significance of $\alpha = 5\%$ was adopted for all analyses.

A 2 (prime type: *a priori* vs. initials) x 2 (target type: other-relevant vs. possessor-relevant) x 3 (block) analysis of variance of priming differences yielded a significant constant effect, that is, the overall priming difference was significantly above zero, $F(1,59) = 7.69, p < .01$. This effect was not significantly moderated by any of the factors or any interaction of factors, all F 's < 1.55 besides the main effect of prime type, $F(1,59) = 2.81, p = .10$. This last result provided neither clear-cut evidence for the claim that the two types of primes (i.e., *a priori* valent vs. initials) do make a difference nor that they do not. In fact, whereas the *a priori* priming effect was significant for other-relevant targets, $t(59) = 2.17, p < .05$, and possessor-relevant targets, $t(59) = 2.26, p < .05$, the effect of initials was only significant for other-relevant targets, $t(59) = 1.70, p < .05$ (one-sided), but not for possessor-relevant targets, $t(59) = -0.25, n.s.$ (see Figure 1).

Individual Differences

We observed a significant correlation of $r = .28, p < .05$, between the self-esteem scale and the priming effect for initials and possessor-relevant targets. The corresponding correlation for other-relevant targets was non-significant, $r = .04, n.s.$ However, this differentiation between possessor-relevant targets and other-relevant targets is underestimated by these analyses. There are two details about the data that should be taken into consideration. First, whereas for the interpretation of mean priming differences (as was the case in the preceding paragraph of the *Results* section) the yoked participants' initials served as a perfect control, it turned out that for the analyses of covariances the alternative control letters (see

Materials) worked slightly better.⁵ Therefore, we calculated a new priming measure using the alternative control for the succeeding analyses. Second, the direct measure of self-esteem correlated highly with hopelessness and moderately with mood which was not the case for the implicit measure (see Table 1). Therefore we calculated the residuals variable of explicit self-esteem, controlling for hopelessness and mood.

As can be easily seen from Table 1, the implicit measure for possessor-relevant targets correlates fairly substantially with explicit self-esteem, especially if the latter is controlled for hopelessness and mood. The corresponding correlations of explicit self-esteem with the priming measure for other-relevant targets were $r = -.14$, *n.s.*, and $r = .07$, *n.s.*, for the raw self-esteem scale and the residualized score, respectively. To substantiate the differentiation between possessor-relevance and other-relevance, we correlated the *difference* between the initials' priming measures for possessor-relevant and other-relevant targets with the explicit self-esteem measures, $r = .39$, $p < .003$, and $r = .31$, $p < .02$, for the raw self-esteem scale and the residualized score, respectively. (Note, that these analyses correspond to the interaction test in a 2 [self-esteem: low vs. high] x 2 [target type: other-relevant vs. possessor-relevant] analysis of variance of initials' priming differences if explicit self-esteem would have been dichotomized by a median-split.)

Regressing the initials' priming effect for possessor-relevant targets on *z*-standardized explicit self-esteem yielded predicted values with a 90%-interval of confidence not including zero for explicit self-esteem values below $z = -.46$ and above $z = 1.10$. That is, for those with an explicit self-esteem of approximately a half standard deviation unit (or more) below the mean, a significant (with $p < .05$, one-tailed) negative implicit self-esteem can be expected. Likewise, for those having an explicit self-esteem of approximately one standard deviation unit (or more) above the mean, a significant positive implicit self-esteem is predicted.

Finally, we have to acknowledge that according to diagnostic standards, the reliability of the implicit measure is very low (see Table 1).⁶ As expected, the *a priori* priming indices did not correlate with any of the questionnaire measures ($-.16 < r < .14$).

Direct effects

In the last phase of the experiment, participants were urged to categorize masked primes as positive or negative, although they reported no conscious impression of primes. For some participants, the parametric signal detection sensitivity measure d' could not be calculated because their hit rate or false alarm rate was either zero or one. Therefore the non-parametric equivalent A' (see Pollack, 1970; Pollack & Norman, 1964) was calculated. A' takes on values in the range from one (perfect discrimination) to zero (perfect discrimination, but reversed use of keys) with 0.5 denoting random responding. The signal detection sensitivity for the valence of the *a priori* positive and negative primes was $A' = .60$ ($SD = .23$), which is significantly above random responding, $t(59) = 3.40, p < .01$. Thus, to a certain extent participants were able to discriminate between a clearly positive (i.e., *love*) and a clearly negative (i.e., *war*) masked stimulus. However, A' does not moderate the relationship between implicit and explicit self-esteem, as indicated by a non-significant product term in a moderated regression with initial's priming effect (possessor-relevant targets) as criterion, self-esteem, A' , and the product term of self-esteem and A' as predictors. This holds for the overall self-esteem scale as well as for the residuals variable of explicit self-esteem, all (absolute) $t(56) < 0.88, n.s.$

An analogous calculation for the discernment of one's own initials vs. yoked partner's initials yielded a sensitivity of $A' = .48$ ($SD = .18$), which was not different from random responding, $t(59) = -0.71, n.s.$ Most importantly, this index does not correlate with any of the variables used for Table 1 (all absolute r 's $< .17, p > .20$).

Discussion

We found two essential results in this experiment: For other-relevant targets, the personal initials' prime was processed as if it was a positive stimulus compared to the control condition. This pattern of results is compatible to the view that the self has positive valence for the average participant ("self-positivity bias") if seen from the other-relevant perspective, or in other words: The self is seen as non-threatening to others and as morally positive.

For possessor-relevant targets, however, a comparable self-positivity bias was clearly missing. Most importantly, for these materials the priming difference correlated significantly with explicit self-esteem. This pattern of results is compatible to the view that individual differences in self-esteem correspond to the positivity or negativity of the self if seen from the possessor-relevant perspective. Why did we expect (and find) a positive correlation of moderate size, although others constantly failed to do so (Bosson et al., 2000; Farnham et al., 1999, Jordan et al. 2003)? We want to give two different potential answers: One by focusing on our implicit measure and its details and one by focusing on the process of responding to the self-report measure

The affective priming task

As mentioned in the introduction, Bosson et al. (2000) found results with a supraliminal as well as a subliminal variant of the affective priming task as tools for the assessment of self-esteem that were largely discouraging. Neither did they find a positivity bias for the average participant, nor a correlation with explicit self-esteem. We briefly want to discuss procedural differences that might be the cause for the differences in results. However, since the supraliminal task departs in almost all aspects from our variant (i.e, besides presentation of primes, level of aggregation, target set size, calculation of measure), turning it into a completely different task, we will focus the discussion on the subliminal task of Bosson and colleagues. Likewise, we do not discuss failures to find substantial correlations between self-report measures and the self-esteem IAT (Bosson et al., 2000; Farnham et al., 1999, Jordan et

al. 2003) because this would open up a completely different discussion (see, e.g., Rothermund & Wentura, 2004).

Choice of self-related primes. Bosson et al. (2000) used pronouns (*me, myself*) as self-related primes whereas we presented participants' initials. Though ultimately this will be a matter for empirical testing (see Spalding & Hardin, 1999, for positive evidence), it is conceivable that first-person pronouns flashed on the screen are not automatically regarded as self-relevant: Typically, a first-person pronoun that was heard or read just points to another person (i.e., the person I'm talking to, the protagonist of a modern novel etc.) and not to myself.

Choice of target words. Bosson et al. (2000) used six pairs of evaluatively loaded antonyms (*good-bad, love-hate, nice-mean, winner-loser, superior-inferior, fine-poor*) that are a mixture of other-relevant, possessor-relevant, and other words that are not related to this distinction. Although usually the type of target was not commonly a point of concern in affective priming research, the present results indicate subtle differences with regard to the other-possessor distinction.

Procedural details. Bosson and colleagues (2000) used the subliminal task of Spalding and Hardin (1999) which has slightly different parameters of presentation than ours (e.g., 17 ms prime presentation [compared to 43 ms], followed by a 17 ms blank screen before masking [compared to direct masking]; SOA of 184 ms [compared to 57 ms]). Most importantly, they did not use a response window technique. Our own experiments (Wentura & Degner, 2004; see also Wentura & Rothermund, 2003) show that masked affective priming effects can be particularly observed with speed instructions. Given accuracy instructions, affective priming effects were diminished or were even negative. Without using strict speed instructions (e.g., with a response window) one cannot be sure whether some participants

were guided by emphasizing speed (thereby sacrificing accuracy) or by emphasizing accuracy (thereby sacrificing speed) with uncontrollable implications for the priming effect.

Self-report measures of self-esteem

There are two different perspectives on self-report measures. First, we can argue that for some variables – amongst them self-esteem – the first-person perspective is privileged (Greve, 1996): Who knows better than I how I feel about myself? In this regard, self-report measures reflect conscious and deliberately reasoned evaluations of the self. Seen from this perspective (and ignoring problems of strategic self-presentations in certain contexts that distort validity), self-reports are the indispensable standard for the assessment of self-esteem. Then, we can ask whether there are correspondences between variations in self-reported self-esteem and variations on a micro-behavioral level. The affective priming paradigm is one paradigm that taps this level. Of course, it is an empirical question whether there is indeed a correspondence between variations in self-reported self-esteem and variance in a certain response-time based paradigm: We might be mistaken about the underlying processes, we might use wrong procedural parameters etc. However, despite failures it remains a plausible assumption that there must be correspondences between the “phenomenological level” (i.e., self-reported self-esteem) and the “functional” level (i.e., the level that is presumably tapped by response-time based paradigms).

Alternatively, we can view self-reports as complex judgmental behavior. For example, to respond to items such as “I feel I do not have much to be proud of,” participants have to try to retrieve instances of “being proud of something” from memory. The character of retrieved instances and the easiness of retrieval, for example, will determine the answer. The psychology of judgment has identified many components of this process. At least some of them will validly reflect stable differences in self-esteem: For example, if we assume that a person high in self-esteem will more often experience situations of being proud, such that

more instances of being proud will come to mind by cueing with the item, a great portion of variance in responses to that item will reflect valid self-esteem differences. How do we construct the correspondence of self-reported self-esteem and implicit self-esteem from this viewpoint? There are several possibilities. For example, a strong association of the self with a positive evaluation (which is presumably assessed by the affective priming paradigm) might be a result of lots of experiences of feeling positive, worthy, proud of oneself etc. Insofar responses on self-report items are driven by the memory of those situations, a correspondence of explicit and implicit measures is expected.

However, we might even give the automatic evaluation component a more prominent role in the process of answering self-report items. A judgment often requires a comparison between a target and a reference standard. For example, the Rosenberg scale includes two items that explicitly refer to a comparison between the self and other people (“I am able to do things as well as most other people.” “I feel that I'm a person of worth, at least on an equal plane with others.”). Other items (e.g., “At times I think I am no good at all”) invite participants to compare themselves to others. According to a recent theoretical account of comparison processes (Mussweiler, 2001, 2003), a comparison process starts either by a similarity hypothesis or a dissimilarity hypothesis. A similarity hypothesis is supposed to increase the accessibility of attributes on which the self and the reference standard are similar, thereby yielding an assimilation effect (i.e., higher self-esteem values with regard to the Rosenberg items). Starting by a dissimilarity hypothesis, attributes are activated on which the self and the standard are dissimilar, thereby yielding a contrast effect (i.e., lower self-esteem values with regard to the Rosenberg items). We might speculate whether the type of the *a priori* hypothesis (i.e., whether the comparison process starts either by a similarity hypothesis – “I’m at least as good as other people” – or a dissimilarity hypothesis – “I’m not as good as

other people”) is determined by the sign – i.e., the positivity or negativity – of the automatic evaluation (assessed by the affective priming paradigm).⁷

Note that this reasoning allows for interesting follow-up hypotheses. For example, it is compatible with the assumption that high explicit self-esteem can assume qualitatively different forms (see, e.g., Jordan et al., 2003). According to this rationale, some people possess secure high self-esteem, whereas others possess defensive high self-esteem, that is, self-views that are fragile and vulnerable to threat. Jones and colleagues assume defensive high self-esteem individuals to be characterized by low implicit self-esteem and high explicit self-esteem. If those individuals experience sometimes their level of implicit self-esteem directly, they respond by self-image maintenance techniques. Using our speculation given above would mean that defensive high self-esteem individuals experience the *a priori* dissimilarity hypothesis (e.g., “I am *not* able to do things as well as most other people”) that was instigated by the automatic evaluation component. However, one defensive manoeuvre might be to actively counteract the automatic tendency to (negatively) contrast the self with others in responding to the Rosenberg items such that high self-esteem ratings are the consequence. It will be worthwhile to follow this line of reasoning in future research.

An interesting alternative explanation for our finding of a substantial correlation between implicit and explicit self-esteem comes from the study by Jones and colleagues (2002) who examined the relation between explicit self-esteem and an implicit measure (i.e., ratings of own initials) dependent on the presence or absence of threat. They found a positive correlation between explicit and implicit self-esteem only if the implicit measure was administered after a mild self-threat condition. Thus, insofar as our priming task with the experience of repeated failure is considered to be a self-threat condition, the relationship between implicit and explicit self-esteem might not reflect stable dispositions but – as Jones et al. have argued – self-regulation processes (see Dodgson & Wood, 1998, for similar

findings). However, in their study explicit reference to the self was made before the implicit measure was administered: Firstly, participants filled out the explicit self-esteem measure, followed, secondly, by the self-threat manipulation which consisted in writing some lines about negative aspects of themselves. That was not the case in our study. Moreover, our priming measure concealed the reference to the participants' self more than a direct rating of letters might do.

A final remark with regard to the self-report measure: Explicit self-esteem is highly correlated with mood and well-being (i.e., negatively correlated with hopelessness). Though the pattern of correlations fits some obvious expectations (i.e., on average low self-esteem is expected to be accompanied by dampened mood and low expectations with regard to the future) the size of the correlations reveals a problem of some explicit measures: Common variance might be inflated by a lack of discriminant validity. The same does not apply to the implicit self-esteem measure.

To give an outlook on future research with the masked affective priming task, the divergence of effects for other-relevant and possessor-relevant evaluation is in our view the most promising finding from the results. With regard to implicit self-esteem, it has to be shown that there is validity in the affective priming task and the other-relevant and possessor-relevant distinction beyond the correlation with explicit self-esteem. Moreover, there is a potential to examine the structure of other implicit attitudes. That is, it might be hypothesized that, for example, implicit prejudice is not equivalent to an association of a social category with an *undifferentiated* negative evaluation. Instead, it might be the case that on the level of automatic activated evaluations, depreciation (e.g., prejudice against elderly people) and imputed hostility (e.g., prejudice against ethnic minorities) might be differentiated. Possibly, a version of the affective priming technique is suited to capture this differentiation. Of course, there are two caveats: First, the distinction between possessor- and other-relevance is slightly

confounded with extremity. However, in other studies (Wentura et al., 2000) it could be shown that this confound does not play a causal role. Second, we have to add that this differentiation hypothesis is not easily deducible from theories on affective priming. It has been shown that priming effects in the evaluation task presumably do not reflect semantic associations but response competition processes (e.g., Klauer, Roßnagel, & Musch, 1997; Wentura, 1999, Fazio, 2001). The categorization task, however, only demands a simple positive or negative response, such that obtaining a differentiation seems to be implausible. However, our experiments (Wentura & Degner, 2004; see also Wentura & Rothermund, 2003) have shown that affective priming effects are sometimes susceptible to type of valence (i.e., whether target and prime are both possessor-relevant or other-relevant). In sum, we see a great potential for the masked affective priming task to capture different types of automatic evaluation.

References

- Beck, A. T., Weissman, A., Lester, D., & Trexler, L. (1974). The measurement of pessimism: The hopelessness scale. *Journal of Consulting and Clinical Psychology, 42*, 861-865.
- Bosson, J. K., Swann, W. B. Jr., & Pennebaker, J. W. (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.
- Dalbert, C. (1992). Subjektives Wohlbefinden junger Erwachsener: Theoretische und empirische Analysen der Struktur und Stabilität [Subjective well-being of young adults: theoretical and empirical analyses of structure and stability]. *Zeitschrift für Differentielle und Diagnostische Psychologie, 13*, 207-220.
- De Houwer, J. (in press). The extrinsic affective Simon task. *Experimental Psychology*.
- De Houwer, J., & Eelen, P. (1998). An affective variant of the Simon paradigm. *Cognition and Emotion, 8*, 45-61.
- Dodgson, P. G., & Wood, J. V. (1998). Self-esteem and the cognitive accessibility of strengths and weaknesses after failure. *Journal of Personality and Social Psychology, 75*, 178-197.
- Draine, S. C., & Greenwald, A. G. (1998). Replicable unconscious semantic priming. *Journal of Experimental Psychology: General, 127*, 286-303.
- Farnham, S., Greenwald, A., & Banaji, M. (1999). Implicit Self-esteem. In D. Abrams & M. Hogg (Eds.), *Social identity and social cognition* (pp. 230-248). Oxford: Blackwell.
- Fazio, R. H. (1990). Multiple processes by which attitudes guide behavior: The MODE model as an integrative framework. In M. P. Zanna (Ed.), *Advances in experimental psychology* (Vol. 23, pp. 75-109). San Diego, CA: Academic Press.

Fazio, R. H. (2001). On the automatic activation of associated evaluations: An overview. *Cognition and Emotion, 15*, 115-141.

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 research: Their meaning and uses. *Annual Review of Psychology, 54*, 297- 327.

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.

Ferring, D., & Filipp, S.-H. (1996). Messung des Selbstwertgefühls: Befunde zu Reliabilität, Validität und Stabilität der Rosenberg-Skala. Skala [Measurement of self-esteem: Reliability, validity, and stability of the Rosenberg-scale]. *Diagnostica, 42*, 284-292.

Frings, C., & Wentura, D. (2003). Who is watching "Big Brother"? TV consumption predicted by masked affective Priming. *European Journal of Social Psychology, 33*, 779-791

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. *Journal of Personality and Social Psychology, 79*, 1022-1038.

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.

Greenwald, A. G., & Nosek, B. A. (2001). Health of the Implicit Association Test at age 3. *Zeitschrift für Experimentelle Psychologie, 48*, 85-93.

Greve, W. (1996). Erkenne Dich selbst? Argumente zur Bedeutung der "Perspektive der ersten Person [Know thyself? Arguments on the relevance of the first-person perspective].

Sprache & Kognition, 15, 104-119.

Hager, W., Mecklenbräuker, S., Möller, H., & Westermann, R. (1985). Emotionsgehalt, Bildhaftigkeit, Konkretheit und Bedeutungshaltigkeit von 580 Adjektiven: Ein Beitrag zur Normierung und zur Prüfung einiger Zusammenhangshypothesen [Emotionality, imageability, concreteness, and meaningfulness of 580 adjectives: A contribution to norming and test of correlations]. *Archiv für Psychologie*, 137, 75-97.

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 high self-esteem. *Journal of Personality and Social Psychology*, 85, 969-978.

Klauer, K. C. (1998). Affective Priming. *European Review of Social Psychology*, 8, 67-103.

Klauer, K. C., & Musch, J. (2003). Affective priming: Findings and theories. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion*. Mahwah, NJ: Lawrence Erlbaum.

Koole, S. L., & Pelham, B. W. (2003). On the nature of implicit self-esteem: The case of the name letter effect. In S. J. Spencer, S. Fein, M. P. Zanna, & J. M. Olson (Eds.), *Motivated*

social perception: The ninth Ontario symposium (pp. 93-116). Mahwah, NJ: Lawrence Erlbaum.

Krampen, G. (1994). *Skalen zur Erfassung von Hoffnungslosigkeit (H-Skalen)* [Scales to measure hopelessness]. Göttingen, Germany: Hogrefe.

Kunda, Z. (1999). *Social cognition. Making sense of people*. Cambridge, MA: MIT Press.

Möller, H., & Hager, W. (1991). Angenehmheit, Bedeutungshaltigkeit, Bildhaftigkeit und Konkretheit-Abstraktheit von 452 Adjektiven [Pleasantness, meaningfulness, imageability and concreteness of 452 adjectives]. *Sprache und Kognition*, 10, 39-51.

Mussweiler, T. (2001). 'Seek and ye shall find': Antecedents of assimilation and contrast in social comparison. *European Journal of Social Psychology*, 31, 499-509.

Mussweiler, T. (2003). Comparison processes in social judgment: Mechanisms and consequences. *Psychological Review*, 110, 472-489.

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

Otten, S., & Wentura, D. (1999). About the impact of automaticity in the Minimal Group Paradigm: Evidence from affective priming tasks. *European Journal of Social Psychology*, 29, 1049-1071.

Peeters, G. (1983). Relational and informational patterns in social cognition. In W. Doise & S. Moscovici (Eds.), *Current Issues in European Social Psychology* (Vol. 1, pp. 201-237). Cambridge: Cambridge University Press.

Peeters, G., & Czapinski, J. (1990). Positive-negative asymmetry in evaluations: The distinction between affective and informational negativity effects. *European Review of Social Psychology*, 1, 33-60.

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

Pollack, I. (1970). A non-parametric procedure for evaluation of true and false positives. *Behavior Research Methods and Instrumentation, 2*, 155-156.

Pollack, I., & Norman, D. A. (1964). A non-parametric analysis of recognition experiments. *Psychonomic Science, 1*, 125-126.

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

Rothermund, K., & Wentura, D. (2004). Underlying processes in the Implicit Association Test (IAT): Dissociating salience from associations. *Journal of Experimental Psychology: General, 133*, 139-165.

Wentura, D. (1999). Activation and inhibition of affective information: Evidence for negative priming in the evaluation task. *Cognition and Emotion, 13*, 65-91.

Wentura, D. & Degner, J. (2005). *Automatic evaluation isn't that crude! Moderation of masked affective priming by type of valence*. Manuscript submitted for publication.

Wentura, D., & Rothermund, K. (2003). The “meddling-in” of affective information: A general model of automatic evaluation effects. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 51-86). Mahwah, NJ: Lawrence Erlbaum.

Wentura, D, Rothermund, K., & Bak, P. (1998). [Ratings for possessor- vs. other-relevance of adjectives]. *Unpublished data*.

Wentura, D, Rothermund, K., & Bak, P. (2000). Automatic vigilance: The attention-grabbing power of approach- and avoidance-related social information. *Journal of Personality and Social Psychology, 78*, 1024-1037.

Author Note

Dirk Wentura, Department of Psychology, University of Jena, Germany; Michael Kulfanek, University of Münster, Germany, Werner Greve, University of Hildesheim, Germany.

The research reported in this article was supported by a grant from the Deutsche Forschungsgemeinschaft (DFG; WE 2284/3).

Correspondence concerning this article should be addressed to Dirk Wentura, Department of Psychology, Friedrich Schiller-University Jena, Wildstr. 1, D-07743 Jena, Germany, e-mail: d.wentura@uni-jena.de

Footnotes

¹ This might be called a virtue of indirect measures if one holds a strong version of the distinction between direct and indirect measures. However, this would be only convincing if there were other virtues of indirect measures that will not not be found in direct ones (e.g., predictive validity).

² Actually, Peeters (1983) used the terms *other-profitability* and *self-profitability*. The positive connotation of *profitability* makes it difficult to grasp the meaning of *negative other-profitable* or *negative self-profitable*.

³ Because this presumably reflects a tendency to particularly devalue traits that transgress ethical standards, this result is not surprising. However, Wentura et al. (2000) have shown that their effects for the relevance distinction were not reducible to slight differences in pleasantness.

⁴ In Germany, the use of a middle initial is rather untypical.

⁵ This might be due to the fact that for some participants the yoked initials overlap in one letter. In fact, removing all participants with an overlap ($n=14$) increases the correlation between the self-esteem scale and the priming effect for initials and possessor-relevant targets to $r = .32, p < .05$.

⁶ Note, that calculating Cronbach's Alpha with the three blocks as "items" is roughly the same as the Spearman-Brown adjusted split-half reliability: Instead of projecting the correlation of two test halves – which is the reliability of the test halves – to the complete test, now the mean correlation of test thirds – which can be considered the mean reliability of the thirds – is projected to the complete test.

⁷ This *post hoc* reasoning is corroborated by the following *a posteriori* analysis. Items of the Rosenberg scale were rated (N=4) according to the degree of instigating a comparison process. Across items, the average rating significantly correlates with the item-implicit self-esteem correlation, $r = .78$ (Spearman-Rho; $p = .007$).

Table 1
Correlations and Reliabilities

Measure	ISE	ESE	ESEres	H	Mood
Implicit Self-Esteem (ISE)	.48	.34**	.41***	-.12	-.02
Explicit Self-Esteem (ESE)		.88	.59***	-.80***	.52***
ESE Residuals (ESEres)			-	.00	.00
Hopelessness (H)				.89	-.52***
Mood					.89

Note: Reliabilities in the diagonal cells (i.e., Cronbach's Alpha; for implicit self-esteem with blocks considered as items; see text for further explanation); implicit self-esteem is the priming difference for possessor-relevant targets, using the alternative control (see *Materials*); explicit self-esteem (residuals) is the Rosenberg Scale, residualized with regard to hopelessness and mood.

** $p < .01$ *** $p < .001$

Figure Captions

Figure 1. Mean priming differences for a priori primes and initials primes as function of target type.

Appendix

The Target Terms and Their Pleasantness Norm Values^a

Other-Relevant				Possessor-Relevant			
Negative		Positive		Negative		Positive	
grausam (cruel)	-84	gütig (kind)	51	einsam (lonely)	-60	begabt (gifted)	49
boshaft (malicious)	-72	human (humane)	57	lustlos (listless)	-52	heiter (serene)	56
gemein (mean)	-66	gerecht (just)	60	leiblos (lifeless)	-52	aktiv (active)	60
geizig (miserly)	-61	ehrlich (honest)	74	unfähig (incapable)	-50	schön (beautiful)	77
gierig (greedy)	-60	zärtlich (affectionate)	80	träge (sluggish)	-50	gesund (healthy)	79
Mean	-69		64		-53		64
SD	10		12		4		13

^aaccording to Hager, Mecklenbräuer, Möller, & Westermann (1985), Möller & Hager (1991).

