Probing the malleability of implicit and explicit self-esteem: An interview approach

Iain S. Glen and Rainer Banse

1. Humboldt-Universität zu Berlin, Germany
2. University of York, UK

Abstract
A resource and a deficit focused interview were used to manipulate explicit and implicit self-esteem in a pre-post design. Implicit self-esteem was assessed using the Implicit Association Test (IAT, Greenwald, McGhee, & Schwartz, 1998) and explicit using standard questionnaires. It was hypothesized that implicit would decrease more than explicit self-esteem in the deficit interview condition as presentation and consistency needs would make explicit self-esteem more resistant to negative self-relevant information. Moderate increases were expected in implicit and explicit self-esteem in the resource interview condition. Overall, the hypotheses received little support. There was an increase in the self-liking aspect of explicit self-esteem but in both interview conditions. In addition, a significantly lower stability of the IAT in the deficit than in the resource interview condition suggests that the deficit interview elicited substantial differential change of implicit self-esteem. The study shows that the self-esteem IAT had theoretically meaningful relations to explicit self-related measures and was relatively robust against manipulation.

Key words: self-esteem, implicit self-esteem, IAT.

Correspondence should be sent to Rainer Banse, Department of Psychology, University of York, Heslington, York, YO10 5DD, UK (e-mail: r.banse@psychology.york.ac.uk)
INTRODUCTION

Self-esteem, the positivity of a person's evaluation of self (Baumeister, 1998), is generally accepted to be an important psychological construct and a central component of an individual's daily experience (e.g., Kernis, 2003). Despite the lack of clear evidence on the benefits of having high self-esteem, and the problems associated with fragile high self-esteem, it is generally seen as being desirable and closely linked to psychological health (e.g., Baumeister, 1998; Jordan, Spencer, & Zanna, 2003; Taylor & Brown, 1988). If it is so important then the question of its malleability automatically arises: Can a person's self-esteem be changed, either positively or negatively, or is it resistant against manipulation?

Explicit and implicit self-esteem

In order to tackle this question it is important to distinguish between explicit and implicit self-esteem (e.g., Bosson, Swann, & Pennebaker, 2000; Greenwald & Banaji, 1995; Koole & Pehlam, 2003). This distinction can be found in recent dual-process, or dual-system theories, that postulate two different modes of information processing, an explicit or reflective mode in which information is processed more intentionally, consciously, and controlled, and an implicit or impulsive mode in which information is processed more automatically, unconsciously, and effortlessly (e.g., Chaiken & Trope, 1999; Epstein, 1990; Smith & DeCoster, 2000). In order to define what is implied here by implicit and explicit self-esteem, it is important to first take a closer look at these two modes. For this purpose, Strack and Deutsch's (in press) Reflective and Impulsive Model (RIM), the most recent and as yet most comprehensive model, seems to be particularly appropriate. They describe the two modes as follows: In the reflective system, behavior is generated as a consequence of an intentional decision process based on declarative propositional knowledge. It requires allocation of attention and cognitive capacity, but the system is very flexible and able to perform complex cognitive operations. In the impulsive system, behavior is elicited through automatically spreading activation between associative links. Its activity may be accompanied by an experiential state of awareness but not necessarily so. It requires little cognitive capacity, but it is relatively
rigid and not able to perform complex operations. The two systems are assumed to act in parallel, both synergistically and antagonistically. Based on this model, we will use the terms explicit and implicit modes of processing to refer to processing within these reflective and impulsive systems, respectively. This distinction is not based on the presence or absence of consciousness (see also, Greenwald & Banaji, 1995; Fazio & Olson, 2003).

Because self-esteem can be regarded as an attitude toward the self, the vivid debate on the conceptual status of implicit attitudes as a single or a dual construct may be informative for delineating the meaning of the construct of implicit self-esteem. Whereas Wilson, Lindsey, and Schooler (2000) suggest a dual-attitude model, whereby a person can simultaneously hold an implicit and an explicit attitude towards an object, other researchers prefer to distinguish between implicit or explicit measures of a single attitude construct (e.g., Fazio & Olsen, 2003), or to use the terms explicit and implicit only to refer to underlying psychological processes that are tapped by different measures, but not to mental contents (e.g., Strack & Deutsch, in press).

Our position is that it is useful to refer to explicit and implicit self-esteem as labels for different constructs when explicit self-esteem is defined as the representation of self-esteem based on evaluative decisions in the reflective system, and implicit self-esteem as the representation of self-esteem based on automatic evaluative associations in the impulsive system, without necessarily being unconscious. At the empirical level, one can distinguish between explicit measures, which are based on information that is intentionally given to inform about the self, and implicit measures which are based on information that is not intentionally given to inform about the self (cf. Asendorpf, Banse, & Muccke, 2002). Explicit measures require by definition the processing of self-evaluation in the reflective system, whereas implicit measures tap only impulsive processes.

Implicit self-esteem measures

A major challenge has been to develop reliable and valid measures of implicit self-esteem. Of all the available measures, the name-letter evaluation task (Nuttin, 1985) and above all the self-esteem IAT (Greenwald & Farnham, 2000) are the most promising (see Bosson et al., 2000;
Greenwald & Farnham, 2000; Jordan et al., 2003). The IAT shows acceptable internal consistency and retest-reliability. With regard to convergent validity, the IAT correlates weakly but positively with explicit self-esteem measures, it has a positive bias which is associated with explicit self-esteem, and it predicts criterion variables predicted by explicit self-esteem with at least modest effect size, such as success and failure and persistence in the face of failure. At this point it is important to make clear that the validity of the IAT as a measure of implicit constructs and its methodology is subject to a vivid debate (see for example the special issues on the IAT in Zeitschrift fuer Experimentelle Psychologie 2001, 48(2) and Journal of Personality and Social Psychology 2001, 81(5)).

Do we need the construct implicit self-esteem?

As Greenwald and Farnham (2000) state “The distinction between explicit and implicit operation of the self is especially interesting if it turns out that the self functions differently in these two modes” (p. 1022). In addition, it is most useful if implicit self-esteem can be shown to be associated with, or be able to predict, behavior and experience that explicit self-esteem cannot and vice versa. Asendorpf et al. (2002) refer to this as a double dissociation strategy. Evidence of such a double dissociation pattern for self-esteem has recently been provided by Spalding and Harding (1999): Implicit self-esteem uniquely predicted nonverbal anxious behavior, whereas explicit self-esteem uniquely predicted self-judgments. There has also been much recent research into the moderating role of implicit self-esteem in distinguishing between secure and fragile high explicit self-esteem (for a review, see Kinnis, 2003).

Malleability of implicit and explicit self-esteem

Of interest in this study is the malleability of implicit and explicit self-esteem in response to a manipulation. According to Heatherton and Polivy (1991), “few studies have found significant changes in self-esteem by using self-report measures that were designed to measure self-esteem” (p. 896). One explanation is that self-esteem is an enduring personality disposition and only measures of state self-esteem are re-
sponsive to fluctuations (e.g., Heatherton & Polivy, 1991). An alter-
native explanation is that explicit measures do not necessarily reflect an
accurate appraisal of oneself, rather how one wants to see and present
oneself. Just as one’s explicit self-concept is a self-construction
(Baumeister, 1998) or self-theory (Epstein, 1973), so is one’s explicit
self-esteem something constructed and not merely experienced.
Farukam, Greenwald, and Banaji (1999) go as far as to suggest that
(explicit) “self-esteem measures may be capturing the wrong construct:
the motive to present a positive attitude toward self rather than genuine
self-esteem”, i.e., implicit self-esteem (p. 231). If self-report measures
do not reflect self-esteem accurately, they may also not reflect changes
in ‘genuine’ self-esteem accurately. We do not consider implicit self-
estee to be more genuine than explicit self-esteem. Our view is that
self-report measures assess explicit self-esteem that may be influenced
by complex, deliberate information processing whereas implicit self-
estee as assessed by the IAT relies mainly on automatic activation of
self-related evaluation. Change in explicit self-esteem may or may not
be associated with change in implicit self-esteem.

In addition to the need for accurate appraisal, there are three self
motives that may influence the effect of a manipulation on self-esteem:
automatic self-enhancement or the positivity bias, strategic or self-
presentational self-enhancement, and self-consistency (see Baumeister,
1998). All three are expected to influence explicit self-esteem but only
automatic self-enhancement should influence implicit self-esteem.

Seeing oneself positively or the tendency to enhance the self (e.g.,
Greenwald, 1980) is, according to Greenwald and Banaji (1995), a part
of both explicit and implicit self-esteem. At least in healthy individuals,
it should thus make both implicit and explicit self-esteem resistant to
negative self-relevant information and sensitive to positive self-relevant
information.

Presenting oneself positively on the other hand is, as mentioned
above, assumed only to influence explicit self-esteem. Paulus (1986) dis-
tinguishes between two self-enhancing presentation strategies, self-
deception, i.e., presenting oneself positively to oneself, as implied
above by Farnham et al. (1999), and impression management, i.e., pre-
senting oneself positively to others. It should thus make explicit but not
implicit self-esteem further resistant to negative self-relevant information
and more sensitive to positive self-relevant information.
Another self-motive that can influence the change in explicit self-esteem is the motive for self-consistency or self-verification, (e.g., Lecky, 1945; Swann, 1983, 1990). This desire to maintain a stable and unchanging sense of self has been the subject of much research and often contrasted with the motive for self-enhancement (e.g., Brown, 1993; Swann, Griffin, Predmore, & Gaines, 1987). As consistency effects have been shown to be slower and require more cognitive processing than enhancement effects, it has been proposed that the self-enhancement response is automatic and "affective" and the self-consistency one controlled and "cognitive" (e.g., Swann, 1990). Self-consistency should thus make explicit self-esteem resistant to any discrepant information, whether positive or negative. It should be noted that Swann (1990) also proposes a strategic self-presentational stage in which the costs and benefits of self-enhancement against self-verification are weighed up. A choice of self-enhancement at this stage would be self-presentational and require cognitive resources.

In summary, whereas automatic self-enhancement, or the positivity bias, is assumed to influence both implicit and explicit self-esteem, self-consistency and self-presentational self-enhancement are deliberative, requiring more cognitive resources, and are thus assumed to influence only explicit self-esteem. In the case of negative self-relevant information, this would imply that explicit self-esteem should be more resistant to change than implicit self-esteem, as explicit self-esteem is subject to automatic self-enhancement, self-presentational self-enhancement and self-consistency, which in this case all function to keep self-esteem positive, whereas implicit self-esteem is subject to only automatic self-enhancement. However, in the case of positive self-relevant information this would imply that both explicit and implicit self-esteem should be increased to some extent, as automatic self-enhancement should make both implicit and explicit self-esteem sensitive to the positive self-relevant information. Explicit self-esteem will of course be more sensitive to self-relevant information due to presentational self-enhancement, but on the other hand self-consistency should make it less sensitive. Any increase in self-esteem may however be attenuated. Firstly, as self-esteem is already biased, positive information may simply confirm one's already positive self-esteem. Secondly, as it is already high, there may be a ceiling effect, preventing it from further increase. Thirdly and relevant only for explicit self-esteem, the self-presenter must strike a balance between the opposing forces of favorability and plausibility.
Malleability of self-esteem

(c.g., Schlenker, 1980, 1986). Any increases due to the desire to make a good impression may be capped by the desire not to make excessive and implausible positive claims.

The view that implicit self-esteem may be more malleable than explicit self-esteem in the case of negative self-relevant information must be contrasted with the prevailing notion that implicit attitudes are often seen as overlearned and more stable than explicit ones (e.g., Wilson et al., 2000; Strack & Deutsch, in press; Fazio & Olson, 2003). However, there is also increasing evidence that not only are implicit attitudes malleable and dependent on context (for a review, see Blair, 2002), but also is implicit self-esteem (Jones, Pehlam, Mirenberg, & Hetts, 2002; Dijkstra et al., in press). It is our aim that this investigation shed further light on the malleability of implicit self-esteem.

The present study: Manipulating self-esteem

There have been many attempts to manipulate self-esteem by making positive or negative self-relevant information available, primarily using a “bogus pipeline”, priming techniques, or social comparisons (for a review, see Heatherton & Polivy, 1991; Riketta & Dauenheimer, 2003). It is important to distinguish between explicit and implicit manipulations (see Gregg, 2003). Explicit manipulations, such as verbal feedback, require the processing of given information in the reflective system, whereas implicit manipulations, such as subliminal priming or nonverbal feedback, limit the processing of information to the impulsive system. We chose to manipulate self-esteem with an explicit manipulation by using an interview focusing either on participants’ successes and strengths or on their failures and weaknesses, labeled resource and deficit interview, respectively (see Method section). This choice was based on three lines of reasoning: First we wanted to manipulate self-esteem in an ecologically valid way; second, due to ethical concerns, we wanted the participants to be aware that the focus was on self so that they could remain in control; and third, we did not want to provide either false or new information about self.

As a manipulation check of this manipulation, two variables were assessed. First we assessed self-efficacy (Bandura, 1977), as the interview dealt not only with self-esteem but also personal effectiveness. Self-efficacy, especially in its generalized form, is according to Tafarodi
and Swann (2001) very close to the construct of self-competence, self-competence being the valuative imprint of general self-efficacy on identity. Second, mood was assessed in order to pick up any temporary, more affective influence of the interview, which might suggest a change in state but not trait self-esteem. If the manipulation is successful, there should be increases in both variables in the resource interview condition and decreases in the deficit interview condition.

Hypotheses. Following our argumentation that self-presentational self-enhancement and self-consistency motives only influence explicit self-esteem and will make it more resistant to negative self-relevant information than implicit self-esteem, we expected that the deficit focused interview would cause a decrease in implicit (Hypothesis 1a) but have no effect on explicit self-esteem (Hypothesis 1b).

As both implicit and explicit self-esteem are subject to automatic self-enhancement whilst self-presentational self-enhancement and self-consistency conflict for explicit self-esteem, both should be sensitive to positive self-relevant information, and thus we expected that the resource focused interview would cause a increase in both implicit (Hypothesis 2a) and explicit self-esteem (Hypothesis 2b).

METHOD

Participants

A total of 64 German speaking psychology students (55 female) participated in the study in exchange for course credits. Age ranged from 19 to 54 (M = 24.0, SD = 6.3). One participant’s data was excluded from analyses because she failed to attend the second measurement occasion, and two participants’ data due to technical problems. Thirty-two participants were randomly assigned to the deficit interview and 29 to the resource interview.

Design

The study was carried out with a randomized pre-post design with two measurement occasions (Time 1 and Time 2) and with the between group variables interview focus (resource interview vs. deficit interview
condition) and order of explicit and implicit measurement at Time 1 and Time 2 (explicit vs. implicit first condition).

Measures

Implicit self-esteem. Implicit self-esteem was assessed by the self-esteem IAT (Greenwald & Farnham, 2000; see also Greenwald et al., 1998). This is a computerized categorization task that measures automatic associations of self-relevant and non-self-relevant target words with pleasant and unpleasant (affective IAT) or positive and negative trait (evaluative IAT) attribute words. We used the evaluative IAT, because it is conceptually akin to the explicit self-esteem measures, and idiographic format (i.e., using individual self-related items such as the name or telephone number of participants). Subjects press a left or right key to rapidly categorize each of a series of stimuli related to me or not me (see Procedure), or to positive or negative trait words.1 The self-esteem IAT consisted of the usual five steps (see Table 1). The IAT effect is computed by subtracting the mean latency for the me + positive block from that for the me + negative block (Step 5 minus Step 3). Thus, scores reflect the ease with which participants associate positive versus negative trait words with the self. Latencies larger than 3000 ms were recoded to 3000 ms and latencies less than 300 ms to 300 ms. Reflecting the improved algorithm from Greenwald, Nosak, and Banaji (2003), error latencies were replaced with the block mean response plus two standard deviations, and the difference between the two combined tasks was divided by the pooled standard deviation of the response latencies of these blocks. The Spearman-Brown corrected split half correlation between the first and the second half of these blocks was .82 at Time 1 and .71 at Time 2. It has to be noted that this reliability estimate may be influenced by systematic time × person effects.

1. The positive trait words were: schlau (clever), klug (intelligent), erfolgreich (successful), geschätzt (valued), stolz (proud), geliebt (loved), kompetent (competent), sympathisch (pleasant), liebenswürdig (friendly). The negative trait words were: dummi (thick), hässlich (ugly), schwach (weak), ungeschickt (inept), verachtet (despised), gehasst (hated), nutzlos (useless), Versager (loser), gemein (nasty), beschämmt (ashamed).
Table 1
IAT procedure and key assignments

<table>
<thead>
<tr>
<th>Sequence</th>
<th>N of trials</th>
<th>Task</th>
<th>Response key assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Left key</td>
</tr>
<tr>
<td>1</td>
<td>40</td>
<td>Attribute discrimination</td>
<td>Negative</td>
</tr>
<tr>
<td>2</td>
<td>40</td>
<td>Target discrimination</td>
<td>Not me</td>
</tr>
<tr>
<td>3</td>
<td>80</td>
<td>Initial combined task</td>
<td>Negative</td>
</tr>
<tr>
<td>4</td>
<td>40</td>
<td>Reversal target discrimination</td>
<td>Me</td>
</tr>
<tr>
<td>5</td>
<td>80</td>
<td>Reversed combined task</td>
<td>Negative</td>
</tr>
</tbody>
</table>

**Self-report measures.** *Explicit self-esteem* was assessed by the Rosenberg Self-esteem Scale (RSES; Rosenberg, 1965; German translation from Ferring & Filipp, 1996) (α = .79) and the Revised Self Liking and Competence Scale (SLSCR, Tafarodi and Swann, 2001; translation by the authors) (self-liking, α = .87; self-competence, α = .83). In order to reduce complexity and to increase reliability, the two self-esteem measures (RSES and SLSCR) were first standardized and then combined into a single aggregate measure, labeled simply explicit self-esteem (α = .89). *Self-efficacy* was assessed by the Generalized Self-Efficacy Scale (GSE; Jerusalem & Schwarzer, 1986) (α = .86). These self-report measures were presented on the computer with participants responding to all scales on a 5-point scale ranging from 1 (strongly disagree) to 5 (strongly agree). *Mood* was assessed using the two parallel forms of the good and bad mood scales of the “Mehrdimensionale Befindlichkeitsfragebogen” (Steyer, Schwenkmezger, Notz, & Eid, 1997) (α = .80). Participants rated how they felt on a 5 point answer scale ranging from 1 (not at all) to 5 (a lot).

**Resource and deficit interviews**

The self-esteem manipulation consisted of two semi-structured interviews on coping strategies and aspects of the self, attention being focused on personal successes and strengths (resource interview condition)
or personal failures and weaknesses (deficit interview condition). It was
developed by the authors and 4 interviewers who were advanced
students in clinical psychology courses. Participants were first asked to
consider situations where they experienced personal success or failure.
They were then asked how and why they coped well or badly and to
consider their role, generalizations, and consequences. Participants were
then asked general questions about their selves, in terms of strengths or
weaknesses. The interview lasted between 30 to 40 minutes. All inter-
viewers were extensively trained to assure parallel interview tech-
niques.2

Procedure

At Time 1, the participants reported to the lab individually, were
seated in a room containing a desktop computer, informed about the
experimental procedure, and asked for some biographical information.
The experimenter stressed the confidentiality of participants’ responses
and asked for a code name to anonymously recognize them at Time 2.
The participants were asked to provide 10 items (out of a list of 16
categories3) strongly related to the self (e.g., forename, surname) and
10 parallel items not related to the self that were then used for the IAT
not-me dimension. Then the participants were administered the implicit
and explicit self-esteem measures (in counterbalanced order). After com-
pleting all tasks, the experimenter scheduled a follow-up session with a
different experimenter and thanked them. The mean length of time
between Times 1 and 2 was 10.44 days (min. 7, max. 27 days).

On returning at Time 2, the participants were informed that they
would be interviewed about their coping strategies, and then again com-
plete a number of tasks on the computer. Before and after the interview
the participants filled in two parallel versions of the mood scale. The
procedure was identical to Time 1. In the deficit interview condition,

2. More detailed information is available from the authors.

3. The 16 target categories were first name, surname, job/course, year of
birth, day of birth, sex, place of birth, nationality, street, telephone number,
town, district, post code, right or left handed, star sign, and hobby.
participants were re-interviewed for 10 minutes, focusing now on their resources. Finally, participants were debriefed about the purpose of the study and thanked for their participation. The procedure and order of assessments is shown in Table 2.

Table 2
Procedure of the Time 1 and Time 2 assessments

<table>
<thead>
<tr>
<th>Time 1</th>
<th>Construction of idiographic IAT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Implicit first condition</td>
<td>Explicit first condition</td>
</tr>
<tr>
<td>IAT, RSES, SLSCR, GSE</td>
<td>RSES, SLSCR, GSE, IAT</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time 2</th>
<th>Mood</th>
<th>Intervention (resource vs. deficit interview)</th>
<th>Mood</th>
</tr>
</thead>
<tbody>
<tr>
<td>Implicit first condition</td>
<td>Explicit first condition</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IAT, RSES, SLSCR, GSE</td>
<td>RSES, SLSCR, GSE, IAT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Debriefing and resource interview for deficit interview condition</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


RESULTS

Correlations

Correlations are shown in Table 3 (Time 1 above the main diagonal, Time 2 below, retest-correlations in bold in the main diagonal). At Time 1, there were weak positive and partly significant correlations between the implicit and explicit self esteem measures (RSES, $r = .20, p = .12$; SL, $r = .25, p = .049$, SC, $r = .19, p = .15$, self-esteem, $r = .25, p = .052$). Implicit self-esteem correlated with mood ($r = .35, p = .01$). All explicit self-esteem measures were highly correlated, and
there were medium to high correlations between the explicit self-esteem measures, self-efficacy, and mood.

At Time 2, the correlations between implicit and explicit self-esteem were no longer significant (RSES, $r = .00$, ns; SL, $r = .07$, ns; SC, $r = .04$, ns, self-esteem, $r = .03$, ns). The difference between implicit-explicit self-esteem correlation at Time 1 and Time 2 almost reached significance ($Z = 1.73, p = .08$). Generally the correlations between the explicit self-report measures increased, whereas the correlations between the implicit and explicit self-esteem measures decreased.

**Order effects on implicit-explicit correlation**

At Time 1, the IAT correlated significantly or marginally significantly with the explicit self-esteem measures in the explicit first condition (RSES, $r = .34$, $p = .07$; SL, $r = .36$, $p = .05$; SC, $r = .33$, $p = .09$; self-esteem, $r = .41$, $p = .03$; $N = 29$) but not in the implicit first condition (RSES, $r = .07$, ns; SL, $r = .14$, ns; SC, $r = .04$, ns; self-esteem, $r = .09$, ns; $N = 32$). This order effect replicates that reported by Bosson et al. (2000). However, the correlation differences did not reach statistical significance ($Z = 1.28$, ns). There were no differences at Time 2.

**Reliability and stability**

As reported before, the internal consistency of the explicit measures were high throughout (ranging from .79 to .89), as were the retest-correlations or stabilities (ranging from .71 to .83). Across experimental conditions, the stability of the IAT was $r_{tt} = .61$, which is comparable to the retest-reliability observed by Bosson et al. (2000). This value is high compared with other IATs, as was the split-half reliability (for a review, see Steffens & Buchner, 2003), especially considering the new scoring algorithm tends to produce lower internal consistencies and retest-reliabilities. Using the old algorithm, the Spearman-Brown corrected split half reliabilities were .94 at Time 1 and .90 at Time 2; $r_{tt} = .76$. However, the stability of the implicit measure, but not of the explicit measures, was significantly lower ($Z = 1.99, p < .05$) in the deficit interview condition ($r_{tt} = .41$) than in the resource interview condition ($r_{tt} = .75$).
Table 3
Correlations between all variables at Times 1 (above the main diagonal) and 2 (below the main diagonal)
Retest-reliabilities are printed in bold italics in the main diagonal

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Self-Esteem IAT</td>
<td>.67</td>
<td>.20</td>
<td>.25*</td>
<td>.19</td>
<td>.26*</td>
<td>.25</td>
<td>.09</td>
<td>.35**</td>
</tr>
<tr>
<td>2. Rosenberg (RSES)</td>
<td>.00</td>
<td>.83</td>
<td>.73**</td>
<td>.47**</td>
<td>.70**</td>
<td>.92**</td>
<td>.50**</td>
<td>.43**</td>
</tr>
<tr>
<td>3. Self-Liking (SL)</td>
<td>.03</td>
<td>.80**</td>
<td>.79</td>
<td>.37**</td>
<td>.79**</td>
<td>.82**</td>
<td>.43**</td>
<td>.35**</td>
</tr>
<tr>
<td>4. Self-Competence (SC)</td>
<td>.04</td>
<td>.60**</td>
<td>.46**</td>
<td>.81</td>
<td>.87**</td>
<td>.73**</td>
<td>.68**</td>
<td>.21</td>
</tr>
<tr>
<td>5. Self-Competence-Self-Liking (SLSCR)</td>
<td>.05</td>
<td>.80**</td>
<td>.81**</td>
<td>.90**</td>
<td>.80</td>
<td>.93**</td>
<td>.68**</td>
<td>.33*</td>
</tr>
<tr>
<td>6. Self-Esteem (aggregate)</td>
<td>.03</td>
<td>.96**</td>
<td>.90**</td>
<td>.76**</td>
<td>.96**</td>
<td>.85</td>
<td>.64**</td>
<td>.41**</td>
</tr>
<tr>
<td>7. Self-Efficacy (GSE)</td>
<td>.10</td>
<td>.47**</td>
<td>.51**</td>
<td>.78**</td>
<td>.77**</td>
<td>.63**</td>
<td>.81</td>
<td>.34**</td>
</tr>
<tr>
<td>8. Mood</td>
<td>.06</td>
<td>.29*</td>
<td>.25</td>
<td>.21</td>
<td>.27*</td>
<td>.29*</td>
<td>.36**</td>
<td>.71</td>
</tr>
</tbody>
</table>

* $p < .05$; ** $p < .01$
Change in self-esteem

The following analyses were carried out using multiple regression analyses of residual change scores (i.e., using change in self-esteem after removing all variance that could be explained by initial values of the variables; Cohen & Cohen, 1975; Rogosa, Brandt, & Zimowski, 1982). Intervention (coded .5 for resource and -.5 for deficit interview) was then used as the predictor of this residual change. The means and variances of the variables at Times 1 and 2, and the change and the variance of change are reported in Table 4. With regard to the between group variable intervention (resource vs. deficit interview), the effect for explicit and implicit self-esteem and the manipulation check variables failed to reach significance. Separate t-tests on the pre-post differences indicated that there was weak positive change in self-liking in the resource ($t(29) = 2.01, p = .054$) and in the deficit interview condition ($t(32) = 2.48, p = .02$). The manipulation check of the two interview conditions showed the expected positive change in mood ($t(29) = 3.14, p < .01$) and a marginally significant positive change in self-efficacy ($t(28) = 1.98, p = .06$) in the resource interview condition, but contrary to expectation no negative change in the deficit interview condition.

Thus, whereas the hypothesis of an increase in explicit self-esteem in the resource interview condition received some support for the self-liking measure (Hypothesis 2b) there was no increase in implicit self-esteem in the resource condition (Hypothesis 2a) nor a significant decrease in implicit self-esteem in the deficit condition (Hypothesis 1a). Contrary to expectations, there was an increase in explicit self-esteem (Hypothesis 1b) in the deficit interview condition.

DISCUSSION

Correlations and order effects

Replicating other studies (e.g., Greenwald & Farnham, 2000; Bosson et al., 2000), the correlation between the self-esteem IAT and explicit self-esteem was positive but weak ($r = .25$), being significant in the explicit first condition but not in the implicit first condition. Bosson et al. (2000) suggested that focusing on explicit self-esteem primes
<table>
<thead>
<tr>
<th>Measure</th>
<th>N</th>
<th>Time 1</th>
<th>Time 2</th>
<th>Diff. T1/T2</th>
<th>Significance of difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>m</td>
<td>m</td>
<td>m</td>
<td>t-test</td>
</tr>
<tr>
<td></td>
<td></td>
<td>sd</td>
<td>sd</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-Esteem IAT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>all</td>
<td>61</td>
<td>1.01</td>
<td>.93</td>
<td>-.08</td>
<td>t(60) = -1.72, p = .09</td>
</tr>
<tr>
<td>resource</td>
<td>29</td>
<td>1.01</td>
<td>.94</td>
<td>-.06</td>
<td>t(28) = -1.01, p = .32</td>
</tr>
<tr>
<td>deficit</td>
<td>32</td>
<td>1.01</td>
<td>.91</td>
<td>-.10</td>
<td>t(31) = -1.38, p = .18</td>
</tr>
<tr>
<td>Rosenberg (RSES)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>all</td>
<td>61</td>
<td>4.20</td>
<td>4.21</td>
<td>.01</td>
<td>t(60) = .25, p = .80</td>
</tr>
<tr>
<td>resource</td>
<td>29</td>
<td>4.20</td>
<td>4.24</td>
<td>.04</td>
<td>t(28) = .77, p = .45</td>
</tr>
<tr>
<td>deficit</td>
<td>32</td>
<td>4.20</td>
<td>4.18</td>
<td>-.02</td>
<td>t(31) = -.33, p = .74</td>
</tr>
<tr>
<td>Self-Liking (SL)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>all</td>
<td>61</td>
<td>3.71</td>
<td>3.89</td>
<td>.18</td>
<td>t(60) = 3.20, p = .00</td>
</tr>
<tr>
<td>resource</td>
<td>29</td>
<td>3.76</td>
<td>3.94</td>
<td>.17</td>
<td>t(28) = 2.01, p = .054</td>
</tr>
<tr>
<td>deficit</td>
<td>32</td>
<td>3.66</td>
<td>3.85</td>
<td>.19</td>
<td>t(31) = 2.48, p = .02</td>
</tr>
<tr>
<td>Self-Competence (SC)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>all</td>
<td>61</td>
<td>3.43</td>
<td>3.40</td>
<td>-.02</td>
<td>t(60) = -.47, p = .64</td>
</tr>
<tr>
<td>resource</td>
<td>29</td>
<td>3.43</td>
<td>3.47</td>
<td>.04</td>
<td>t(28) = .62, p = .54</td>
</tr>
<tr>
<td>deficit</td>
<td>32</td>
<td>3.42</td>
<td>3.34</td>
<td>-.08</td>
<td>t(31) = -1.09, p = .23</td>
</tr>
<tr>
<td>Self-Esteem (aggregate)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>all</td>
<td>61</td>
<td>.90</td>
<td>.92</td>
<td>.00</td>
<td>t(28) = .46, p = .65</td>
</tr>
<tr>
<td>resource</td>
<td>29</td>
<td>.93</td>
<td>.08</td>
<td>.04</td>
<td>t(31) = -.40, p = .70</td>
</tr>
<tr>
<td>deficit</td>
<td>32</td>
<td>-.03</td>
<td>.07</td>
<td>-.04</td>
<td>t(31) = .49, p = .49</td>
</tr>
<tr>
<td>Self-Efficacy (GSE)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>all</td>
<td>61</td>
<td>3.38</td>
<td>3.44</td>
<td>.06</td>
<td>t(60) = 1.49, p = .14</td>
</tr>
<tr>
<td>resource</td>
<td>29</td>
<td>3.46</td>
<td>3.58</td>
<td>.13</td>
<td>t(28) = 1.98, p = .06</td>
</tr>
<tr>
<td>deficit</td>
<td>32</td>
<td>3.31</td>
<td>3.32</td>
<td>.56</td>
<td>t(31) = .11, p = .9</td>
</tr>
<tr>
<td>Mood</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>all</td>
<td>61</td>
<td>4.17</td>
<td>4.28</td>
<td>.11</td>
<td>t(60) = 2.12, p = .04</td>
</tr>
<tr>
<td>resource</td>
<td>29</td>
<td>4.18</td>
<td>4.37</td>
<td>.19</td>
<td>t(28) = 3.14, p = .00</td>
</tr>
<tr>
<td>deficit</td>
<td>32</td>
<td>4.16</td>
<td>4.20</td>
<td>.05</td>
<td>t(31) = .54, p = .60</td>
</tr>
</tbody>
</table>
thoughts and influences a subsequent test of implicit self-esteem, increasing the correlation (see also Joncs et al., 2002). An alternative explanation is that when there is little motivation or capacity for reflection, explicit self-esteem measures will largely reflect implicit self-esteem, whereas the likelihood of a dissociation increases when there is motivation and capacity to override the effect of implicit self-esteem (Fazio & Olson, 2003; Koole, Dijkstra, & van Knippenberg, 2001; Koole & Pehlam, 2003; Wilson et al., 2000). In consequence, assessing implicit self-esteem first may lead to more self-reflection when completing the explicit measures, causing these responses to dissociate from their implicit responses. In line with this explanation, explicit first correlations were higher only at Time 1. At Time 2 the participants were not only being assessed for the second time, but had also had the self-reflecting interview.

Dissociation between implicit and explicit self-esteem

The correlations between implicit self-esteem and the explicit self-esteem measures decreased between Times 1 and 2, whilst the correlations between the explicit self-esteem measures increased. In addition the positive correlation between implicit self-esteem and mood at Time 1 disappeared at Time 2. Again, it is plausible to assume that the interview prompted participants to self-reflect and build a more consistent explicit self-esteem, but which dissociated from implicit self-esteem.

Malleability of explicit self-esteem

Overall there was less change in explicit self-esteem than expected and only for self-liking, which increased in both interview conditions. That self-liking also increased in the deficit interview condition was unexpected. We hypothesized that due to self-presentational self-enhancement and self-consistency, explicit self-esteem would be resistant to the negative self-relevant information and not decrease. The increase in the deficit interview condition may have a number of reasons. Firstly, the increases only for explicit self-esteem, in both interviews conditions, could have been due to self-presentational self-enhancement overriding any self-consistency needs, the act of self-reflection prompting the participants to self-enhance in the explicit but not in the implicit mode.
Furthermore, the threat to self-esteem in the deficit interview may have prompted self-presentational self-enhancement as a (over)compensation for the negative self-relevant information.

Secondly, for ethical reasons, in the deficit interview the interviewers were instructed to focus on participants' deficits but not to be critical towards the person. As a result, the participants could reflect and talk about themselves, and yet feel accepted as a person. It has been documented that talking about oneself in an atmosphere of being unconditionally positively regarded has positive effects on self-esteem (e.g., Rogers, 1959). This interpretation is supported by verbal feedback provided from the participants in informal post experimental interviews. The positive effect on self-liking may therefore not have been due to the content of the interview rather the form, that is, that someone was interested and willing to listen to someone being open about themselves. Furthermore, the interviewers themselves may have involuntarily compensated for the negative content of the deficit interview with more positive nonverbal behavior. Whether these findings challenge the validity of the deficit orientated interview as a method of decreasing self-esteem will be discussed below.

In summary, the results suggest that, at least for self-liking, explicit self-esteem is malleable, implying that given the opportunity to self-reflect people will self-enhance at the explicit level. However, whereas participants could present a better picture of how they like themselves, they were more consistent in their presentation of their competencies between Times 1 and 2. This seems to have been independent of the interview focus. These differential results for self-liking and self-competence underlie the rational for assessing both components of self-esteem.

Malleability of implicit self-esteem

On one hand there was no group level change in implicit self-esteem in either the resource or deficit interview condition, suggesting that the IAT-measure was resistant to the manipulation. This finding supports the argument that implicit self-esteem is overlearned and stable rather than being sensitive to the activation of positive or negative self-relevant information in the form of an interview as we hypothesized. That the self-liking aspect of explicit self-esteem increased in both interview
conditions implies further that implicit self-esteem is less malleable than explicit self-esteem at the group level.

However, the lower stability of implicit self-esteem in the deficit interview condition, which was not observed for explicit self-esteem, suggests that the deficit interview did cause differential change in implicit self-esteem, that is, change at the individual level. An alternative explanation of this result is that the correlation difference was due to a decrease of the implicit self-esteem variance in the deficit interview condition. As can be seen in Table 4, the variance for the IAT was in fact lower at time 2, but in both interview conditions. A test of the homogeneity of variances in the deficit interview was marginally significant ($\rho < .10$) and hence inconclusive in eliminating the alternative explanation. However, when the problem of different variance differences is avoided by using the Spearman rank correlation, the results remain virtually identical to those calculated with Pearson correlations: $r_{uu} = .46$ in the deficit, and $r_{dd} = .73$ in the resource interview condition. This result strongly supports the interpretation that there was more differential change in the IAT in the deficit than in the resource interview condition. Thus, although there was no expected decrease in implicit self-esteem in the deficit interview condition at the group level as hypothesized, there is evidence for substantial change at the individual level. This result suggests that implicit self-esteem may indeed be more malleable when negative self-relevant information is made available. Of course it is necessary to replicate this finding. Moreover, it seems to be an interesting avenue of further research to investigate the person factors that moderate such change when people are confronted with negative self-relevant information.

The manipulation

As the malleability findings are dependent on the manipulation used, the suitability of the two interviews as a method of self-esteem manipulation needs to be considered. Firstly, our manipulation was explicit. Participants were aware about the self-related content of the interview, and encouraged to reflect upon themselves, thus making this process susceptible to self-verifying and self-schema-driven processes of selective perception and interpretation (e.g., Markus, 1977; Swann, 1990). Therefore, consistency effects may not only have occurred during the
assessment stage but also during the interview itself. Implicit manipulations, such as subliminal priming or evaluative conditioning, can circumvent these processes, as participants are unaware of being manipulated and have no call to reflect. There is also some recent evidence that implicit manipulations can change both explicit (state) (e.g., Riketta & Dauenheimer, 2003) and implicit self-esteem (e.g., Dijksterhuis, 2004).

Secondly, our manipulation focused on actual and available aspects of self rather than giving bogus feedback. The latter may be more effective as it necessarily shocks the individual with new and unexpected information, being both false yet ‘objective’, but it stretches ethics to the limit. Thirdly, our manipulation was interpersonal. Not only may the nonverbal behavior of the interviewers have been influential, but the participants may have been influenced by the interpersonal situation itself (e.g., Mead, 1934). Alternatively, participants could be asked to write about positive and negative aspects of themselves (e.g., Jones et al., 2002).

Finally, our manipulation was highly ecologically valid, being similar to a diagnostic interview or a counseling session, and was the most extreme positive and negative manipulation that we judged to be ethical. Attempting to reduce self-esteem would pang the conscience of any psychologist, but attempting to raise self-esteem, particularly in a way that would hardly ever happen in real life, may also be problematic, particularly as high self-esteem is not without its dangers (e.g., Baumeister, 1998; Kernis, 2003).

It may be suggested that our interviews were just not strong or valid enough, particularly the deficit interview. However, in the interviews the participants did have to confront themselves with and reflect on either positive or negative self-relevant information and there is some evidence that the manipulation did influence self-esteem. Firstly, there was positive change in self-liking independent of the interview focus. Secondly, there was positive change in the manipulation check variables, mood and self-efficacy in the resource interview condition. This suggests that the interviews, and particularly the resource interview, had a weak positive effect on the self-report measures. The increases in self-liking and self-efficacy reflect increases in trait measures, which are by definition difficult to change and more susceptible to self-consistency effects as participants are asked to give a general picture of their selves. We would expect larger increases in state measures, particularly as there
Malleability of self-esteem

was also an increase in mood (e.g., Heatherton & Polivy, 1991). Thirdly, the lower stability of the IAT in the deficit interview suggests there was more differential change in implicit self-esteem in the deficit than in the resource interview condition.

Together these results indicate that the two interviews not only had an effect, but also had differential effects. Change in both implicit and explicit self-esteem was nevertheless small at both the group and individual level, suggesting either that neither are very malleable, or alternatively that the manipulation was too weak. However, if only stronger manipulations can substantially change implicit and explicit self-esteem, this would only underline that implicit and explicit self-esteem are relatively stable. Without a direct comparison of different manipulations with the same self-esteem measures under otherwise equal conditions the question of whether the results are due to the lacking strength of the manipulation or the robustness of implicit and explicit self-esteem cannot be conclusively answered.

Another important issue is that of the validity and reliability of the measures used, particularly the IAT. Farnham et al. (1999) suggested that (explicit) self-esteem measures may be capturing the wrong construct and this suspicion has indeed motivated implicit measurement of self-esteem. However, how sure are we that the IAT is measuring the right construct? Our interpretations on the malleability of implicit self-esteem, and the comparison to the malleability of explicit self-esteem, are of course dependent on the IAT’s reliability and validity. The split-half reliability as well as the retest-reliability of the IAT reached levels of established questionnaire measures. The present results also indicate convergent validity of the IAT as it did correlate significantly with explicit self-esteem measures, although the correlation was dependent on the order of implicit and explicit measures. Overall, the available evidence suggests that the self-esteem IAT has almost satisfactory reliability and there is evidence confirming its convergent validity with explicit self-esteem measures as well as its predictive behavior. Strong evidence for the incremental validity of the self-esteem IAT as demonstrated for a different implicit self-esteem measure (Spalding & Harding, 1999) has yet to be convincingly shown.

To summarize, there are two main interpretations for the general result that there was only weak change in implicit and explicit self-esteem. Firstly, explicit and implicit self-esteem are largely resistant against manipulation, i.e., they are both stable constructs and not very
malleable. Secondly, one or both are malleable, but the manipulation and particularly the deficit interview was too weak to substantially change one or both. However, if a stronger manipulation as used here is necessary to successfully change self-esteem, this again would indicate that self-esteem is fairly robust against manipulation.

CONCLUSION

As self-esteem is so fundamental to an individual’s psychological functioning, so are the conditions under which it changes. Clearly, malleability is dependent on the manipulation used and may be moderated by other variables. Thus, in addition to asking whether implicit or explicit self-esteem is more malleable, a further interesting question is under which boundary conditions, by which means, and to what extent are they malleable. The comparison of implicit and explicit self-esteem with regard to different manipulations and moderators is an interesting avenue of future research, particularly if dissociations can be shown to be dependent on the type of manipulation, and thus would add to the increasing support for the construct validity of implicit self-esteem.

ACKNOWLEDGEMENTS

We would like to thank Franziska Becker, Anett Dehn, Stephanie Lemberg, Andrea Naß, Stefanie Reichert, Franziska Jentsch, Konrad Schnabel and Dorota Uzdawinis for their help with this project and to Bertram Gawronski and the anonymous reviewers for helpful comments on an earlier version of the manuscript.

RÉSUMÉ

Dans une plan expérimental prête test – post-test, l’estime de soi implicite et l’estime de soi explicite sont manipulées grâce à une interview centrée sur les ressources ou sur les déficits des participants. L’estime de soi implicite est mesurée avec le Test d’Association Implicite (IAT, Greenwald, McGhee et Schwartz, 1998) et l’estime de soi explicite par
l'intermédiaire de questionnaires standard. Nous supposons que l'estime de soi implicite diminuera plus que l'estime de soi explicite dans la condition d'interview « déficits » parce que les besoins de présentation et de consistance de soi rendront l'estime de soi explicite plus résistante à l'information négative auto-pertinente. Une augmentation modérée de l'estime de soi explicite aussi bien qu'implicite est attendue dans la condition d'interview « ressources ». Globalement, ces hypothèses n'ont été que peu validées. Il y a eu une augmentation de la dimension « amour de soi » de l'estime de soi explicite, mais dans les deux conditions d'interview. De plus, la stabilité de l'IAT significativement plus faible dans la condition d'interview « déficits » par rapport à la condition d'interview « ressources » suggère que l'interview « déficits » provoque un changement différentiel substantiel de l'estime de soi implicite. L'étude montre que l'estime de soi IAT a théoriquement des relations significatives avec les mesures explicites du soi et est relativement résistante aux manipulations.

REFERENCES


Received 17 November, 2003
Accepted 15 February, 2004