



Universiteit
Leiden
The Netherlands

Neural correlates of the motivation to be moral

Nunspeet, F. van

Citation

Nunspeet, F. van. (2014, May 27). *Neural correlates of the motivation to be moral*. Kurt Lewin Institute Dissertation Series. Ridderprint B.V., Ridderkerk. Retrieved from <https://hdl.handle.net/1887/25829>

Version: Corrected Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/25829>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/25829> holds various files of this Leiden University dissertation.

Author: Nunspeet, Félice van

Title: Neural correlates of the motivation to be moral

Issue Date: 2014-05-27

Part II

**The importance of being
perceived as moral by others**

Chapter 4

Evaluation by an in- or outgroup member differentially affects moral task performance and the underlying cognitive mechanisms

This chapter is based on: Van Nunspeet, F., Derks, B., Ellemers, N., & Nieuwenhuis, S. Moral impression management: Evaluation by an ingroup member during a moral IAT enhances perceptual attention and conflict monitoring. *Manuscript under review.*

According to the Oxford dictionary being moral means “holding high principles for proper conduct”. But what is considered ‘proper’? Of course, individuals can have their own principles of what is good and bad. Nevertheless, the groups to which we belong (teams, organizations, or societies), 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: People are motivated to adjust their own behavior to moral (compared to competence) ingroup norms (Ellemers, Pagliaro, Barreto, & Leach, 2008), as a way to earn respect from fellow ingroup 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: When asked to form an impression about other individuals, people are more inclined to gather information concerning morality than concerning competence or sociability (Brambilla, Rusconi, Sacchi, & Cherubini, 2011). Even when an impression has to be made within milliseconds, trustworthiness judgments are made faster than judgments of sociability and competence (Willis & Todorov, 2006). Moreover, people monitor their own behavior to maintain a moral self-image (Jordan & Monin, 2008). 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. This is not only evident in self-report measures (Pagliaro et al., 2011). For example, Van Nunspeet, Ellemers, Derks and Nieuwenhuis (2014) have shown that people adapt their implicit behavior when this is perceived as indicative of their morality: During an Implicit Association Test (IAT) participants were more inclined to control their negative bias towards Muslim women when they thought

the test measured their morality than when they thought the test measured their competence.

The reasoning that the significance of morality derives from its implications for people's social identity, leads to the prediction that the motivation to be moral should be particularly relevant in an ingroup context. Thus, we hypothesize that when participants are evaluated by an ingroup, rather than an outgroup member, they are more motivated to control their bias during performance on an IAT indicating morality (compared to competence).

Event-Related Brain Potentials and Moral Performance

The desire to be moral may elicit socially desirable answers pertaining to morality. This 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.

In prior research, Van Nunspeet et al. (2014) revealed the cognitive processes that were associated with performance on a morally framed IAT. When test implications were presented in terms of morality compared to competence, participants' perceptual attention and response monitoring were enhanced during task performance. More specifically, event-related brain potentials (ERPs) suggested that participants paid more attention to the group membership of the photographed individuals presented in the IAT. This so-called social categorization was evident in modulations of the N1 and P150, two ERP components occurring around 100 and 200 ms after stimulus onset, that typically are larger when viewing ingroup vs. outgroup faces (e.g., Ito & Urland, 2003; Kubota & Ito, 2007). Van Nunspeet et al. (2014) argued that perceptual attention to the group membership of the women in the IAT was enhanced to enable participants to perform in line with their moral values.

Additionally, when morality instead of competence was emphasized in the IAT instruction, participants showed enhanced brain responses to the difference between incongruent and congruent trials and to errors. Specifically, the N450 and error-related negativity (ERN) modulations were larger when moral test implications were emphasized (Van Nunspeet et al., 2014). The N450, a negative

deflection around 400-500ms after stimulus-onset, is a component associated with conflict-monitoring, e.g. in language incongruencies (e.g., Nigam, Hoffman, & Simons, 1992), the Stroop task (e.g., Rebai, Bernard, & Lannou, 1997), and the IAT (Williams & Themanson, 2011). The ERN on the other hand, is a negative deflection within 100ms after a response is given. It is known to be larger for incorrect than correct responses (e.g., Gehring, Goss, Coles, Meyer, & Donchin, 1993; Nieuwenhuis, Blom, Band, & Kok, 2001), and for significant compared to non-significant errors (Hajcak, Moser, Yeung, & Simons, 2005). The findings of Van Nunspeet et al. (2014) thus suggest that the incongruity between the different IAT trials, as well as incorrect responses were perceived as more significant when the IAT was presented as a moral test. Additionally, their ERN results suggested that people are more concerned to show immoral than incompetent behavior.

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 (an ingroup rather than an outgroup member). 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 outgroup member (in the current research, a Muslim woman; Lowery, Hardin, & Sinclair, 2001; Richeson & Ambady, 2003). This is why we introduced minimal categories: Based on a questionnaire ostensibly assessing personality styles, participants were evaluated by a non-Muslim individual who was presented as someone with the same (ingroup) or another personality type (outgroup).

We thus predict that participants will show a weaker IAT bias when the moral (compared to competence) test implications are emphasized, especially when they are evaluated by an ingroup (vs. outgroup) member. Extending the research of Van Nunspeet et al. (2014), we anticipate that participants who are evaluated by an ingroup member and to whom the moral test implications are emphasized will show increased perceptual attention towards pictures of Muslim versus non-Muslim women (indexed by N1 and/or P150 modulations) and enhanced conflict- and response-monitoring as indicated by the N450 and/or ERN. We tested these

hypotheses in two studies; an initial behavioral study (Study 4.1) and a follow-up study in which we recorded an electroencephalogram (EEG) during IAT performance (Study 4.2).

Study 4.1

Method

Participants and design.

Ninety-five non-Muslim students (3 males, $M_{\text{age}} = 19.2$ years, $SD = 2.0$) participated for money or course credits. One participant was excluded from the analyses, because s/he 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) X 2 (evaluator: ingroup/outgroup 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, head phone 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'- or 'O'-type. After completing the questionnaire and a short pause, participants 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 either with a member of the same 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 him or herself and told that s/he 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 either read about the moral 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 fifty minutes.

Instruments.

The Implicit Association Test. 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 5 pictures of positive and 5 pictures of negative scenes, selected from the International Affective Picture System (IAPS; Lang et al., 2005).

For congruent trials, pictures of non-Muslim women shared the same response key as positive pictures and pictures of Muslim women the same response key as negative pictures. For incongruent trials, this was the case for non-Muslim women and negative pictures, and Muslim women and positive pictures. The order of the (in)congruent blocks was counterbalanced across participants. Training blocks 1, 2 and 4 consisted of 26 trials, 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 (see Figure 4.1). The feedback screen consisted of a movie clip (1250ms) of an evaluator showing either positive or negative feedback. To ensure that participants were aware that the evaluator was an in- or an outgroup member, two text displays indicated the group memberships of the participant and the evaluator. In case participants did not respond in time (i.e., within 680 ms), the feedback screen showed the words “too late”.

Morality vs. competence 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.

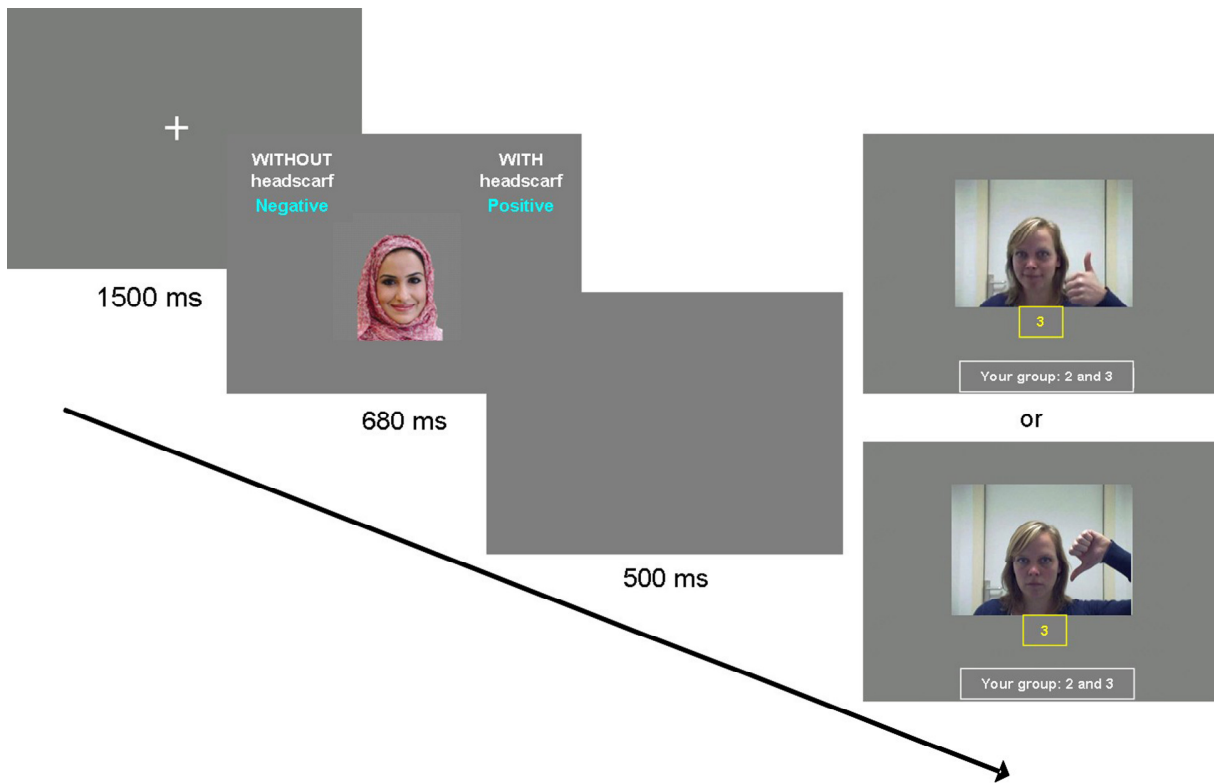


Figure 4.1. An IAT trial. The feedback screen was a movie display (1250ms) in which the confederate (here displayed as ingroup member) gave positive or negative feedback.

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 et al., 2003; see also Van Nunspeet et al., 2014).

Checks. To check that the perceived validity of the IAT did not differ between the 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 two 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 and Discussion

Checks.

As intended, participants in the four experimental conditions did not differ in their ability to identify with the experimentally created ingroup (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; $M = 3.64$, $SD = 1.62$; $F(3,90) < 1$. In line with prior findings, participants in the morality condition indicated positive impression management to be more important than participants in the competence condition; $M_{\text{morality}} = 4.83$, $SD = 1.01$; $M_{\text{competence}} = 4.28$, $SD = 1.04$; $F(1,90) = 6.58$, $p = .01$, $\eta^2 = .07$. There was no effect of evaluator nor an interaction effect; F 's < 1.49 , p 's $> .23$, indicating the importance of the moral task was enhanced, independently of whether participants were evaluated by an in- or an outgroup member.

IAT effect.

Overall, participants showed the standard IAT effect, indicating a negative implicit bias towards Muslim women; $t(93) = 6.83$, $p < .001$. More errors were made on incongruent than on congruent trials; respectively $M = 9.35$, $SD = 7.01$ and $M = 6.46$, $SD = 5.40$; $t(93) = 4.50$, $p < .001$; this was not affected by task domain or evaluator, all F 's < 1.87 , p 's $> .18$. Consistent with previous research (Van Nunspeet et al., 2014), an ANOVA with the D score as dependent variable and domain and evaluator as independent factors, revealed a significant main effect of domain; $F(1,90) = 5.57$, $p = .02$, $\eta^2 = 0.06$. Overall, participants in the morality condition showed a smaller IAT effect than participants in the competence condition, $M(\text{morality}) = 0.18$, $SD = 0.34$; $M(\text{competence}) = 0.36$, $SD = 0.39$ ⁴. Additionally, we found the predicted interaction effect between domain and evaluator; $F(1,90) = 4.26$, $p = .04$, $\eta^2 = 0.05$ (see Figure 4.2), indicating that participants who were evaluated by an ingroup member showed significantly less bias in the morality than in the competence condition; $M(\text{morality}) = 0.10$, $SD = 0.32$; $M(\text{competence}) = 0.43$, $SD = 0.33$; $F(1,90) = 9.82$, $p < .01$, $\eta^2 = .10$, while

⁴ 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; $M(\text{morality}) = 494.85$, $SD = 20.10$; $M(\text{competence}) = 480.65$, $SD = 16.61$; $F(1,93) = 13.62$, $p < .001$, $\eta^2 = .13$.

this was not the case when evaluated by an outgroup member; $M(\text{morality}) = 0.26$, $SD = 0.34$; $M(\text{competence}) = 0.29$, $SD = 0.44$; $F < 1$. 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 ingroup is relatively minimal). In Study 4.2 we examine what cognitive processes are associated with the tendency to conform to moral values in group contexts.

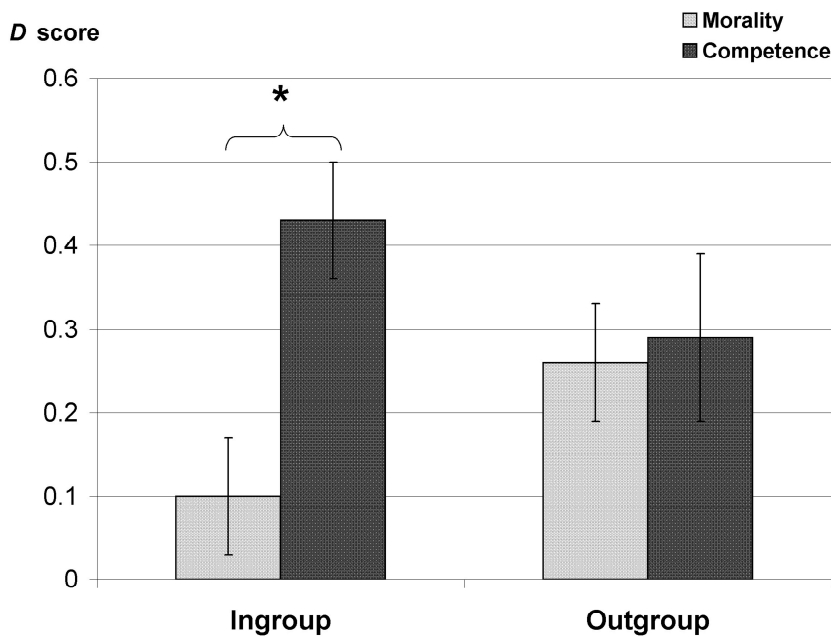


Figure 4.2. Condition means relevant to the interaction effect on the D scores.

Study 4.2

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) X 2 (evaluator: ingroup/outgroup member) between-participants design.

Procedure.

The procedure and measures were similar to Study 4.1, 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 680ms to 550ms, and the total number of test trials increased to 600 (300 congruent and 300 incongruent trials).

EEG acquisition.

The EEG was recorded from 19 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, offline 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 eye-blink 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 -300ms to 1000ms after the event. These epochs were subjected to artifact rejection, then averaged and baseline-corrected by subtracting the average signal value between 200-0ms pre-stimulus or between 300-50ms 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 midline 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 post-stimulus (N1: 90-110ms; P150: 100-250ms; N450: 325-500ms), whereas the ERN was quantified as the peak amplitude of the signal between -50 and 150ms around the response. Each average ERN was based on at least 10 trials⁵. Peak amplitude values of the N1, P150, and N450 were submitted to a 2 (electrode site: FCz/Cz or CPz/Pz [N450]) x 2 (target: Muslim/non-Muslim women) x 2 (congruency: congruent/incongruent) mixed-model ANOVA. Peak amplitude values of the ERN were submitted to a 2 (electrode site) x 2 (accuracy: correct/error) x 2 (congruency) mixed-model ANOVA. In every analysis, domain (morality/competence) and evaluator (ingroup/outgroup) were included as between-participants factors⁶.

Results and Discussion

Behavioral results.

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

IAT effect. Overall, participants showed the standard IAT effect, indicating a negative implicit bias towards Muslim women; $t(63) = 5.46$, $p < .001$. IAT effects were not systematically affected by evaluator or task domain; $F^2s < 1$, $p^2s > .1$, indicating that the emphasis on morality or competence and the group membership of the evaluator was not visible in task performance. This likely is due to the changes we made to optimize the task for ERP recordings: To ensure enough

⁵ Some participants made less than 10 errors, explaining different degrees of freedom between the stimulus- and response-locked ERP analyses.

⁶ Electrode site was not of interest for the current research, see Appendix B for significant interaction effects with this factor.

errors to reliably estimate the ERN, the maximum response time was reduced. In Study 4.1 –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 ingroup/outgroup evaluators using a response window of 680 ms (as in Study 4.1 and prior research), the IAT bias was significantly lower in the morality compared to the competence condition, when participants were evaluated by a minimal ingroup member (Van Nunspeet, Ellemers, & Derks, *manuscript under review*).

Nonetheless, the identity of the evaluator did affect behavioral responses in the current data. Besides the fact that more errors were made on incongruent ($M = 34.4$, $SD = 18.4$) than congruent trials ($M = 25.6$, $SD = 15.5$); $t(63) = 4.87$, $p < .001$, participants in the ingroup evaluator condition made fewer errors ($M = 50.3$, $SD = 24.9$) than participants in the outgroup evaluator condition ($M = 70.4$, $SD = 33.4$); $F(1,60) = 7.28$, $p = .01$, $\eta^2 = .11$. This is consistent with our reasoning that participants are generally more motivated to perform well when evaluated by an ingroup member.

ERP results.

Perceptual attention.

N1. The N1 results revealed the expected evidence of categorization: The N1 was larger for pictures of Muslim women ($M = -7.18 \mu\text{V}$, $SE = 0.35$) than non-Muslim women ($M = -6.91 \mu\text{V}$, $SE = 0.35$); $F(1,56) = 3.52$, $p = .07$, $\eta^2 = .06$ (see Figure 4.3). The predicted interaction between target, domain and evaluator was significant; $F(1,56) = 4.36$, $p = .04$, $\eta^2 = .07$. Separate analyses for ingroup vs. outgroup evaluators revealed a marginally significant interaction between target and task domain in case of an ingroup evaluator; $F(1,29) = 3.53$, $p = .07$, $\eta^2 = .11$, but not in case of an outgroup evaluator; $F(1,27) = 1.02$, $p = .32$. Separate analyses per task domain revealed a significant target by evaluator interaction in the moral domain; $F(1,31) = 6.69$, $p = .02$, $\eta^2 = .18$, but not in the competence domain; $F < 1$. As a result, categorization was significantly enhanced in the morality/ingroup

condition ($F[1,56] = 11.35, p = .001, \eta^2 = .17$), but not in the other conditions ($F^s < 1$; see Figure 4.4).

P150. Analyses of the P150 only revealed the expected main effect of target: The P150 was larger for pictures of Muslim women ($M = 5.44 \mu\text{V}, SE = 0.48$) than non-Muslim women ($M = 3.77 \mu\text{V}, SE = 0.43$); $F(1,56) = 93.13, p < .001, \eta^2 = .62$ (see Figure 4.3).

This suggests that enhanced social categorization of (non-)Muslim women in case of moral task performance under ingroup evaluation, only occurs in initial stages of perceptual attention (N1).

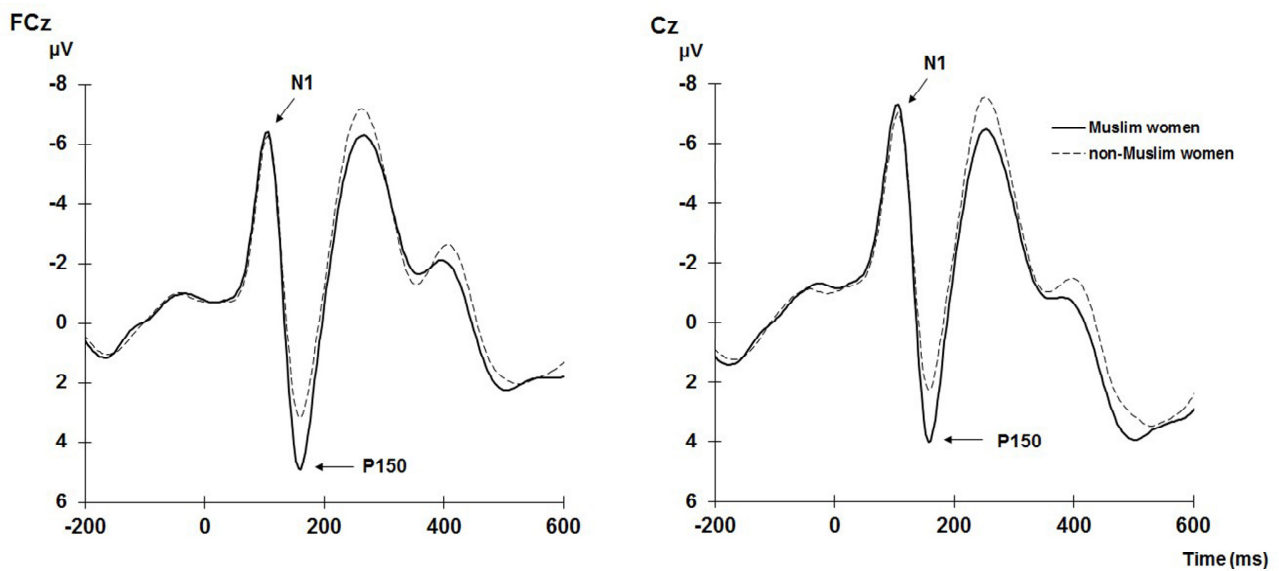


Figure 4.3. Differences in N1 and P150 amplitudes for pictures of Muslim and non-Muslim women. Only the N1 modulation interacted with task domain and evaluator.

Conflict- and response-monitoring.

N450. Results showed the anticipated effect of congruency: The N450 was larger for incongruent ($M = -0.13 \mu\text{V}, SE = 0.33$) compared to congruent ($M = 0.64 \mu\text{V}, SE = 0.40$) trials; $F(1,56) = 5.92, p = .02, \eta^2 = .10$ (see Figure 4.5).

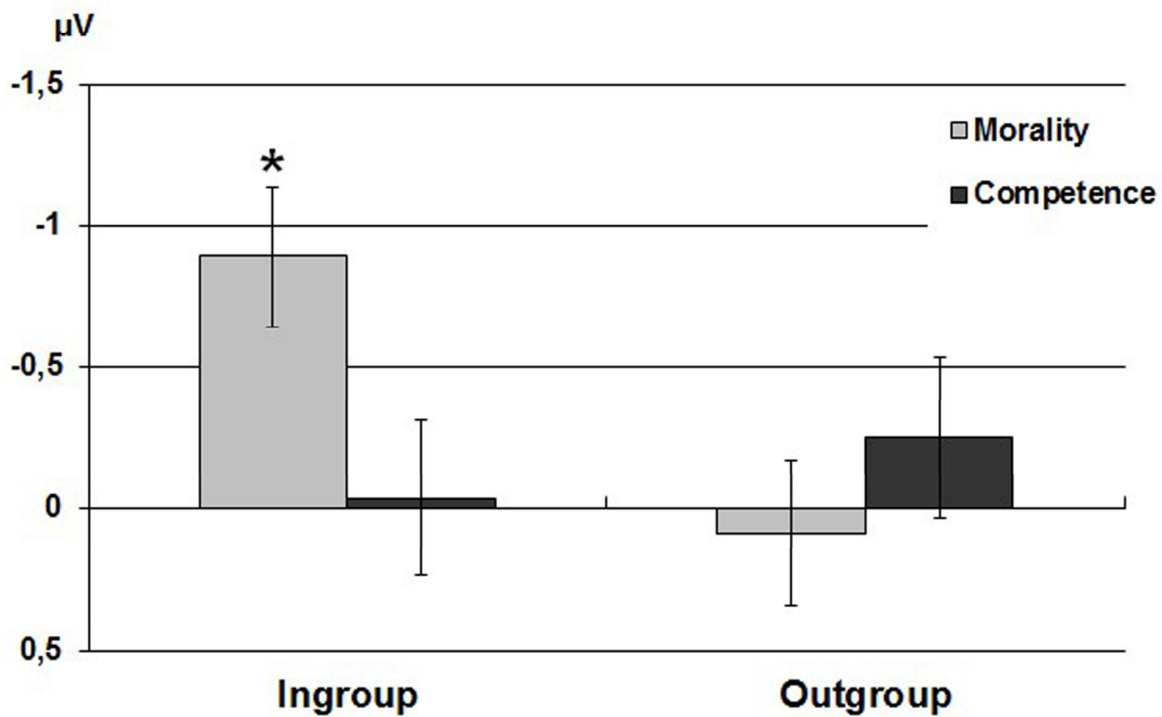


Figure 4.4. The mean differences in N1 amplitude between Muslim vs. non-Muslim targets for each condition.

There was also a main effect of target: The N450 was larger for non-Muslim ($M = -0.28 \mu\text{V}$, $SE = 0.37$) compared to Muslim women ($M = 0.79 \mu\text{V}$, $SE = 0.33$); $F(1,56) = 24.06$, $p < .001$, $\eta^2 = .30$. Importantly, both main effects were qualified by a significant four-way interaction between congruency, target, domain and evaluator; $F(1,56) = 5.75$, $p = .02$, $\eta^2 = .09$. Separate analyses for the task domain conditions revealed a significant interaction between congruency, target and evaluator in the morality condition; $F(1,31) = 5.36$, $p < .03$, $\eta^2 = .15$, but not in the competence condition; $F(1,25) = 1.30$, $p = .27$.

Furthermore, in the morality condition, there was an interaction between congruency and target in the ingroup evaluator condition; $F(1,16) = 10.26$, $p = .006$, $\eta^2 = .39$, but not in the outgroup evaluator condition; $F < 1$. The N450 modulation on incongruent compared to congruent trials in the morality/ingroup condition was significant when viewing pictures of non-Muslim women; $F(1,16) =$

6.45, $p = .02$, $\eta^2 = .29$, and not when viewing Muslim women; $F < 1$ ^{7,8} (see Figure 4.6). These results suggests that conflict-monitoring was enhanced (on non-Muslim trials) when moral test implications were stressed and participants were evaluated by an ingroup member.

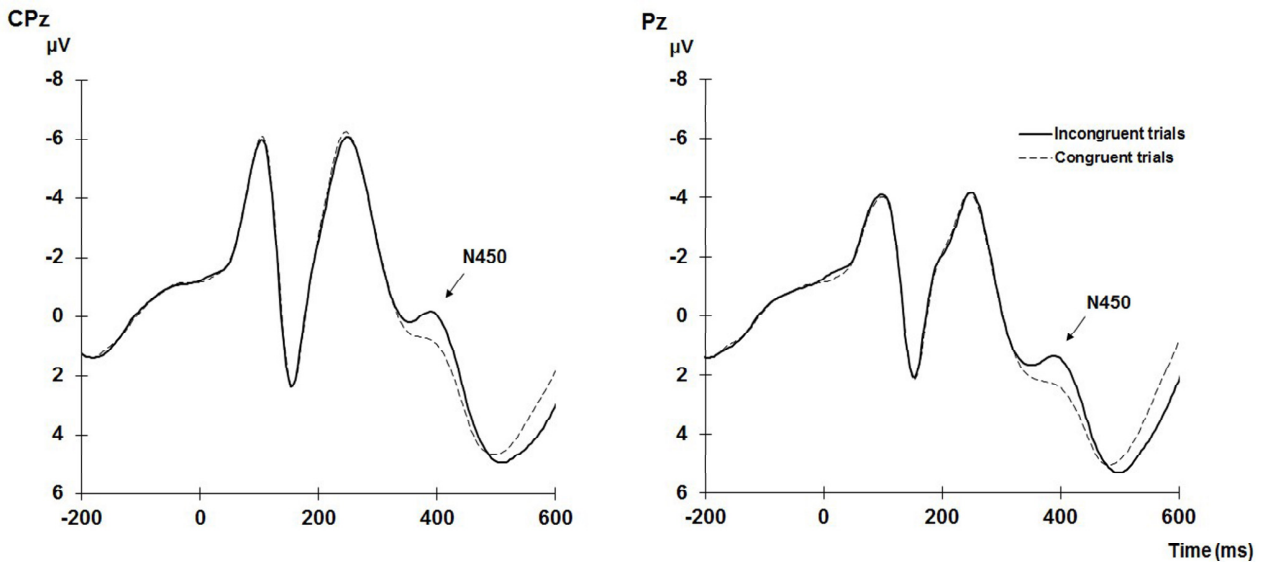


Figure 4.5. Differences in N450 amplitudes for incongruent and congruent trials.

ERN. As anticipated, results showed that the ERN was larger for incorrect ($M = -6.90 \mu\text{V}$, $SE = 0.69$) than correct trials ($M = 2.95 \mu\text{V}$, $SE = 0.46$); $F(1,44) = 173.52$, $p < .001$, $\eta^2 = .80$. There was a marginally significant interaction effect between accuracy and task domain; $F(1,44) = 3.37$, $p = .07$, $\eta^2 = .07$, indicating that the ERN modulation was somewhat larger in the competence ($M_{\text{difference}} = -1.22 \mu\text{V}$, $SE = 1.12$; $F[1,44] = 100.07$, $p < .001$, $\eta^2 = .70$) than the morality condition ($M_{\text{difference}} = -8.48 \mu\text{V}$, $SE = 0.99$; $F[1,44] = 73.49$, $p < .001$, $\eta^2 = .63$). More importantly however, the ERN modulation in the morality and competence

⁷ Findings of N450 modulations for targets can be found in Appendix B.

⁸ This may reflect the specific nature of our paradigm, in which these trials confronted participants with pictures of a non-Muslim target, while receiving feedback from a non-Muslim evaluator, arguably increasing the need for conflict-monitoring.

conditions differed depending on evaluator type: There was a marginally significant between-subjects interaction effect of task domain and evaluator; $F(1,44) = 3.59, p = .07, \eta^2 = .08^9$. Even though the simple contrasts were not significant (F 's $< 2.32, p$'s $> .14$), the means pattern indicates a reversal of the effect. Response monitoring was enhanced under ingroup evaluation in the morality ($M = -2.67 \mu\text{V}, SE = 0.83$) compared to the competence condition ($M = -1.36 \mu\text{V}, SE = 0.94$), but enhanced in the competence ($M = -2.96 \mu\text{V}, SE = 0.98$) compared to the morality condition ($M = -0.88 \mu\text{V}, SE = 0.86$) under outgroup evaluation (see Figure 4.7).

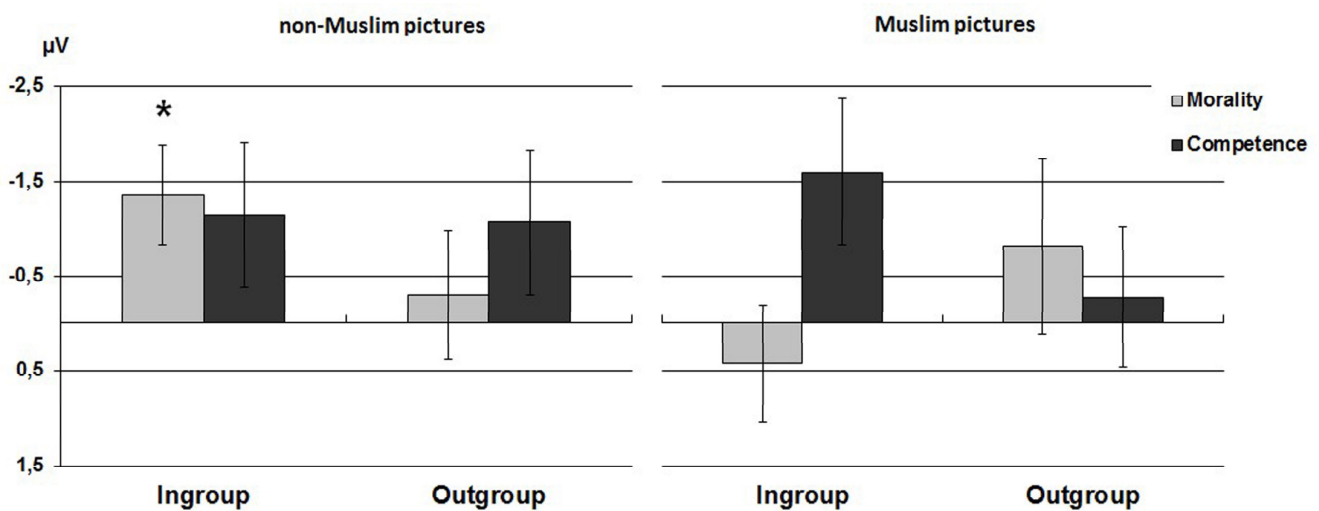


Figure 4.6. The mean differences in N450 amplitude between incongruent vs. congruent trials for each condition.

⁹ To clarify this marginally significant effect, we conducted separate analyses for FCz and Cz. Results showed that both interaction effects were only significant at Cz (accuracy*domain: $F[1,44] = 3.98, p = .05, \eta^2 = .08$; domain*evaluator: $F[1,44] = 4.07, p = .05, \eta^2 = .09$).

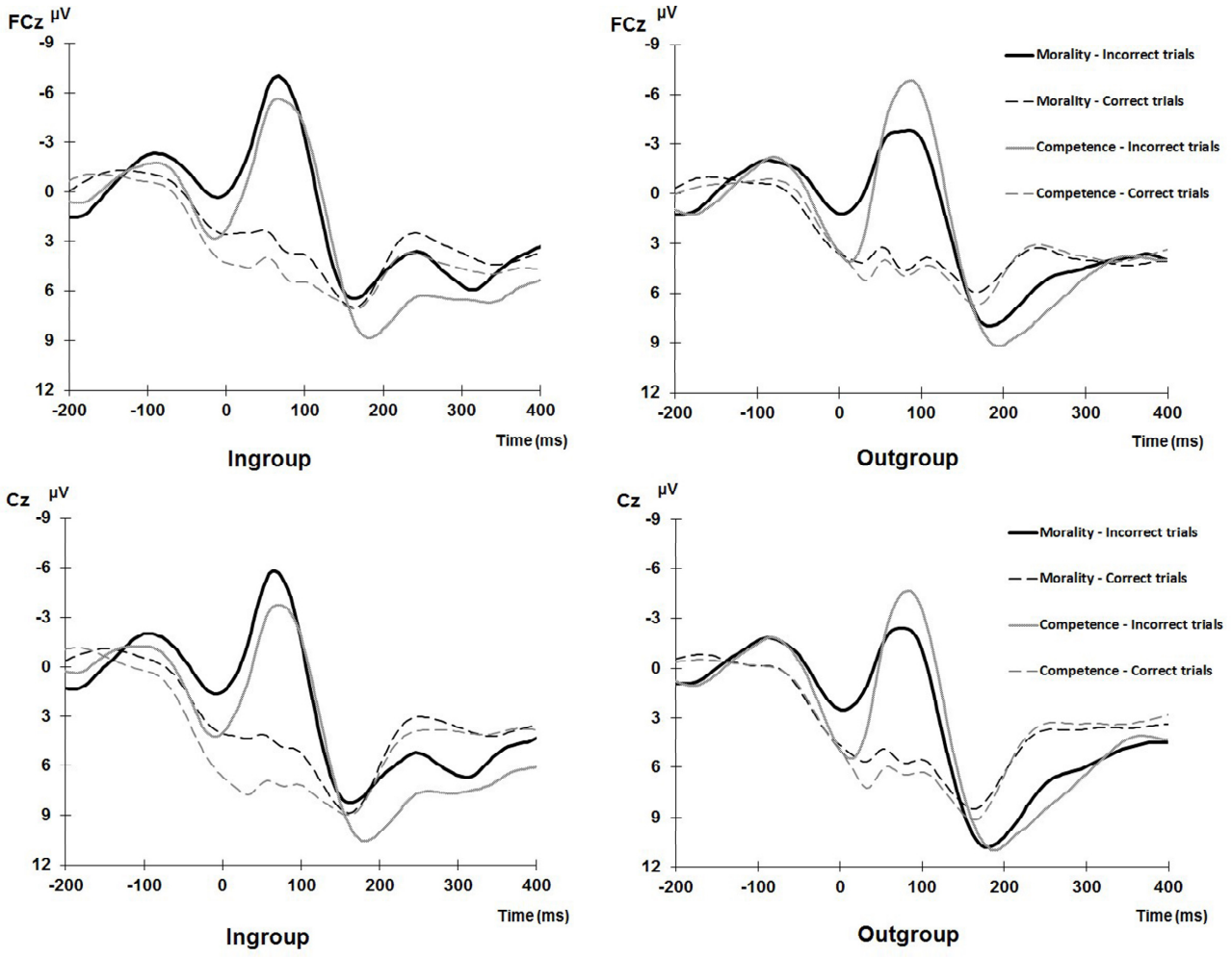


Figure 4.7. The error-related negativity in the morality and competence conditions in case of an ingroup (left) or outgroup evaluator (right).

General Discussion

The current studies extend previous research on the motivation to comply with moral ingroup 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 a (self-relevant) ingroup member. In Study 4.1, participants responded more slowly on congruent IAT trials, suggesting inhibition of prepotent reaction tendencies possibly revealing prejudice. Complementing prior research (Van Nunspeet et al., 2014), ERP results in Study 4.2 revealed that this was associated with enhanced perceptual attention and social categorization of the target women in the IAT (as indicated by the N1). Participants were thus more focused on the identity of the different targets presented, which is needed to control biased responses. Thus, emphasizing the moral implications of the IAT does not make people insensitive to social categorizations. Instead, it triggers increased perceptual attention in order to adjust behavior. Indeed, previous ERP research has revealed that similar early attentional processes can be moderated by motivational states (e.g., Amodio, 2010; Cunningham, Van Bavel, Arbuckle, Packer, & Waggoner, 2012). Thus, our findings help understand how people control their prejudice towards Muslim women, to show they are moral in front of self-relevant others. Note however that this is different from the attempts to appear unprejudiced towards target group representatives, as revealed by Lowery et al. (2001) and Richeson and Ambady (2003). Complementing this prior work, we reveal that bias control can also be affected by the importance of sharing moral norms with one's ingroup.

Conflict- and response monitoring (indicated by the N450 and ERN) were also affected by the moral or competence implications of the IAT and the ingroup vs. outgroup evaluator. That is, the detection of incongruent compared to congruent trials (N450 modulation) was enhanced when participants in the morality/ingroup condition viewed non-Muslim women. Moreover, whereas response monitoring (ERN on correct and incorrect trials) seemed to be enhanced in the morality compared to the competence condition when the evaluator was an ingroup member, this pattern was reversed when the evaluator was an outgroup member. Since ERN amplitudes have been found to be related to estimates of

control (Amodio, et al., 2004), this suggests increased motivation to control bias towards Muslim women in the moral ingroup condition. We did not anticipate participants to be particularly sensitive to competence task instructions when evaluated by an outgroup member. However, a similar (non-significant) reversal of the importance of competence vs. morality depending on the group membership of the evaluator was observed in the behavioral results of Study 4.1 and the N1 results in Study 4.2: Whereas behavioral bias on the moral IAT was reduced in the ingroup evaluation condition, bias on the competence IAT was diminished in the outgroup evaluation condition. Likewise, the N1 modulation was greater in the morality/ingroup than in the morality/outgroup condition, while it was somewhat larger in the competence/outgroup than the competence/ingroup condition. Although (probably due to limited statistical power) these effects did not reach significance, they could suggest that whereas moral impression management is more important in the ingroup, displaying competence is more relevant towards the outgroup. Future research could further examine this. The current findings demonstrate the importance of morality for self and social identity, by revealing that people are especially motivated to adjust their moral task performance when monitored by a self-relevant group; this is associated with increased perceptual attention and conflict monitoring.

Acknowledgements

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.