Social Psychological and Personality Science
2015, Vol. 6(2) 183-192

DOI: 10.1177/1948550614548076
spps.sagepub.com
SAGE

Moral Impression Management: Evaluation by an In-Group Member During a Moral IAT Affects Perceptual Attention and Conflict and Response Monitoring

Félice van Nunspeet[1,2], Belle Derks[1,2], Naomi Ellemers[1,2], and Sander Nieuwenhuis[2,3]

Abstract

Previous research revealed that emphasizing morality increases motivational processes that improve people's task performance. Here we examined whether this emphasis differentially affects people's performance in the presence of an in-group compared to an out-group member. Ostensibly while being evaluated by another person, participants performed an Implicit Association Test that was framed as a test of either their morality or their competence. Results showed a smaller bias toward Muslim women in the morality compared to the competence condition, but this effect was more pronounced when participants were evaluated by a member of their minimal in-group. Moreover, in that same condition, event-related potentials revealed increased perceptual attention (N1) and affected conflict and response monitoring (N450 and error-related negativity). These findings suggest that being moral is especially important when monitored by the in-group and reveal the cognitive processes associated with controlling intergroup bias in a social situation.

Keywords

attitudes, categorization, control, morality, intergroup processes, self/identity, motivation and performance, prejudice/stereotyping, social identity

According to the Oxford dictionary, being moral means ''holding high principles for proper conduct.'' But what is considered ''proper''? Individuals can have their own principles of what is good and bad. Nevertheless, the groups to which we belong and the group members to whom we feel connected often define relevant standards of morality (see also Ellemers & Van den Bos, 2012). Behaving according to those standards is perceived as important, that is, people are motivated to adjust their behavior to moral (compared to competence) in-group norms (Ellemers, Pagliaro, Barreto, & Leach, 2008), as a way to earn respect from fellow in-group members (Pagliaro, Ellemers, & Barreto, 2011). Moreover, people identify more strongly with a moral than a competent group and are more proud to be a member of that group (Leach, Ellemers, & Barreto, 2007). People's willingness to belong to moral groups and their pride in being a moral group member can be explained by Social Identity Theory which proposes that people's self-views depend upon the groups to which they belong (Tajfel, 1978). Indeed, moral characteristics convey important social information, that is, when people form an impression of others, they are more inclined to gather information concerning morality than about competence or sociability, both explicitly (Brambilla, Rusconi, Sacchi, & Cherubini, 2011) and implicitly (i.e., when an impression has to be made within milliseconds; Willis & Todorov, 2006). Moreover, people monitor their own behavior to maintain a moral self-image (Jordan & Monin, 2008). Thus, due to the identity-defining function of morality—especially in group contexts—being moral is what we consider important in others and ourselves (Ellemers & Van den Bos, 2012). The motivation to be moral elicits the tendency to adjust one's behavior to moral norms, even on less explicit measures. van Nunspeet, Ellemers, Derks, and Nieuwenhuis (2014) have shown that people adapt their behavior when this is perceived as indicative of their morality. During an Implicit Association

[1] Social and Organizational Psychology Unit, Institute of Psychology, Leiden University, the Netherlands
[2] Leiden Institute for Brain and Cognition, Leiden, the Netherlands
[3] Cognitive Psychology Unit, Institute of Psychology, Leiden University, the Netherlands

Corresponding Author:
Félice van Nunspeet, Social and Organizational Psychology Unit, Institute of Psychology, Leiden University, P.O.B. 9555, 2300 RB, the Netherlands.
Email: nunspeetfvan@fsw.leidenuniv.nl

Test (IAT) participants were more inclined to control their negative bias toward Muslim women when they thought the test measured their morality (when compared to their competence).

Since the significance of morality derives from its implications for people's social identity, we expect the motivation to be moral to be particularly relevant in an in-group context. We thus hypothesize that when participants are evaluated by an in-group member, rather than an out-group member, they are more motivated to control their bias during performance on a ''moral'' IAT.

Moral Performance and Event-Related Brain Potentials (ERPs)

The desire to be moral may elicit socially desirable answers, which complicates the interpretation of self-reports on the importance of morality. Additionally, it remains unclear *how* people control their behavior to appear moral. Examining the cognitive processes involved in displaying moral behavior can elucidate how this is achieved.

van Nunspeet et al. (2014) previously revealed that three different types of cognitive processes were affected when the implications of the IAT were presented in terms of morality compared to competence, namely perceptual attention toward stimulus information, the processing of stereotype-incongruent versus stereotype-congruent information, and response monitoring. ERPs suggested that participants paid more attention to the group membership of the targets presented in the IAT. This so-called social categorization was evident in modulations of the N1 and P150, two ERP components occurring around 100–200 ms after stimulus onset, that typically are larger when viewing in-group versus out-group faces (e.g., Amodio, 2010; Ito & Urland, 2003; Kubota & Ito, 2007). van Nunspeet et al. (2014) argued that perceptual attention to the group membership of the IAT targets was enhanced to enable participants to perform in line with their moral values.

When morality instead of competence was emphasized in the IAT instruction, participants also showed enhanced brain responses to the difference between stereotype-incongruent and stereotype-congruent trials. This process of conflict monitoring can be observed in the N450, a negative deflection around 400–500 ms after stimulus onset which is evident in language incongruencies (e.g., Nigam, Hoffman, & Simons, 1992), the Stroop task (e.g., Rebai, Bernard, & Lannou, 1997), and the IAT (Williams & Themanson, 2011). Furthermore, participants seemed to be more concerned about displaying immoral than incompetent behavior, as the participants who had read the moral test implications also showed enhanced error monitoring. A process associated with the error-related negativity (ERN), that is, a negative deflection within 100 ms after a response is given which is known to be larger for incorrect than correct responses (e.g., Gehring, Goss, Coles, Meyer, & Donchin, 1993) and for significant compared to nonsignificant errors (Hajcak, Moser, Yeung, & Simons, 2005).

Moral Performance in Group Contexts

In the current research, we hypothesize that participants are more motivated to perform in line with moral values when they are being evaluated by a self-relevant other. To examine this, we need to exclude alternative motivations to control bias, such as the wish to avoid offending the IAT target group in the presence of an ethnic out-group member (Lowery, Hardin, & Sinclair, 2001; Richeson & Ambady, 2003). We therefore introduced minimal categories. Based on a questionnaire ostensibly assessing personality styles, non-Muslim participants were evaluated by a non-Muslim individual who was presented as someone with the same (in-group) or another (out-group) personality type.

We thus predicted that participants show a weaker IAT bias when the moral (compared to competence) test implications are emphasized, especially when they are evaluated by an in-group (vs. out-group) member. Extending prior research (van Nunspeet, Ellemers, Derks, & Nieuwenhuis, 2014), we also anticipated that participants who are evaluated by an in-group member and to whom the moral test implications are emphasized show enhanced ERP modulations associated with early perceptual attention (indexed by N1 and/or P150 modulations) and conflict- and error monitoring (indicated by the N450 and ERN). We tested these hypotheses in two studies by examining whether the morality/in-group condition differed from the other three conditions. Specifically, we conducted an initial behavioral study (Study 1) and a follow-up study in which we recorded an electroencephalogram (EEG) during IAT performance (Study 2).

Study 1

Method

Participants and Design

Ninety-five non-Muslim students (3 males, $M_{age} = 19.2$ years, $SD = 2.0$) participated for money or course credits. One participant was excluded from the analyses because he or she responded too late on more than 25% of trials, indicating lack of attention. Participants were randomly assigned to conditions in the 2 (task domain: morality/competence) $\times$ 2 (evaluator: in-group/out-group member) between-participants design.

Procedure

After participants signed an informed consent in which it was explained that their participation could be recorded on video, they were seated in an individual room with a webcam, headphone, and a camera placed in a top corner of the cubicle. Participants were told they would be paired with another participant based on questionnaire scores (ostensibly) assessing their personality styles and indicating whether they were either a so-called ''P'' type or a so-called ''O'' type. After completing the questionnaire and a short pause, participants

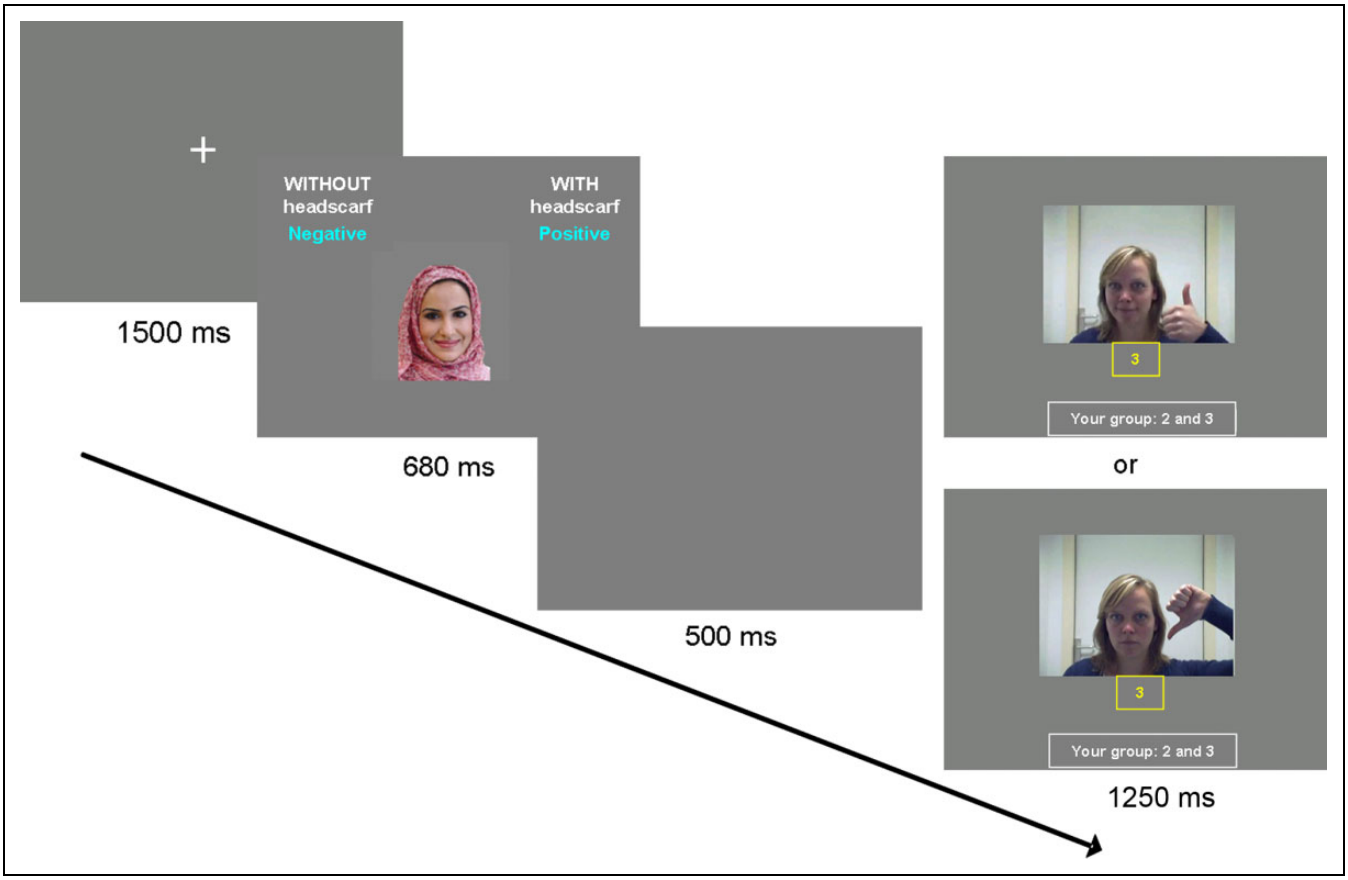

Figure 1. Example of an Implicit Association Test (IAT) trial.

saw their own and other participants' scores (i.e., participant numbers were presented in combination with the P- and O-personality styles). Participants were then informed that they would cooperate with a member either of the same group or a different group (as determined by their personality style). Then the IAT was introduced as a reaction time task during which the other person (i.e., the evaluator) would observe and give them feedback on every trial. After that, a webcam connection was simulated. The evaluator introduced himself or herself and told that he or she would observe and provide feedback to the participant. A smile and thumbs up would follow a correct trial; frowning and pointing thumbs down an incorrect trial. Participants then read about either the moral implications or competent implications of the upcoming task and started with the IAT.

In reality, all participants were said to have a P-personality style and were introduced to a (same gender) confederate whose introduction was prerecorded. After the IAT, participants completed additional questions, were debriefed, and thanked. The experiment lasted approximately 50 min.

Instruments

The IAT. Participants performed an IAT as designed by Greenwald, McGhee, and Schwartz (1998). Stimuli representing the target concepts consisted of 10 pictures of non-Muslim and 10 pictures of Muslim women (faces without and with a headscarf, respectively). Stimuli that represented positive and negative attributes consisted of five pictures of positive and five pictures of negative scenes selected from the International Affective Picture System (IAPS; Lang, Bradley, & Cuthbert, 2005).

For congruent trials, pictures of non-Muslim women shared the same response key as positive pictures, and pictures of Muslim women shared the same response key as negative pictures. For incongruent trials, this was the case for non-Muslim women and negative pictures and for Muslim women and positive pictures. The order of the incongruent blocks was counterbalanced across participants. Training Blocks 1, 2, and 4 consisted of 26 trials, and Test Blocks 3 and 5 of 156 trials each. Every trial started with a fixation point, followed by a stimulus, a blank screen, and a feedback screen (Figure 1). To ensure that participants were aware that the evaluator (presented during the feedback) was an in- or an out-group member, two text displays indicated the group memberships of the participant and the evaluator. In case participants did not respond in time, the feedback screen showed the words "too late."

Task domain. Task domain was introduced using the instructions described in van Nunspeet et al. (2014). Without

mentioning the IAT design, or how performance would be measured, participants read the test would indicate their moral values concerning egalitarianism in the *morality* condition or their ability to learn new tasks in the *competence* condition. In both conditions, participants were instructed to respond as quickly and accurately as possible. The test implications were repeated before each test block.

The IAT effect. The dependent measure was the IAT effect (i.e., the *D*-score), which was calculated as the difference in reaction times on incongruent and congruent trials divided by a pooled *SD* of all correct trials (Greenwald, Nosek, & Banaji, 2003; van Nunspeet et al., 2014).

Checks. To check that the perceived validity of the IAT did not differ between conditions, we asked participants to respond to the statement, "My test score can assess what kind of person I am." Furthermore, we asked to what extent participants hoped to have made a good impression on the evaluator, "I hope the evaluator has the impression that I am competent/kind/moral" (3 items, $\alpha = .90$). Identification with the P-type group was checked with 2 items ("I identify strongly with the P group" and "I feel equal to the other group members in terms of general attitudes and beliefs"; $r = .41$, $p < .001$). Participants could respond on a 7-point scale (1 = *completely disagree* to 7 = *completely agree*).

Results

Checks

As intended, participants in the four experimental conditions did not differ in their ability to identify with the experimentally created in-group (grand-average $M = 3.77$, $SD = 1.20$), $F(3, 90) = 1.37$, $p = .26$, and did not think differently about the perceived validity of the test, $M = 3.64$, $SD = 1.62$; $F < 1$. In line with prior findings, participants in the morality condition indicated positive impression management to be more important than participants in the competence condition, $M_{\text{morality}} = 4.83$, $SD = 1.01$; $M_{\text{competence}} = 4.28$, $SD = 1.04$; $F(1, 90) = 6.58$, $p = .01$, $\eta^2_p = .07$. There was neither an evaluator effect nor an interaction effect, Fs < 1.49, ps > .23.

IAT Effect

More errors were made on incongruent ($M = 9.35$, $SD = 7.01$) than congruent trials ($M = 6.46$, $SD = 5.40$), $t(93) = 4.50$, $p < .001$, which was not affected by task domain or evaluator, Fs < 1.87, ps > .18. Overall, participants showed the standard IAT effect, indicating a negative implicit bias toward Muslim women, $t(93) = 6.83$, $p < .001$. Consistent with previous research (van Nunspeet et al., 2014), an analysis of variance (ANOVA) with the *D*-score as dependent variable and domain and evaluator as independent factors revealed a significant main effect of domain, $F(1, 90) = 5.57$, $p = .02$, $\eta^2_p = 0.06$. Overall, participants in the morality condition showed a smaller IAT effect than participants in the competence condition, $M_{\text{morality}} = 0.18$, $SD = 0.34$; $M_{\text{competence}} = 0.36$, $SD = 0.39$.[1] We also found an interaction effect between domain and evaluator, $F(1, 90) = 4.26$, $p = .04$, $\eta^2_p = 0.05$ (Figure 2). As expected, participants who were evaluated by an in-group member while they performed the "moral" IAT showed a significant weaker negative bias ($M_{\text{morality/in-group}} = 0.10$, $SD = 0.32$) than participants in the other conditions ($M_{\text{morality/out-group}} = 0.26$, $SD = 0.34$; $M_{\text{competence/in-group}} = 0.43$, $SD = 0.33$; $M_{\text{competence/out-group}} = 0.29$, $SD = 0.44$), $t(93) = -2.65$, $p = .01$.

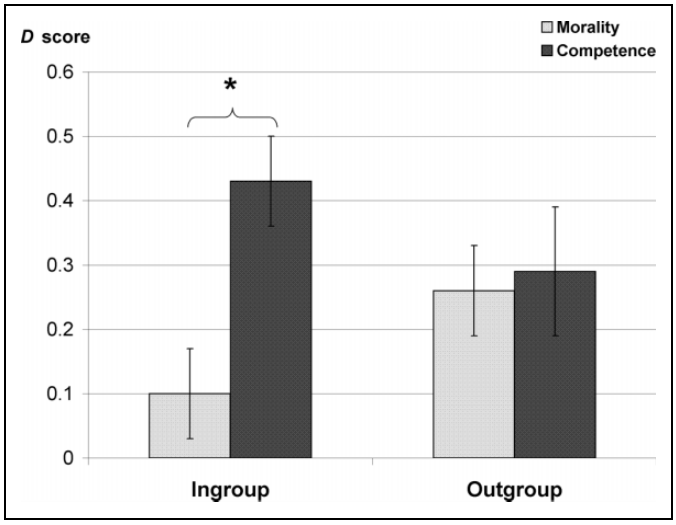

Figure 2. Condition means relevant to the interaction effect on the *D*-scores.

These findings extend previous research by showing that moral impression management is particularly important in an intragroup context (even if the broader significance of the in-group is relatively minimal). In Study 2, we examine the cognitive processes associated with the tendency to conform to moral values in group contexts.

Study 2

Method

Participants and Design

Sixty-seven non-Muslim, right-handed, healthy students (18 males, $M_{\text{age}} = 20.6$ years, $SD = 2.1$) participated for money or course credits. Three participants were excluded from all analyses because they responded too late on more than 25% of the trials, indicating lack of attention. Two other participants could not be included in the analysis of self-report data because they failed to complete the questions, and four participants had to be excluded from the ERP analyses, because of technical problems during the EEG acquisition. Remaining participants were randomly distributed across conditions of the 2 (domain: morality/competence) × 2 (evaluator: in-group/out-group member) between-participants design.

Procedure

The procedure and measures were similar to Study 1, with the following exceptions. Participants completed the questionnaire to ostensibly determine personality style before they came to the EEG lab. The feedback screens in the IAT consisted of a photograph of the confederate instead of a movie display. Finally, to elicit a sufficient number of errors to reliably estimate the ERN, the maximum duration of the stimulus presentation was reduced from 680 to 550 ms, and the total number of test trials increased to 600 (300 congruent and 300 incongruent trials).

EEG Acquisition

The EEG was recorded from 19 silver/silver chloride (Ag/AgCl) scalp electrodes mounted in an elastic cap and, from the left and right mastoids, using a 19-channel Biosemi active-electrode recording system (sampling rate 256 Hz). To assess horizontal and vertical eye movements, electrodes were placed on the outer canthi of the eyes and approximately 1 cm above and below the participant's right eye. EEG activity was recorded using ActiView software, off-line data analyses were performed using brain vision analyzer, and the experiment was presented with E-prime software. The EEG signal was referenced off-line to the average mastoid signal, corrected for ocular and eyeblink artifacts using the method of Gratton, Coles, and Donchin (1983), and filtered (1–15 Hz). Single-trial stimulus-locked and response-locked epochs were extracted, ranging from −300 to 1,000 ms after the event. These epochs were subjected to artifact rejection, then averaged and baseline corrected by subtracting the average signal value between 200 and 0 ms prestimulus or between 300 and 50 ms prior to the response. Separate stimulus-locked ERP epochs were created for correct congruent and incongruent trials with pictures of Muslim and non-Muslim women. Separate response-locked ERP epochs were created for correct and error trials.

ERP Analyses

Visual inspection of the data indicated that the N1, P150, and ERN components were most evident at electrode sites FCz and Cz. The N450 was most evident at CPz and Pz. The stimulus-locked ERP components were quantified as the peak amplitude within a time window poststimulus (N1: 90–110 ms, P150: 100–250 ms, and N450: 325–500 ms), whereas the ERN was quantified as the peak amplitude of the signal between −50 and 150 ms around the response. Each average ERN was based on at least 10 trials.[2] Peak amplitude values of the N1, P150, and N450 were submitted to a 2 (electrode site: FCz/Cz or CPz/Pz) × 2 (target: Muslim/non-Muslim women) × 2 (congruency: congruent/incongruent) mixed-model ANOVA. Peak amplitude values of the ERN were submitted to a 2 (electrode site) × 2 (accuracy: correct/error) × 2 (congruency) mixed-model ANOVA.[3] In every analysis, domain (morality/competence) and evaluator (in-group/out-group) were included as between-participant factors.

Results

Behavioral Results

Checks. As intended, identification with the in-group (2 items, $r = .50$, $p < .001$) was equal across experimental conditions (grand-average $M = 3.53$, $SD = 1.36$), $F < 1$, as was the perceived validity of the test; $M = 3.58$, $SD = 1.56$; $F < 1$. Again, participants in the morality condition indicated more concern about impression management than in the competence condition, $M_{\text{morality}} = 5.25$, $SD = 0.83$; $M_{\text{competence}} = 4.63$, $SD = 0.82$; $F(1, 58) = 8.39$, $p = .01$, $\eta^2_p = .13$.

IAT effect. More errors were made on incongruent ($M = 34.4$, $SD = 18.4$) than congruent trials ($M = 25.6$, $SD = 15.5$), $t(63) = 4.87$, $p < .001$, and participants in the in-group evaluator condition made fewer errors ($M = 50.3$, $SD = 24.9$) than participants in the out-group evaluator condition ($M = 70.4$, $SD = 33.4$), $F(1, 60) = 7.28$, $p = .01$, $\eta^2_p = .11$. Overall, participants showed the standard IAT effect, indicating a negative implicit bias toward Muslim women, $t(63) = 5.46$, $p < .001$. However, IAT effects were not affected by evaluator or task domain (Fs < 1), nor did the morality/in-group condition differ from the other three conditions, $t(63) = 0.12$, $p = .90$. This is likely due to the changes we made to optimize the task for ERP recordings: To ensure enough errors to reliably estimate the ERN, the maximum response time was reduced. In Study 1—and in previous research (van Nunspeet et al., 2014)—participants controlled their bias by delaying responses on congruent trials, which may have been impossible in this study, given the tight response deadline. A follow-up study corroborates this explanation. When we examined behavioral effects of task instruction and in-group/out-group evaluators using a response window of 680 ms (as in Study 1 and prior research), the IAT bias was significantly lower in the morality compared to the competence condition, when participants were evaluated by a minimal in-group member (van Nunspeet, Ellemers, & Derks, 2014).

ERP Results

Perceptual attention

N1. The N1 results revealed the expected evidence of social categorization: The N1 was larger for pictures of Muslim women ($M = -7.18$ μV, $SE = 0.35$) than non-Muslim women ($M = -6.91$, standard error [SE] = 0.35), $F(1, 56) = 3.52$, $p = .07$, $\eta^2_p = .06$ (Figure 3). As predicted, the discrepancy between these targets differed depending on the interaction between task domain and evaluator, $F(1, 56) = 4.36$, $p = .04$, $\eta^2_p = .07$: The N1 modulation of social categorization was significantly increased in the morality/in-group condition when compared to the other conditions, $t(59) = -2.74$, $p = .01$ (Figure 4).[4]

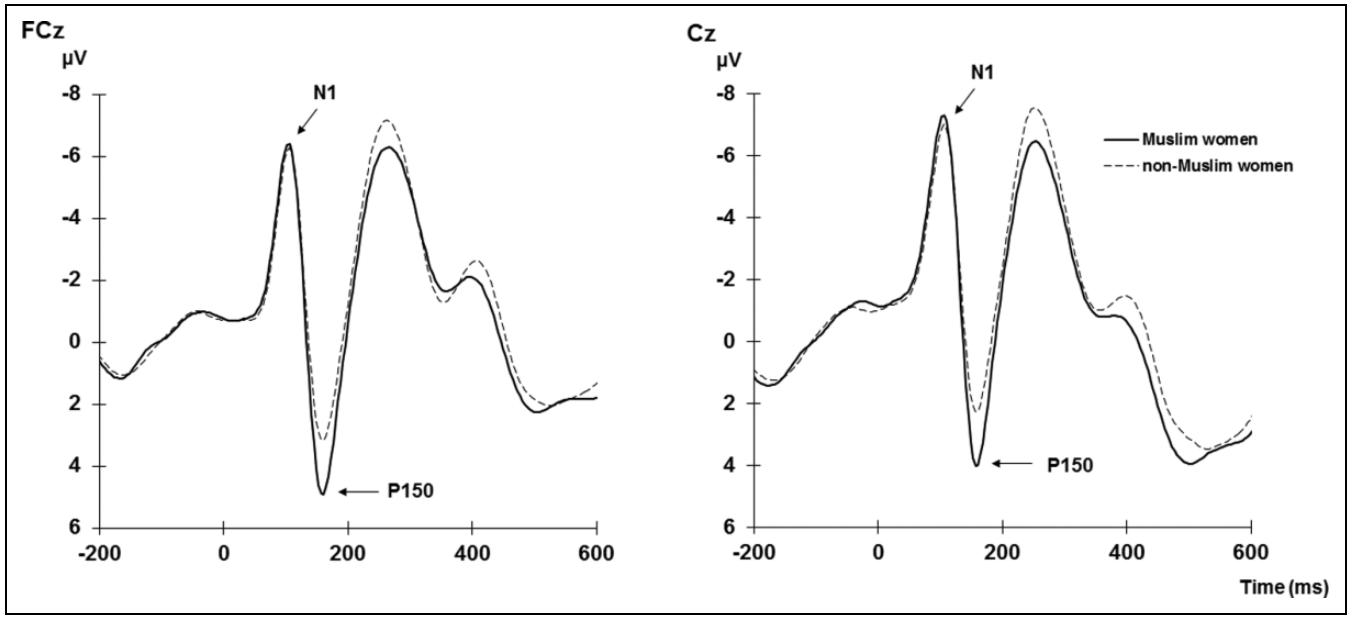

Figure 3. Differences in N1 and P150 amplitudes for pictures of Muslim and non-Muslim women. Only the N1 modulation interacted with task domain and evaluator.

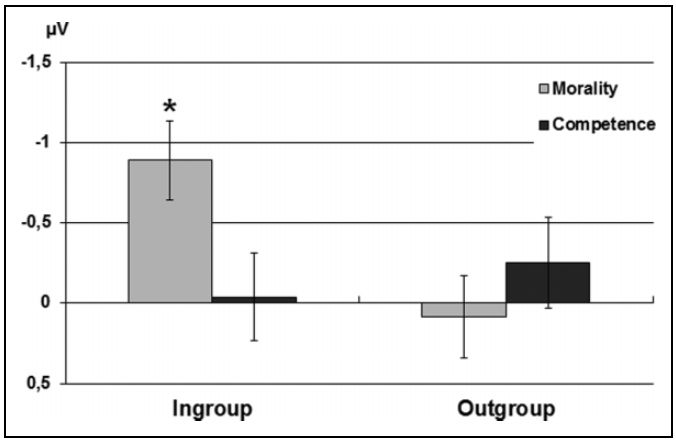

Figure 4. The mean differences in N1 amplitude between Muslim and non-Muslim targets for each condition.

P150. Analyses of the P150 only revealed the expected main effect of target: The P150 was larger for pictures of Muslim women ($M = 5.44$ μV, $SE = 0.48$) than for pictures of non-Muslim women ($M = 3.77$, $SE = 0.43$), $F(1, 56) = 93.13$, $p < .001$, $\eta^2_p = .62$ (Figure 3). No effects of task domain or evaluator were found (Fs < 1).

Conflict and response monitoring

N450. Results showed the anticipated effect of congruency: The N450 was larger for incongruent ($M = -0.13$ μV, $SE = 0.33$) than congruent ($M = 0.64$, $SE = 0.40$) trials, $F(1, 56) = 5.92$, $p = .02$, $\eta^2_p = .10$ (Figure 5). There was also a main effect of target: The N450 was smaller for Muslim ($M = 0.79$ μV, $SE = 0.33$) than non-Muslim trials ($M = -0.28$, $SE = 0.37$), $F(1, 56) = 24.06$, $p < .001$, $\eta^2_p = .30$. The predicted three-way interaction between congruency, domain, and evaluator was not significant ($F < 1$). However, there was a significant four-way interaction between congruency, domain, evaluator, and target, which revealed that conflict monitoring (i.e., the difference between the N450 on incongruent vs. congruent trials) differed depending on the interaction between task domain, evaluator, and (unexpectedly) target, $F(1, 56) = 5.75$, $p = .02$, $\eta^2_p = .09$. Separate analyses for the task domain conditions revealed that the interaction between target and evaluator was significant in the morality condition, $F(1, 31) = 5.36$, $p < .03$, $\eta^2_p = .15$, but not in the competence condition, $F(1, 25) = 1.30$, $p = .27$. This interaction indicated that the difference in conflict monitoring between Muslim and non-Muslim targets was evident in the in-group evaluator condition, $F(1, 31) = 6.61$, $p < .02$, $\eta^2_p = .18$, but not in the out-group evaluator condition, $F < 1$. Specifically, conflict monitoring differed in the morality/in-group condition when compared to the other three conditions, $t(59) = 2.31$, $p = .03$, but—unexpectedly—this occurred because these participants showed *less* conflict monitoring for Muslim than non-Muslim targets (Figure 6).

ERN. As anticipated, results showed that the ERN was larger for incorrect ($M = -6.90$ μV, $SE = 0.69$) than correct trials ($M = 2.95$, $SE = 0.46$), $F(1, 44) = 173.52$, $p < .001$, $\eta^2_p = .80$. There was a marginally significant interaction effect between accuracy and task domain, $F(1, 44) = 3.37$, $p = .07$, $\eta^2_p = .07$, indicating that the ERN modulation was somewhat larger in the competence ($M_{difference} = -11.22$ μV, $SE = 1.12$), $F(1, 44) = 100.07$, $p < .001$, $\eta^2_p = .70$, than the morality condition ($M_{difference} = -8.48$ μV, $SE = 0.99$), $F(1, 44) = 73.49$, $p < .001$, $\eta^2_p = .63$. The predicted three-way interaction between congruency, domain, and evaluator was not significant ($F < 1$), and we thus did not

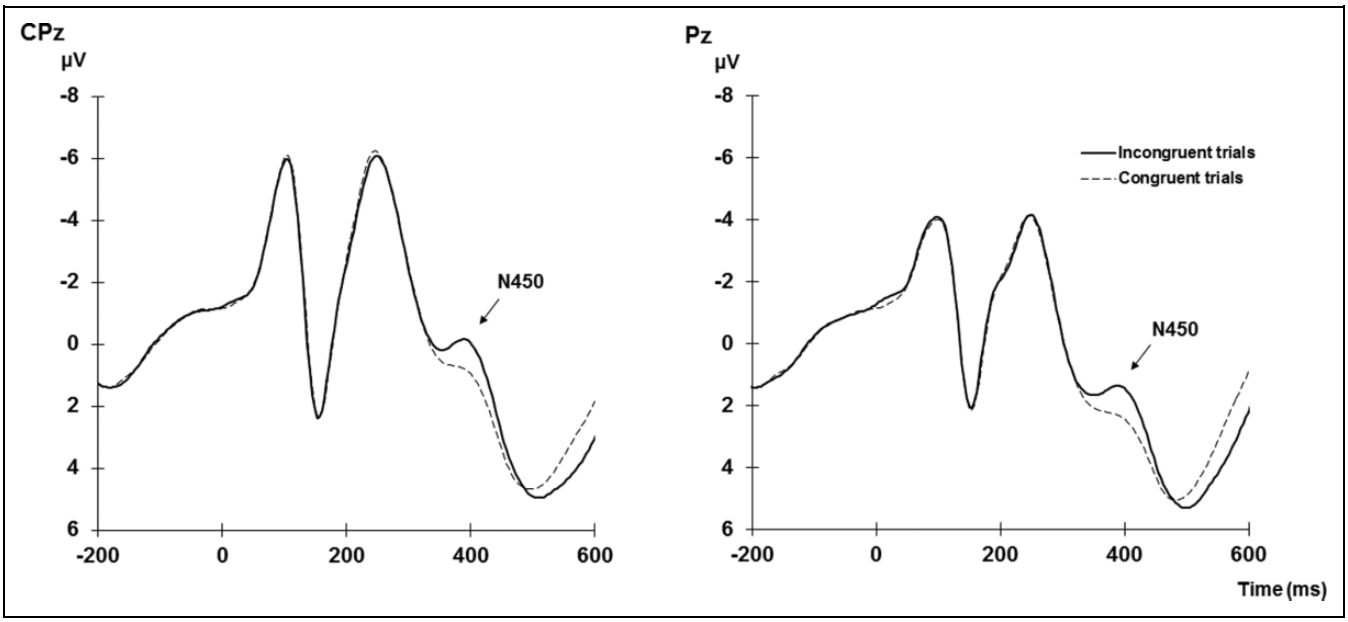

Figure 5. Differences in N450 amplitudes for incongruent and congruent trials.

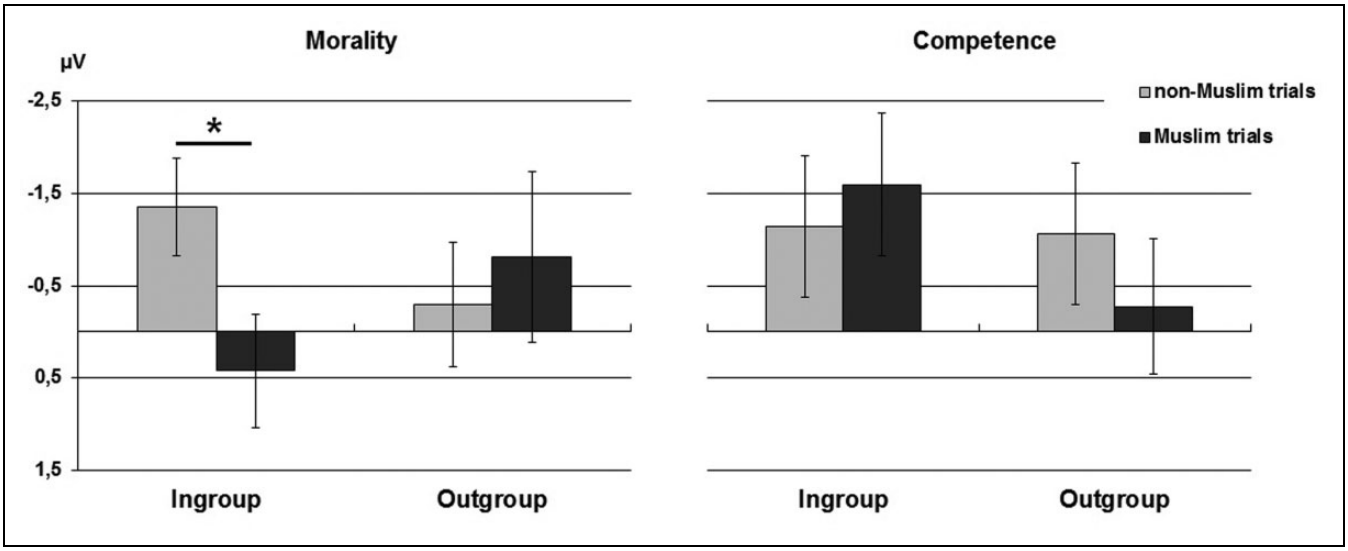

Figure 6. The mean differences in N450 amplitude between incongruent and congruent trials for each condition.

compute planned contrasts. Interestingly, however, the overall interaction effect between task domain and evaluator was marginally significant, $F(1, 44) = 3.59$, $p = .07$, $\eta^2_p = .08$, and significant at Cz, $F(1, 44) = 4.07$, $p = .05$, $\eta^2_p = .09$. Although the simple contrasts were not significant across the two electrode sites (Fs < 2.54, ps $> .12$), there is a reversal in the means pattern, indicating that the ERN modulation in the morality and competence conditions depends on the evaluator (Figure 7). Response monitoring (on incorrect and correct trials) under in-group evaluation was enhanced in the morality ($M = -2.70$ μV, $SE = 0.83$) compared to the competence condition ($M = -1.36$, $SE = 0.94$), but under out-group evaluation, it was enhanced in the competence ($M = -2.96$, $SE = 0.98$) compared to the morality condition ($M = -0.88$, $SE = 0.86$; Figure 8).

General Discussion

The current studies extend previous research on the motivation to comply with moral in-group norms (Ellemers et al., 2008; Pagliaro et al., 2011). We discovered that participants controlled implicit bias when the moral implications of an IAT were emphasized and when they were evaluated by an (minimal) in-group member. In Study 1, participants responded more slowly on congruent IAT trials, suggesting inhibition of prepotent reaction tendencies possibly revealing prejudice. These findings are consistent with previous research concerning people's *moral* motivation to inhibit prejudice (van Nunspeet et al., 2014). Importantly, however, such a motivation is different from the attempts to appear

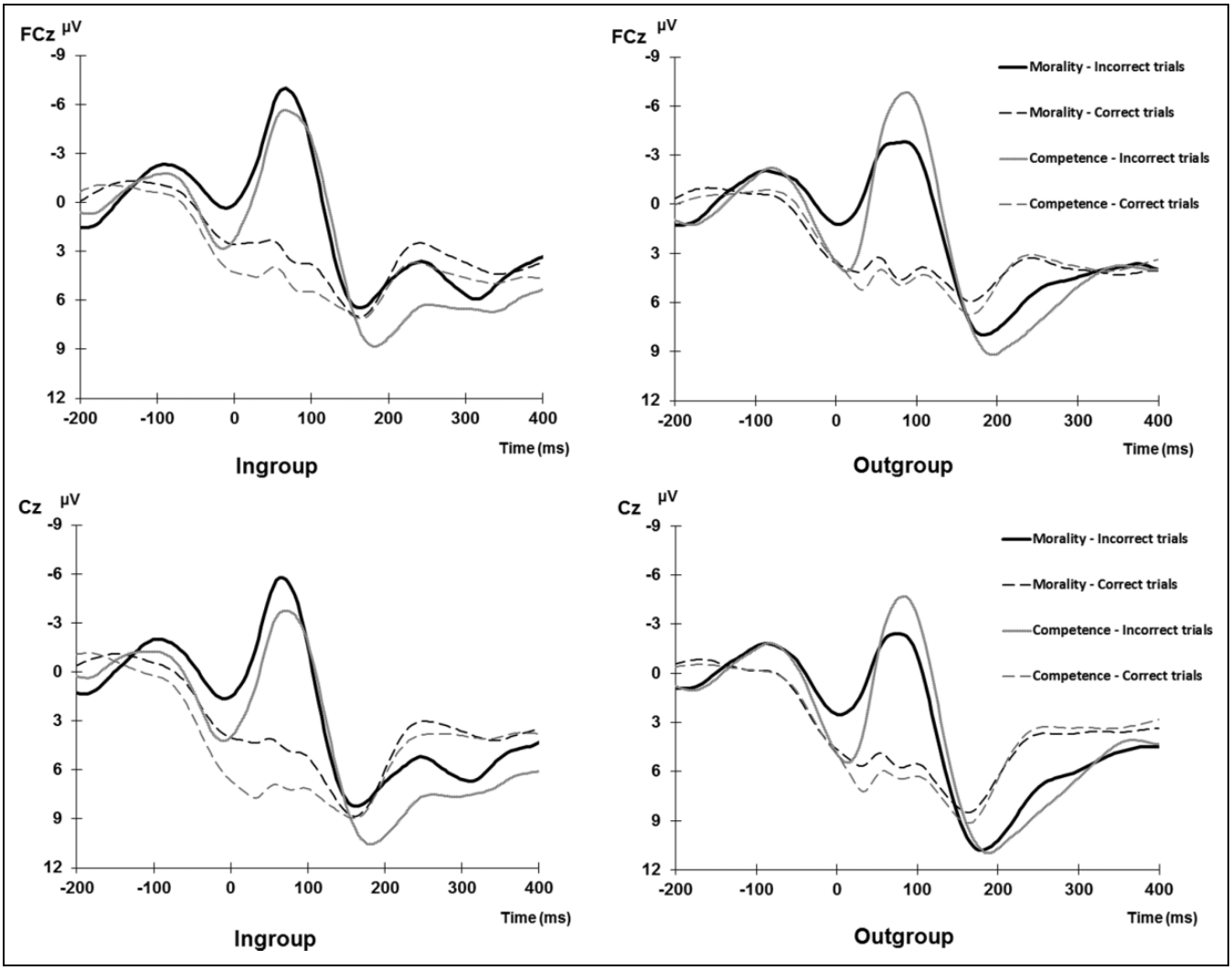

Figure 7. The error-related negativity in the morality and competence conditions in case of an in-group (left) or out-group evaluator (right).

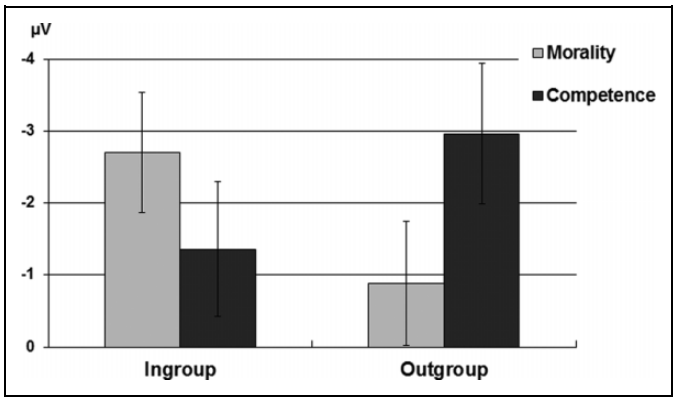

Figure 8. Mean error-related negativity (ERN) amplitudes (averaged over correct/incorrect trials, and electrode sites FCz/Cz) as a function of task domain and evaluator.

unprejudiced toward target group representatives (e.g., Lowery et al., 2001; Richeson & Ambady, 2003). We argue that an in-group categorization based on membership of the IAT target groups raises fundamentally different concerns than the ones we investigated in the current research. Specifically, the desire to avoid offending a representative of the target group (a Muslim woman or a Black individual) is quite different from the desire to demonstrate one's morality in front of self-relevant others (based on minimal group characteristics). Complementing prior work, we thus reveal that bias control can also be induced by the concern to appear moral in the eyes of others who are relevant to one's self-views.

Complementing previous research (van Nunspeet et al., 2014) and consistent with our hypothesis, ERP results in Study 2 clearly showed that the emphasis on morality and the evaluation by an in-group member was associated with enhanced perceptual attention and social categorization of the target women in the IAT (indicated by the N1). These results imply that emphasizing the moral implications of the IAT do not make people insensitive to social categorizations. Instead, participants were more focused on the identity of the different targets presented. Their perceptual attention increased, presumably to be able to adjust behavior, especially in the presence of in-group members. In other words, the

motivation to behave morally in front of self-relevant others helped participants to approach the task differently and to initiate processes that would help them to accomplish their goal.

We also found that conflict monitoring of stereotype-incongruent versus stereotype-congruent trials (indexed by the N450) differed in the morality/in-group condition when compared to the other conditions. This is in line with our reasoning that appearing moral is especially important in front of in-group members, as this particular social context affected conflict monitoring under moral task instructions. Unexpectedly, however, this effect was moderated by the IAT targets. Conflict monitoring of participants in the morality/in-group condition was clearly diminished for Muslim, as compared to non-Muslim, targets. Although we did not anticipate such an effect, this may indicate that participants in the morality/in-group condition adopt a different performance strategy compared to participants in the other conditions. Specifically, this observation seems to reveal that these participants are less likely to determine whether Muslim trials are stereotype congruent or stereotype incongruent, presumably because they are less inclined to rely on societal stereotypes.

The introduction of a social context (i.e., an evaluator) also affected the process of response monitoring, albeit somewhat differently and less strongly than expected (effects were only marginally significant). Although we had anticipated that error monitoring (indicated by the ERN) might be enhanced in the case of an emphasis on morality and evaluation by an in-group member, participants in this condition revealed enhanced monitoring of their incorrect *as well as* correct responses. Likewise, overall response monitoring was enhanced for participants who had been confronted with the test implications concerning their competence and who were evaluated by an out-group member.

Thus, results concerning different ERP components are consistent with our general prediction that task motivation and strategy are adapted when moral implications of the IAT are emphasized, and performance is evaluated by an in-group member. When interpreting our—predicted and unanticipated—results, it is important to keep in mind that we addressed different types of processes that occur at separate time points. Future research could more specifically examine whether the nature of specific adaptations that are made in different phases of the process actually indicates different types of motivation. For instance, being reminded of the implications of one's task performance in terms of moral in-group values might cause an intrinsic motivation to approach the task differently resulting in early attention to stimulus information in order to influence one's responses. On the other hand, when the desire to adapt one's behavior is driven more by an external motivation, for instance, when performance is evaluated by an out-group member or when the competence implications of the task are emphasized, this may only show up (relatively) later, during the processing of one's incorrect responses.

The current findings revealed that and how an emphasis on the moral implications of one's behavior, in combination with evaluation by an in-group member, reduces people's implicit prejudice and affects their perceptual attention and conflict—and response monitoring during task performance. These results extend previous research in several ways. First, they demonstrate the importance of morality for the self and social identity by revealing that people are especially motivated to adjust their moral task performance when monitored by a self-relevant other—even if this is indicated by a shared minimal group membership. Second, the examination of ERPs—with their high temporal precision—made it possible to reveal *how* people's motivation to appear moral affects unconscious cognitive processes: Evaluation by an in-group member not only increases people's moral task performance but also affects the way they *attend to* and *process information* that is relevant for behavioral adjustment. Furthermore, the current findings could be helpful for the development of interventions aimed at prejudice control that do not require the presence of target group representatives.

Acknowledgments

We thank Tamar van Herk, Laura de Reus, Ilona Domen, and Alma Vermeulen for their help with data collection; Marieke Visser and Wouter Steijn for their contribution to the experiment; and David Amodio for his advice concerning data analyses.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: 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 (KNAW) and the Netherlands Organization for Scientific Research (NWO); and a NWO-VENI grant [451-08-022] awarded to Belle Derks.

Notes

1. Consistent with previous research (van Nunspeet et al., 2014), this difference was related to increased response latencies on congruent trials in the morality compared to the competence condition; $M_{\text{morality}} = 494.85$, $SD = 20.10$; $M_{\text{competence}} = 480.65$, $SD = 16.61$; $F(1, 93) = 13.62$, $p < .001$, $\eta^2_p = .13$.
2. Some participants made less than 10 errors, explaining different degrees of freedom between the stimulus- and response-locked event-related brain potential analyses.
3. See the Supplementary Data (see Online Supplemental Material found at http://spps.sagepub.com/supplemental) for interaction effects with electrode site.
4. Simple main effects, following the higher order interaction, are reported in the Supplementary Data (see Online Supplemental Material found at http://spps.sagepub.com/supplemental).

Supplemental Material

The online supplementary material is available at http://spps.sagepub.com/supplemental.

References

Amodio, D. M. (2010). Coordinated roles of motivation and perception in the regulation of intergroup responses: Frontal

cortical asymmetry effects on the P2 event-related potential and behavior. *Journal of Cognitive Neuroscience*, *22*, 2609–2617.

Brambilla, M., Rusconi, P., Sacchi, S., & Cherubini, P. (2011). Looking for honesty: The primary role of morality (vs. sociability and competence) in information gathering. *European Journal of Social Psychology*, *41*, 135–143. doi:10.1002/ejsp.744

Ellemers, N., & Van den Bos, K. (2012). Morality in groups: On the social-regulatory functions of right and wrong. *Social and Personality Psychology Compass*, *6*, 878–889. doi:10.1111/spc3.12001

Ellemers, N., Pagliaro, S., Barreto, M., & Leach, C. W. (2008). Is it better to be moral than smart? The effects of morality and competence norms on the decision to work at group status improvement. *Journal of Personality and Social Psychology*, *95*, 1397–1410. doi:10.1037/a0012628

Gehring, W. J., Goss, B., Coles, M. G. H., Meyer, D. E., & Donchin, E. (1993). A neural system for error detection and compensation. *Psychological Science*, *4*, 385–390.

Gratton, G., Coles, M. G., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography & Clinical Neurophysiology*, *55*, 468–484.

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.

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. doi:10.1037/0022-3514.85.2.197

Hajcak, G., Moser, J. S., Yeung, N., & Simons, R. F. (2005). On the ERN and the significance of errors. *Psychophysiology*, *42*, 151–160. doi:10.1111/j.1469-8986.2005.00270.x

Ito, T. A., & Urland, G. R. (2003). Race and gender on the brain: Electrocortical measures of attention to the race and gender of multiply categorizable individuals. *Journal of Personality and Social Psychology*, *85*, 616–626. doi:10.1037/0022-3514.85.4.616

Jordan, A. H., & Monin, B. (2008). From sucker to saint: Moralization in response to self-threat. *Psychological Science*, *19*, 809–815. doi:10.1111/j.1467-9280.2008.02161.x

Kubota, J. T., & Ito, T. A. (2007). Multiple cues in social perception: The time course of processing race and facial expression. *Journal of Experimental Social Psychology*, *43*, 738–752. doi:10.1016/j.jesp.2006.10.023

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2005). International affective picture system (IAPS): Digitized photographs, instruction manual and affective ratings. *Technical Report A-6*. University of Florida, Gainesville, FL.

Leach, C. W., Ellemers, N., & Barreto, M. (2007). Group virtue: The importance of morality (vs. competence and sociability) in the positive evaluation of in-groups. *Journal of Personality and Social Psychology*, *93*, 234–249. doi:10.1037/0022-3514.93.2.234

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.

Nigam, A., Hoffman, J. E., & Simons, R. F. (1992). N400 to semantically anomalous pictures and words. *Journal of Cognitive Neuroscience*, *4*, 15–22.

Pagliaro, S., Ellemers, N., & Barreto, M. (2011). Sharing moral values: Anticipated ingroup respect as a determinant of adherence to morality-based (but not competence-based) group norms. *Personality and Social Psychology Bulletin*, *37*, 1117–1129. doi:10.1177/0146167211406906

Rebai, M., Bernard, C., & Lannou, J. (1997). The Stroop's test evokes a negative brain potential, the N400. *International Journal of Neuroscience*, *91*, 85–94.

Richeson, J. A., & Ambady, N. (2003). Effects of situational power on automatic racial prejudice. *Journal of Experimental Social Psychology*, *39*, 177–183. doi:10.1016/S0022-1031(02)00521-8

Tajfel, H. (1978). *Differentiation between social groups*. London, England: Academic Press.

van Nunspeet, F., Ellemers, N., & Derks, B. (2014). *Reducing implicit prejudice against Muslim women: The effects of moral concerns, intra- and intergroup motives*. Manuscript under editorial consideration.

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. doi:10.1093/scan/nss118

Williams, J. K., & Themanson, J. R. (2011). Neural correlates of the implicit association test: Evidence for semantic and emotional processing. *Social Cognitive and Affective Neuroscience*, *6*, 468–476.

Willis, J., & Todorov, A. (2006). First impressions: making up your mind after a 100-ms exposure to a face. *Psychological Science*, *17*, 592–598. doi:10.1111/j.1467-9280.2006.01750.x

Author Biographies

Félice van Nunspeet was a PhD student at the Social and Organizational Psychology Unit of Leiden University during the time of submission of the original manuscript. She completed her PhD degree and currently works as a postdoctoral researcher at the same department and institution, where she continues to examine people's (intrinsic) motivation to be moral using behavioral and neurobiological measures such as electroencephalogram and functional magnetic resonance imaging.

Belle Derks is an associate professor in Social and Organizational Psychology at Leiden University. Her research focuses on the psychological (motivation, performance) and physiological (cardiovascular activity, electroencephalogram) impact of stigma and group-based devaluation.

Naomi Ellemers is a full professor of the social psychology of organizations at Leiden University. Her work on group processes and intergroup relations includes experimental studies as well as applied research in organizations.

Sander Nieuwenhuis is a full professor in cognitive psychology at Leiden University. In his research, he uses electroencephalogram, functional magnetic resonance imaging, psychofarmacology, computational models, and various behavioral methods to examine how the brain regulates and monitors information processing to optimize behavioral performance.