



Contents lists available at ScienceDirect

## Journal of Experimental Social Psychology

journal homepage: [www.elsevier.com/locate/jesp](http://www.elsevier.com/locate/jesp)

## Reports

Understanding patterns of attitude change: When implicit measures show change, but explicit measures do not<sup>☆</sup>Bertram Gawronski<sup>\*</sup>, Etienne P. LeBel

Department of Psychology, University of Western Ontario, Social Science Centre, London, Ont., Canada N6A 5C2

## ARTICLE INFO

## Article history:

Received 13 June 2007

Revised 26 March 2008

Available online 25 April 2008

## Keywords:

Attitude change

Evaluative conditioning

Implicit measures

Introspection

Mere ownership

## ABSTRACT

A common assumption in research on attitudes is that indirect measures assess relatively stable implicit attitudes, whereas traditional self-report measures assess more recently acquired explicit attitudes that coexist with old, presumably stable implicit attitudes. This assumption seems difficult to reconcile with research showing experimentally induced changes on implicit but not explicit measures. The present research tested a process-account of such asymmetrical patterns. Specifically, we argue that implicit measures show experimental effects that do not emerge on explicit measures when (a) the pairing of an attitude object with positive or negative valence creates new automatic associations in memory, and, at the same time, (b) the consideration of additional information about the attitude object eliminates the impact of automatic associations on self-reported evaluative judgments. Results from three studies support these predictions. Implications for research on attitude change are discussed.

© 2008 Elsevier Inc. All rights reserved.

Since Thurstone (1928) declared that “attitudes can be measured,” the attitude construct has become one of the most important concepts in social psychology. More recently, research on attitudes has experienced a second measurement revolution with the development of a new class of indirect measures (for reviews, see Fazio & Olson, 2003; Wittenbrink & Schwarz, 2007).<sup>1</sup> These measures differ from traditional self-report measures, in that they do not require explicit evaluations of an attitude object. Rather, evaluations inferred from these measures are based on participants’ performance on experimental paradigms, such as sequential priming (Neely, 1977) or response interference tasks (Kornblum, Hasbroucq, & Osman, 1990). Examples of these measures include, among others, the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998), affective priming (Fazio, Jackson, Dunton, & Williams, 1995),

and the Affect Misattribution Procedure (AMP; Payne, Cheng, Govorun, & Stewart, 2005).

A common assumption in research on attitudes is that indirect measures provide access to relatively stable implicit attitudes that have their roots in early socialization experiences. In contrast, traditional self-report measures are often assumed to assess more recently acquired explicit attitudes that coexist with the old, presumably stable implicit attitude (e.g., Gregg, Seibt, & Banaji, 2006; Petty, Tormala, Brinol, & Jarvis, 2006; Rudman, Phelan, & Heppen, 2007; Rydell & McConnell, 2006; Wilson, Lindsey, & Schooler, 2000). Even though these assumptions are consistent with research showing that self-reported evaluations are sometimes easier to change than indirectly assessed evaluations (e.g., Gawronski & Strack, 2004; Gregg et al., 2006; Petty et al., 2006; Rydell & McConnell, 2006), there is a significant body of research showing exactly the opposite pattern (e.g., Dasgupta & Greenwald, 2001; Karpinski & Hilton, 2001; Olson & Fazio, 2006; for a review, see Gawronski & Bodenhausen, 2006). The latter findings seem difficult to reconcile with the assumption that self-report measures reflect more recently acquired explicit attitudes that coexist with older, highly stable implicit attitudes tapped by indirect measures.

The main goal of the present research was to test a process-account of experimental effects on implicit but not explicit measures. In line with contemporary dual-process models (e.g., Fazio & Olson, 2003; Gawronski & Bodenhausen, 2006), we argue that implicit measures provide a proxy for automatic associations in memory, which may or may not influence verbal judgments reflected in self-report measures (for a review, see Hofmann, Gschwendner, Nosek, & Schmitt, 2005). According to these models, the influence

<sup>☆</sup> The present research was supported by Canada Research Chairs Program Grant 202555, Social Sciences and Humanities Research Council of Canada Grant 410-2005-1339, and Academic Development Fund of the University of Western Ontario Grant 05-303. Portions of this research were presented at the 7th Annual Conference of the Society for Personality and Social Psychology, Palm Springs, CA; the 78th Annual Meeting of the Midwestern Psychological Association, Chicago, IL; the 2006 Duck Conference on Social Cognition, Corolla, NC; and the 7th Attitudes Preconference at the 8th Annual Conference of the Society for Personality and Social Psychology, Memphis, TN. We would like to thank Keith Payne for providing the stimulus materials for the Affect Misattribution Procedure employed in Experiment 3; and Xenia Avvakumova and Arjun Sharma for their help in collecting the data.

<sup>\*</sup> Corresponding author.

E-mail address: [bgawrons@uwo.ca](mailto:bgawrons@uwo.ca) (B. Gawronski).

<sup>1</sup> In the present article, we use the terms *direct* and *indirect* to describe features of measurement procedures, and the terms *explicit* and *implicit* to depict the measurement scores obtained with direct and indirect measures.

of automatic associations on evaluative judgments depends, among other factors, on the consideration of additional information over and above automatic associations. Thus, if people have the motivation and the opportunity to deliberate about additional information, this information may either reduce the perceived validity of automatic associations (Gawronski & Bodenhausen, 2006) or simply dilute their impact on evaluative judgments (Fazio & Olson, 2003). Thus, if a given factor, such as the pairing of an attitude object with positive or negative valence, creates new automatic associations in memory, these changes in automatic associations may not be picked up by verbal self-reports if additionally considered information eliminates the impact of automatic associations on evaluative judgments.<sup>2</sup> From this perspective, implicit measures may show experimental effects that do not emerge on explicit measures when (a) the pairing of an attitude object with positive or negative valence creates new automatic associations in memory, and, at the same time, (b) the consideration of additional information about the attitude object eliminates the impact of automatic associations on self-reported evaluative judgments.

Even though correlational research supports the assumption that enhanced motivation to deliberate reduces the impact of automatic associations on evaluative judgments (for a review, see Hofmann et al., 2005), the significance of the proposed processes has rarely been investigated in the context of attitude change (for notable exceptions, see Gawronski & Strack, 2004; Schuette & Fazio, 1995). The present research goes beyond earlier studies by orthogonally manipulating (a) automatic associations of an attitude object, and (b) the consideration of additional information about that object. Our prediction was that explicit measures should be as susceptible to our manipulation of automatic associations as implicit measures, if participants do not consider additional information about the attitude object. If, however, participants are encouraged to consider additional information, our manipulation of automatic associations should influence only implicit measures, but not explicit measures. Moreover, in line with the assumption that automatic associations influence evaluative judgments unless their impact is reduced by other information, implicit and explicit measures should be highly correlated when participants do not consider additional information about the attitude object. However, correlations should be significantly lower when participants are encouraged to consider additional information.

## Experiment 1

The main goal of Experiment 1 was to pre-test the effectiveness of an experimental procedure designed to manipulate the impact of automatic associations on evaluative judgments. Previous research on the correspondence between automatic associations and evaluative judgments emphasized the moderating role of individual differences or object characteristics (see Hofmann et al., 2005). However, there is surprisingly little research that employed an experimental approach (for notable exceptions, see Gawronski & Strack, 2004; Nier, 2005). In the present research, we adapted a procedure from Wilson and Dunn (1986) in which participants are asked to introspect on either their feelings for the attitude object or on reasons why they like or dislike the attitude object. We expected that a focus on feelings would reduce the consideration of additional information over and above automatic associations. In contrast, a focus on reasons should encourage participants to consider additional information, which often deviates from the evalu-

ation implied by automatic associations (Wilson, Dunn, Kraft, & Lisle, 1989). Hence, correlations between implicit and explicit measures were expected to be higher when participants are instructed to introspect on their feelings than when they are instructed to introspect on reasons for their preferences.

## Methods

### Participants and design

Hundred and sixty-four undergraduates (107 female) participated in a study on personal preferences in exchange for course credit. Participants were randomly assigned to one of the two introspection conditions (feelings vs. reasons). Due to a computer malfunction, data from seven participants were only partially recorded and were thus excluded from analyses.

### Procedure and measures

As a measure of automatic associations, we used an IAT (Greenwald et al., 1998) designed to assess implicit preferences for Coke over Pepsi. In the first block, pictures of Pepsi and Coke had to be assigned to the categories *Pepsi* (left) or *Coke* (right). In the second block, participants were presented with positive and negative words (e.g., *paradise*, *sickness*) that had to be classified according to the categories *negative* (left) and *positive* (right). In the third block, target and attribute trials were presented in alternating order, with Pepsi pictures and negative words on the left key and Coke pictures and positive words on the right key. In the fourth block, participants again had to categorize individual pictures of *Coke* and *Pepsi*, now with a reversed key assignment. Finally, the fifth block again combined the two types of trials, with Coke pictures and negative words on the left key and Pepsi pictures and positive words on the right key. Blocks 1, 2, and 4 each consisted of a total of 20 trials; blocks 3 and 5 each comprised 80 trials. The inter-trial interval following correct responses was 250 ms. Incorrect responses were indicated with the word "ERROR!" appearing for 1000 ms in the center of the screen.

After the IAT, participants were instructed to consider either (a) how they feel about Coke versus Pepsi or (b) why they prefer either Coke or Pepsi. Participants in both conditions were asked to write down their responses in a text-box on the computer screen. Immediately after the introspection task, participants were asked to indicate how much they like Coke and Pepsi on two scales ranging from 1 (*not at all*) to 7 (*very much*).

## Results

IAT data were aggregated using the D-600 algorithm (Greenwald, Nosek, & Banaji, 2003). Scores were calculated such that higher values indicate a stronger preference for Coke over Pepsi (Cronbach's  $\alpha = .72$ ). A corresponding index was calculated for self-reported evaluations by subtracting likeability ratings for Pepsi from likeability ratings for Coke. One-way ANOVAs revealed that implicit preference scores ( $M_s = .19$  vs.  $.28$ ),  $F(1, 155) = 1.79$ ,  $p = .18$ ,  $\eta^2 = .011$ , as well as explicit preference scores ( $M_s = .14$  vs.  $.43$ ),  $F(1, 155) = 0.59$ ,  $p = .44$ ,  $\eta^2 = .004$ , did not differ across introspection conditions. However, consistent with our predictions, correlations between the two measures did differ across introspection conditions,  $z = 2.27$ ,  $p = .02$ . Whereas implicit and explicit preference scores were significantly correlated when participants were asked to introspect on their feelings ( $r = .51$ ,  $p < .001$ ), the two scores were only weakly related when participants were asked to introspect on reasons ( $r = .19$ ,  $p = .08$ ).

## Discussion

Results from Experiment 1 support the effectiveness of different introspection foci in influencing the impact of automatic associa-

<sup>2</sup> Note that the consideration of additional information may not always reduce the impact of automatic associations on verbal self-reports. Rather, a reduction may occur only when the additionally considered information is inconsistent, but not when it is consistent with the evaluation implied by automatic associations (Gawronski & Bodenhausen, 2006).

tions on evaluative judgments. Consistent with this claim, implicit and explicit measures were highly correlated when participants were asked to introspect on their feelings for the attitude objects. However, correlations were significantly lower when participants were asked to think about reasons for their preference. In Experiment 2, we applied this finding to the present question of asymmetrical effects on implicit and explicit measures by orthogonally manipulating automatic associations and focus of introspection.

## Experiment 2

The objective of Experiment 2 was to test our main prediction that implicit measures should show effects that do not emerge on explicit measures when (a) the pairing of an attitude object with positive or negative valence creates new automatic associations, and (b) the consideration of additional information eliminates the impact of automatic associations on self-reported evaluative judgments. To manipulate automatic associations, Experiment 2 employed an evaluative conditioning (EC) paradigm. EC effects are often explained by processes of associative learning, in which repeated pairings of a conditioned stimulus (CS) with positive or negative unconditioned stimuli (US) create new automatic associations pertaining to the CS, thereby leading to an associative transfer of US valence to the CS (for reviews, see De Houwer, Thomas, & Baeyens, 2001; Walther, Nagengast, & Traselli, 2005). In the present study, participants were repeatedly presented with pairings of familiar CS with positive or negative US. Immediately afterwards, participants were asked to consider either (a) their feelings for the CS, or (b) their knowledge about the CS. Finally, all participants completed an explicit and an implicit measure of CS evaluations. We expected that our EC manipulation would show effects on the implicit measure irrespective of whether participants focused on their feelings or their knowledge. In contrast, the explicit measure should be influenced by the EC manipulation only when participants were asked to focus on their feelings, but not when they were asked to think about their knowledge about the attitude object. Moreover, implicit and explicit measures were expected to be highly correlated when participants focused on their feelings, whereas correlations should be significantly lower when participants focused on their knowledge.

## Methods

### Participants and design

Eighty undergraduates (50 female) participated in a study on personal preferences in exchange for course credit. Participants were randomly assigned to one of the four conditions of a 2 (introspection: feelings vs. knowledge)  $\times$  2 (CS-US pairings: Europe-positive/Asia-negative vs. Europe-negative/Asia-positive) between-subjects design.

### Evaluative conditioning

The EC paradigm was adapted from Dijksterhuis (2004) and involved subliminal presentations of the CS which were followed by supraliminal presentations of positive or negative US. As CS, we used the words *Europe* and *Asia*; as US we used 20 positive and 20 negative adjectives (e.g., *nice*, *ugly*). The conditioning trials were included in a lexical-decision task that involved a categorization of the US adjectives and meaningless non-words as either meaningful or meaningless. On each trial of the task, participants were first presented with a masking stimulus (XXXXXX) that appeared for 500 ms on the screen. The masking stimulus was then replaced by the word *Europe* or *Asia* (CS) which was presented for 17 ms. Immediately afterwards, a positive or negative adjective (US) appeared on the screen. Depending

on the experimental condition, *Europe* was consistently followed by a positive (negative) word whereas *Asia* was consistently followed by a negative (positive) word. The critical conditioning trials (20 for each of the two CS) were interspersed with 40 trials that included a 17 ms presentation of a control stimulus (*Chair*) which was followed by a non-word. Participants' task was to indicate as quickly as possible whether the stimulus presented on the screen was a meaningful word or a meaningless non-word. The inter-trial interval was 1000 ms.

### Introspection

Participants were asked to either (a) take a moment to think about their feelings for Europe and Asia, or (b) take a moment to think about what they know about Europe and Asia. Participants in both conditions were instructed to write down their responses in a text-box on the computer screen.

### Measures

An IAT was used as to assess implicit preferences for Europe over Asia. The general procedure of the IAT was identical to the one employed in Experiment 1; the only exception being that participants were required to categorize names of 10 European and 10 Asian countries instead of pictures of Coke and Pepsi (e.g., *Germany*, *Japan*). The self-report measure was also identical to Experiment 1; the only difference being that participants were asked to indicate how much they like Europe and Asia rather than Coke and Pepsi. The order of the two measures was counterbalanced across participants.

## Results

### Implicit measure

IAT data were aggregated using the D-600 algorithm (Greenwald et al., 2003). Scores were calculated such that higher values indicated a stronger preference for Europe over Asia (Cronbach's  $\alpha = .76$ ). Submitted to a 2 (CS-US pairings)  $\times$  2 (introspection) ANOVA, IAT scores revealed a significant conditioning effect,  $F(1,76) = 4.05$ ,  $p < .05$ ,  $\eta^2 = .051$  (see Fig. 1). Specifically, participants showed a stronger preference for Europe over Asia when the word *Europe* was repeatedly paired with positive words and the word *Asia* was repeatedly paired with negative words than when *Europe* was repeatedly paired with negative words and *Asia* was repeatedly paired with positive words. No other main or interaction effect reached statistical significance (all  $F$ s  $< 1$ ).

### Explicit measure

In keeping the scoring direction of the IAT, explicit preference scores were calculated by subtracting likeability ratings for Asia

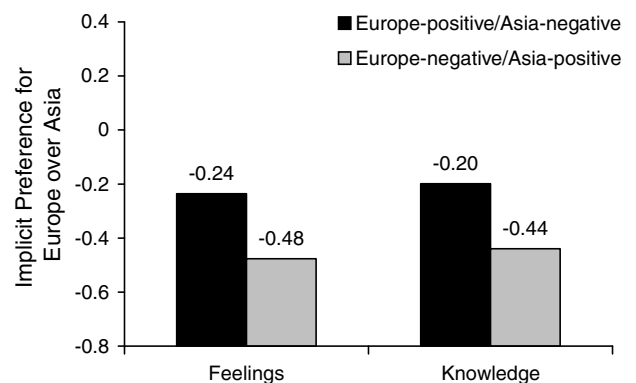


Fig. 1. Mean values of implicit preference for Europe over Asia as a function of evaluative conditioning (Europe-positive/Asia-negative vs. Europe-negative/Asia-positive) and introspection focus (feelings vs. knowledge), Experiment 2.

from likeability ratings for Europe. Submitted to a 2 (CS–US pairings)  $\times$  2 (introspection) ANOVA, these scores revealed a significant two-way interaction,  $F(1,76) = 5.84$ ,  $p = .02$ ,  $\eta^2 = .071$ . Consistent with our predictions, conditioning effects emerged only when participants focused on their feelings, but not when they focused on their knowledge (see Fig. 2). Specifically, participants in the feelings condition tended to show a stronger preference for Europe over Asia when Europe was repeatedly paired with positive words and Asia was repeatedly paired with negative words than when Europe was repeatedly paired with negative words and Asia was repeatedly paired with positive words,  $F(1,38) = 3.62$ ,  $p = .06$ ,  $\eta^2 = .087$ . In contrast, participants in the knowledge condition showed a non-significant tendency in the opposite direction,  $F(1,38) = 2.46$ ,  $p = .12$ ,  $\eta^2 = .061$ .

#### Implicit–explicit correlations

Overall, explicit and implicit preference scores showed a non-significant positive correlation ( $r = .18$ ,  $p = .11$ ). However, consistent with our predictions, correlations significantly differed across the two introspection conditions,  $z = 3.87$ ,  $p < .001$ . Whereas participants in the feelings condition showed a significant positive correlation between preference scores ( $r = .48$ ,  $p = .002$ ), participants in the knowledge condition showed a significant negative correlation ( $r = -.36$ ,  $p = .02$ ).

#### Discussion

Results from Experiment 2 support our process-account of asymmetrical effects on implicit and explicit measures. In the present study, repeated pairings of CS with positive or negative US influenced an explicit measure of CS evaluations only when participants focused on their feelings regarding the CS, but not when they focused on their knowledge about the CS. In contrast, our implicit measure was influenced by the EC manipulation regardless of whether participants focused on feelings or knowledge. Moreover, implicit and explicit measures showed a significant positive correlation when participants focused on their feelings. However, correlations between the two measures were reversed when participants focused on their knowledge. These results suggest that EC effects on evaluative judgments are contingent upon whether EC-related changes in automatic associations further influence self-reported evaluative judgments. Such an influence seems particularly likely when participants focus on their feelings for the attitude object. However, the impact of automatic associations on evaluative judgments seems to be lower when participants consider additional information about the attitude object.

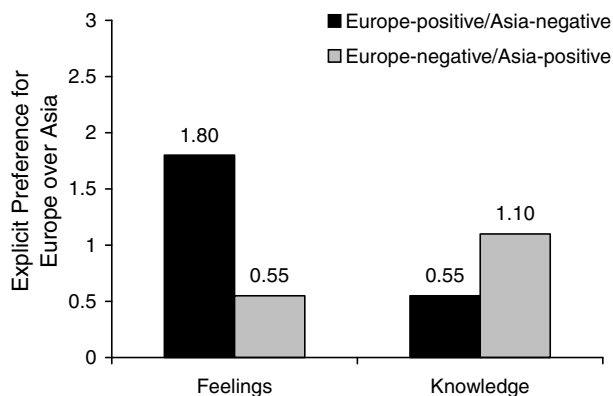


Fig. 2. Mean values of explicit preference for Europe over Asia as a function of evaluative conditioning (Europe-positive/Asia-negative vs. Europe-negative/Asia-positive) and introspection focus (feelings vs. knowledge), Experiment 2.

### Experiment 3

Walther et al. (2005) recently argued that the self can function as a US, such that objects that are “paired” with the self automatically acquire the valence of the self. Given that most people have a positive evaluation of themselves (e.g., Bosson, Swann, & Pennebaker, 2000; Greenwald & Farnham, 2000; Koole, Dijksterhuis, & Van Knippenberg, 2001), this positive evaluation may automatically transfer to any object that becomes associated with the self. This idea was recently applied to ownership effects (Beggan, 1992) by Gawronski, Bodenhausen, and Becker (2007). Specifically, Gawronski et al. argued that ownership creates an automatic association between owned objects and the self, thereby leading to an associative transfer of automatic self-evaluations to owned objects. In line with these assumptions, Gawronski et al. found that automatic evaluations of newly acquired objects became more favorable as a result of ownership, with automatic object evaluations being positively related to automatic evaluations of the self.

Drawing on Gawronski et al.’s (2007) associative account of ownership effects, Experiment 3 investigated the emergence of asymmetrical ownership effects on implicit and explicit measures. For this purpose, participants were randomly given one of two different pictures as a special token of appreciation for their participation. Immediately afterwards, half of the participants were asked to introspect on their feelings toward the two pictures. The remaining half were asked to think about why they prefer one picture over the other (Wilson & Dunn, 1986). Finally, all participants completed an implicit and an explicit measure of evaluations of the two pictures. In line with our earlier findings, we expected that ownership leads to more positive evaluations of owned compared to non-owned pictures on the implicit measure irrespective of whether participants introspected on feelings or reasons. In contrast, ownership effects on the explicit measure should occur only when participants were asked to introspect on their feelings, but not when they were asked to introspect on reasons for their preference. Moreover, implicit and explicit measures were expected to be highly correlated when participants were asked to focus on their feelings, whereas correlations should be significantly lower when participants were asked to think about reasons for their preference.

#### Methods

##### Participants and design

Eighty-one undergraduates (44 female) participated in a study on attention, attitudes, and social judgment in return for course credit. The experiment consisted of a 2 (object: owned vs. not owned)  $\times$  2 (introspection: feelings vs. reasons) mixed-model design, with the first variable as a within-subjects factor and the second as a between-subjects factor. Data from six participants who did not contact the experimenter to receive their picture (see below) were excluded from analyses.

##### Ownership

The ownership manipulation closely followed the one employed by Gawronski et al. (2007, Experiment 4). After participants completed several tasks that were unrelated to the present study, they were told that they would be given one of two color prints as a gift in gratitude for their participation. For this purpose, we chose two large color postcards (14  $\times$  20 cm) from the series *Earth from Above* by Yann Arthus Bertrand (published by AMI-Images). Based on pretests, we selected two relatively similar, yet sufficiently distinct pictures of aerial shots of a dromedary caravan in the desert (i.e., *Dromedary caravans near Nouakchott, Mauritania* and *Dromedary caravan in the dunes near Nouakchott, Mauritania*). Which of



the two picture participants received was determined randomly by the experimenter by rolling a dice. After the picture was determined, participants were given their postcard and asked to continue with the experiment.

### Introspection

The introspection manipulation was identical to the one employed in Experiment 1. Half of the participants were asked to consider their feelings toward the two pictures; the remaining half were asked to consider why they preferred a particular picture. During the instruction, the two pictures were presented on the screen to facilitate the introspection task. Participants in both conditions were asked to write down their thoughts in a text-box on the screen.

### Measures

As a measure of automatic associations, we used Payne et al.'s (2005) AMP. On each trial of the task, participants were first presented with a fixation cross for 1000 ms, which was replaced by one of the two pictures for 75 ms. Control trials involved a presentation of a grey square. The presentation of the prime stimuli was followed by a blank screen for 125 ms, after which a Chinese character appeared for 100 ms. The Chinese character was then replaced by a black-and-white pattern mask, and participants had to indicate whether they considered the Chinese character as more pleasant or less pleasant than the average Chinese character. The pattern mask remained on the screen until participants had given their response. Following the instructions employed by Payne et al. (2005), participants were told that the pictures can sometimes bias people's responses to the Chinese characters, and that they should try their absolute best not to let the pictures bias their judgments of the Chinese characters. The task included a total of 30 trials for each of the two pictures and 30 control trials using a grey square as prime stimulus. As target stimuli, we used a pool of 90 distinct Chinese characters adapted from Payne et al. (2005). Self-reported evaluations of the two pictures were assessed with two likeability rating scales, ranging from 1 (not at all) to 7 (very much).

### Results

#### Implicit measure

Implicit evaluation scores were created by calculating the mean proportion of more pleasant responses for each of the two pictures (Cronbach's  $\alpha = .74$  and  $.76$ ). Submitted to a 2 (ownership)  $\times$  2 (introspection) mixed-model ANOVA, these indices revealed a significant main effect of ownership,  $F(1,73) = 8.14$ ,  $p = .006$ ,  $\eta^2 = .100$ , indicating that evaluations were more favorable for owned as compared to non-owned pictures (see Fig. 3). No other main or interaction effect reached statistical significance (all  $F_s < 1.58$ ).

#### Explicit measure

To test the impact of ownership and introspection on explicit evaluation scores, likeability ratings of the two pictures were submitted to the same 2 (ownership)  $\times$  2 (introspection) ANOVA. This analysis revealed a significant two-way interaction,  $F(1,73) = 6.01$ ,  $p = .02$ ,  $\eta^2 = .076$  (see Fig. 4). Consistent with our predictions, participants showed more favorable evaluations of owned compared to non-owned pictures when they focused on their feelings,  $F(1,37) = 8.02$ ,  $p = .007$ ,  $\eta^2 = .178$ . However, self-reported evaluations were unaffected by ownership when participants focused on reasons for their preference,  $F(1,36) = 0.44$ ,  $p = .51$ ,  $\eta^2 = .012$ .

#### Implicit–explicit correlations

To investigate the impact of introspection on the relation between implicit and explicit measures, we subtracted likeability ratings for non-owned pictures from likeability ratings

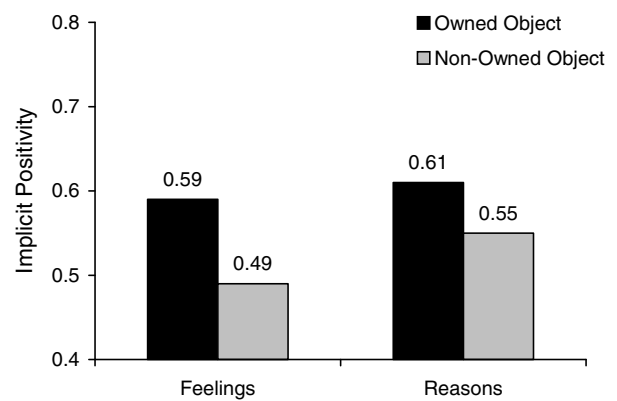


Fig. 3. Mean values of implicit object evaluations as a function of ownership (owned vs. non-owned) and introspection focus (feelings vs. reasons), Experiment 3.

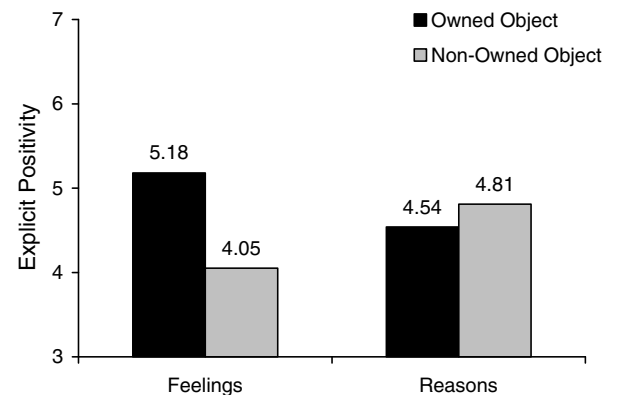


Fig. 4. Mean values of explicit object evaluations as a function of ownership (owned vs. non-owned) and introspection focus (feelings vs. reasons), Experiment 3.

for owned pictures, resulting in an index of explicit preference for the owned over the non-owned picture. A corresponding index was calculated for implicit preferences. Overall, implicit and explicit preference scores were positively correlated ( $r = .22$ ,  $p = .06$ ). More importantly, correlations significantly differed between the two introspection conditions,  $z = 2.06$ ,  $p = .04$ , such that preference scores were significantly correlated when participants focused on their feelings ( $r = .40$ ,  $p = .01$ ), but not when they focused on reasons ( $r = -.07$ ,  $p = .69$ ).

### Discussion

Results from Experiment 3 support our assumption that ownership effects on explicit measures (Beggan, 1992) depend on the relative impact of automatic associations on self-reported evaluative judgments. Moreover, the present study shows that ownership can influence implicit measures without affecting explicit measures. Specifically, ownership led to a preference for owned over non-owned objects on the explicit measure only when participants focused on their feelings for these objects, but not when they thought about reasons for their preference. In contrast, ownership effects on the implicit measure were unaffected by introspection foci, in that participants showed a preference for owned over non-owned objects regardless of whether they focused on feelings or reasons. Moreover, implicit and explicit measures were highly correlated when participants focused on their feelings, but not when they focused on reasons.

## General discussion

The main goal of the present research was to test a process-account of previous studies showing changes on implicit but not explicit measures (e.g., Karpinski & Hilton, 2001; Olson & Fazio, 2006). Drawing on contemporary dual-process models (Fazio & Olson, 2003; Gawronski & Bodenhausen, 2006), we argued that such asymmetrical patterns should emerge when (a) the pairing of an attitude object with positive or negative valence creates new automatic associations in memory, and (b) the consideration of additional information about the attitude object eliminates the impact of automatic associations on self-reported evaluative judgments. In the present research, we tested these predictions by manipulating automatic associations via repeated pairings of neutral CS with positive or negative US (Experiment 2) and ownership (Experiment 3); the impact of automatic associations on evaluative judgments was manipulated by means of participants' introspection foci on feelings versus reasons (Experiments 1 and 3) or feelings versus knowledge (Experiment 2). Consistent with our predictions, implicit measures were influenced by the pairings of the attitude object with positive or negative valence irrespective of participants' introspection foci. In contrast, explicit measures were affected by evaluative pairings only when participants focused on their feelings, but not when they focused on their knowledge or on reasons for their preference. Moreover, implicit and explicit measures were positively correlated when participants introspected on their feelings, but they were uncorrelated (or even negatively correlated) when participants focused on knowledge or reasons.

These results help us understand earlier findings showing experimental effects on implicit but not explicit measures. For example, Olson and Fazio (2006) found that repeated pairings of Black and White faces with positive and negative stimuli influenced only implicit but not explicit measures of racial attitudes. Similarly, Karpinski and Hilton (2001) found that repetitive combinations of the words *youth* and *elderly* with positive and negative words influenced only implicit but not explicit measures of preferences for young over old. Drawing on the present findings, these results may be explained in terms of the relative impact of automatic associations on evaluative judgments. Specifically, the conditioning manipulations employed by Olson and Fazio (2006) and Karpinski and Hilton (2001) may have influenced automatic associations of the respective target groups. However, to the degree that participants considered other information about these groups, EC-related changes were reflected only on implicit but not explicit measures.

Even though the primary goal of the present research was to investigate patterns that involve effects on implicit but not explicit measures, our process-account also helps us to understand the opposite pattern: effects on explicit but not implicit measures (e.g., Gawronski & Strack, 2004; Gregg et al., 2006). One example is a study by Gregg et al. (2006) in which participants were presented with evaluative information about two hypothetical groups. For one group, this information was consistently positive; for the other group, this information was consistently negative. After participants completed an implicit and an explicit measure of evaluations of the two groups, the experimenter indicated that there had been an error in the presentation of the information about the two groups, such that the particular pairings of positive and negative statements about the two groups was intended to be opposite to the one presented to the participant. The experimenter then asked the participant to imagine a reversal of the presented information and to complete the two measures again. Results indicate that the experimenter's instructions led to a reversal on the explicit measure, with the implicit measure remaining in line with the original

descriptions. Even though this finding is often interpreted as evidence for the presumed higher stability of implicit compared to explicit attitudes, it can be easily explained by the process-account proposed in the present research. Specifically, one could argue that the presentation of positive and negative information influenced automatic associations of the two groups, with reversal instructions simply eliminating the impact of automatic associations on evaluative judgments. From this perspective, the mechanisms underlying Gregg et al.'s (2006) findings may be regarded as identical to the ones investigated in the present studies, even though the two lines of research are framed in completely opposite ways.

## Conclusion

The main goal of the present research was to test a process-account of previous findings showing changes on implicit but not explicit measures. Drawing on contemporary dual-process models (Fazio & Olson, 2003; Gawronski & Bodenhausen, 2006), we argued that such asymmetrical changes emerge when (a) the pairing of an attitude object with positive or negative valence creates new automatic associations in memory, and (b) the consideration of additional information about the attitude object eliminates the impact of automatic associations on self-reported evaluative judgments. From this perspective, the relative impact of a given manipulation on implicit and explicit measures is not determined by the relative stability of two distinct attitude representations in memory (cf. Wilson et al., 2000). Rather, whether a given manipulation leads to changes on implicit or explicit measures depends on two factors: (a) the creation of new automatic associations and (b) the impact of these associations on verbal judgments. These factors explain not only the pattern addressed in the present research: changes on implicit but not explicit measures. They also integrate the opposite pattern: changes on explicit but not implicit measures. From this perspective, the proposed process-account provides a parsimonious framework for a variety of different findings in the attitude change literature.

## References

- Beggs, J. K. (1992). On the social nature of nonsocial perception: The mere ownership effect. *Journal of Personality and Social Psychology*, 62, 229–237.
- Bosson, J. K., Swann, W. B., & Pennebaker, J. W. (2000). Stalking the perfect measure of implicit self-esteem: The blind men and the elephant revisited? *Journal of Personality and Social Psychology*, 79, 631–643.
- Dasgupta, N., & Greenwald, A. G. (2001). On the malleability of automatic attitudes: Combating automatic prejudice with images of admired and disliked individuals. *Journal of Personality and Social Psychology*, 81, 800–814.
- De Houwer, J., Thomas, S., & Baeyens, F. (2001). Associative learning of likes and dislikes: A review of 25 years of research on human evaluative conditioning. *Psychological Bulletin*, 127, 853–869.
- Dijksterhuis, A. (2004). I like myself but I don't know why: Enhancing implicit self-esteem by subliminal evaluative conditioning. *Journal of Personality and Social Psychology*, 86, 345–355.
- Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology*, 69, 1013–1027.
- Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and use. *Annual Review of Psychology*, 54, 297–327.
- Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin*, 132, 692–731.
- Gawronski, B., Bodenhausen, G. V., & Becker, A. (2007). I like it, because I like myself: Associative self-anchoring and post-decisional change of implicit evaluations. *Journal of Experimental Social Psychology*, 43, 221–232.
- Gawronski, B., & Strack, F. (2004). On the propositional nature of cognitive consistency: Dissonance changes explicit, but not implicit attitudes. *Journal of Experimental Social Psychology*, 40, 535–542.
- Greenwald, A. G., & Farnham, S. D. (2000). Using the Implicit Association Test to measure self-esteem and self-concept. *Journal of Personality and Social Psychology*, 79, 1022–1038.
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. K. L. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology*, 74, 1464–1480.

- Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, 85, 197–216.
- Gregg, A. P., Seibt, B., & Banaji, M. R. (2006). Easier done than undone: Asymmetry in the malleability of implicit preferences. *Journal of Personality and Social Psychology*, 90, 1–20.
- Hofmann, W., Gschwendner, T., Nosek, B. A., & Schmitt, M. (2005). What moderates implicit–explicit consistency? *European Review of Social Psychology*, 16, 335–390.
- Karpinski, A., & Hilton, J. L. (2001). Attitudes and the Implicit Association Test. *Journal of Personality and Social Psychology*, 81, 774–788.
- Koole, S. L., Dijksterhuis, A., & Van Knippenberg, A. (2001). What's in a name: Implicit self-esteem and the automatic self. *Journal of Personality and Social Psychology*, 80, 669–685.
- Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional overlap: Cognitive basis for stimulus-response compatibility—a model and taxonomy. *Psychological Review*, 97, 253–270.
- Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General*, 106, 226–254.
- Nier, J. A. (2005). How dissociated are implicit and explicit racial attitudes? A bogus pipeline approach. *Group Processes and Intergroup Relations*, 8, 39–52.
- Olson, M. A., & Fazio, R. H. (2006). Reducing automatically activated racial prejudice through implicit evaluative conditioning. *Personality and Social Psychology Bulletin*, 32, 421–433.
- Payne, B. K., Cheng, S. M., Govorun, O., & Stewart, B. D. (2005). An inkblot for attitudes: Affect misattribution as implicit measurement. *Journal of Personality and Social Psychology*, 89, 277–293.
- Petty, R. E., Tormala, Z. L., Brinol, P., & Jarvis, W. B. G. (2006). Implicit ambivalence from attitude change: An exploration of the PAST Model. *Journal of Personality and Social Psychology*, 90, 21–41.
- Rudman, L. A., Phelan, J. E., & Heppen, J. B. (2007). Developmental sources of implicit attitudes. *Personality and Social Psychology Bulletin*, 33, 1700–1713.
- Rydell, R. J., & McConnell, A. R. (2006). Understanding implicit and explicit attitude change: A systems of reasoning analysis. *Journal of Personality and Social Psychology*, 91, 995–1008.
- Schuetter, R. A., & Fazio, R. H. (1995). Attitude accessibility and motivation as determinants of biased processing: A test of the MODE model. *Personality and Social Psychology Bulletin*, 21, 704–710.
- Thurstone, L. L. (1928). Attitudes can be measured. *American Journal of Sociology*, 33, 529–554.
- Walther, E., Nagengast, B., & Traselli, C. (2005). Evaluative conditioning in social psychology: Facts and speculations. *Cognition and Emotion*, 19, 175–196.
- Wilson, T. D., & Dunn, D. S. (1986). Effects of introspection on attitude–behavior consistency: Analyzing reasons versus focusing on feelings. *Journal of Personality and Social Psychology*, 22, 249–263.
- Wilson, T. D., Dunn, D. S., Kraft, D., & Lisle, D. J. (1989). Introspection, attitude change, and attitude–behavior consistency: The disruptive effects of explaining why we feel the way we do. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 22, pp. 287–343). New York: Academic Press.
- Wilson, T. D., Lindsey, S., & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review*, 107, 101–126.
- Wittenbrink, B., & Schwarz, N. (Eds.). (2007). *Implicit measures of attitudes*. New York: Guilford Press.