

# Implicit Racial Stereotypes may Temporarily Shift after Reading a Story

**Maddalena Marini** ([maddalena.marini@unimore.it](mailto:maddalena.marini@unimore.it))

Dipartimento di Comunicazione e Economia, Viale Allegri 9  
42100 Reggio Emilia, ITALY

**Sandro Rubichi** ([sandro.rubichi@unimore.it](mailto:sandro.rubichi@unimore.it))

Dipartimento di Comunicazione e Economia, Viale Allegri 9  
42100 Reggio Emilia, ITALY

**Giuseppe Sartori** ([giuseppe.sartori@unipd.it](mailto:giuseppe.sartori@unipd.it))

Dipartimento di Psicologia, Via Venezia 8  
35100 Padova, ITALY

## Abstract

This study demonstrated that implicit preferences for Whites compared to Blacks can be temporarily (i.e., less than a week) reduced or reversed by reading a counter-stereotypical scenario about an assault. Explicit racial associations (i.e., White = Bad, Black = Good) appear to be not a crucial component of the scenario for shifting the bias but seem to play a role in increasing the malleability effect of implicit racial attitudes. When they were not included in the scenario, a reduction of the implicit preferences for Whites compared to Blacks (Experiment 2) was observed, whereas when they were included we found a reverse of the implicit racial bias indicating a preference for Blacks than to Whites (Experiment 1).

**Keywords:** Implicit Association Test; racial bias; malleability; explicit racial associations; persistence over time.

## Introduction

One of the most used implicit measure of social cognition is the Implicit Association Test (IAT; Greenwald, McGhee & Schwartz, 1998; Nosek, Greenwald & Banaji, 2007; Nosek, Smyth, Hansen et al., 2007). In a classic IAT paradigm (Greenwald et al., 1998), participants are asked to classify four types of stimuli by pressing one of two keys. In particular, they are asked to discriminate stimuli belonging to two target categories and stimuli belonging to two evaluative dimensions. Faster responses are made when the interest category shares the same motor response with the stereotypic evaluative dimension (congruent condition) than when it shares the same motor response with the non-stereotypic evaluative dimension (incongruent condition).

Originally, implicit measures were assumed to be inflexible and resistant to change and insensitive to external pressures and strategic processes (Bargh, 1999; Dovidio & Fazio, 1992; Fazio, Jackson, Dunton, & William, 1995). In contrast to this assumption, recent research has shown that the implicit measures are highly sensitive to contextual influences (for a review, see Blair, 2002; Dasgupta & Asgari, 2004; Mitchell, Nosek & Banaji, 2003).

Multiple interventions were found to elicit shifts of implicit biases as measured by IAT (e.g., Dasgupta & Greenwald, 2001; Blair, Ma, & Lenton, 2001; Foroni & Mayr, 2005). Foroni and Mayr (2005), for example, have demonstrated that reading a counter-stereotypical

scenario produces a reduction in implicit preferences for flowers compared to insects. In their scenario, participants were asked to imagine a fictional post-apocalyptic world in which, due to radiation effects, flowers have become highly noxious, whereas insects serve as a crucial, indirect food resource for humans. Moreover, explicit counter-stereotypical associations between the two categories of interest were included. Insects were positive and flowers were negative, contrary to the normal established association flowers are good and insects are bad (Greenwald et al., 1998).

Even though it has been demonstrated that reading a scenario influences evaluations in the IAT, the underlying mechanisms of this malleability effect are not sufficiently well understood. Foroni and Mayr (2005) suggested that shift of the implicit preferences can be attributed by embedding the relevant categories (e.g., in their case flowers and insects) within a novel and coherent conceptual network that entails a newly knowledge structure and a revaluation of the categories.

We conducted the present research to investigate the underlying mechanisms and the persistence over time of the malleability of implicit attitudes as measured by the IAT, focusing on the Foroni and Mayr's scenario paradigm (2005). Specifically, we investigated whether implicit racial stereotypes could be shifted after reading a counter-stereotypical scenario and whether this change could be detected after one week.

## Experiment 1

In Experiment 1, according to Foroni and Mayr (2005), we used a counter-stereotypical scenario in which the to-be-evaluated categories (i.e., in our case White and Black) were embedded within a meaningful conceptual network that entails a revaluation of the categories. In addition, explicit associations were included. Specifically, we used a pro-black scenario describing an assault, in which a positive action was assigned to a Black individual (i.e., rescuer role) and a negative action to a White individual (i.e., aggressor role).

## Method

**Participants** Twenty undergraduate students (age 20-29; mean=21.25; 12 women and 8 men) participated in the experiment. They were all white participants enrolled at the University of Modena and Reggio Emilia. They were divided into two groups. Ten participants read a pro-black scenario, while the other ten read a pro-white scenario.<sup>1</sup>

**Procedure and stimuli** Participants took part in three sessions.<sup>2</sup> The first and the second sessions were carried out the same day, while the third session was carried out one week later. In all sessions, participants performed the five blocks of the standard race IAT (Greenwald et al., 1998). They were asked to respond as quickly and accurately as possible to the stimuli, by pressing one of two labelled keys. The stimuli consisted of 16 words (8 good words and 8 bad words) and 12 pictures (6 White individuals and 6 Black individuals). The good evaluative dimension included words like peace and love, while the bad evaluative dimension included words like evil and hurt. The White category included pictures of White European individuals while the Black category included pictures of Black African individuals. The stimuli were the same as in the race IAT on the website (<https://implicit.harvard.edu/implicit/italy/>).

In the second session participants were asked to read a scenario describing an assault prior to the race IAT. Two different versions of the same scenario were used. In the pro-black scenario, the aggressor was a white man and the rescuer was a black man; in the pro-white scenario, the aggressor was a black man and the rescuer was a white man. In the pro-black condition participants were instructed to imagine that the following task (i.e., the race IAT) was a tutoring program that was supposed to establish these associations firmly in people's minds: "White = Bad, Black = Good"; the reverse association was presented in the pro-white condition.

**Results and Discussion** Analyses were conducted on the average reaction times (RTs) in blocks 3 and 5. Incorrect responses were excluded from RT analysis as were responses faster than 300 ms and slower than 3000 msec. Repeated measures analysis of variance (ANOVA) was conducted with scenario group (pro-black vs. pro-white) as a between-subjects factor and congruency (congruent vs. incongruent) and session (first, second, third) as within-subjects factors. A main effect of congruency was detected,  $F(1, 17) = 26.979, p < .0001, \eta^2 = .613$ , with RTs faster in the congruent than in the incongruent block (677 ms vs. 892 ms). The main effect of session,  $F(2, 34) = 14.446, p < .0001, \eta^2 = .459$ , and the interaction between congruency and session were significant,  $F(2, 34) = 13.550, p < .0001, \eta^2 = .444$ . RTs decreased from the first to the third session (first session = 839 ms, second session = 798 ms, third session = 716 ms) and although the IAT effect for RTs (i.e., the RT difference between congruent and incongruent blocks) was significant for all

three sessions [first session:  $F(1, 19) = 39.864, p < .0001, \eta^2 = .677$ ; second session:  $F(1, 19) = 4.412, p < .05, \eta^2 = .188$ ; third session:  $F(1, 18) = 31.058, p < .0001, \eta^2 = .633$ ], it was smaller in the second than the first session and the third session (126 vs. 332 vs. 193 ms). Also, a significant three-way interaction was found between congruency, scenario group and session,  $F(2, 34) = 9.809, p < .0001, \eta^2 = .366$ . Separate ANOVAs for scenario group showed a significant interaction between congruency and session for the black scenario group,  $F(2, 16) = 25.845, p < .0001, \eta^2 = .764$ , but not for the pro-white scenario group,  $F(2, 18) = 1.764, p = .200, \eta^2 = .164$ . The IAT effect for RTs was significant in all three sessions for the pro-white scenario group [first session:  $F(1, 9) = 15.028, p < .01, \eta^2 = .625$ , second session:  $F(1, 9) = 19.737, p < .01, \eta^2 = .687$ ; third session:  $F(1, 9) = 17.436, p < .01, \eta^2 = .660$ ]; while for the pro-black scenario group a significant difference was evident in the first,  $F(1, 9) = 25.524, p < .001, \eta^2 = .739$ , and the third sessions,  $F(1, 8) = 12.697, p < .01, \eta^2 = .613$ .

In addition, we computed the D index<sup>3</sup> (Greenwald, Nosek, & Banaji, 2003) for each subject and submitted it to an ANOVA with scenario group (pro-black, pro-white) as a between-subjects factor and session (first, second, third) as a within-subjects factor. Overall, a significant effect of session was detected,  $F(2, 34) = 15.930, p < .0001, \eta^2 = .484$ . The D index showed a significant difference between the first and the second sessions (1.00 vs. .42;  $p < .0001$ ) and between the second and the third sessions (.42 vs. .83;  $p < .01$ ), while no significant difference was found between the first and the third sessions (1.00 vs. .83;  $p = .203$ ). The interaction between session and scenario group was significant,  $F(2, 34) = 21.029, p < .0001, \eta^2 = .553$ . The average D index did not differ across the three sessions for the pro-white scenario group ( $F(2, 18) = .472, p = .631, \eta^2 = .050$ ). By contrast, a significant difference between sessions was obtained for the pro-black scenario group,  $F(2, 16) = 23.217, p < .0001, \eta^2 = .744$ . In particular, pairwise comparisons showed that D values did not differ between the first and third sessions ( $p = .189$ ) but significant differences were found both between the first and second sessions ( $p < .01$ ) and between the second and third sessions ( $p < .01$ ).

In sum, we found that all participants showed a pro-white bias both in the first (pro-black: IAT effect = 331 ms,  $D = 1.09$ ; pro-white: IAT effect = 334 ms,  $D = .91$ ) and in third session (pro-black: IAT effect = 169 ms,  $D = .77$ ; pro-white: IAT effect = 215 ms,  $D = .88$ ). An interesting result emerged in the second session after participants read a pro-black scenario. Here, they showed a -39 ms reversal IAT effect and a negative D index ( $D = -.14$ ), indicating a pro-black bias. No change was detected in the group who read the pro-white scenario, which continued to show a pro-white bias (IAT effect = 291 ms,  $D = .98$ ; see Figure 1). The malleability of the pro-white bias emerged only immediately after the reading of the pro-black scenario but was not further present at retest after one week.

<sup>1</sup> In the present study the pro-white scenario was used for control purposes.

<sup>2</sup> Nineteen out of twenty students returned for the third session.

<sup>3</sup> A positive D index indicated a pro-white bias in the participants while a negative D index indicated a pro-black bias.

These results demonstrate that implicit racial biases can be reversed by a counter-stereotypical scenario with explicit racial associations and that such effect lasts for less than a week. It is important to consider that the malleability effect found in the present experiment cannot be attributed to tendency to respond in a socially desirable manner. To note, the IAT is an implicit (automatic) measure of attitudes (for a review, see Greenwald, Banaji, Rudman, Farnham, Nosek, & Mellot, 2002) developed in order to overcome the problem of social desirability in the explicit (self-reported) measures (Levy, 1981; Peltier & Walsh, 1990; Simon & Simon, 1975; Zerbe & Paulhus, 1987). Whereas attitudes measured by explicit measures are conscious and controllable, implicit measures allow us to detect unintentional and automatic attitudes (Wegner & Bargh, 1998), and therefore are less influenced by the tendency of people to deceive their views in order to present themselves in a socially acceptable manner. Accordingly, several studies have revealed that participants cannot control voluntarily IAT performances (Asendorpf, Banse, & Mücke, 2002; Banse, Seise, & Zerbes, 2001; Egloff, & Schmukle, 2002; Kim, 2003; Fiedler, & Bluemke, 2005; Steffens, 2004). For example, Kim (2003) showed that participants could not fake the IAT effectively when merely instructed to respond with a pro-black attitude.

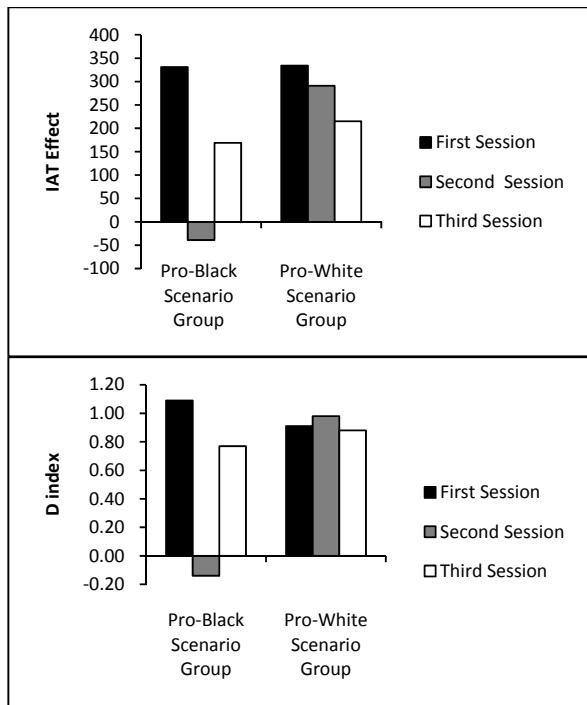


Figure 1: IAT effect and D index for Experiment 1.

## Experiment 2

One possible explanation for the shift of the implicit racial bias observed in the Experiment 1 is that the scenario explicitly associated the Whites and Blacks categories with a positive or a negative evaluative dimension. In fact, participants were asked to imagine that the goal IAT was to establish in people's minds the

associations "White = Bad, Black = Good". Therefore the present experiment tested whether excluding the explicit racial associations the scenario shifted still the implicit preferences for Whites over Blacks as in Experiment 1 and whether this change persisted for more than a week. For this reason, we modified the scenario used in Experiment 1 eliminating the explicit racial associations.

## Method

**Participants** Twenty-four undergraduate students (age 18-25; mean=22.5; 20 women and 4 men) participated in the experiment. They were all white participants enrolled at the University of Modena and Reggio Emilia. Twelve read the pro-black scenario, while the other twelve read the pro-white scenario. None of them participated in the previous experiment.

**Procedure and stimuli** They were the same as in Experiment 1 with the only difference that the explicit racial associations (e.g., that "White = Bad, Black = Good") were removed.

**Results and Discussion** The same analyses as in Experiment 1 were conducted on RTs and the D index.

For RTs, a main effect of congruency was found,  $F(1, 22) = 58.721, p < .0001, \eta^2 = .727$ , with RTs faster in the congruent than in the incongruent block (690 ms vs. 893 ms). Also the main effect of session,  $F(2, 44) = 27.372, p < .0001, \eta^2 = .554$ , and the interaction between congruency and session were significant,  $F(2, 44) = 17.520, p < .0001, \eta^2 = .443$ . RTs decreased from the first to the third session (first session = 856 ms, second session = 801 ms, third session = 718 ms) and although the IAT effect for RTs was significant for all the three sessions [first session:  $F(1, 23) = 76.198, p < .0001, \eta^2 = .768$ ; second session:  $F(1, 23) = 14.212, p < .001, \eta^2 = .382$ ; third session:  $F(1, 23) = 33.463, p < .0001, \eta^2 = .593$ ], it was smaller in the second (140 ms) than the first (303 ms) and the third sessions (165 ms). Also, a significant three-way interaction between congruency, scenario group and session,  $F(2, 44) = 6.951, p < .01, \eta^2 = .240$ , was found. Separate ANOVAs for scenario group showed a significant interaction between congruency and session for the black scenario group,  $F(2, 22) = 30.605, p < .0001, \eta^2 = .736$ . There was a tendency to significance for the pro-white scenario group,  $F(2, 22) = 3.060, p = .067, \eta^2 = .218$ . The IAT effect was significant in all three sessions only for the pro-white condition [first session:  $F(1, 11) = 33.115, p < .0001, \eta^2 = .751$ , second session:  $F(1, 11) = 29.059, p < .0001, \eta^2 = .725$ ; third session:  $F(1, 11) = 13.871, p < .01, \eta^2 = .558$ ]; while for the pro-black condition a significant difference was evident only in the first,  $F(1, 11) = 40.962, p < .0001, \eta^2 = .788$ , and in the third sessions,  $F(1, 11) = 27.328, p < .0001, \eta^2 = .713$ .

For the D index, a significant effect of session,  $F(2, 44) = 12.86, p < .0001, \eta^2 = .368$  (first session = .85, second session = .47, third session = .73), and of scenario group were found,  $F(1, 22) = 5.044, p < .05, \eta^2 = .187$  (pro-black scenario group = .56, pro-white scenario group = .81).

Moreover, a significant interaction between session and condition was found,  $F(2, 44) = 10.764, p < .0001, \eta^2 = .329$ . The average D index did not differ across the three sessions for the pro-white scenario group. By contrast, a significant difference between sessions was obtained for the pro-black condition,  $F(2, 22) = 18.210, p < .0001, \eta^2 = .623$ .

In particular, D values did not differ between the first and third sessions but a significant difference was found between the first and second sessions ( $p < .001; .82$  vs.  $.73$ ) and between the second and third sessions ( $p < .001; .13$  vs.  $.73$ ).

In sum, we found a significant pro-white bias for either of two scenario groups both in the first (pro-black: IAT effect = 298 ms, D = .82; pro-white: IAT effect = 308 ms; D = .89) and in the third sessions (pro-black: IAT effect = 141 ms, D = .73; pro-white: IAT effect = 189 ms, D = .74). In the second session a reduction of the pro-white bias after reading a pro-black scenario was found (IAT effect = 31, D = .13). No change in attitude scenario were observed in the group who read the pro-white scenario (IAT effect = 249 ms, D = .80; see Figure 2).

These results indicate that implicit racial biases can be reduced by a counter-stereotypical scenario without explicit racial associations and confirm the temporariness of the malleability effect. Thus, the shift of the implicit bias cannot be attribute to the component of the scenario that contains the explicit associations per se. Rather, the effect seems depend on the new evaluative network within the relevant categories (i.e., in our case White and Black categories) were embedded, as suggested by Foroni and Mayer (2005).

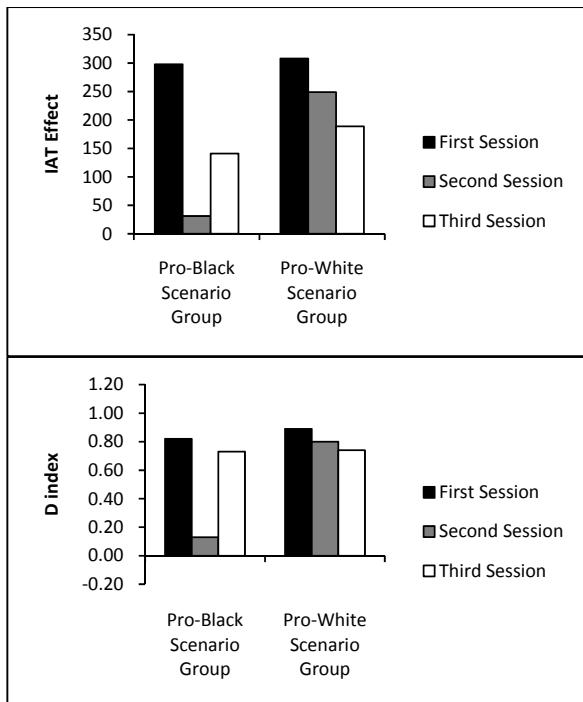


Figure 2: IAT effect and D index for Experiment 2.

## General Discussion

We investigated the role and persistence over time of the effect of a counter-stereotypical scenario on the malleability of implicit racial attitudes in groups of white students. We used different counter-stereotypical scenarios about an assault, in which a positive action was assigned to a black person (i.e., rescuer role) and a negative action to a white person (i.e., aggressor role). In Experiment 1, explicit racial associations (e.g., "White = Bad, Black = Good") were included in the scenario. Conversely, in Experiment 2, no explicit racial associations were given.

We demonstrated that implicit racial biases may be shifted (i.e., reduced or reversed) by reading a counter-stereotypical scenario. Moreover, we showed that the presence of explicit racial associations (e.g., "White = Bad, Black = Good") in the scenario is not critical to shift the implicit bias. In fact, a reduction of implicit racial bias was observed even when the explicit racial associations were removed from the scenario. However, it is important to note that when the explicit racial associations were included in our scenario a greater malleability effect was observed. In particular we obtained a reverse of the implicit racial bias, indicating a counter-stereotypical racial attitude (i.e., pro-black).

Taken together, these results demonstrate that reading a counter-stereotypical scenario can shift the implicit racial attitudes and seem to confirm that this change may be attributed to new evaluation of the relevant categories, as suggested by Foroni and Mayer (2005). One alternative explanation to shift of bias could be that in the scenarios participants were asked to perform a mental imagery task; in particular, they were asked to imagine themselves as victims of an assault. According to this view, research within the social psychology has demonstrated that mental imagery has relevant benefits on the malleability of attitudes. For instance, it has been shown that when participants are asked to imagine encountering an out-group member (e.g., young participants imagined talking to an elderly person), before reporting their attitude towards the out-group, reductions of the explicit intergroup bias (Turner, Crisp, & Lambert, 2007; see also Stathi & Crisp, 2008) and implicit prejudice as measured by the IAT were observed (Turner & Crisp, 2010). Similarly, reductions of implicit attitudes are also found when participants are asked to imagine a counter-stereotypical mental image before going through an IAT (Blair, et al., 2001). More specifically, when participants are asked to create a counter-stereotypical mental image of strong woman before to perform a gender IAT a reduction in the automatic association between females and weakness are obtained. To note, also in the scenario used by Foroni and Mayer (2005) participants were asked to imagine themselves as protagonists of the counter-stereotypical story. However, it is necessary to consider also that in our scenario participants were asked to imagine themselves as victims of assault. We suggest that the imagined self-involvement task evoked emotions toward social groups that can shifted the implicit bias. In particular, a negative emotion toward in-group members (i.e., Whites) and a positive emotion toward out-group members (i.e., Blacks). Mental imagery has been found

to elicit similar emotional and motivational responses as the real experience (Dadds, Bovbjerg, Redd, & Cutmore, 1997) and neuropsychological studies have shown that mental imagery shares the same neurological bases as perception and employ similar neurological mechanisms as memory, emotion, and motor control (Kosslyn, Ganis, & Thompson, 2001). In addition, research on emotions and intergroup bias has provided initial evidence that specific negative emotional states are capable of creating an implicit prejudice toward out-group from thin air (DeSteno, Dasgupta, Barlett, & Cajdric, 2004). It is important to note that in this study participants were not induced to experience negative emotions toward out-group but simply to experience a negative emotional state by writing an autobiographical event from the past that had made them very angry. We extend this finding suggesting that evoking negative emotions towards in-group members and positive emotions towards out-group members can elicit a shift of implicit bias, that is, a decreasing of the implicit negative evaluation of the out-group. Finally, our results demonstrate that the malleability of implicit bias is only a temporary effect: shifts in racial attitudes produced by reading a counter-stereotypical scenario were not observed after one week. To our knowledge this is the only study that investigated the persistence over time of the malleability effect on implicit bias over such an extended period. For example, Dasgupta and Greenwald (2001) showed that the malleability effect obtained by exposing White participants to admired Black and disliked White individuals may persist in a follow-up test session 24 hours later. Our results extend this finding by providing evidence that the malleability effect was no longer present after one week.

Future works may help to investigate and disentangle the role of imagined self-involvement and of emotions on the shift of implicit bias and to clarify exactly how the malleability effect lasts over time. At present, however, the current findings extend previous research on the malleability of implicit prejudices by suggesting that emotions evoked by imagined self-involvement scenario may be capable to reduce temporally implicit negative attitudes.

## References

Asendorpf, J. B., Banse, R., & Mucke, D. (2002). Double dissociation between implicit and explicit personality self-concept: The case of shy behavior. *Journal of Personality and Social Psychology, 83*, 380-393.

Banse, R., Seise, J., & Zerbes, N. (2001). Implicit attitudes toward homosexuality: Reliability, validity, and controllability of the IAT. *Zeitschrift für Experimentelle Psychologie, 48*(11), 45-60.

Bargh, J. A. (1999). The cognitive monster: The case against the controllability of automatic stereotype effects. In S. Chaiken & Y. Trope (Eds.), *Dual process theories in social psychology* (pp. 361-382). New York: Guilford.

Blair, I. V. (2002). The malleability of automatic stereotypes and prejudice. *Personality and Social Psychology Review, 6*, 242-261.

Blair, I. V., Ma, J. E., & Lenton, A. P. (2001). Imagining stereotypes away: The moderation of implicit stereotypes through mental imagery. *Journal of Personality & Social Psychology, 81*, 828-841.

Dadds, M.R., Bovbjerg, D.H., Redd, W.H., & Cutmore, T.R. (1997). Imagery in human classical conditioning. *Psychological Bulletin, 122*, 89-103.

Dasgupta, N., & Asgari, S. (2004). Seeing is believing: Exposure to counterstereotypic women leaders and its effect on automatic gender stereotyping. *Journal of Experimental Social Psychology, 40*, 642-658.

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.

DeSteno, D., Dasgupta, N., Bartlett, M.Y., & Cajdric, A. (2004). Prejudice From Thin Air: The Effect of Emotion on Automatic Intergroup Attitudes. *Psychological Science, 15*, 319-324.

Dovidio, J. F., & Fazio, R. H. (1992). New technologies for the direct and indirect assessment of attitudes. In J. Tanur (Ed.), *Questions about questions: Inquiries into the cognitive bases of surveys* (pp. 204-237). New York: Russell Sage Foundation.

Egloff, B., & Schmukle, S. C. (2002). Predictive validity of an implicit association test for assessing anxiety. *Journal of Personality and Social Psychology, 83*, 1441-1455.

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.

Fiedler, K., & Bluemke, M. (2005). Faking the IAT: Aided and unaided response control on the implicit association tests. *Basic and Applied Social Psychology, 27*(4), 307-316.

Foroni, F., & Mayr, U. (2005). The power of a story: New, automatic associations from a single reading of a short scenario. *Psychonomic Bulletin & Review, 12*, 133-144.

Greenwald, A.G., Banaji, M.R., Rudman, L.A., Farnham, S.D., Nosek, B.A., & Mellot, D.S. (2002). A unified theory of implicit attitudes, beliefs, self-esteem and self-concept. *Psychological Review, 109*(1), 3-25.

Greenwald, T. 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.

Kosslyn, S.M., Ganis, G., & Thompson, W.L. (2001). Neural foundations of imagery. *Nature Reviews Neuroscience, 2*, 635-642.

Levy, S.J. (1981). Interpreting Consumer Mythology: A Structural Approach to Consumer Behavior. *Journal of Marketing, 45*, 49-61.

Kim, D-Y. (2003). Voluntary controllability of the Implicit Association Test (IAT). *Social Psychology Quarterly, 66*, 83-96.

Mitchell, J. P., Nosek, B. A., & Banaji, M. R. (2003). Contextual variations in implicit evaluation. *Journal of Experimental Psychology: General, 132*, 455-469.

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2007). The Implicit Association Test at age 7: A methodological and conceptual review. In J. A. Bargh (Ed.), *Social Psychology and the Unconscious: The Automaticity of Higher Mental Processes* (pp. 265-292). New York: Psychology Press.

Nosek, B. A., Smyth, F. L., Hansen, J. J., Devos, T., Lindner, N. M., Ranganath, K. A., Smith, C. T., Olson, K. R., Chugh, D., Greenwald, A. G., & Banaji, M. R. (2007). Pervasiveness and correlates of implicit attitudes and stereotypes. *European Review of Social Psychology, 18*, 36-88.

Peltier, B.D., & Walsh, J.A. (1990). An Investigation of Response Bias in the Chapman Scales. *Educational and Psychological Measurement, 50*, 803-815.

Simon, J., & Simon R. (1975). The Effect of Money Incentives on Family Size: A Hypothetical-Question Study. *Public Opinion Quarterly, 38*, 585-595.

Stathi, S., & Crisp, R.J. (2008). Imagining intergroup contact promotes projection to outgroups. *Journal of Experimental Social Psychology, 44*, 943-957.

Steffens, M.C. (2004). Is the Implicit Association Test immune to faking? *Experimental Psychology, 51*, 165-179.

Turner, R.N., Crisp, R.J., & Lambert, E. (2007). Imagining intergroup contact can improve intergroup attitudes. *Group Processes and Intergroup Relations, 10*, 427-441.

Turner, R.N., & Crisp, R.J. (2010). Imagining intergroup contact reduces implicit prejudice. *British Journal of Social Psychology, 49*, 129-142.

Zerbe, W.J., & Paulhus, D.L. (1987). Socially Desirable Responding in Organizational Behavior: A Reconceptualization. *Academy of Management Review, 12*, 250-264.