

# Social Contexts and Personal Moral Motives Reduce Implicit Prejudice: A Direct Comparison

Félice van Nunspeet, Naomi Ellemers, and Belle Derks  
Utrecht University

The current research extends previous work (e.g., Blair, 2002; Lai et al., 2014; van Nunspeet, Ellemers, & Derks, 2015) by directly comparing the effectiveness of different contextual factors and personal motives on the reduction of implicit bias toward Muslim women in a  $2 \times 2 \times 2$  research design. Non-Muslim participants performed an implicit association test (IAT) for which the implications for their morality or competence were emphasized, and while their performance was evaluated by a minimal ingroup or outgroup member, who was either presented as a member of the devalued group (a woman with a headscarf) or not (a woman without a headscarf). Study 1 revealed that performance monitoring by the evaluator with a headscarf resulted in significant bias reduction. Interestingly, in case of performance monitoring by an evaluator without a headscarf, emphasizing the moral implications of participant's task performance was as effective. In Study 2, we replicated the first finding and found no significant bias toward Muslim women when participants' IAT performance was evaluated by a woman with a headscarf. Additionally, the prolonged version of the IAT used in this study resulted in a further decrease in bias when the evaluator with a headscarf was presented as a minimal ingroup member. Important implications of the effectiveness of these contextual factors and personal motives are discussed with regard to their difference in level of applicability and implementation.

*Keywords:* implicit bias, morality, intergroup relations, cross-categorization

The study of attitudes, stereotypes, and prejudice is often complicated by social desirability issues: People may adjust their explicit attitudes to appear unbiased (e.g., Crosby, Bromley, & Saxe, 1980). However, even while people explicitly endorse egalitarian views, they may dis-

play biases on an implicit level (e.g., Dovidio, Kawakami, & Beach, 2001). This suggests that discrimination, at least in part, has “gone underground.” Implicit biases are the “automatic” evaluative and/or stereotypical associations people make with their own and other groups. These associations are more difficult to control, but nevertheless impact upon and (unconsciously) affect intergroup behavior (e.g., Amodio & Devine, 2006). In order to address the issues related to prejudice and discrimination, it has therefore become crucial to focus on this implicit level and to examine how to reduce *implicit* social biases.

A popular and widely used implicit measure of prejudice is the Implicit Association Test (Greenwald, McGhee, & Schwartz, 1998). The IAT is based on the notion that it is easier to associate one's ingroup (or majority group members) with positive attributes and an outgroup (or minority group members) with negative attributes than vice versa. Interestingly, though the IAT is frequently presented as an implicit measure—facilitating in the need to examine “automatic” biased responses—by now several studies have shown that be-

---

Félice van Nunspeet, Naomi Ellemers, and Belle Derks, Faculty of Social Sciences, Utrecht University.

This work was supported by the SNS-Reaal KNAW-Merian prize and the NWO-Spinoza prize, awarded to Naomi Ellemers, by the Royal Netherlands Academy of Arts and Sciences and the Netherlands Organization for Scientific Research; and an NWO-VENI grant [451-08-022] awarded to Belle Derks. The authors declare there are no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. We thank Johannes Parzonka, Eva Schildkamp, and Titia van Malenstein for help with data collection; Bianca van Nunspeet for her contribution to the experiment; and Kees Verduin for help with data processing. This research was conducted at the Social and Organizational Psychology Unit, Institute of Psychology, Leiden University, the Netherlands.

Correspondence concerning this article should be addressed to Félice van Nunspeet, Social, Health, & Organizational Psychology, Utrecht University, P.O. Box 80140, 3508 TC Utrecht, the Netherlands. E-mail: [f.vannunspeet@uu.nl](mailto:f.vannunspeet@uu.nl)

havioral performance on this task is malleable (e.g., Blair, 2002; Fiedler & Bluemke, 2005). Important for the current research, this thus offers the opportunity to use the IAT to examine what kind of factors are effective in reducing implicit bias.

Previous research has revealed that implicit biases can be influenced by personal motives and social contexts (for overviews see Blair, 2002; Lai et al., 2014). *Personal motives* may either increase or decrease people's implicit evaluative bias toward outgroup members. For example, inducing stereotype threat among Whites—by triggering the stereotype that they are racists—increased implicit biases toward Blacks (Frantz, Cuddy, Burnett, Ray, & Hart, 2004; Rudman, Dohn, & Fairchild, 2007). In contrast, emphasizing the implications of people's performance on an IAT in terms of moral values—rather than task competence—led participants to inhibit their negative bias against Muslims (van Nunspeet, Ellemers, Derks, & Nieuwenhuis, 2014). Furthermore, *social contexts* may affect people's implicit biases. For instance, when a Black experimenter is present during participants' performance on an IAT, Whites are able to inhibit their pro-White bias (e.g., Lowery, Hardin, & Sinclair, 2001). Interestingly however, such a decrease in bias may also be evident when people are in the presence of their own (e.g., White) ingroup members (Castelli & Tomelleri, 2008). Moreover, such a social context may be especially effective in reducing people's implicit negative bias when they know this is an opportunity to reveal their moral intentions. Specifically, framing the implications of individuals' IAT performance in terms of moral (vs. competence) values reduces participants' negative bias toward Muslims, especially when their performance is evaluated by a (minimal) ingroup, rather than an outgroup member (van Nunspeet, Derks, Ellemers, & Nieuwenhuis, 2015).

Both personal motives and particular social contexts have thus been shown to affect implicit bias, and may therefore be implemented in interventions aimed at reducing (implicit) prejudice. However, their ease of implementation may differ as some contexts or motives may be more easy to apply or induce than others. For instance, intergroup contact has been shown to be highly effective in reducing implicit bias, but such social interactions may occur less frequently than is desirable because members of different (ethnic or religious) groups are often segregated. Fortunately, even indirect types of contact (e.g., having

friends who interact with outgroup members) or imagined interactions may cause similar decreases in intergroup bias (see Lemmer & Wagner, 2015, and Miles & Crisp, 2014, for meta-analyses). Moreover, results from a meta-analysis comparing IAT bias reduction effects (Lai et al., 2014), revealed that implementation-intention strategies or counterstereotypical exemplars were highly effective in reducing implicit bias. However, whereas such cognitive strategies may be easier to implement than facilitating intergroup contact, they have their own limitations like an effortful training or a repeated exposure to different exemplars. It is thus as yet unclear how they compare to alternative mechanisms affecting bias such as contextual factors or personal motives. Furthermore, there is no previous research that directly compared the effects of different kinds of social context and personal motives in *one study design*. It is thus unclear whether the different mechanisms that reduce bias, which differ in their ease of real life implementation, have parallel or accumulating effects on the inhibition of people's implicit bias against outgroup members.

In the current research, we therefore examined and directly compared the effects of different social contexts and personal motives on people's evaluative bias against an outgroup. Specifically, using one study design, we tested to what extent people's evaluative bias against Muslim women is affected by (a) emphasizing the moral (compared to competence) test implications of the IAT to induce *personal motives*, as well as by different *social contexts* in which participants are being evaluated by (b) someone who is presented either as a minimal ingroup or outgroup member, and (c) who is a representative of the devalued outgroup in the IAT (i.e., who does or does not appear to be Muslim, because of [not] wearing a headscarf). Extending previous research (e.g., Blair, 2002; Lai et al., 2014), we combined these different contexts and motives in the current study to directly compare their effects on implicit bias reduction, and to examine whether and how they may interact with each other.

Additionally, we conducted exploratory analyses to examine *how* a reduction in implicit bias is achieved. In prior studies on the effects of intergroup contact or cognitive strategies on people's evaluative biases, little attention has been devoted to the way in which such a bias (i.e., IAT performance) was reduced. In an IAT, bias is reduced when the difference between response latencies on

prejudice-incongruent trials (i.e., trials in which ingroup members are associated with negative stimuli and outgroup members with positive stimuli) and prejudice-congruent trials (i.e., where ingroup members are associated with positive stimuli, and outgroup members with negative stimuli) is diminished. However, this can be accomplished in two ways: Either by responding more quickly to incongruent trials, or by responding more slowly on congruent trials. Studies that did examine specific response time patterns for prejudice-congruent and—incongruent trials revealed that an intergroup context can both enhance fast responding on incongruent trials (e.g., Lowery et al., 2001), and reduce fast responding on congruent trials (e.g., Richeson & Ambady, 2003). Moreover, personal motives (such as the motivation to appear moral, van Nunspeet et al., 2014; and chronic egalitarian personal standards, Devine, Plant, Amodio, Harmon-Jones, & Vance, 2002) often seem to delay responses on congruent trials. However, supplementing egalitarian goals with implementation intentions has also been shown to enhance faster responding on prejudice-incongruent trials (Webb, Sheeran, & Pepper, 2012).

The current research consists of two studies. In Study 1, we directly compared the effects of three manipulations: (a) an emphasis on the moral (vs. competence) implications of the IAT, which participants completed (b) while their performance was evaluated by a minimal ingroup (vs. outgroup) member, (c) who was (vs. was not) a representative of the devalued group (i.e., a woman with or without a headscarf, respectively). Based on previous research, we expected that an emphasis on the moral task implications (van Nunspeet et al., 2014), as well as a combination of such an emphasis and evaluation by a minimal ingroup member (van Nunspeet, Derks, et al., 2015), would significantly decrease implicit prejudice. Additionally, we expected the evaluation by a representative of the devalued group (i.e., a woman with a headscarf) to also diminish negative implicit bias (e.g., Lowery et al., 2001), though its effect relative to the other two intervention strategies is what we aimed to determine—thereby extending the existing literature. In Study 2, we further tapped into these relative effects: Here, we focused on participants' (un)shared minimal group membership with the evaluator and an emphasis on the moral (vs. competence) task implications, while all participants were evaluated by a representative of the devalued group (i.e., a

woman with a headscarf). Moreover, the IAT in Study 2 was prolonged, which extended participants exposure to their evaluator. Previous research has revealed that people can learn to control or change their implicit bias score (e.g., Fiedler & Bluemke, 2005; French, Franz, Phelan, & Blaine, 2013; Kawakami, Dovidio, Moll, Hermsen, & Russin, 2000; Olson & Fazio, 2006), and that an increase in IAT trials can facilitate learning effects and consequently affect bias scores (van Nunspeet, Derks, et al., 2015, 2014). We therefore explored whether such a longer version of the IAT, in which the evaluator was always a representative of the devalued group, might enhance the effects of the minimal group context and the emphasis on morality observed in Study 1.

## Study 1

### Method

**Participants and design.** Only female, non-Muslim, participants were recruited to take part in the experiment to exclude the possibility that another alternative categorization (i.e., gender) would further complicate the results observed. Two-hundred and 25 Dutch university students ( $M_{\text{age}} = 20.5$  years,  $SD = 2.6$ ) participated in the study and received either money or course credit for their participation. Two participants were excluded from analyses: One due to technical problems, another because she responded outside of the time limit on *all* trials of the IAT. Participants were randomly assigned to one of the eight experimental conditions of the 2 (Personal Motive: Morality/Competence)  $\times$  2 (Evaluator's *Minimal* Group Membership: Ingroup/Outgroup)  $\times$  2 (Evaluator's *Religious* Group Membership: Evaluator With/Without a Headscarf) between-participants design.

**Procedure.** Participants were seated in an individual computer room with a webcam on top of the computer screen, and a camera behind them in a top corner of the cubicle. They were told that they would be working together with another participant. To manipulate the minimal group membership of the other person (the evaluator), participants completed a questionnaire ostensibly assessing cooperation and decision-making styles, with items such as "I think it is important to solve problems in an honest manner," "I think it is justified to always consider

how something will benefit myself,” and “You have to be strategic to be successful.” Participants’ score on the questionnaire was said to result in either a so-called “P- or O-type” personality style. They were further informed that because O- and P-type individuals differ in their way of thinking, solving problems, and cooperation styles, P-type individuals get along better with other P-types, whereas O-type individuals get along better with similar O-types.

After a short waiting period, participants learned about their own alleged personality style and the styles of the other participants and they were informed whom they would be working with during the experiment. The other person either was said to have the same personality style as the participant (to convey a shared minimal group membership), or she allegedly had the other personality style (to indicate this individual belonged to a different group membership). Participants then read that they would perform a computer task. During the first part of the experiment, the other person would supposedly observe and give them feedback and the roles would be reversed in the second part. Thereafter, a web-cam connection was simulated. To manipulate the religious group membership, the other person (who, in reality, was the same individual at all times) was a woman who was, or was not, wearing a headscarf (see Figure 1). During the web-cam feed, she introduced herself and said that she would observe and provide visual feedback on every trial of the task. Then, participants were directed to the IAT. To manipulate participants’ personal motive, they either read that the task they were going to perform could indicate their endorsement of *moral values* concerning egalitarianism and discrimination (the morality condition), or their *ability* to process new information and to learn new tasks (the competence condition). After that, the IAT started and all participants were instructed to respond as quickly and accurately as possible. The test implications were repeated before the start of each test block (see also van Nunspeet, Derks, et al., 2015, 2014).

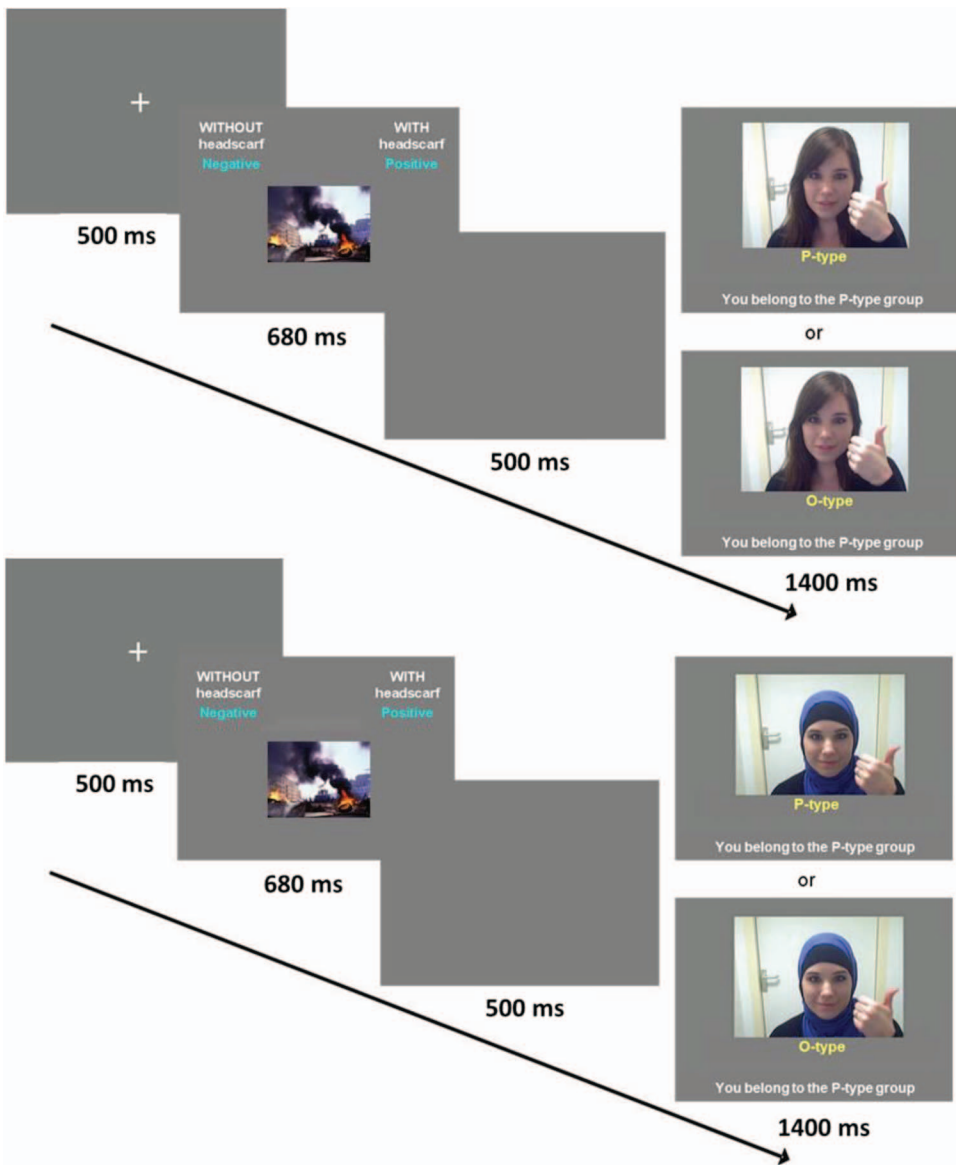
In reality, all participants were said to have a “P”- personality style and were introduced to a confederate whose movies were prerecorded. Feedback displays during the IAT were related to participants’ actual responses (i.e., positive feedback when they responded correctly, negative feedback when they responded incorrectly).

After the IAT, participants completed some self-report items (described below) and were properly debriefed.

### Instruments.

**The Implicit Association Test.** Participants performed a modified version of the classic IAT (as designed by Greenwald et al., 1998), which was similar to the IAT used in previous studies on implicit prejudice toward Muslim women (van Nunspeet, Derks, et al., 2015, 2014). Specifically, stimuli representing the target concepts consisted of 10 pictures of Muslim women (wearing a headscarf) and 10 pictures of non-Muslim women (not wearing a headscarf). Stimuli that represented positive and negative attributes consisted of five pictures of positive scenes, and five pictures of negative scenes, selected from the International Affective Picture System (Lang et al., 2005).

Participants completed five IAT blocks in total. In (training) Block 1, they were asked to respond to the pictures of women by pressing a left key for Muslim women and a right key for non-Muslim women. In (training) Block 2 they were asked to use the same two keys to respond to the negative and positive pictures. In Block 3 (a test block) both picture types were presented and participants responded with one key to pictures of both Muslim women and negative scenes, and with the other key to pictures of both non-Muslim women and positive scenes (i.e., congruent trials). In (training) Block 4, the response keys for the pictures of Muslim and non-Muslim women were switched and in Block 5 (a test block), participants had to respond to pictures of both non-Muslim women and negative scenes with one key, and to pictures of both Muslim women and positive scenes with one other key (i.e., incongruent trials). Blocks 1, 2, and 4 consisted of 20 trials, Blocks 3 and 5 of 70 trials each. Every trial started with a fixation point (500 ms), followed by stimulus presentation (680 ms), a blank screen (500 ms), and a feedback screen (1,400 ms). The feedback screen consisted of a movie clip of the evaluator showing either positive (smiling and holding “thumbs up”) or negative (frowning and pointing “thumbs down”) feedback. To ensure that participants were aware of the minimal group membership of their evaluator, we inserted a text display below the movie indicating the personality type of the evaluator, and a text display at the bottom of the screen indicating the personality



*Figure 1.* An (incongruent) IAT trial, depicting positive feedback. The (same) evaluator was either a woman without or with a headscarf, who was presented as a minimal ingroup member (indicated by the shared P-type personality group membership) or outgroup member (different personality type). See the online article for the color version of this figure.

type group of the participant (see Figure 1). In case participants did not respond in time, they saw the words “too late.”

**The IAT effect.** The dependent measure was the IAT effect, indicated by the  $D$  score, and measured as the difference in RTs on incongruent and congruent trials divided by a pooled  $SD$  of all

correct trials (Greenwald et al., 2003). We included all correct (in)congruent trials, replaced error latencies with a replacement value ( $M + 2 SD_{\text{correct}}$ ) and replaced latencies exceeding the maximum response time with the maximum response time of 680 ms (see also van Nunspeet, Derks, et al., 2015, 2014). The resulting positive  $D$

scores are an indication of people's evaluative bias against Muslim women.

**Checks.** Directly after the IAT, participants were asked to indicate whether the test assessed how well they were able to process information and to learn new tasks, or their moral values concerning egalitarianism and discrimination. We also checked the evaluator's minimal group manipulation by asking participants to indicate whether their evaluator was a member of the same or another minimal group. Furthermore, we tested participants' perceptions of the validity of the test (i.e., "My test score can assess what kind of person I am"), and their overall impression of their evaluator ("I think the participant who gave me feedback is competent/kind/moral," three items). Participants could respond on a 7-point Likert scale (1 = *completely disagree*, 7 = *completely agree*).

## Results

**Checks.** Results showed that 96% ( $N = 105$ ) of participants in the morality condition correctly indicated the moral implications of the task; 97% ( $N = 110$ ) of participants in the competence condition correctly indicated the task implications regarding their competence; 95% ( $N = 103$ ) of participants whose evaluator was an ingroup member correctly answered that their evaluator was a member of their own group; and 100% ( $N = 115$ ) of participants whose evaluator was an outgroup member answered correctly that their evaluator was a member of the other group. Excluding the participants who answered one of the checks incorrectly ( $N = 10$ ) did not alter the pattern of the means, so we included these participants in all analyses.

As intended, there were no reliable effects of experimental condition on participants' perceived validity of the test (overall  $M = 3.32$ ,  $SD = 1.48$ ;  $F$ 's  $\leq 2.71$ ,  $p$ 's  $\geq .10$ ). Regarding participants' impression of their evaluator, there were some marginally significant effects: Though the overall impression of their evaluator was quite positive ( $M_{\text{competent}} = 5.17$ ,  $SD = 1.14$ ;  $M_{\text{kind}} = 5.70$ ,  $SD = 0.87$ ;  $M_{\text{moral}} = 5.24$ ,  $SD = 0.97$ ), the evaluator with a headscarf was rated somewhat more kind ( $M = 5.81$ ,  $SD = 0.90$ ) than the evaluator without a headscarf ( $M = 5.58$ ,  $SD = 0.82$ ),  $F(1, 219) = 3.39$ ,  $p = .07$ , and the evaluator (irrespective of both group memberships) was rated as marginally

more competent in the competence condition ( $M = 5.31$ ,  $SD = 1.20$ ) than in the morality condition ( $M = 5.02$ ,  $SD = 1.05$ ),  $F(1, 219) = 3.17$ ,  $p = .08$ . There were no other effects,  $F$ 's  $\leq 2.23$ ,  $p$ 's  $\geq .11$ .

**IAT effect.** An ANOVA with personal motive, and evaluator's minimal and religious group memberships as independent factors revealed a significant main effect of evaluator's religion,  $F(1, 215) = 11.68$ ,  $p = .001$ ,  $\eta_p^2 = .05$ . Whereas participants whose performance was evaluated by a woman without a headscarf showed significant bias against Muslim women ( $M = 0.16$ ,  $SD = 0.45$ ),  $t(108) = 3.73$ ,  $p < .001$ , this bias was reduced to nonsignificance when participants were evaluated by a woman with a headscarf ( $M = -0.04$ ,  $SD = 0.45$ ),  $t(113) = -0.90$ ,  $p = .37$ . Additionally, the interaction between personal motive and evaluator's religion was marginally significant,  $F(1, 215) = 2.88$ ,  $p = .09$ ,  $\eta_p^2 = .01$ . Analysis of simple main effects indicated that when evaluated by a woman with a headscarf the IAT bias no longer emerged in the morality or the competence condition ( $M = -0.03$ ,  $SD = 0.50$ ,  $M = -0.04$ ,  $SD = 0.41$ , respectively;  $F < 1$ ). However, when evaluated by a woman without headscarf, participants for whom the moral task implications were emphasized showed a significantly weaker (nonsignificant) negative bias ( $M = 0.07$ ,  $SD = 0.46$ ),  $t(58) = 1.22$ ,  $p = .23$ ) than participants to whom the competence task implications were emphasized ( $M = 0.27$ ,  $SD = 0.42$ ),  $t(49) = 4.46$ ,  $p < .001$ ,  $F(1, 215) = 4.99$ ,  $p = .03$ ,  $\eta_p^2 = .02$ .<sup>1</sup>

**Additional exploratory analyses: Inspection of RTs.** The analysis of response latencies on correct *congruent* trials (reflecting prejudice-congruent associations between ingroup and positivity, and between outgroup

<sup>1</sup> van Nunspeet, Derks, et al. (2015) showed that emphasizing morality versus competence reduced implicit bias in the presence of an evaluator without a headscarf who was presented as a minimal ingroup (vs. outgroup) member. Although this interaction effect was not significant in the current study ( $F < 1$ ), we did compute simple main effects to explore whether the pattern of results was consistent with this previous finding: Indeed, the effect of personal motive was stronger for performance evaluation by a minimal ingroup ( $M_{\text{morality}} = 0.04$ ,  $SD = 0.52$ ;  $M_{\text{competence}} = 0.27$ ,  $SD = 0.52$ ),  $F(1, 105) = 3.73$ ,  $p = .06$ ,  $\eta_p^2 = .03$ , compared with a minimal outgroup member ( $M_{\text{morality}} = 0.11$ ,  $SD = 0.40$ ;  $M_{\text{competence}} = 0.26$ ,  $SD = 0.31$ ),  $F(1, 105) = 1.49$ ,  $p = .23$ .

and negativity) revealed significant effects of our manipulations in line with the observed pattern of implicit bias reduction reported above. Parallel to the effect of evaluator's religious group on the implicit bias score, this also significantly affected RTs on congruent trials,  $F(1, 215) = 7.09, p = .008, \eta_p^2 = .03$ . Participants whose evaluator wore a headscarf responded more slowly on congruent trials ( $M = 503.97, SD = 24.24$ ) than participants whose evaluator was a woman without a headscarf ( $M = 495.45, SD = 27.13$ ). Moreover, replicating previous work (van Nunspeet et al., 2014), participants to whom the moral ( $M = 502.81, SD = 24.33$ ) rather than the competence ( $M = 496.88, SD = 27.30$ ) task implications were emphasized, responded significantly slower on congruent trials,  $F(1, 215) = 3.92, p = .05, \eta_p^2 = .02$ . Finally, participants responded marginally more slowly on congruent trials when their evaluator was a minimal ingroup ( $M = 502.89, SD = 26.64$ ) than a minimal outgroup member ( $M = 496.91, SD = 25.14$ ),  $F(1, 215) = 2.73, p = .10, \eta_p^2 = .01$ .

Although there were no significant interaction effects;  $F$ 's  $\leq 1.84, p \geq .18$ , to enable a more direct comparison with the analyses for implicit bias, we analyzed RTs on congruent trials per evaluator's religious group membership condition. Replicating the previous pattern, when participants were evaluated by a woman with a headscarf there were no significant effects of personal motive or evaluator's minimal group on congruent response latencies,  $F$ 's  $\leq 2.44, p$ 's  $\geq .12$ . However, when evaluated by a woman without a headscarf, participants responded significantly more slowly on congruent trials in the morality ( $M = 500.61, SD = 23.66$ ) than in the competence condition ( $M = 489.36, SD = 29.83$ ),  $F(1, 105) = 4.67, p = .03, \eta_p^2 = .04$ .

Analysis of response latencies on the correct incongruent trials (reflecting the prejudice-incongruent associations between ingroup and negativity, and between outgroup and positivity) revealed no main effects of personal motive or evaluator's minimal or religious group memberships, nor the interaction between evaluator's religious group and personal motive (all  $F$ 's  $\leq 1.04, p \geq .31$ ). Thus, the experimental manipulations that resulted in a reduction of

implicit bias did not cause participants to adjust their responses on incongruent trials.<sup>2</sup>

## Discussion

Extending previous research (e.g., Blair, 2002; Lai et al., 2014; van Nunspeet, Derks, et al., 2015, 2014) the results of Study 1 reveal that IAT performance evaluation by an evaluator with a headscarf is an impactful way of reducing non-Muslims' implicit bias toward Muslim women. Moreover, consistent with previous results (van Nunspeet, Derks, et al., 2015, 2014), in the presence of an evaluator without a headscarf, the emphasis on participants' morality also significantly reduced bias. In both cases, bias reduction was associated with the inhibition of prejudice confirming responses (i.e., slower responses on prejudice-congruent trials). An emphasis on the moral (as compared with the competence) implications of the task did not reduce implicit bias any further in case participants' performance was evaluated by a representative of the devalued target group. Nor was presenting this evaluator as a minimal ingroup member a cause for further bias reduction.

In Study 2 we focused on the effects of the minimal group context and the emphasis on the moral task implications, while all participants were evaluated by a representative of the devalued target group (i.e., a woman with a headscarf). Here, we increased the number of IAT trials—thereby prolonging the task, and extending participants' exposure to the evaluator, who was again presented as a minimal ingroup (or outgroup) member. Previous research has revealed that people are able to *learn* how to adjust their IAT performance (e.g., Fiedler & Bluemke, 2005), and that learning participants positive associations with the outgroup helps them to reduce their

<sup>2</sup> We found an unexpected interaction between personal motive and evaluator's minimal group membership,  $F(1, 215) = 4.02, p = .05, \eta_p^2 = .02$ . There was no difference between evaluator's minimal group type in the morality condition ( $M_{\text{ingroup}} = 495.28, SD = 24.98; M_{\text{outgroup}} = 498.33, SD = 22.66, F < 1$ ), but participants in the competence condition responded faster on incongruent trials when the evaluator was presented as a minimal outgroup ( $M = 490.10, SD = 21.74$ ) versus ingroup member ( $M = 499.61, SD = 22.78$ ),  $F(1, 215) = 4.63, p = .03, \eta_p^2 = .02$ .

implicit bias (by means of evaluative conditioning; e.g., French et al., 2013; Kawakami et al., 2000; Olson & Fazio, 2006). Additionally, previous IAT studies have revealed how an increase in IAT trials may cause learning effects and consequently affect bias scores (although this was not the focus of that research; see van Nunspeet, Derks, et al., 2015, 2014). We therefore explored in Study 2 whether we might observe an effect of the minimal group membership which the evaluator did or did not share with participants, on a longer version of the IAT.

## Study 2

### Method

**Participants.** As in Study 1, only female, non-Muslim, Dutch university students ( $N = 102$ ;  $M_{\text{age}} = 21.3$  years,  $SD = 3.1$ ) participated in the study for money or course credits. One participant was excluded from the analyses because she did not respond within the time limit on more than 25% of the IAT trials, suggesting lack of attention.

**Procedure.** The IAT and the procedure were highly similar to those described in Study 1. However, in Study 2, all participants received feedback from a woman with a headscarf, and participants were randomly assigned to one of the four experimental conditions of the 2 (Personal Motive: Morality/Competence)  $\times$  2 (Evaluator's Minimal Group Membership: Ingroup/Outgroup) between-participants design. Moreover, the number of trials in the two IAT test blocks was increased to 120 (instead of 70) trials per block.

### Results

**Checks.** Ninety-eight percent ( $N = 49$ ) of participants in the morality and 96% ( $N = 49$ ) of participants in the competence condition correctly reported the task implications. Ninety-two percent ( $N = 47$ ) of participants whose evaluator was an ingroup member and 98% ( $N = 49$ ) of participants whose evaluator was an outgroup member reported their evaluators' minimal group correctly. Because exclusion of the participants who answered one of the checks incorrectly ( $N = 6$ ) did not alter the pattern of means, we included those participants in all analyses.

As intended, participants in all conditions indicated that the test was able to assess what kind of person they are to a similar degree (overall  $M = 3.44$ ,  $SD = 1.57$ ),  $F$ 's  $\leq 1.23$ ,  $p$ 's  $\geq .27$ . Moreover, there were no effects of personal motive or evaluator's minimal group on participants' impression of their evaluator, which was quite positive overall ( $M_{\text{competent}} = 5.36$ ,  $SD = 1.09$ ;  $M_{\text{kind}} = 5.85$ ,  $SD = 0.84$ ;  $M_{\text{moral}} = 5.54$ ,  $SD = 0.98$ ), all  $F$ 's  $\leq 1.68$ ,  $p$ 's  $\geq .20$ ).

**IAT effect.** Consistent with Study 1, now that all participants were evaluated by a woman with a headscarf, on average the implicit bias against Muslim women was non-significant ( $M = -.02$ ,  $SD = .32$ ),  $t(100) = -.53$ ,  $p = .60$ . Additionally, an ANOVA with personal motive and evaluator's minimal group as independent factors revealed a main effect of group membership: Participants whose performance was evaluated by a minimal ingroup member showed significantly less bias ( $M = -0.08$ ,  $SD = 0.27$ ) than participants who were evaluated by a minimal outgroup member ( $M = 0.05$ ,  $SD = 0.35$ ),  $F(1, 97) = 5.02$ ,  $p = .03$ ,  $\eta_p^2 = .05$ . The effect of personal motive was marginally significant,  $F(1, 97) = 2.89$ ,  $p = .09$ ,  $\eta_p^2 = .03$ : Implicit bias was also reduced in case of an emphasis on the moral ( $M = -0.07$ ,  $SD = 0.33$ ) compared with the competence ( $M = 0.03$ ,  $SD = 0.29$ ) task implications.

**Additional exploratory analyses: Inspection of RTs.** We proceeded by examining whether RTs on correct congruent and incongruent trials differed across experimental conditions. Consistent with the results of Study 1 in which we examined RTs on congruent trials for participants who received feedback from a woman with a headscarf, we observed a general tendency to slow down on congruent trials. This indicates that participants' inclination to inhibit prejudice conforming responses did not depend on the evaluator being an in- or an outgroup member or on personal motive ( $F$ 's  $\leq 2.66$ ,  $p$ 's  $\geq .11$ ). Additionally, participants showed reduced response latencies on incongruent trials when the evaluator with a headscarf was presented as a minimal ingroup ( $M = 478.87$ ,  $SD = 23.90$ ) rather than outgroup member ( $M = 493.26$ ,  $SD =$



23.08),  $F(1, 97) = 9.47, p = .003, \eta_p^2 = .09,$ <sup>3</sup> indicating that participants responded faster to associations between outgroup members and positivity and between ingroup members and negativity in this condition.

## Discussion

Replicating the results of Study 1, we found in Study 2 that participants whose performance was evaluated by a representative of the devalued target group in the IAT (i.e., a woman with a headscarf), showed no significant bias toward Muslim women. Interestingly and extending the findings of Study 1, on this prolonged version of the IAT where exposure to the evaluator was extended, we found that presenting this evaluator as a minimal ingroup (vs. outgroup) member further reduced implicit bias. A similar, though less pronounced, effect was found for an emphasis on the moral implications of the task. Additionally, the findings in Study 2 reveal that these effects were associated with participants' responses to prejudice-incongruent (rather than prejudice-congruent) trials. These findings thus suggest that the effectiveness of performance evaluation by a member of the devalued group in the IAT, can even be further enhanced—by introducing another, but shared, (minimal) group membership, or an emphasis on the moral implications of one's behavior.

## General Discussion

This is the first research in which the effectiveness of different social contexts and personal motives on the reduction of implicit bias (measured using an IAT) are compared in one study design. Importantly, and extending previous research (e.g., Blair, 2002; Lai et al., 2014; van Nunspeet, Ellemers, et al., 2015), we directly compared motives and contexts that differ in their level of applicability and implementation. That is, we examined participants' negative bias against Muslim women (a) after emphasizing the moral implications of their task performance; (b) in a social context where their performance was evaluated by a self-relevant other (i.e., a minimal ingroup member); and (c) in a social context where their task performance was evaluated by a representative of the devalued group (i.e., a woman with a headscarf). Additionally, we examined how these social

contexts and personal motives affected participants' response patterns on the IAT, revealing not only their relative effectiveness, but also (at least in part) why this may be the case.

## Theoretical and Practical Implications

Results of Study 1 revealed that participants showed no sign of bias against Muslim women when they were being evaluated by a woman representing the target group (i.e., wearing a headscarf). Interestingly, the significant bias reduction was associated with the inhibition of prejudice: Participants slowed down their responses on congruent trials, suggesting that they aimed to inhibit their prepotent responses to rapidly associate Muslim women with negativity and non-Muslim women with positivity. When the evaluator did not wear a headscarf, we found the same pattern of inhibition of prejudice-conforming responses when the moral (rather than the competence) implications of the test were emphasized. This is especially important because emphasizing one's morality thus proves to be an effective way to facilitate bias reduction, and may therefore be an alternative strategy when intergroup contact is not feasible. Moreover, it may help to increase the reduction of bias in the presence of a representative of the devalued group, as we have seen that an emphasis on the moral task implications also had a (small) effect in Study 2 where the evaluator was always a woman with a headscarf.

Furthermore, in Study 2, we found that presenting the evaluator with a headscarf as an ingroup member on a minimal group dimension reduced participants' bias toward Muslim wom-

<sup>3</sup> To directly test the effect of the increase in trials, we combined Study 2 ( $N = 101$ ) with the condition in which the evaluator was a woman with a headscarf in Study 1 ( $N = 114$ ). Results of an ANOVA with RTs on incongruent trials as dependent variable and number of trials, personal motive, and evaluator's minimal group as independent factors showed a main effect of number of trials: Participants responded faster on incongruent trials when the number of trials was increased ( $M_{120\text{trials}} = 485.99, SD = 24.47, M_{70\text{trials}} = 496.90, SD = 22.66$ ),  $F(1, 207) = 11.82, p = .001, \eta_p^2 = .05$ . Moreover, there was a number of Trials  $\times$  Evaluator's Minimal Group interaction;  $F(1, 207) = 7.50, p = .007, \eta_p^2 = .04$ , indicating that for the increased number of trials, participants only responded faster on incongruent trials when they were evaluated by a minimal ingroup member ( $M_{120\text{trials}} = 478.86, SD = 23.90, M_{70\text{trials}} = 498.75, SD = 22.64$ ),  $F(1, 207) = 18.82, p < .001, \eta_p^2 = .08$ .

en. Interestingly, these effects—found for the prolonged IAT in which the exposure to the evaluator was increased—were associated with faster responding on prejudice-incongruent trials. The findings of Study 2 are consistent with research of Van Bavel and Cunningham (2009) who examined the effects of cross-categorization on a face priming task, and extend prior research which revealed that shared (minimal) group membership(s) can override people's explicit evaluative bias against outgroup members (e.g., Crisp et al., 2001; Urada, Stenstrom, & Miller, 2007), by showing related findings for people's *implicit* bias.

There are thus different ways in which implicit prejudice can be reduced, making it possible to choose one that seems most viable in a specific situation. The presence of a member of the devalued target group seems to have the greatest impact on bias reduction. However, this is not always feasible in real life and we should not solely rely on this effect in everyday interactions: Social groups that are the focus of prejudice research are generally minority groups in society that are often segregated from the majority in education, housing, and work, preventing extensive intergroup interactions. The current research reveals alternative ways to achieve bias reduction that can be applied even in contexts where groups are segregated. Our finding thus contribute to insights on prejudice reduction by demonstrating the potential impact of emphasizing personal moral motives and of providing people with evaluations by others who share the same ingroup membership—also when no representative of the devalued outgroup is present.

Nevertheless, an alternative explanation for our findings could be the more general process of “social tuning” (e.g., Sinclair, Lowery, Hardin, & Colangelo, 2005). If this is the case, one may argue that evaluation by a woman with a headscarf raises concern about one's interaction with Muslim women, and this caused participants to influence their performance. Countering this explanation, however, prior research has revealed that the adjustments in participants' performance on the IAT can indeed be traced to moral motives specifically (van Nunspeet, Derks, et al., 2015, 2014). In fact, in the present study too, the evaluation by a (minimal) ingroup member did not have the same effect on bias reduction when the implications of partic-

ipants' behavior were emphasized in terms of competence. The increased impact of a moral task frame speaks to the validity of our current analysis, and cannot be accounted for by a more general “social tuning” explanation.

### Limitations and Future Research

We do note that specific circumstances were in place in the current research. First, the addition of movie clips to induce social feedback on every trial, makes the interpretation of the findings concerning our bias scores not directly comparable to scores derived from more standard versions of the IAT (in which, for instance, only incorrect responses are indicated; Greenwald et al., 1998). However, our aim was to compare different manipulations using a single IAT, which was similar to the IAT previously used in related studies (van Nunspeet, Derks, et al., 2015, 2014). Second, it remains unclear which aspect of our manipulations concerning the evaluator with a headscarf caused the effect of faster responses on prejudice-incongruent trials. We only revealed this response pattern when the number of IAT trials was increased. However, we did not test the effect of prolonging the IAT in a single study. Further research is thus needed to establish whether the effect was solely caused by this specific increase, and to examine whether experimental learning effects extend beyond IAT task performance.

Furthermore, participants received feedback on every trial and because they made relatively few errors, they received almost continuous positive feedback. This repeated exposure to a positive ingroup exemplar may have contributed to the impact of the presence of a group representative. It stands to reason that this effect may be less strong when there is no feedback provided by that person, or when the task is more difficult and participants are exposed to as much or more negative rather than positive instances of intergroup contact.

The fact that feedback was presented on every trial throughout the experiment may also have contributed to our current results: The greater effect of the presence of the evaluator may reflect more frequent exposure to this manipulation as compared to the effect of the emphasis on the moral task implications which was presented only three times during the IAT (i.e., before the IAT started, and before each test

block). On the one hand, this methodological difference may make it difficult to directly compare the effectivity of the social contexts and personal moral motives. On the other hand, the methodological difference also reflects, at least to some extent, fundamental differences that are likely to exist in the real world. That is, recurrent contact with a member of an outgroup may constitute a relatively effective way to decrease bias, but (as mentioned previously) realizing such extensive intergroup interactions may not always be feasible. By comparison, periodically emphasizing the moral implications of people's behavior may be more easily achieved.

Moreover, the importance people attach to the different group memberships manipulated in this research (i.e., minimal categorization based on a brief personality test vs. strong religious beliefs) may be rather different in itself. Religion could be considered more salient for one's social identity than cooperation and decision-making styles. Although different types of cross-cutting categorizations are also likely to differ in their salience in real life, this could explain the difference in effect sizes between the social contexts in our research. A follow-up study in which participants' performance is evaluated by someone with whom they (do not) share two equally relevant group memberships could address the question whether such contexts would have an equal impact on bias reduction. Because the inclusion of only female participants is also a limitation of our research (limiting the generalizability of our findings) such a study could compare the effects of an evaluator with the same or a different religion and gender. These two group memberships are more equal in their societal relevance, and including a male evaluator would make up a more balanced design in which both male and female participants should be included. Nevertheless, even though the difference in the strength of our social context effects may be related to the way they were introduced in our research, we think it is important to recognize the effects of focusing people on their moral values, and of introducing a (minimal) shared group dimension—especially because we showed how these have additional (although somewhat less pronounced) effects.

Finally, future research may extend our findings by examining the cognitive processes underlying the impact of personal motives and

social context factors on prejudice reduction (using, for instance, event-related brain potentials). Moreover, computational models may be applied to disentangle the processes underlying automatic evaluations and control, such as the process-dissociation model (e.g., Payne, 2001) and the Quad-model (Conrey, Sherman, Gawronski, Hugenberg, & Groom, 2005).

## Conclusion

We have shown that evaluative bias against Muslim women can be reduced by several means. Being evaluated by a representative of the Muslim target group (i.e., a woman who wears a headscarf) significantly reduces bias. Additionally, under particular circumstances, introducing a shared minimal group membership with the evaluator may have an additional effect. However, when such intergroup contact is not feasible, prejudice reduction can also be instigated by emphasizing the moral implications of people's displays of implicit bias.

## References

- Amodio, D. M., & Devine, P. G. (2006). Stereotyping and evaluation in implicit race bias: Evidence for independent constructs and unique effects on behavior. *Journal of Personality and Social Psychology, 91*, 652–661. <http://dx.doi.org/10.1037/0022-3514.91.4.652>
- Blair, I. V. (2002). The malleability of automatic stereotypes and prejudice. *Personality and Social Psychology Review, 6*, 242–261. [http://dx.doi.org/10.1207/S15327957PSPR0603\\_8](http://dx.doi.org/10.1207/S15327957PSPR0603_8)
- Castelli, L., & Tomelleri, S. (2008). Contextual effects on prejudiced attitudes: When the presence of others leads to more egalitarian responses. *Journal of Experimental Social Psychology, 44*, 679–686. <http://dx.doi.org/10.1016/j.jesp.2007.04.006>
- Conrey, F. R., Sherman, J. W., Gawronski, B., Hugenberg, K., & Groom, C. J. (2005). Separating multiple processes in implicit social cognition: The quad model of implicit task performance. *Journal of Personality and Social Psychology, 89*, 469–487. <http://dx.doi.org/10.1037/0022-3514.89.4.469>
- Crisp, R. J., Hewstone, M., & Rubin, M. (2001). Does multiple categorization reduce intergroup bias? *Personality and Social Psychology Bulletin, 27*, 76–89. <http://dx.doi.org/10.1177/0146167201271007>
- Crosby, F., Bromley, S., & Saxe, L. (1980). Recent unobtrusive studies of Black and White discrimi-

- nation and prejudice: A literature review. *Psychological Bulletin*, 87, 546–563. <http://dx.doi.org/10.1037/0033-2909.87.3.546>
- Devine, P. G., Plant, E. A., Amodio, D. M., Harmon-Jones, E., & Vance, S. L. (2002). The regulation of explicit and implicit race bias: The role of motivations to respond without prejudice. *Journal of Personality and Social Psychology*, 82, 835–848. <http://dx.doi.org/10.1037/0022-3514.82.5.835>
- Dovidio, J. F., Kawakami, K., & Beach, K. R. (2001). Implicit and explicit attitudes: Examination of the relationship between measures of intergroup bias. In R. Brown & S. Gaertner (Eds.), *Blackwell handbook of social psychology: Intergroup processes* (pp. 175–197). Maiden, MA: Blackwell.
- Fiedler, K., & Bluemke, M. (2005). Faking the IAT: Aided and unaided response control on the Implicit Association Test. *Basic and Applied Social Psychology*, 27, 307–316.
- Frantz, C. M., Cuddy, A. J., Burnett, M., Ray, H., & Hart, A. (2004). A threat in the computer: The race implicit association test as a stereotype threat experience. *Personality and Social Psychology Bulletin*, 30, 1611–1624. <http://dx.doi.org/10.1177/0146167204266650>
- French, A. R., Franz, T. M., Phelan, L. L., & Blaine, B. E. (2013). Reducing Muslim/Arab stereotypes through evaluative conditioning. *The Journal of Social Psychology*, 153, 6–9. <http://dx.doi.org/10.1080/00224545.2012.706242>
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74, 1464–1480. <http://dx.doi.org/10.1037/0022-3514.74.6.1464>
- 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. <http://dx.doi.org/10.1037/0022-3514.85.2.197>
- Kawakami, K., Dovidio, J. F., Moll, J., Hermsen, S., & Russin, A. (2000). Just say no (to stereotyping): effects of training in the negation of stereotypic associations on stereotype activation. *Journal of Personality and Social Psychology*, 78, 871–888.
- Lai, C. K., Marini, M., Lehr, S. A., Cerruti, C., Shin, J. E. L., Joy-Gaba, J. A., . . . Nosek, B. A. (2014). Reducing implicit racial preferences: I. A comparative investigation of 17 interventions. *Journal of Experimental Psychology: General*, 143, 1765–1785. <http://dx.doi.org/10.1037/a0036260>
- Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2005). International affective picture system (IAPS): *Digitized photographs, instruction manual and affective ratings*. Tech. Rep. No. A-6. University of Florida, Gainesville, FL. Retrieved from <https://www2.uni-fesp.br/dpsicobio/adap/instructions.pdf>
- Lemmer, G., & Wagner, U. (2015). Can we really reduce ethnic prejudice outside the lab? A meta-analysis of direct and indirect contact interventions. *European Journal of Social Psychology*, 45, 152–168. <http://dx.doi.org/10.1002/ejsp.2079>
- Lowery, B. S., Hardin, C. D., & Sinclair, S. (2001). Social influence effects on automatic racial prejudice. *Journal of personality and social psychology*, 81, 842–855.
- Miles, E., & Crisp, R. J. (2014). A meta-analytic test of the imagined contact hypothesis. *Group Processes & Intergroup Relations*, 17, 3–26. <http://dx.doi.org/10.1177/1368430213510573>
- Olson, M. A., & Fazio, R. H. (2006). Reducing automatically activated racial prejudice through implicit evaluative conditioning. *Personality and Social Psychology Bulletin*, 32, 421–433. <http://dx.doi.org/10.1177/0146167205284004>
- Payne, B. K. (2001). Prejudice and perception: The role of automatic and controlled processes in misperceiving a weapon. *Journal of Personality and Social Psychology*, 81, 181–192.
- Richeson, J. A., & Ambady, N. (2003). Effects of situational power on automatic racial prejudice. *Journal of Experimental Social Psychology*, 39, 177–183. [http://dx.doi.org/10.1016/S0022-1031\(02\)00521-8](http://dx.doi.org/10.1016/S0022-1031(02)00521-8)
- Rudman, L. A., Dohn, M. C., & Fairchild, K. (2007). Implicit self-esteem compensation: Automatic threat defense. *Journal of Personality and Social Psychology*, 93, 798–813. <http://dx.doi.org/10.1037/0022-3514.93.5.798>
- Sinclair, S., Lowery, B. S., Hardin, C. D., & Colangelo, A. (2005). Social tuning of automatic racial attitudes: The role of affiliative motivation. *Journal of Personality and Social Psychology*, 89, 583–592. <http://dx.doi.org/10.1037/0022-3514.89.4.583>
- Urada, D., Stenstrom, D. M., & Miller, N. (2007). Crossed categorization beyond the two-group model. *Journal of Personality and Social Psychology*, 92, 649–664. <http://dx.doi.org/10.1037/0022-3514.92.4.649>
- Van Bavel, J. J., & Cunningham, W. A. (2009). Self-categorization with a novel mixed-race group moderates automatic social and racial biases. *Personality and Social Psychology Bulletin*, 35, 321–335. <http://dx.doi.org/10.1177/0146167208327743>
- van Nunspeet, F., Derks, B., Ellemers, N., & Nieuwenhuis, S. (2015). Moral impression management: Evaluation by an ingroup member during a moral IAT affects perceptual attention and conflict- and response monitoring. *Social Psychological & Personality Science*, 6, 183–192. <http://dx.doi.org/10.1177/1948550614548076>

- van Nunspeet, F., Ellemers, N., & Derks, B. (2015). Reducing implicit bias: How moral motivation helps people refrain from making 'automatic' prejudiced associations. *Translational Issues in Psychological Science, 1*, 382–391. <http://dx.doi.org/10.1037/tps0000044>
- van Nunspeet, F., Ellemers, N., Derks, B., & Nieuwenhuis, S. (2014). Moral concerns increase attention and response monitoring during IAT performance: ERP evidence. *Social Cognitive and Affective Neuroscience, 9*, 141–149. <http://dx.doi.org/10.1093/scan/nss118>
- Webb, T. L., Sheeran, P., & Pepper, J. (2012). Gaining control over responses to implicit attitude tests: Implementation intentions engender fast responses on attitude-incongruent trials. *British Journal of Social Psychology, 51*, 13–32. <http://dx.doi.org/10.1348/014466610X532192>

Received August 15, 2016

Revision received June 29, 2017

Accepted July 21, 2017 ■



AMERICAN  
PSYCHOLOGICAL  
ASSOCIATION

APA JOURNALS®

Publishing on the Forefront of Psychology

## ORDER INFORMATION

Start my 2018 subscription to

***Group Dynamics: Theory, Research, and Practice***®

ISSN: 1089-2699

### PRICING

APA Member/Affiliate	\$79
Individual Nonmember	\$146
Institution	\$651

Call **800-374-2721** or **202-336-5600**

Fax **202-336-5568** | TDD/TTY **202-336-6123**

**Subscription orders must be prepaid.** Subscriptions are on a calendar year basis. Please allow 4-6 weeks for delivery of the first issue.

Learn more and order online at:  
[www.apa.org/pubs/journals/gdn](http://www.apa.org/pubs/journals/gdn)

GDNA18