I Like Myself but I Don’t Know Why: Enhancing Implicit Self-Esteem by Subliminal Evaluative Conditioning

Ap Dijksterhuis
University of Amsterdam

On the basis of a conceptualization of implicit self-esteem as the implicit attitude toward the self, it was predicted that implicit self-esteem could be enhanced by subliminal evaluative conditioning. In 5 experiments, participants were repeatedly presented with trials in which the word I was paired with positive trait terms. Relative to control conditions, this procedure enhanced implicit self-esteem. The effects generalized across 3 measures of implicit self-esteem (Experiments 1–3). Furthermore, evaluative conditioning enhanced implicit self-esteem among people with low-temporal implicit self-esteem and among people with high-temporal implicit self-esteem (Experiment 4). In addition, it was shown that conditioning enhanced self-esteem to such an extent that it made participants insensitive to negative intelligence feedback (Experiments 5a and 5b). Various implications are discussed.

The fact is indubitable that one’s own children always pass for the prettiest and brightest, the wine from one’s own cellar for the best—at least for its price,—one’s own house and horses for the finest. (Horwicz, as cited in James, 1890, p. 326)

Our preference for our own children, wine, house, and for some of us, even our own horses, over those of others is rooted in a pervasive and universal human need: the need to feel good about oneself. The scientific investigation of this need or motive dates back to theorizing on self-esteem, and the extreme form it can take that James called “self-love” (1890, pp. 309–329). The attention the topic of self-esteem received from the scientific community was, and still is, enormous.

The need for self-esteem is often regarded as a core concern of humans, and various models and theories point out that high self-esteem benefits people in important ways (see Baumeister, 1998). Self-affirmation theory (Steele, 1988) posits that high self-esteem is a buffer against stress and experiences of failures. Likewise, Taylor and Brown (1988) have argued that a positive view of the self promotes happiness and mental health in general. Terror-management theory (e.g., Greenberg et al., 1992) proposes that high self-esteem helps people to deal with existential threats. In addition, research on the sociometer model (e.g., Leary & Baumeister, 2000) shows that high self-esteem is a sign of social acceptance and liking. In general, a moderate or high level of self-esteem seems to be a prerequisite for healthy human functioning. It should come as no surprise, then, that the vast majority of people indeed view themselves rather positively. That is, they have moderate or high self-esteem (Banaji & Prentice, 1994; Baumeister, 1998; Greenwald, 1980; Taylor & Brown, 1988).

Implicit Self-Esteem: Important, but Elusive

In the past 20 years, social psychology has witnessed an important shift. At first, social psychology’s core concepts were largely viewed as being the result of conscious processes. Today it is recognized that automatic or unconscious processes play a major role in almost all social psychological processes (Bargh, 1984; Bargh & Chartrand, 1999; Devine, 1989; Dijksterhuis & Bargh, 2001; Fazio, 1990; Greenwald & Banaji, 1995; Wegner & Bargh, 1998). The scientific investigation of self-esteem is no exception. Whereas self-esteem used to be seen as the result of conscious self-evaluative processes, much recent work has emphasized the role of unconscious or implicit self-evaluative processes (e.g., Brown, 1993; Greenwald & Banaji, 1995; Kitayama & Karasawa, 1997; Koole, Dijksterhuis, & van Knippenberg, 2001; Pelham, Mirenberg, & Jones, 2002). As a result, these days researchers differentiate between explicit self-esteem and implicit self-esteem. The former is based on conscious processes, whereas the latter is the result of automatic self-evaluative processes. It reflects unconscious associations with the self (Greenwald & Banaji, 1995; Koole & Pelham, in press).

Greenwald and Banaji (1995) treat implicit self-esteem as an attitudinal construct and define it as “the introspectively unidentified (or inaccurately identified) effect of the self-attitude on evaluation of self-associated and self-dissociated objects” (p. 11). According to this definition, self-esteem goes hand in hand with a positive evaluation of self-associated stimuli. As was previously suggested in the introductory quote, there is some evidence for this tendency to like self-associated objects. People like various mundane objects (such as pens or mugs) more as soon as they own them (Kahneman, Knetsch, & Thaler, 1990). People like people more when they are members of the same group or social category (Tajfel, 1970), and people display a preference for the letters in their own name over other letters and a preference for their birthday numbers over other numbers (Kitayama & Karasawa, 1997; Nuttin, 1987).

This research was supported by Nederlandse Organisatie voor Wetenschappelijk Onderzoek Grant 016.025.030 (“Vernieuwingsimpuls”) I thank Rifka van der Meer, Barbara Gradeneegger, Maaike Valk, Peter Laurentius, and Hugo Bulder for their help in this project and Brett Pelham and Pamela Smith for helpful comments on an earlier version of this article.

Correspondence concerning this article should be addressed to Ap Dijksterhuis, Social Psychology Program, University of Amsterdam, Roetersstraat 15, 1018 WB, Amsterdam, the Netherlands. E-mail: a.j.dijksterhuis@uva.nl
Despite the fact that the study of implicit self-esteem has only just begun, we already have good reasons to believe that implicit self-esteem is of considerable importance. Low-implicit self-esteem has negative consequences that are to a large extent comparable to those of low-explicit self-esteem. Just as threats to the self-concept lead to decreased explicit self-esteem, they also lead to lower implicit self-esteem (Jones, Pelham, Mirenberg, & Hetts, 2002; Koole, Smeets, van Knippenberg, & Dijksterhuis, 1999). Furthermore, the “buffer” function of self-esteem against threatening experiences has been demonstrated for implicit self-esteem. Relative to people with high-implicit self-esteem, people with low-implicit self-esteem show diminished levels of aspiration after failure (Greenwald & Farnham, 2000) and more anxiety during a very personal interview than high-implicit self-esteem individuals (Spalding & Hardin, 1999). It has been argued that in some specific cases the buffer function of implicit self-esteem could well be more important than that of explicit self-esteem (Hetts & Pelham, 2001). Existential threats have been shown to operate most strongly at unconscious levels (Arndt, Greenberg, Pyszczynski, & Solomon, 1997), rendering it likely that high-implicit self-esteem rather than high-explicit self-esteem is the main prerequisite for coping with such threats. In addition, Hetts and Pelham (2001) discussed preliminary evidence suggesting that high-implicit self-esteem may be more important for coping with stigmatization than high-explicit self-esteem.

Despite these encouraging findings, the study of implicit self-esteem is not without problems. Whereas it is known to some extent what implicit self-esteem does (see previous paragraphs), we do not yet quite know what it is. The major reason is that researchers have used many different measures of implicit self-esteem. Spalding and Hardin (1999) used a priming measure based on work by Fazio and colleagues (e.g., Fazio, Jackson, Dunton, & Williams, 1995), Greenwald and Farnham (2000) used a self-esteem Implicit Association Test (IAT), and various others (Koole et al., 2001; Pelham et al., 2002) used name-letter preferences (Nuttin, 1987). In a recent contribution, Bosson, Swann, and Pennebaker (2000) took a close look at various measures of implicit self-esteem and had to draw rather sobering conclusions. Most important, they showed that the seven measures they investigated did not correlate with each other (and, as an aside, not with measures of explicit self-esteem). In addition, they showed that the predictive validity of these measures was rather poor and that the test-retest reliability of only two out of the seven measures was acceptable (see also Greenwald & Farnham, 2000; Koole, Dijksterhuis, & van Knippenberg, 2001; Koole et al., 1999). The conclusions Bosson et al. drew were discouraging. The bottom line is that researchers are faced with a number of interesting and potentially important findings that are lumped together under the header of effects of implicit self-esteem, while at the same time implicit self-esteem remains very elusive at a conceptual level.

To the Core of the Phenomenon

The main goal of this article was to shed some more light on what implicit self-esteem is. My approach is aimed at trying to grasp and manipulate what I think constitutes the essence of implicit self-esteem. Because it has been convincingly shown that the many measures of implicit self-esteem are uncorrelated (or at least hardly correlated), it is worthwhile to pursue a different strategy. Rather than comparing measures, in this article I try to manipulate implicit self-esteem at a very basic level. With such a manipulation, I hope to achieve two goals. First, I want to show that such a manipulation will show effects on several (unrelated) measures of implicit self-esteem. The second goal is to investigate the effects of this same manipulation in an area that is traditionally associated with the study of self-esteem, namely the way people deal with (negative) personality or intelligence feedback.

The aim to manipulate implicit self-esteem was translated into an aim to enhance implicit self-esteem, because I deemed it ethically inappropriate to try to reduce implicit self-esteem with evaluative conditioning.1 Now, how can implicit self-esteem be enhanced? How can people be induced to like their name letters or their horses more? Greenwald and Banaji’s (1995) and others’ conceptualization of implicit self-esteem as an attitudinal construct opens an interesting possibility. Attitudes can be formed or changed in a basic structural fashion through evaluative conditioning (Levey & Martin, 1975; for a review, see also De Houwer, Thomas, & Baeyens, 2001). In evaluative-conditioning research, an attitude object (conditioned stimulus, or CS) is paired with a positively or negatively valenced stimulus (unconditioned stimulus, or US). After repeated pairings, the attitude object takes on the valence of the US. Pairing with a negative stimulus leads to a more negative attitude, whereas pairing with a positive stimulus leads to a more positive attitude. This effect has been shown numerous times, with a range of negative (a shock, a nasty odor, negative words or pictures) and positive (a nice odor, positive pictures or words, a free lunch) conditioning stimuli. Interestingly, recent evidence also shows that evaluative conditioning occurs even when the CS or US is presented subliminally (see De Houwer, Hendrickx, & Baeyens, 1997; Krosnick, Betz, Jussim, & Lynn, 1992; Niedenthal, 1990; see also Dijksterhuis, Aarts, & Smith, in press).

Can this knowledge be applied to improving self-esteem? As the vast majority of experiments in the evaluative-conditioning domain have been done with novel or neutral attitude objects, the self cannot easily be compared to the attitude objects typical for evaluative-conditioning work. After all, the self is neither novel nor neutral, and it is indeed an “object” toward which we already have an attitude. However, although most experiments in the realm of evaluative conditioning have used novel objects, there are some exceptions. For example, Cacioppo and colleagues (Cacioppo, Marshall-Godell, Tassinary, & Petty, 1992) attempted to condition attitudes toward both random letter strings and real words. They obtained effects of evaluative conditioning for both, the difference being that the effects were bigger for random letter strings. In addition, Stuart, Shimp, and Engle (1987) applied conditioning techniques to consumer attitudes, and they conditioned attitudes toward both new brands and existing brands (e.g., Coca Cola) in various experiments. They consistently found that attitudes toward both types of brands were affected by evaluative conditioning.

1 In Experiments 4 and 5, implicit self-esteem was reduced, but this was achieved by giving participants negative feedback. These effects are relatively easy to alleviate (e.g., by telling participants the feedback was false). It cannot be known for sure that the effects of conditioning are as easy to alleviate, therefore self-esteem was not lowered with this manipulation.
conditioning. As conditioning techniques have already been successfully applied to objects toward which we already have an attitude, it may be possible to condition self-esteem.

In summary, the knowledge obtained in the evaluative conditioning domain indicates that it is at least theoretically possible to enhance implicit self-esteem. If a representation of the self (such as the words I or me) is repeatedly paired with positively valenced stimuli (such as positive words), it should be possible to enhance implicit self-esteem. Furthermore, results from the evaluative-conditioning domain suggest that this can be achieved subliminally. Hence, it follows that self-esteem can be structurally changed implicit in a very direct way that does not require conscious intervention.

Overview of the Experiments

In all experiments, participants were repeatedly presented with the word I (the Dutch word ik) on a computer screen. Immediately after each presentation of I, a positive trait term (e.g., nice, smart, warm) was presented. This procedure, in which the US follows the CS (called forward-conditioning), is the procedure most commonly used in evaluative-conditioning research. With this procedure, the aim was to enhance implicit self-esteem, based on the hypothesis that the self would take on the valence of the positive trait terms. The number of pairings (15 or 16 in the present experiments) was based on earlier research on evaluative conditioning (see De Houwer et al., 2001).

Experiments 1–3

In the first three experiments, the hypothesis that evaluative conditioning could change implicit self-esteem was tested. The goal was to show effects of the manipulation on various measures of implicit self-esteem that were shown earlier to be uncorrelated (see Bosson et al., 2000). The two measures were chosen that according to the analysis of Bosson et al. (2000) are the most promising: initial preference and the self-esteem IAT (Greenwald & Farnham, 2000). Not only are these presumably the measures that are used most often but also they are the only measures tested by Bosson et al. (2000) that showed acceptable reliability (see also Greenwald & Farnham, 2000; Koole et al., 2001). In Experiment 1, the effect of the manipulation on preference for initials (and on name–letter liking in general) was tested. In Experiment 2, an alternative explanation of the manipulation was ruled out. In this experiment, implicit self-esteem was again assessed by measuring preferences for initials. In Experiment 3, the effects of the manipulation on the self-esteem IAT (Greenwald & Farnham, 2000) was tested.

Experiment 4

With Experiment 4, I wanted to shed light on the potency of the self-esteem manipulation. Whether evaluative conditioning could be applied to enhance self-esteem among both people whose self-esteem was made low (by negative intelligence feedback) and people whose self-esteem was made high (by positive intelligence feedback) was tested.

Experiments 5a and 5b

Experiments 5a and 5b were to some extent exploratory. They were designed to test some of the consequences of enhanced implicit self-esteem. Again, participants whose implicit self-esteem was enhanced were compared with control participants whose implicit self-esteem was not manipulated. In Experiment 5a, effects of negative intelligence feedback on mood were investigated, whereas in Experiment 5b effects of negative feedback on task persistence were investigated.

Experiment 1

Method

Participants and design. Seventy-eight undergraduate students (48 women and 30 men) from the University of Amsterdam participated in the experiment in return for course credits or for Dfl. 7.50 (approximately $4 US). All participants were randomly assigned to a conditioned self-esteem condition or a control condition.

Procedure and materials. Upon entering the laboratory, participants were greeted by a female experimenter and seated in an individual cubicle in front of an iMAC computer. The experimenter explained that the instructions would be given by the computer program. She started the computer, and left the cubicle.

The first stage was the conditioning stage. The task was a primed lexical decision task with 30 trials that were presented in random order. Each trial started with a row of Xs presented in the center of the computer screen for 500 ms. On 15 of the 30 trials, the row was immediately followed by the presentation of the word I (the Dutch word ik) for 17 ms (that is, one refresh rate on a 60 Hz screen). In the conditioned self-esteem condition, a positive trait word was presented immediately after the word I. Different traits were used for each of the 15 trials.2 In the control condition, the word I was always followed by a mundane, evaluatively neutral word (such as chair or bike). During the 15 remaining trials, the row of Xs was followed by the presentation of a single X for 17 ms, immediately followed by a random letter string. All words (and rows of Xs) were presented in black on a white computer screen; the font used was Chicago 14. The 15 random letter strings and the 15 positive words (conditioned self-esteem condition) or 15 neutral words (control condition) served as targets for the lexical decision task. Participants had to decide as fast as possible whether the target was a real word by pressing one of two keys. As soon as participants pressed a key, the target disappeared. After a 1-s delay, the next trial started.

After completion of the lexical decision task, the initial-preference task (IPT) was introduced. This procedure was developed by Nuttin (1987), who found that letters are evaluated more favorably by people with names that contain similar letters relative to people with names that do not. Moreover, this effect appears to be strongest for people’s initials. In recent years, the name–letter effect (NLE), or IPT, has often been used as a measure of implicit self-esteem (see Greenwald & Banaji, 1995; Jones et al., 2002; Kitayama & Karasawa, 1997; Koole et al., 2001; Koole & Pelham, in press). Participants were asked to evaluate the attractiveness of all 26 letters of the (Dutch) alphabet on a 7-point scale (1 = not at all beautiful, 7 = extremely beautiful). The letters were presented individually and in random order on a computer screen, and participants were asked to evaluate each letter by pressing the key corresponding to their evaluation. Finally, participants were debriefed carefully. They were asked whether they had seen anything unusual during the lexical decision task. No participant indicated having seen anything strange. Moreover, participants were asked whether they had seen words (other than the target words) flashing on the screen. Again, no participant had seen any flashes. After the

2 The trait terms (or traitlike terms) used were warm (warm), lief (sweet), aardig (nice), oprecht (sincere), eerlijk (honest), mooi (beautiful), vrolijk (cheerful), sliem (smart), sterk (strong), wijs (wise), gezond (healthy), leuk (funny), blij (happy), prettig (nice), and positief (positive).
debriefing, participants were asked to write their name on a list in return for course credits or money. Subsequently, they were thanked and dismissed.

**Results**

The NLE was calculated according to a procedure used earlier by Kitayama and Karasawa (1997; see also Koole et al., 2001). First, the baseline attractiveness of all letters was established by calculating the mean attractiveness of each letter for participants whose names did not contain this letter. Subsequently, for each participant the NLE was calculated by computing the difference between their own evaluation of their name letters and the baseline evaluation of these same letters. Two scores were calculated. The first score was based on participants’ initials only; the second score was based on participants’ full names.  

The means are listed in Table 1. As can be seen, the NLE was more pronounced for participants in the conditioned self-esteem condition compared with participants in the control condition. For both initials and full names, this difference was reliable, $F(1, 76) = 5.15, p < .03$; and, $F(1, 76) = 4.47, p < .04$, respectively. Hence, the results confirmed the hypothesis that evaluative conditioning enhanced implicit self-esteem.

**Experiment 2**

The conclusion that the manipulation enhanced implicit self-esteem is somewhat premature, as the nature of the control condition allows for an alternative explanation. Note that participants in the conditioned self-esteem condition were presented with positive words, whereas participants in the control condition were presented with neutral words. It is possible that presentation of positive words enhanced participants’ mood rather than their self-esteem per se. A positive mood, in turn, could have led participants to evaluate letters more positively. This may have inflated their NLEs, because these effects were based on a baseline calculated for all (conditioned self-esteem and control) participants. Therefore, a different control condition was used in Experiment 2. All participants were presented with positive words, the difference being that the positive words in the control condition were not preceded by the word $I$.

In Experiment 2, I also tried to obtain additional evidence for the prediction by administering the name–letter task both before and after the conditioning treatment.

<p>| Table 1 |</p>
<table>
<thead>
<tr>
<th>NLEs for Initials and Full Names as a Function of Experimental Condition (Experiment 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Variable</strong></td>
</tr>
<tr>
<td>Conditioned SE</td>
</tr>
<tr>
<td>$SD$</td>
</tr>
<tr>
<td>Control</td>
</tr>
<tr>
<td>$SD$</td>
</tr>
</tbody>
</table>

*Note. Higher scores represent higher implicit self-esteem. NLEs = name-letter effects; SE = self-esteem.*

**Method**

**Participants and design.** Thirty-five undergraduate students (23 women and 12 men) from the University of Amsterdam participated in the experiment in return for course credits or for Dfl. 7.50. All participants were randomly assigned to a conditioned self-esteem condition or to a control condition.

**Procedure and materials.** Upon entering the laboratory, participants were greeted by a female experimenter and seated in an individual cubicle in front of an iMAC computer. The experimenter explained that the instructions would be given by the computer program. After starting the computer program, she left the cubicle.

First, participants completed the letter evaluations, using the same method as in Experiment 1. Second, participants performed a lexical decision task. The conditioned self-esteem condition was exactly the same as in Experiment 1. However, the control condition now contained the same target words as the conditioned self-esteem condition: 15 positive trait terms and 15 random letter strings. Whereas in the conditioned self-esteem condition the positive traits were preceded by the subliminal presentation of the word $I$, for control participants the positive words (and the random letter strings) were preceded by the subliminal presentation of the letter $x$. Upon finishing the lexical decision task, participants were asked to evaluate the letters of the alphabet.

At the end of the experiment, participants were debriefed carefully. First, they were asked whether they had seen anything unusual during the lexical decision task. Second, participants were asked whether they had seen words (other than the target words) flashing on the screen. As in Experiment 1, no participants indicated any suspicion, and no participant reported having seen any flashes. After the debriefing, participants were asked to write their name on a list in return for course credits or money. Subsequently, they were dismissed.

**Results**

The NLEs were calculated the same way as in Experiment 1. This time two NLEs were obtained: one before the treatment, the other after the treatment. Because the effects in Experiment 1 were most pronounced for initials, the NLE was calculated only for initials in this experiment.

The resulting NLEs were subjected to a 2 (condition: conditioned self-esteem vs. control) $\times$ 2 (timing of measurement: before vs. after treatment) mixed-model analysis of variance (ANOVA). This analysis yielded the predicted two-way interaction, $F(1, 33) = 4.32, p < .05$. The means are listed in Table 2. Participants in the conditioned self-esteem condition showed higher implicit self-esteem after the treatment than before the treatment, relative to control participants.

**Experiment 3**

In the first two experiments, I relied on name–letter and initial preferences to assess implicit self-esteem. In Experiment 3, I wanted to replicate the effects of the manipulation by using a different measure of implicit self-esteem. I chose to use the self-esteem IAT, as developed by Greenwald and Farnham (2000).

I also made two minor changes in the manipulation. First, I aimed to test the possibility that the entire self-esteem manipulation could be presented subliminally. Whereas in Experiments 1

---

3 In all experiments in which the name–letter task was used, it was first confirmed that the experimental manipulation(s) did not affect liking for baseline (i.e., nonname) letters.
Conditioned SE disappeared, the word subliminally. Each trial consisted of the following sequence: First, a row of items 1 and 2, the aim here was to present both the CS and the US program, she left the cubicle.

Control condition.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Time of measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Premake</td>
</tr>
<tr>
<td>Conditioned SE</td>
<td>0.61ab</td>
</tr>
<tr>
<td>SD</td>
<td>1.07</td>
</tr>
<tr>
<td>Control</td>
<td>0.40b</td>
</tr>
<tr>
<td>M</td>
<td>1.06</td>
</tr>
</tbody>
</table>

Note. Higher scores represent higher implicit self-esteem (SE). Values with different subscripts differ significantly (p < .05).

and 2 only the CS (the word I) was presented subliminally, in Experiment 3, both the CS and the US (positive trait terms) were presented subliminally for 17 ms.

Second, I made a minor procedural change in response to the data of Experiment 2. In Experiment 2, the critical two-way interaction was caused not only by an increase in self-esteem among conditioned self-esteem participants but also by a small decrease in self-esteem by control participants. There is a possibility that something in the procedure used was responsible for the effects. During the lexical decision tasks, participants were told to respond as quickly as possible. This caused some participants to make mistakes (in Experiment 2, the error rate was 4.8%), and these mistakes may have caused mild frustration, and thus lower self-esteem. I concede this reasoning is speculative, and it is outside the scope of this article to explicitly test for this possibility. Still, it did prompt a change in the procedure. In Experiment 3, the targets in the lexical decision task were always random letter strings. Participants simply decided whether a string started with a vowel or a consonant. In addition, participants were not explicitly instructed to respond as quickly as possible.

Method

Participants and design. Sixteen undergraduate students (12 women and 4 men) from the University of Amsterdam participated in the experiment in return for course credits or for Dfl. 7.50. All participants were randomly assigned to either a conditioned self-esteem condition or to a control condition.

Procedure and materials. Upon entering the laboratory, participants were welcomed by a female experimenter and seated in an individual cubicle in front of an iMAC. The experimenter explained that the instructions would be given by the computer program. After starting the computer program, she left the cubicle.

The first task was the evaluative-conditioning procedure. Unlike Experiments 1 and 2, the aim here was to present both the CS and the US subliminally. Each trial consisted of the following sequence: First, a row of Xs appeared on the screen for 500 ms. Immediately after the row of Xs disappeared, the word t was presented for 17 ms, immediately followed by a positive trait (in the conditioned self-esteem condition) or a neutral word (in the control condition). These positive-versus-neutral conditioning words were also presented for 17 ms and were masked by the presentation of a target word. The targets were random letter strings, and all participants had to do was to decide whether each letter string started with a vowel or a consonant. The trials were presented in random order.

Immediately after this task, participants’ implicit self-esteem was assessed with a Dutch version of the self-esteem IAT (Greenwald & Farnham, 2000; see also Greenwald, McGhee, & Schwartz, 1998). The task consisted of five blocks. As usual, the third and fifth block were the two critical blocks. Here, participants were presented with the six positive words, the six negative words, the six self-related words, and the six non-self-related words. Each word appeared only once. During the third block, participants were requested to press a key on the left of the keyboard whenever a positive or a self-related word appeared. Whenever a negative word or a non-self-related word appeared, participants were requested to press the right key. Throughout the task, the words self-related and positive remained on the left side of the screen, and the words nonself-related and negative remained on the right side of the screen. Because positive words and self-related words were assigned to the same key, this block was the congruent IAT block. During the fifth block, the positive words and the non-self-related words were assigned to the left key, whereas negative and self-related words were assigned to the right key. Throughout the task, the words nonself-related and positive remained on the left side of the screen, and the words self-related and negative remained on the right side of the screen. Because positive words and self-related words were assigned to different keys, this block was the incongruent IAT block.

After participants finished the IAT (Greenwald & Farnham, 2000), they were thoroughly debriefed. One participant indicated she had seen flashes on the screen during the lexical decision task, but she did not report having seen any words. Subsequently, they were thanked, paid, and dismissed.

Results

For each participant, the mean response latency was calculated on all 24 trials of the two critical blocks. Response latencies for incorrect classifications (3.9%) were not included. Because outliers were practically absent, trimming procedures were not used. Also, analyses on log-transformed response latencies confirmed the analysis below.

The mean response latencies for the two blocks were subjected to a 2 (condition: conditioned self-esteem vs. control) × 2 (IAT block: congruent vs. incongruent) mixed-model ANOVA. This analysis revealed the predicted two-way interaction, F(1, 14) = 8.84, p < .01. As can be seen in Table 3, the difference in response latency between blocks was much more pronounced in the conditioned self-esteem condition, indicating higher self-esteem.

In the first three experiments, evidence was obtained that evaluative-conditioning techniques can be applied to enhance implicit self-esteem. Furthermore, effects of the manipulation were obtained on what is considered (see also Bosson et al., 2000) the...
two most promising measures of implicit self-esteem. In addition, with Experiment 3 it was demonstrated that the self-esteem manipulation has an effect when presented completely subliminally. That is, when both the CS and the US were presented subliminally, the procedure still resulted in higher self-esteem among experimental participants.⁷

### Experiment 4

In the first three experiments, it was shown that implicit self-esteem could be enhanced by evaluative conditioning. Because we know that low self-esteem comes at certain costs, it is justifiable to ask the question whether the evaluative-conditioning procedure could enhance self-esteem among people whose self-esteem is made low. The major aim of Experiment 4 was to investigate the potential moderating role of initial self-esteem. Did the manipulation increase self-esteem among all participants, or were people with low- and high-initial self-esteem affected differentially by the manipulation? In concrete terms, I investigated whether participants’ self-esteem can increase irrespective of whether this self-esteem has just been lowered (by negative intelligence feedback) or has just been boosted (by positive intelligence feedback).

**Method**

**Participants and design.** Eighty-three undergraduate (52 women and 31 men) students from the University of Amsterdam participated in the experiment in return for course credits or for €5 (approximately $6 US). All participants were randomly assigned to the cells of a 2 (condition: conditioned self-esteem vs. control) × 2 (intelligence feedback: negative vs. positive) design.

**Procedure and materials.** Upon entering the laboratory, participants were welcomed by a female experimenter and seated in an individual cubicle in front of an iMAC computer. The experimenter explained that the instructions would be given by the computer program. After starting the computer program, she left the cubicle. The first task was conducted to manipulate self-esteem by giving participants intelligence feedback. It was explained to participants that they had to do a task that was developed by the Personality Department. Performance on this task, it was explained, had been shown to correlate highly with the capacity for analytical thinking and even with intellectual abilities in general. It was stressed that participants should take the task seriously. The task was based on a procedure developed by Ouwerkerk, de Gilder, and de Vries (2000). Participants were repeatedly presented with a row of four circles on the screen. During each trial, one of the circles changed color (from black to red). Participants’ task was to press a button as soon as one of the circles changed color. Participants used a row of four keys immediately above the spacebar (V, B, N, and M). Each key corresponded to a different circle. For example, whenever the second circle changed color, participants had to press the second key (i.e., the B key). Whenever the third circle changed color, participants had to press the third (N) key. These trials were relatively easy. Trials involving the other two circles were more difficult. Whenever the first circle changed color, participants had to press the fourth key (M), and whenever the fourth circle changed color, participants had to press the first key (V). Participants received 60 total trials, each separated by a 1-s pause. It was stressed that participants should perform the task as quickly and as accurately as possible.⁸

After this task, participants received feedback. First, all participants were presented with the same score (supposedly their average reaction time, 640 ms). Subsequently, participants saw a graph in which the scores of a large number of students were allegedly depicted. The graph was a normal curve. In the positive-feedback condition, the mean was much higher than the score of the participant, and the graph indicated that the participant had scored among the best 10% of the student population. Conversely, in the negative-feedback condition, the mean was much lower than the score of the participant, and the graph indicated that the participant had a score among the worst 10% of the student population.

The next task was the evaluative-conditioning procedure. This task was exactly the same as in Experiment 3. Finally, implicit self-esteem was assessed. As in Experiments 1 and 2, participants’ preference for initials was assessed. This was done the same way as in Experiments 1 and 2. At the end of the experiment, participants were debriefed carefully. Two participants indicated they had seen flashes on the screen during the lexical decision task. However, no participants reported having seen any real words. Participants were then told that the intelligence feedback manipulation was false and that actual reaction times were not even measured. After the debriefing, participants were asked to write their name on a list in return for course credits or money. Subsequently, they were thanked, paid, and dismissed.

**Results**

First, preferences for initials were calculated for all participants. These scores were subjected to a 2 (condition: conditioned self-esteem vs. control) × 2 (intelligence feedback: positive vs. negative) between-participants ANOVA. This analysis yielded the two predicted main effects. First, participants in the conditioned self-esteem condition exhibited higher implicit self-esteem than participants in the control condition, \( F(1, 79) = 7.45, p < .01 \). Second, participants in the positive-feedback condition showed higher implicit self-esteem than participants in the negative-feedback condition, \( F(1, 79) = 4.02, p < .05 \). The two-way interaction was very far from reliable, \( F(1, 79) = 0.21 \), indicating that participants in both the negative-feedback condition and the positive-feedback condition were affected by the conditioning manipulation. The means are listed in Table 4. Interestingly, participants in the negative-feedback condition whose self-esteem was not conditioned reported on average a negative preference for initials score.

---

⁷ To my knowledge (see also De Houwer et al., 2001), this is the first demonstration of evaluative conditioning whereby both the CS and the US were presented subliminally.

⁸ This task may seem relatively simple, but it is actually rather frustrating. It is impossible to avoid making mistakes if one is seriously trying to be as fast as possible. The experience of completing this task comes close to the experience of completing an inconsistent IAT block (see Greenwald et al., 1998; see also Experiment 3 of this article).

---

### Table 3

<table>
<thead>
<tr>
<th>Variable</th>
<th>Congruent (self-positive)</th>
<th>Incongruent (self-negative)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conditioned SE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( M )</td>
<td>617.68</td>
<td>979.94</td>
</tr>
<tr>
<td>( SD )</td>
<td>113</td>
<td>154</td>
</tr>
<tr>
<td>Control</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( M )</td>
<td>608.31</td>
<td>762.36</td>
</tr>
<tr>
<td>( SD )</td>
<td>87</td>
<td>113</td>
</tr>
</tbody>
</table>

**Note.** Values with different subscripts differ significantly (\( p < .05 \)). IAT = Implicit Association Test; SE = self-esteem.
This underscores the success of the feedback manipulation. The results of Experiment 4 also document the independent role of explicit feedback and unconscious conditioning in shaping people’s implicit self-esteem.

Experiments 5a and 5b

Experiments 5a and 5b deal with potential consequences of enhanced implicit self-esteem. Now that it has been established that evaluative-conditioning techniques can enhance implicit self-esteem, it is fruitful to investigate potential benefits of enhanced implicit self-esteem. The choice for the benefits under consideration here was based on recent research on implicit self-esteem, but even more on research on effects of explicit self-esteem. Presumably, the most widely documented consequence of self-esteem pertains to the different way low self-esteem and high self-esteem individuals deal with negative feedback or threats to the self in general. As noted in the introduction, low self-esteem individuals usually have more problems coping with negative feedback. Negative feedback can have negative emotional consequences as well as negative behavioral consequences, such as diminished persistence. High self-esteem, however, can function as a buffer against such negative experiences (e.g., Greenwald & Farnham, 2000; Shrauger & Rosenberg, 1970; Steele, 1988).

In Experiments 5a and 5b, however, low and high self-esteem were not compared. Rather, participants whose self-esteem is not manipulated (but may still be reasonably high) were compared with participants whose self-esteem is enhanced. Therefore, the experiments do not allow conclusions about the consequences of low self-esteem. Instead, the aim was to investigate the consequences of very high self-esteem. As a consequence, Experiments 5a and 5b are to some extent exploratory. Whereas the consequences of low self-esteem are well documented, potential consequences of very high self-esteem are largely unknown.

In Experiment 5a, effects of negative personality feedback on mood were tested. First, some participants’ self-esteem was conditioned, whereas others was not. Subsequently, participants all completed a general knowledge questionnaire that was presented as one of “average difficulty.” In reality, however, some were presented with easy questions, whereas others were given very difficult questions. After participants completed the questionnaire, they were given positive or negative feedback. Finally, a mood measure was administered.

In Experiment 5b, the effects of negative personality feedback were tested on task persistence. First, again self-esteem was enhanced among half the participants. Subsequently, participants were given negative (or positive) feedback. Next, participants were given the opportunity to complete a task with the chance to do well, and therefore to restore the negative effects of negative feedback.

Method: Experiment 5a

Participants and design. Ninety-three undergraduate students (62 women and 31 men) from the University of Amsterdam participated in the experiment in return for course credits or for € 5. All participants were randomly assigned to the cells of a 2 (condition: conditioned self-esteem vs. control) × 2 (intelligence feedback: negative vs. positive) between-participants design.

Procedure and materials. Upon entering the laboratory, participants were welcomed by a female experimenter and seated in an individual cubicle in front of an IMAC. The experimenter explained that the instructions would be given by the computer program. After starting the computer program, she left the cubicle.

The second task was the evaluative-conditioning procedure. This task was exactly the same as in Experiment 3. After participants finished this task, the feedback task was introduced. Participants were told that a general knowledge scale was being developed that was designed especially for assessing general knowledge of university students. The scale was said to consist of various subscales of different difficulty. Participants were then told that they would receive a subscale of “average difficulty.” Subsequently, participants indeed completed 15 multiple-choice questions on the computer.

Unbeknownst to the participants, some were presented with a series of 15 easy questions, whereas others were presented with 15 difficult questions. The answers participants gave indeed confirmed this. Of the participants who received the easy questions, everyone answered at least 11 questions correctly. Conversely, of the participants who received the difficult questions, no one gave more than 5 correct answers. After participants finished the questionnaire, their scores were presented on the computer screen (“You answered 12 out of 15 questions correctly”).

Finally, the mood measure was administered. Participants were asked to indicate their mood by answering the question “How good do you feel?” on a 9-point scale ranging from 1 (very bad) to 9 (very good).

After participants finished the mood item, they were thoroughly debriefed. No participants indicated having seen flashes on the screen during the lexical decision task. Subsequently, they were thanked, paid, and dismissed.

Method: Experiment 5b

Participants and design. Fifty-five undergraduate students (38 women and 17 men) from the University of Amsterdam participated in the experiment in return for course credits or for € 5. All participants were randomly assigned to one of the cells of a 2 (condition: conditioned self-esteem vs. control) × 2 (intelligence feedback: negative vs. positive) between-participants design.

Procedure and materials. Upon entering the laboratory, participants were welcomed by a female experimenter and seated in an individual cubicle in front of an IMAC. The experimenter explained that the instructions would be given by the computer program. After starting the computer program, she left the cubicle.

The second task was the evaluative-conditioning procedure. This task was exactly the same as in Experiments 3–5a. After participants finished this task, the feedback task was introduced. Participants were asked to do an association task, based on the Remote Associate Test (Mednick & Mednick, 1967). They were presented with three words on the screen and...
were asked to come up with a 4th word that was associated with the other three (e.g., food, red, and very spicy) would be presented and the correct association would be chili pepper. All participants completed this task five times. An association task was introduced that was strongly indicative of people’s creativity. All participants were told that this association task contained items of varying difficulty and that they would be presented with five items that were of average difficulty. In reality, however, half the participants received five items that were pilot tested as easy (indeed, on average, a little over 70% were answered correctly), whereas others received items that were pilot tested as extremely difficult (on average, only 10% were answered correctly). Each item remained on the screen until a participant either typed an answer or when a participant pressed a designated key to go to the next item. Participants received immediate feedback on each item (“wrong” or “right” appeared on the screen).

After participants completed the five items, they were told that another set of five could be completed in case they would like to but that it was not required. Also, they were allowed to start and finish the task whenever they wanted. Participants were told that the second series of five consisted of rather difficult items. Indeed, these items were difficult. Of interest was the degree of persistence. Therefore, the time participants would spend on this second series of five was recorded.

After participants finished the second series of five associations, they were thoroughly debriefed. No participants indicated having seen flashes on the screen during the lexical decision task. Subsequently, they were thanked, paid, and dismissed.

**Results**

*Experiment 5a.* The mean scores on the mood measure were subjected to a 2 (condition: conditioned self-esteem vs. control) × 2 (intelligence feedback: negative vs. positive) between-participants ANOVA. This analysis revealed a main effect of intelligence feedback reported to be in a worse mood, $F(1, 89) = 5.57, p < .02$. As predicted, this main effect was qualified by the two-way interaction, $F(1, 89) = 4.91, p < .03$. As can be seen in Table 5, control participants responded to feedback the way one would predict: they felt better after receiving positive feedback than after receiving negative feedback. Participants whose implicit self-esteem was enhanced, however, showed no effects of feedback on mood.

*Experiment 5b.* The time each participant spent on the second series of five association items was subjected to a 2 (condition: conditioned self-esteem vs. control condition) × 2 (intelligence feedback: negative vs. positive) between-participants ANOVA. The only reliable effect was the predicted two-way interaction, $F(1, 51) = 4.74, p < .03$. As can be seen in Table 5, control participants behaved the way one would expect. After having received negative feedback, they persist longer and try to do what they can to restore self-esteem. High self-esteem participants did not spend much time on the task, regardless of feedback.

In summary, in both Experiments 5a and 5b, individuals whose self-esteem was enhanced seemed to be insensitive to personality feedback, whereas control participants whose self-esteem was not enhanced did show effects of the intelligence feedback.

**General Discussion**

In our experiments, implicit self-esteem was enhanced through subliminal evaluative conditioning. Pairing the self-depicting word I with positive trait terms consistently improved implicit self-esteem. In addition, by presenting stimuli subliminally, it was demonstrated that the evaluative-conditioning procedure affects implicit self-esteem while bypassing consciousness altogether. Furthermore, evaluative conditioning enhanced self-esteem both among people with temporarily low self-esteem as well as among people with temporarily high self-esteem.

The thought-provoking article by Bosson et al. (2000) demonstrated that research on implicit self-esteem suffers from conceptual clarity. Bosson et al. showed that the (seven) measures that were used in the literature to assess implicit self-esteem did not correlate. In the present study, the way to deal with this conceptual puzzle was to try to manipulate implicit self-esteem and to show effects of the manipulation on various measures of implicit self-esteem and on other psychological measures that are known to be related to self-esteem. The evaluative-conditioning technique produced effects on preferences for initials (see Experiments 1, 2, and 4), on general name–letter liking (see Experiment 1), and on the self-esteem IAT (Greenwald & Farnham, 2000; see Experiment 3). Finally, the manipulation made participants insensitive to negative intelligence feedback (see Experiments 5a and 5b).

**Toward Refining the Concept of Implicit Self-Esteem**

The manipulation used in the present studies has consequences for how implicit self-esteem could be defined. Evaluative condi-

---

9 However, this is likely not necessary to achieve the effect. In a recent experiment, self-esteem of half the participants was increased with the conditioning technique as described in Experiments 3–5b. Apart from these conditions, conditions were created in which participants were explicitly told that the goal of the procedure was to enhance self-esteem. It was even explained to some extent how evaluative conditioning worked. The results showed (initial preference was assessed) that among these participants, the technique still worked.
tioning is a technique that changes attitudes at a very basic level. It changes attitude structurally in that the strength of the association between the attitude object and (in our case) positive affect is altered. That is, it changes the evaluation that is automatically activated upon encountering the attitude object. If a definition of implicit self-esteem is based on the present data in combination with the logic behind evaluative conditioning, it would mean that implicit self-esteem could be defined as the implicit attitude toward the self (see also Bosson et al., 2000). It is important to realize that this definition diverges somewhat from at least some other definitions. Although Greenwald and Banaji’s (1995) definition (“the introspectively unidentified [or inaccurately identified] effect of the self-attitude on evaluation of self-associated and self-dissociated objects,” p. 11) also treats implicit self-esteem as an attitudinal construct, their definition entails not only the attitude itself but also some of its consequences. According to their definition, evaluation of self-associated objects is part of the definition. However, what is proposed here is that implicit self-esteem is the attitude, and that a positive (or negative) evaluation of self-associated objects (such as name letters) is a consequence of high (or low) self-esteem (see also Jones et al., 2002).

Without wanting to spend too much time on discussing variations in possible definitions, I stress that conceptual clarity is essential. The lack of correlations between different measures of implicit self-esteem that Bosson et al. (2000) found may have been due to the fact that the different measures essentially measured different things. Whereas some measures directly measure the implicit attitude toward the self (e.g., the IAT [Greenwald & Farnham, 2000], or priming measures based on work by Fazio et al., 1995), other measures are tapping consequences of this attitude (such as preferences for initials), and are therefore essentially more indirect.

Feedback Insensitivity?

It is well known that explicit self-esteem moderates effects of threats to the self. High self-esteem individuals are better able to “repair” their self-esteem after a failure experience (Steele, 1988). They show more persistence in the face of failure, at least when persisting makes sense (Di Paula & Campbell, 2002; McFarlin, Baumeister, & Blascovich, 1984; Shrauger & Rosenberg, 1970), and they are better able to deal with the emotional consequences of existential threats (Greenberg et al., 1992). Recently, researchers have begun to investigate potential consequences of low- or high-implicit self-esteem. The first findings suggest that effects of implicit self-esteem are much like the effects of explicit self-esteem: People with low-implicit self-esteem show diminished levels of aspiration after failure (Greenwald & Farnham, 2000), and low-implicit self-esteem individuals show more anxiety during a confrontational interview than high-implicit self-esteem individuals (Spalding & Hardin, 1999).

The results of Experiments 5a and 5b are in keeping with earlier work. They emphasize the more general finding that high self-esteem individuals suffer less from threats to the self. The results were extreme, however, in that participants with high-implicit self-esteem were completely oblivious to personality feedback. It is interesting to speculate on how such “feedback insensitivity” ensues. When people process negative feedback about themselves, the self is, by definition, activated. The conceptualization of implicit self-esteem as the automatic attitude toward the self entails that this attitude is activated too during the processing of feedback. It is possible that for people with very high-implicit self-esteem, the positive affect that results from activating this self-attitude “overpowers” the negative affect produced by negative feedback. In a more general sense, it is possible that individuals with very high-implicit self-esteem regulate negative self-related experiences better because of the release of positive affect when the self is activated. Further research may shed light on potential mediating mechanisms.

The effects obtained in the present study are reminiscent of the effects obtained by Spalding and Hardin (1999) and by Greenwald and Farnham (2000). Spalding and Hardin (1999) measured anxiety after either a confronting self-relevant interview or after a self-irrelevant interview. In their experiment, they also showed that high self-esteem individuals were insensitive to their manipulation. They found low levels of anxiety among high self-esteem individuals, regardless of whether the interview was self-relevant or not. Greenwald and Farnham (2000) measured implicit self-esteem and used a regression analysis to study the relation between self-esteem and reactions to feedback. The interactions they obtained did demonstrate that low self-esteem individuals showed more pronounced effects of feedback than high self-esteem individuals. Their approach makes it hard to say whether or not high-implicit self-esteem individuals showed effects of feedback, but their regression slopes suggest that high self-esteem individuals did not show effects of success or failure on mood, perceived task importance, or future aspiration, whereas low self-esteem individuals did show effects on the latter two measures.

The present results are remarkable in their relative extremity. If high-implicit self-esteem indeed makes people insensitive to feedback, it would not only protect them from the negative emotional and behavioral consequences of negative feedback but it would also mean that high-implicit self-esteem individuals can refrain from some of the undesirable strategies people sometimes adopt to repair the effects of failure experiences. It has been shown that failure feedback can lead to increased prejudice (Fein & Spencer, 1997) and out-group derogation (Beauregard & Dunning, 1998). Also, whereas researchers know that high-explicit self-esteem individuals cope better with failure experiences, the question arises whether this is true for high-implicit self-esteem as well. It is possible that the latter individuals cope better, but it is also possible that they do not cope at all. If they are truly oblivious, there is simply nothing to repair or to cope.

A Final Word

It is worthwhile to explicitly mention an intriguing aspect of the present work. Implicit self-esteem can be enhanced, at least temporarily, subliminally in about 25 seconds. It is not yet known how long the effects of this manipulation last. In addition, it is not yet known whether people who could really benefit from enhanced self-esteem (i.e., people with problematically low levels of self-esteem) can benefit from subliminal conditioning techniques. Still, the present work may be useful in its own right. If, for some reason, one does not like one’s horses, one could always acquire new ones. An alternative strategy is offered that is cheaper, faster, and much nicer toward the horses: coming to feel better about the
self in 25 seconds. Presumably, better feelings about one’s horses will soon follow.

References


Received October 8, 2002
Revision received February 10, 2003
Accepted March 18, 2003