

Social identity modifies face perception: an ERP study of social categorization

Belle Derks,^{1,2} Jeffrey Stedeouder,³ and Tiffany A. Ito⁴

¹Department of Social and Organizational Psychology, Leiden University, and ²Leiden Institute for Brain and Cognition, Leiden, The Netherlands,

³Department of Psychiatry, Erasmus Medical Center, Rotterdam, The Netherlands, and ⁴Department of Psychology and Neuroscience, University of Colorado, Boulder, CO, USA

Two studies examined whether social identity processes, i.e. group identification and social identity threat, amplify the degree to which people attend to social category information in early perception [assessed with event-related brain potentials (ERPs)]. Participants were presented with faces of Muslims and non-Muslims in an evaluative priming task while ERPs were measured and implicit evaluative bias was assessed. Study 1 revealed that non-Muslims showed stronger differentiation between ingroup and outgroup faces in both early (N200) and later processing stages (implicit evaluations) when they identified more strongly with their ethnic group. Moreover, identification effects on implicit bias were mediated by intergroup differentiation in the N200. In Study 2, social identity threat (vs control) was manipulated among Muslims. Results revealed that high social identity threat resulted in stronger differentiation of Muslims from non-Muslims in early (N200) and late (implicit evaluations) processing stages, with N200 effects again predicting implicit bias. Combined, these studies reveal how seemingly bottom-up early social categorization processes are affected by individual and contextual variables that affect the meaning of social identity. Implications of these results for the social identity perspective as well as social cognitive theories of person perception are discussed.

Keywords: social categorization; group identification; social identity threat; ERP; N200

One way in which people make sense of their complex social world is through social categorization, perceiving themselves and the people they meet as members of social categories (e.g. men vs women; Europeans vs Americans). Social categorization induces people to think of themselves as group members (Tajfel and Turner, 1986; Turner *et al.*, 1987) and is the starting point for a wide range of positive and negative intergroup phenomena (e.g. social identity, intragroup helping, prejudice, stereotype threat). Rather than focusing on social categorization itself, most research to date has tested how social categorization affects downstream processes like person perception and intergroup bias. In practice, it can be difficult to separately assess processes related to identifying an individual as belonging to a particular group independent of the effects of that categorization on downstream processes like evaluation and stereotype activation. In the current research, we use event-related brain potentials (ERPs) to zoom in on the attentional processes that occur during initial stages of categorization and examine it separately from later evaluative processing. We test whether two different social identity motives that have been found to amplify more downstream intergroup responses (i.e. high group identification and threats to group value) amplify attention to category membership already in the early stages of person perception.

TWO PERSPECTIVES ON SOCIAL CATEGORIZATION

Social psychological research on social categorization has been done in two separate fields: the social identity perspective and the social cognition tradition. The former sees social categorization as the foundation on which social identity is based (Tajfel and Turner, 1986; Turner *et al.*, 1987). Accordingly, people base their identity partly on the groups to which they belong and are therefore motivated to evaluate

groups with which they self-categorize as positively distinctive. The first step in achieving a positively distinctive social identity is social categorization, establishing who is part of the ingroup and who is not. Furthermore, social identity theory highlights a number of factors that increase people's tendency to distinguish between 'us and them'. For example, individuals will only act on a social categorization to the degree that they psychologically identify with it (Tajfel and Turner, 1986; Ellemers *et al.*, 2002). Moreover, establishing clear ingroup/outgroup distinctions becomes more important to people when their social identity is threatened, for instance, when their group has low status or when it lacks distinctiveness from other groups (Branscombe *et al.*, 1999; Jetten *et al.*, 1999). Despite social categorization's central role in the social identity perspective, most social identity research has examined affective, evaluative and behavioral consequences of social categorization (e.g. Tajfel *et al.*, 1971; Haslam *et al.*, 1995; Yzerbyt *et al.*, 2003). This is achieved by manipulating (minimal) social categorizations as an independent variable, rather than measuring the categorization process itself.

By contrast, social cognition research has examined social categorization more directly as part of the person perception process (Brewer, 1988; Fiske and Neuberg, 1990). Here the focus has been on examining the automaticity vs conditionality with which people pay attention to social category information when they encounter others, and the effects of social categorization on impression formation through the activation of associated stereotypes (Macrae and Bodenhausen, 2000; Hugenberg *et al.*, 2010; Quinn and Rosenthal, 2012). This research has conceptualized social categorization as an efficient information processing strategy that occurs unintentionally and outside of our awareness. However, apart from research on target characteristics, contextual factors and interventions that increase or diminish the use of social categories in person perception (e.g. Fiske and Neuberg, 1990; Wheeler and Fiske, 2005; Macrae and Cloutier, 2009; Quinn and Rosenthal, 2012), the social cognition perspective is relatively silent about possible motivational differences between people to differentially attend to ingroup and outgroup faces once they are in categorization mode.

Previous work on social categorization typically relied on downstream processes to make inferences about whether social

Received 16 December 2013; Revised 23 June 2014; Accepted 13 August 2014

The authors thank Peter de Heus for statistical advice, Henk van Steenberg for advice on EEG-data processing, Thijs Schrama for help with programming the experiment and Nassira Akkabi, Alexandra Hajema and Floor van de Water for assistance with data collection. This work was supported by a Veni grant (451-08-022) from the Netherlands Organization for Scientific Research awarded (to B.D.).

Correspondence should be addressed to Belle Derks, Social and Organizational Psychology Unit, Leiden University - Institute of Psychology, P.O. Box 9555, 2300 RB Leiden, The Netherlands. E-mail: derks@fsw.leidenuniv.nl.

categorization has taken place. For example, relatively poorer memory for an outgroup face, or for which outgroup member said what is taken as evidence for social categorization (Taylor *et al.*, 1978; Hugenberg *et al.*, 2010). Often prompted rather than spontaneous categorization is measured by asking participants to explicitly categorize targets and measure decisions and response latencies (Dotsch *et al.*, 2011; Ma *et al.*, 2011). In the current studies, our goal was to measure the degree to which participants attend to social category information directly, without explicitly prompting them or without relying on its downstream indicators. Therefore, we measured ERPs indicating differential attention to ingroup and outgroup faces. Research using ERPs has discovered that people distinguish between ingroup and outgroup faces within the first 300 ms of social perception (Ito and Bartholow, 2009). Furthermore, this online indicator of social categorization has been shown to predict later processes like implicit intergroup bias (Correll *et al.*, 2006; Dickter and Bartholow, 2007).

SOCIAL IDENTITY MOTIVES AND IMPLICIT SOCIAL CATEGORIZATION

In the current investigation, we combined predictions from the social identity perspective with social neuroscience measures associated with social categorization (ERPs) to establish whether social identity motives that have been found to affect intergroup differentiation on downstream processes (e.g. evaluative ingroup bias) also affect the degree to which people implicitly attend to category information in the first stages of person perception. In Study 1, we hypothesized that individuals would more strongly differentiate between ingroup and outgroup targets, as reflected in earlier (ERPs) and later processing (implicit intergroup bias), to the degree that they feel connected and committed to their group (Ellemers *et al.*, 2002). It has been shown that high ingroup identification increases the degree to which individuals distinguish between ingroup and outgroup in their implicit and explicit attitudes and behaviors (Turner *et al.*, 1987; Doosje *et al.*, 1999; Lalonde, 2002; Ashburn Nardo *et al.*, 2003; Jetten *et al.*, 2004). Moreover, in an explicit categorization task, high identifiers have been shown to classify less targets as ingroup members (Castano *et al.*, 2002). Additionally, the Categorization-Identification model (Hugenberg *et al.*, 2010) predicts that identification affects attention to ingroup and outgroup faces, by increasing motivation to individualize ingroup relative to outgroup faces. However, no research up to date has directly assessed the differential attention and increased differentiation among high identifiers in the first stages of the perceptual process.

Secondly, in Study 2 we hypothesized that early attention to social categories is enhanced by threats to the value of one's social identity. Again, research in the social identity tradition has focused on how social identity threats impact on downstream affective, behavioral and coping responses (Crocker and Major, 1989; Ellemers *et al.*, 2002). Moreover, although previous work has found that experienced threat increases social categorization as measured with downstream indicators or in promoted categorization tasks (Miller *et al.*, 2010; Maner *et al.*, 2012), whether social identity threat affects spontaneous attention to social categories in early social perception is unclear. In the current study we tested whether threats to social identity actually change people's perception of their social world, inducing group members to differentiate between ingroup and outgroup faces more strongly.

MEASURING SOCIAL CATEGORIZATION WITH ERPS

Apart from allowing us to measure indicators of early social categorization, examining our hypotheses with ERPs also allows us to examine the social categorization process in more temporal detail. ERP research

on the perception of ingroup and outgroup faces (see Ito and Bartholow, 2009 for a review) has found evidence of early social categorization processes particularly in two components: the P200 (a positive parietal component peaking between 150 and 250 ms) and the N200 (a negative frontocentral component peaking between 200 and 400 ms). These peaks have been associated with selective attention, with higher amplitudes indicating greater attention (Luck, 2005). Previous work has found that when people process faces that differ according to a social category (e.g. race, ethnicity), they typically first devote more attention to outgroup faces before turning their attention to ingroup faces: the early P200-component typically shows higher amplitudes to outgroup targets than ingroup targets, whereas the later N200 component is commonly larger for ingroup targets than for outgroup targets (Dickter and Bartholow, 2007; Ito, 2013; Ito and Bartholow, 2009). In previous work, the P200 has been associated with vigilance for threatening as well as novel cues (e.g. negative facial expressions; Bar Haim *et al.*, 2005). In the context of social categorization, this has been interpreted as reflecting vigilance for the outgroup. The N200, which peaks slightly later, has been interpreted as deeper attention for unique individuating information (Lucas *et al.*, 2013). Previous work has found larger N200s to familiar faces and ingroup faces (Ito and Bartholow, 2009).

Importantly, differential amplitudes to faces of different social groups in P200 and N200 components depend on the perceiver's group membership: whereas Whites show greater P200-amplitudes to Black rather than White faces and larger N200s to White rather than Black faces, Black participants show the exact opposite pattern (Dickter and Bartholow, 2007). Interestingly, for Asian Americans, who present a smaller ethnic group, N200-amplitudes have been found to differ according to context (Willadsen-Jensen and Ito, 2008). In a context of mostly Asian faces, Asians display the common pattern of larger N200s for ingroup (Asian) faces than outgroup (White) faces. However, when viewing faces in a context of mostly White faces, N200s of Asian participants were larger to outgroup (White) faces than ingroup faces. This suggests that Asian Americans attend predominantly to faces from the group that is contextually emphasized and can flexibly adopt the perspective of the group that is in the numerical majority. Importantly, this context effect was not found among White Americans.

In the current studies, we examine effects of social identity on both ERPs to explore whether social identification and social identity threat enhance differentiation in the P200 (suggesting the increased detection of outgroup members as a threat), the N200 (suggesting deeper processing of faces belonging to one group over the other), or both. Alternatively, social identity motives could have their affect only after initial categorization processes have taken place (e.g. when an evaluative response is required). In this case, neither the P200 nor N200 may be affected by social identity motives. Finally, we examine how these initial categorization processes relate to downstream evaluative processes of the ingroup *vs* the outgroup, in this case by measuring ERPs in the context of an evaluative priming task that assesses implicit intergroup evaluation (Fazio *et al.*, 1995). We specifically chose this task over others (e.g. the Implicit Association Test (IAT)) because it requires no explicit responses to ingroup and outgroup targets. Instead, participants categorize positive and negative words that follow the ingroup/outgroup primes (see below for more details). This allows us to assess spontaneous rather than prompted categorization of ingroup and outgroup targets.

OVERVIEW

We performed two studies to test for relations between social identity processes and differentiation between social categories at early (ERPs)

and later (implicit evaluation) stages of information processing. In the first (correlational) study we measured ethnic identification among non-Muslim participants and recorded electroencephalogram (EEG) while they were presented with primes of ingroup (non-Muslim) and outgroup members (Muslims) in an implicit bias task. Study 2 was designed to provide experimental evidence for the link between social identity motives and social categorization in ERPs, and to extend our argument to a research population different from White majority participants. Therefore, we manipulated the presence (*vs* absence) of social identity threat among an ethnic minority group at our university who we expected to have experience with social identity threat, namely Muslim students.

STUDY 1

Methods

Participants

Twenty-one right-handed Dutch non-Muslim students (15 females, 4 males, 2 unspecified; $M_{\text{age}} = 21.48$) participated in return for a fee or partial fulfillment of course requirements. The study was conducted with approval of the Psychology Ethics Committee.

Stimuli

Sixteen grayscale pictures of Dutch-looking women without a headscarf (8) and non-Dutch Muslim women wearing a headscarf (8) were used as stimuli. Pictures were equated for attractiveness as determined by a pilot test ($N = 48$) and edited to have a uniform gray background. We chose to focus on Muslims, as they are a highly salient ethnic/religious outgroup for non-Muslim Dutch, as their societal integration currently dominates the public debate (Phalet and Ter Wal, 2004).

Procedure

After obtaining informed consent and applying the EEG-electrodes, ethnic identification was measured with a paper-and-pencil questionnaire that was presented as a separate study. Participants indicated their religion or belief system, and were asked to describe their ethnicity. Referring to this self-described group, we measured ethnic identification (six items on seven-point scales, $\alpha = 0.94$, $M = 4.60$, $s.d. = 1.29$, i.e. 'Belonging to this group is important to me'; Derks *et al.*, 2009).

Subsequently, following the design from Fazio's evaluative priming task (Fazio *et al.*, 1995), the second study was introduced as examining the automaticity of word comprehension and face recognition. This task measures the automatic activation of evaluations when positive and negative words are primed with exemplars from particular ethnic groups. In phase 1, unprimed word latencies were measured as participants categorized words as positive or negative (positive words: *love, safe, healthy, pleasant, luck, nice, success, tasty*; negative words: *cancer, accident, death, fight, emergency, loss, pain, ugly*; all presented twice in random order). Phases 2 and 3 served to bolster the cover story, by giving participants a facial memorization task and then a memory test. In phase 4, response latencies to words that were primed by Muslim *vs* non-Muslim faces were measured. In each trial (64 in total, presented in random order), a Muslim or non-Muslim face was presented for 315 ms, followed by a blank screen (135 ms). Then, the target word was presented until participants responded. Each face was presented four times, twice preceding a positive word and twice preceding a negative word. Participants were instructed to memorize the faces and categorize the words as quickly as possible. In phase 5, participants again received a facial memory test.

Finally, the electrodes were removed and participants were debriefed, thanked and rewarded for participation.

Scoring

To quantify implicit bias, separate indices were calculated of participants' reaction times to positive and negative words when primed with Muslim or non-Muslim faces. First, trials were dropped in which participants gave an incorrect response ($M = 2.3\%$) and with latencies longer than participants' mean + 3 s.d. ($M = 1.4\%$). Then, we calculated baseline latency scores per word in phase 1, and for each word per prime type (Muslim *vs* non-Muslim) in phase 4. After log transformations, facilitation scores were computed per word and per target and averaged per Target \times Valence combination by subtracting unprimed latencies in phase 1 from primed latencies in phase 4 (positive scores indicate greater facilitation). Because previous work has revealed a positive/negative asymmetry in intergroup bias (more discrimination on positive rather than negative attributes, Wenzel and Mummendey, 1996), we calculated an ingroup positivity score (non-Muslim positive–Muslim positive; positive scores indicate greater facilitation of positive words when primed with non-Muslim than Muslim faces), an outgroup negativity score (Muslim negative–non-Muslim negative; positive scores indicate greater facilitation of negative words when primed with Muslim than non-Muslim faces) and an implicit bias score combining the two.

EEG recording and component processing

Continuous EEG was recorded from 18 scalp sites (Fz, F3, F4, F7, F8, Cz, C3, C4, T7, T8, Pz, P3, P4, P7, P8, Oz, O1 and O2), as well as two electrodes placed on the left and right mastoids, using an ECI Electrocap and the ActiveTwo BioSemi system (BioSemi, Amsterdam, Netherlands). Vertical and horizontal eye movements were recorded with electrodes placed supra- and infra-orbitally at the left eye and on the left *vs* right orbital rim. The ground electrode during acquisition was formed by the Common Mode Sense active electrode and the Driven Right Leg passive electrode. EEG and EOG activity were sampled at 256 Hz and digitized on a laboratory computer using ActiView software (BioSemi).

The EEG data were processed and analyzed using Brain Vision Analyzer software (Brain Products). Off-line, data were re-referenced to the average activity of the mastoid electrodes, band-pass filtered with cutoffs of 0.1 and 30 Hz and corrected for ocular artifacts using the Gratton *et al.* (1983) method. EEG data were segmented into epochs from 200 ms before stimulus onset to 450 ms after stimulus onset. Separate ERP averages were computed for Muslim and non-Muslim faces. Epochs exceeding a 100 μV amplitude difference, a voltage step difference of 50 μV between sample points or with activity $< 0.5 \mu\text{V}$ were excluded from these averages. The resulting averages were baseline corrected for the 200 ms before stimulus onset (see Figure 1). The P200 ($M_{\text{latency}} = 161$ ms) was largest at the midline electrodes Cz and Pz, and scored there by locating maximum positive deflection between 120 and 250 ms. The N200 ($M_{\text{latency}} = 280$ ms) was largest at Fz and Cz, and was scored as the maximum negative deflection between 200 and 350 ms. Peak amplitudes were analyzed with mixed model regression analyses (Electrode \times Target Group [Muslim/non-Muslim] \times Ethnic Identification).

Results

ERP Data

P200. Analyses revealed no reliable Target Group or Identification effects on P200-amplitudes.

N200. N200s were larger to non-Muslim faces than to Muslim faces, but only for participants who reported high identification. Analyses

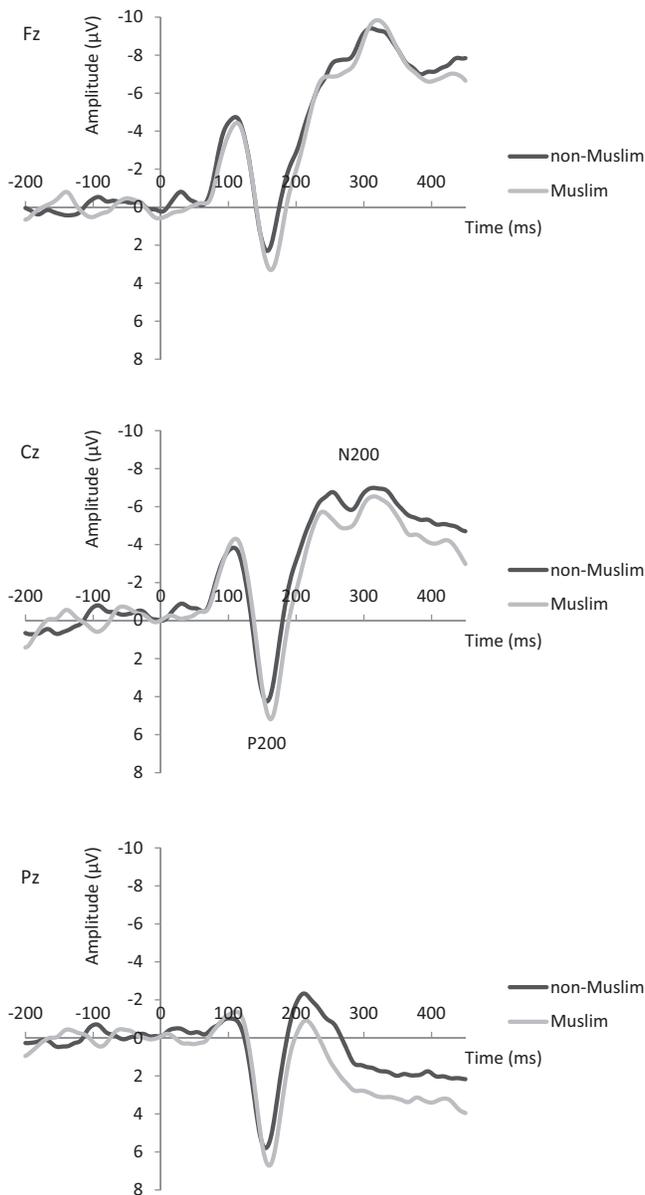


Fig. 1 ERP-waveforms at frontal, central and parietal midline electrodes for non-Muslim and Muslim target faces (Study 1: non-Muslim participants).

showed a significant three-way interaction between Target Group, Electrode and Identification, $F(1,19) = 6.84, P = 0.017$, partial $\eta^2 = 0.27$.¹ Separate analyses per electrode revealed significant effects on Cz but not Fz. On Cz there was a marginally significant Target Group main effect, $F(1,19) = 3.39, P = 0.081$, partial $\eta^2 = 0.15$, which was qualified by the predicted Target Group \times Identification interaction, $F(1,19) = 6.06, P = 0.024$, partial $\eta^2 = 0.24$ (see Figure 2). To assess whether the degree to which Muslim and non-Muslim faces are differentiated was predicted by identification, we compared N200s to the two groups the two groups at high ($M + 1$ s.d.) and low ($M - 1$ s.d.) levels of ethnic identification (Aiken and West, 1991). Among participants who felt relatively strong ties with their ethnic ingroup ($M + 1$ s.d.), N200s were significantly larger to non-Muslim than Muslim targets, $F(1,19) = 9.29, P = 0.007$, partial $\eta^2 = 0.33$.

¹ There was also a significant main effect of electrode, $F(1,19) = 6.29, p = .021$, partial $\eta^2 = .25$, with larger N200s on Fz than Cz.

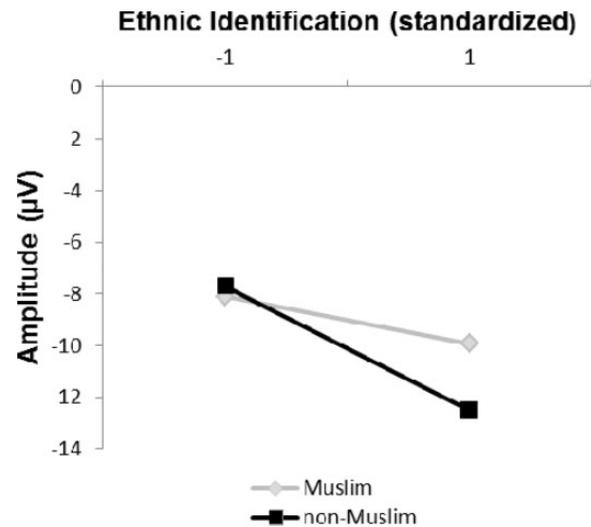


Fig. 2 Mean N200-amplitudes (Cz) to Muslim and non-Muslim targets for non-Muslim participants with low (-1 s.d.) and high ($+1$ s.d.) ethnic identification (Study 1).

Table 1 Correlations between event-related potentials and behavior in study 1

ERP target main effects	Implicit ingroup positivity	Implicit outgroup negativity	Implicit bias (total)
P200 Cz (outgroup-ingroup)	-0.09	-0.13	-0.18
P200 Pz (outgroup-ingroup)	0.01	-0.18	-0.12
N200 Fz (outgroup-ingroup)	0.20	-0.15	0.10
N200 Cz (outgroup-ingroup)	0.52*	-0.36	0.26

Note: * $P < 0.05$.

However, participants who identified less strongly with their ethnic group ($M - 1$ s.d.) showed no significant categorization in N200s between non-Muslim and Muslim targets, $F < 1$.

Implicit intergroup bias

As was found in previous work (Ashburn Nardo et al., 2003), the more participants identified with their ethnic group, the more intergroup bias they showed, $r(21) = 0.48, P = 0.03$. This effect was mainly due to higher ingroup positivity among participants with high ethnic identification, $r(21) = 0.43, P = 0.05$. Identification did not predict increased outgroup negativity, $r(21) = 0.05, P = 0.82$.

Correlations between ERPs and behavior

To reflect the target group main effects, we calculated contrast scores between amplitudes to Muslim and non-Muslim targets on the P200 (positive scores indicate more attention to outgroup) and N200 (positive scores indicate more attention to ingroup) and examined their correlation with implicit intergroup bias (see Table 1).

The N200 contrast was positively related to greater implicit ingroup positivity. Thus, the more participants differentiated between ingroup and outgroup in the N200, the quicker they were to respond to positive words that were primed with ingroup rather than outgroup faces.

Subsequently, we tested whether increased group differentiation in the N200 mediated the relationship between ethnic identification and implicit ingroup positivity, with bootstrapping analysis (Preacher and Hayes, 2008). Results revealed that the effect of ethnic identification on implicit ingroup positivity was reduced to non-significance when the

N200 difference score was included in the model ($B = 0.018$, $t = 1.03$, $P = 0.32$). The indirect effect of ethnic identification on ingroup positivity through differentiation in the N200 was significant ($B = 0.017$; confidence interval = 0.001–0.05). This indicates that the tendency of high ethnically identified participants to implicitly evaluate the ingroup more positively than the outgroup is predicted by their preferential attention to ingroup over outgroup faces as reflected in the N200.

STUDY 2

Method

Participants

Twenty-nine right-handed Muslim students (19 female, 8 male, 2 undisclosed gender; $M_{\text{age}} = 21.45$) participated in the study, which was conducted with approval of the Psychology Ethics Committee.

Procedure

Apart from the social identity threat manipulation, the procedure was identical to Study 1. To manipulate social identity threat, participants first read an introduction about the Netherlands being a multicultural country with inhabitants of different ethnicities and religions. In the high threat condition, participants were asked to describe their experience, thoughts and feelings during an incident in which they had experienced discrimination or disrespect owing to their religion. In the low threat condition, we aimed to offset any social identity threat by asking participants to describe an occurrence in which they had experienced respect from others owing to their religion. The participants performed the evaluative priming task (errors $M = 4.4\%$; too slow $M = 1.2\%$).²

EEG recording and component processing

EEG was recorded and processed in the same way as in Study 1. N200 ($M_{\text{latency}} = 286$ ms) was scored at Fz and Cz (200–350 ms, see Figure 3). The P200 ($M_{\text{latency}} = 156$ ms) was scored at Cz and Pz (120–250 ms). Peak amplitudes were analyzed with mixed model regression analyses (Electrode \times Target Group [Muslim/non-Muslim] \times Condition [$-1 =$ low threat/ $1 =$ high threat]).

Results

ERP data

P200. Again, analyses revealed no significant effects on the P200 amplitude.

N200. Recall that in previous research among members of ethnic majority groups and for the non-Muslim participants in Study 1, N200s were larger to ingroup than outgroup faces. For the Muslim participants in the current study, and in line with previous results among Asian Americans (Willadsen-Jensen and Ito, 2008), this effect was reversed. Overall, participants showed larger N200s to non-Muslim than Muslim faces, $F(1,25^3) = 4.19$, $P = 0.051$, partial $\eta^2 = 0.14$. This effect was qualified, however, by a significant Target Group \times Electrode \times Threat interaction, $F(1,25) = 4.50$, $P = 0.044$, partial $\eta^2 = 0.15$. Follow-up analyses revealed a Target Group main effect, $F(1,25) = 4.87$, $P = 0.037$, partial $\eta^2 = 0.16$, as well as a significant Target Group \times Threat effect, $F(1,25) = 4.54$, $P = 0.043$, partial $\eta^2 = 0.16$ (see Figure 4) on Fz. In the high threat condition, N200s

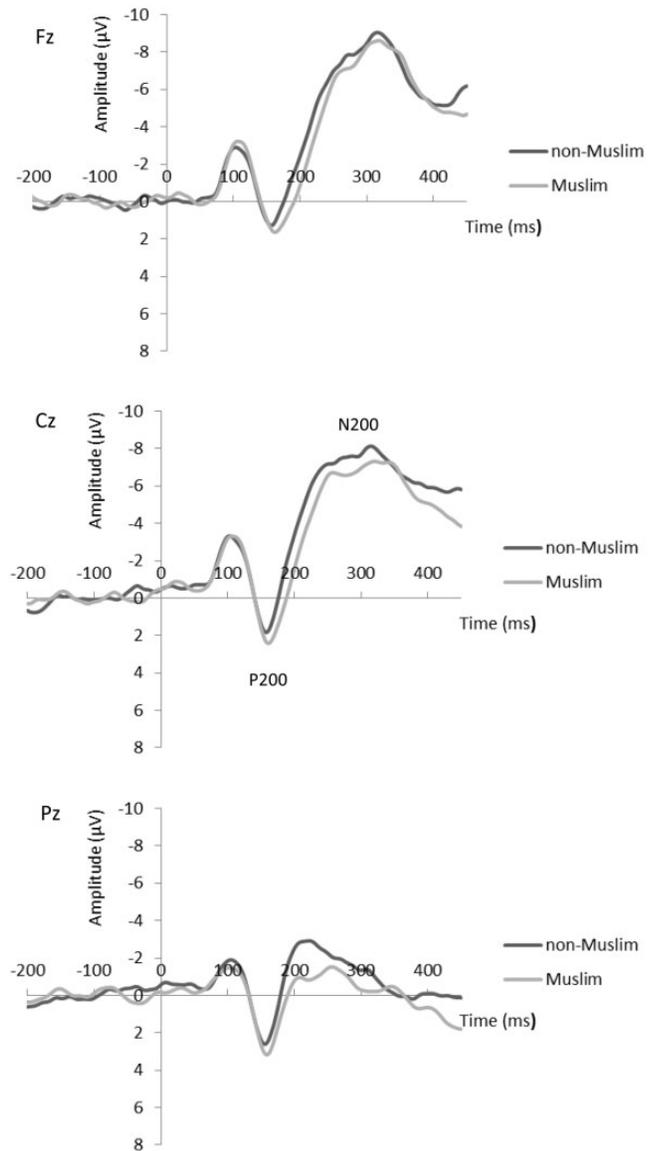


Fig. 3 ERP-waveforms at frontal, central and parietal midline electrodes for non-Muslim and Muslim target faces (Study 2: Muslim participants).

were significantly larger for non-Muslim outgroup faces than Muslim ingroup faces, $F(1,12) = 10.72$, $P = 0.007$, partial $\eta^2 = 0.47$. In the low threat condition, however, N200s did not differ between the two target groups ($F < 1$).

Implicit intergroup bias

Replicating previous work on identity threat and implicit coping (Rudman, Feinberg, and Fairchild, 2002), high threat resulted in significantly more intergroup bias than low social identity threat, $M_{\text{low threat}} = -0.02$, s.d. = 0.083; $M_{\text{high threat}} = 0.06$, s.d. = 0.09; $F(1,26) = 5.37$, $P = .03$, partial $\eta^2 = .17$. However, in contrast to Study 1, the social identity threat manipulation increased outgroup negativity, $F(1,26) = 4.49$, $P = 0.04$, partial $\eta^2 = 0.15$, but did not affect ingroup positivity ($F = 1.27$, $P = 0.27$). This confirms previous research showing that, although intergroup bias mostly takes the form of ingroup favoritism on positive dimensions rather than outgroup derogation on negative dimensions, aggravating conditions like

² One participant was not included in this analysis due to low task accuracy. Because this participant did recognize the faces in recognition phase 5, we concluded that this participant did look at the face targets in the task and therefore included his/her EEG-data.

³ Data from two participants were excluded due to noisy ERP-data.

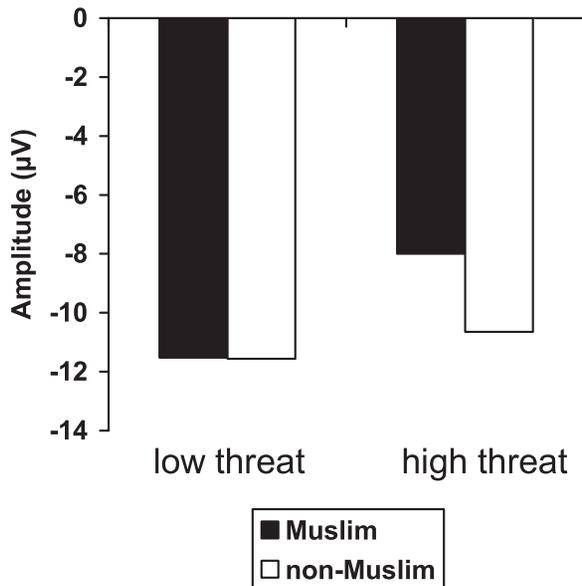


Fig. 4 Mean N200-amplitudes (Fz) to Muslim and non-Muslim targets for Muslim participants under low and high social identity threat (Study 2).

Table 2 Correlations between event-related potentials and behavior in study 2

ERP target main effects	Implicit ingroup positivity	Implicit outgroup negativity	Implicit bias (total)
Complete design			
P200 Cz (outgroup-ingroup)	0.18	-0.17	-0.03
P200 Pz (outgroup-ingroup)	0.05	-0.14	-0.08
N200 Fz (outgroup-ingroup)	0.13	-0.07	0.02
N200 Cz (outgroup-ingroup)	0.02	0.11	0.09
High threat condition			
P200 Cz (outgroup-ingroup)	0.36	-0.31	-0.11
P200 Pz (outgroup-ingroup)	0.25	-0.28	-0.13
N200 Fz (outgroup-ingroup)	0.55*	0.13	0.36
N200 Cz (outgroup-ingroup)	0.41	0.20	0.37
Low threat condition			
P200 Cz (outgroup-ingroup)	-0.01	0.14	0.06
P200 Pz (outgroup-ingroup)	-0.11	-0.25	-0.21
N200 Fz (outgroup-ingroup)	0.06	0.04	0.06
N200 Cz (outgroup-ingroup)	-0.26	-0.04	-0.22

Note: * $P = 0.05$.

identity threat can lead people to protect their collective esteem with outgroup derogation (Otten et al., 1996).

Correlations between ERPs and behavior

First, difference scores were calculated per ERP component reflecting the relevant target main effects (i.e. positive scores indicating more attention to outgroup in P200 and ingroup in N200). Examination of correlations between ERPs and behavior for both conditions simultaneously did not reveal any significant correlations (see Table 2). Examining correlations per condition, however, revealed a significant negative correlation between the N200 contrast on Cz and implicit ingroup positivity that was identical to the one found in Study 1: Participants who paid more attention to Muslim (ingroup) than non-Muslim (outgroup) targets in the N200 were 'more' likely to implicitly associate positive traits with their ingroup rather than their outgroup. However, unlike participants in Study 1, Muslim

participants in the high-threat condition paid more attention to outgroup (non-Muslim) rather than ingroup (Muslim) faces. Correspondingly they actually showed less ingroup positivity to the degree that they paid preferential attention to outgroup over ingroup faces in the N200. In other words, Muslims who resembled highly identified non-Muslim participants in Study 1 and preferentially processed non-Muslim over Muslim faces, also shared their implicit associations (more positivity associated with non-Muslim faces).

DISCUSSION

The goal of the current studies was to test whether early social categorization, as indicated by differential attention to ingroup and outgroup in early ERP components, is a stable process that occurs automatically, or whether this process is affected by people's motivation to express and promote their social identity. In two studies, we indeed found that people attend to category membership more strongly to the degree that they are motivated to promote their social identity. This stronger motivation resulted either from strong psychological attachment to their group (Study 1) or from an induced threat to social identity (Study 2). Both of these social identity motives increased the differentiation between ingroup and outgroup targets in the N200 component. As such, these studies are the first to show that individual or contextual differences that boost people's motivation to promote or protect their social identity affect the extent to which they implicitly categorize others in terms of us vs them.

These results lead to important insights for the social identity perspective. Up till now, social identity researchers have generally studied social categorization as a fixed starting point, a 'cold' cognitive perceptual process that takes place automatically or that can be prompted with a minimal group paradigm. The psychological meaning of the group to the individual or the content of group membership within a specific context was thought to follow from social categorization and moderate downstream consequences (e.g. ingroup favoritism, identity threat, coping; Ellemers et al., 2002; Tajfel and Turner, 1986). The current results suggest however that the relationship between social categorization on one hand and social identification and threat on the other is more dynamic: Social identification and social identity threat not only affect our behavior toward ingroup and outgroup members given that social categorization has taken place, but also feed back into spontaneous forms of person perception to determine the degree to which we perceive our social world through the lens of our social identity.

Importantly, the high temporal resolution of ERPs not only enabled the assessment of early attentional processes, but even allowed us to distinguish two separate stages of social categorization: the P200, which is often interpreted as attention for threatening cues, and the N200, indicating deeper processing of some faces over others. In both studies, significant social identity effects were found on the N200 but not the P200. This suggests that social identification and threats to social identity do not change which groups we process as threatening, but determine which targets receive more attentional resources. Study 1 reveals that as early as 280 ms after stimulus presentation, high identifiers preferentially processed ingroup over outgroup targets. Moreover, the deeper processing of ingroup over outgroup faces at this early stage predicted high identifiers' stronger implicit ingroup favoritism in subsequent biased responses, hereby revealing the impact of this attentional process on intergroup responses.

Study 2 adds to this picture by showing that which group receives the most attentional resources also depends on the direct context: Muslim participants who were reminded of threats to their religious identity processed non-Muslim outgroup targets over targets belonging to their ingroup. This finding resembles results from

Willadsen-Jensen and Ito (2008) among Asian-American participants and suggests that depth of processing as indicated by N200-amplitude is not simply a factor of group membership. Instead, processing depth is affected by the social context and the standing of the groups within that context. Similar to Asian participants who 'can adopt a perspective that matches that of White perceivers when they find themselves in a context that involves primarily White individuals' (Willadsen-Jensen and Ito, 2008, p. 193), Muslim participants in Study 2 also seemed to adopt the perspective of the non-Muslim majority when experiencing social identity threat, processing non-Muslim faces more deeply than Muslim faces. Perhaps because members of minority or marginalized groups are much more likely to interact with, depend on and even identify with the majority group, outgroup faces may under certain circumstances receive even more attention than ingroup faces.

As such, the results of Study 2 emphasize the complex effects of social identity threat on the perception, intergroup evaluations and coping strategies of members of marginalized groups. On the one hand, social identity threat seemed to trigger implicit coping strategies, leading Muslims to show implicit negative outgroup evaluations. However, social identity threat also led Muslim participants to perceive their social world in a way similar to the non-Muslim majority, with enhanced processing of ethnic majority faces. Notably, just as with non-Muslim participants in Study 1, this processing pattern predicted a tendency to implicitly evaluate non-Muslims more positively than Muslims, suggesting that enhanced processing of outgroup faces is detrimental to Muslims' implicit collective self-esteem. One interpretation of these combined ERP and behavioral responses may be Muslims' disidentification from their religious minority group and assimilation into the majority perspective. An obvious avenue for future research would be to measure Muslims' religious identification as well as manipulate social identity threat. It has been shown that high group identification can serve as a buffer and that high identifiers are better able to cope with social identity threat (e.g. Hansen and Sassenberg, 2006). Perhaps highly identified Muslims are more resilient to social identity threat and will not show the N200 patterns found here, but instead show deeper processing of ingroup targets and more positive implicit ingroup evaluations similar to the non-Muslim participants in Study 1.

One unexpected finding in the current studies was the absence of significant categorization effects in the P200. This could be owing to the lower number of trials per target group (32) relative to previous work (e.g. Dickter and Bartholow, 2007: 320 per target race; Ito and Urland, 2003: 160 per target race). Additionally, most previous studies that report intergroup differentiation in the P200 have used targets differing by race (Black vs White faces; e.g. Dickter and Bartholow, 2007; Ito and Urland, 2003, 2005; Amodio, 2010) rather than the distinction in religion (women with and without a headscarf) made in the current study. It is possible that this distinction does not create as much attentional differences as distinctions based on race.

The current results add to a growing body of work indicating that even seemingly bottom-up neural processes associated with the first steps in information processing are influenced by top-down motivational goals (Amodio, 2013; Amodio et al., 2003; Balcetis and Dunning, 2006; Van Bavel and Cunningham, 2010). For example, Van Bavel and colleagues (2008) revealed that differentiation in amygdala activity to Black compared with White targets can be altered by changing people's momentary self-categorizations from a focus on race to membership of a minimal group. The ERP-methodology of the current studies adds to this work by providing a fine-grained insight into how early in person perception social identity effects take place (i.e. around 280 ms) and which cognitive processes are and are not affected by it (i.e. N200 vs P200). The current results show that above and beyond the effects of self-categorization (making one group

membership salient, but not the other), the strength of people's chronic identification with this social category, and a contextually induced need to cope with identity threat has a top-down influence on early social categorization, leading people to more strongly perceive others in line with their social identity.

Conflict of Interest

None declared.

REFERENCES

- Aiken, L.S., West, S.G. (1991). *Multiple Regression: Testing and Interpreting Interactions*. Newbury Park, CA: Sage.
- 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, 2609–17.
- Amodio, D.M. (2013). Implicit prejudice and the regulation of intergroup responses: theoretical contributions of the social neuroscience approach. In: Derks, B., Scheepers, D., Ellemers, N., editors. *Neuroscience of Prejudice and Intergroup Relations*. New York, NY: Psychology Press, pp. 167–89.
- Amodio, D.M., Harmon Jones, E., Devine, P.G. (2003). Individual differences in the activation and control of affective race bias as assessed by startle eyeblink response and self-report. *Journal of Personality and Social Psychology*, 84(4), 738–53.
- Ashburn Nardo, L., Knowles, M.L., Monteith, M.J. (2003). Black Americans' implicit racial associations and their implications for intergroup judgment. *Social Cognition*, 21(1), 61–87.
- Balcetis, E., Dunning, D. (2006). See what you want to see: motivational influences on visual perception. *Journal of Personality and Social Psychology*, 91(4), 612–25.
- Bar Haim, Y., Lamy, D., Glickman, S. (2005). Attentional bias in anxiety: behavioral and ERP study. *Brain and Cognition*, 59(1), 11–22.
- Branscombe, N.R., Ellemers, N., Spears, R., Doosje, B. (1999). The context and content of social identity threat. In: Ellemers, N., Spears, R., Doosje, B., editors. *Social Identity: Context, Commitment, Content*. Oxford, England: Blackwell, pp. 35–58.
- Brewer, M.B. (1988). A dual process model of impression formation. In: Srull, T.K., Wyer, R.S.J., editors. *A Dual Process Model of Impression Formation*. *Advances in Social Cognition*, Vol. 1, Hillsdale, NJ: Lawrence Erlbaum Associates, pp. 1–36.
- Castano, E., Yzerbyt, V., Bourguignon, D., Seron, E.O. (2002). Who may enter? The impact of in-group identification on in-group/out-group categorization. *Journal of Experimental Social Psychology*, 38(3), 315–22.
- Correll, J., Urland, G.R., Ito, T.A. (2006). Event-related potentials and the decision to shoot: the role of threat perception and cognitive control. *Journal of Experimental Social Psychology*, 42(1), 120–8.
- Crocker, J., Major, B. (1989). Social stigma and self-esteem: the self-protective properties of stigma. *Psychological Review*, 96(4), 608–630.
- Derks, B., Van Laar, C., Ellemers, N. (2009). Working for the self or working for the group: how self-versus group affirmation affects collective behavior in low-status groups. *Journal of Personality and Social Psychology*, 96(1), 183.
- Dickter, C.L., Bartholow, B.D. (2007). Racial ingroup and outgroup attention biases revealed by event-related brain potentials. *Social Cognitive and Affective Neuroscience*, 2(3), 189–98.
- Doosje, B., Ellemers, N., Spears, R. (1999). Commitment and intergroup behavior. In: Ellemers, N., Spears, R., Doosje, B., editors. *Social Identity: Context, Commitment, Content*. Oxford, England: Blackwell, pp. 84–106.
- Dotsch, R., Wigboldus, D., Van Knippenberg, A. (2011). Biased allocation of faces to social categories. *Journal of Personality and Social Psychology*, 100, 999–1014.
- Ellemers, N., Spears, R., Doosje, B. (2002). Self and social identity. *Annual Review of Psychology*, 53, 161–186.
- Fazio, R.H., Jackson, J.R., Dunton, B.C., Williams, C.J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: a bona fide pipeline? *Journal of Personality and Social Psychology*, 69(6), 1013–27.
- Fiske, S.T., Neuberg, S.L. (1990). A continuum of impression-formation, from category-based to individuating processes - influences of information and motivation on attention and interpretation. *Advances in Experimental Social Psychology*, 23, 1–74.
- Gratton, G., Coles, M.G.H., Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology*, 55(4), 468–84.
- Hansen, N., Sassenberg, K. (2006). Does social identification buffer or harm? The impact of social identification on anger after social discrimination. *Personality and Social Psychology Bulletin*, 32, 983–996.
- Haslam, S.A., Oakes, P.J., Turner, J.C., McGarty, C. (1995). Social categorization and group homogeneity: changes in the perceived applicability of stereotype content as a function of comparative context and trait favourableness. *British Journal of Social Psychology*, 34(2), 139–60.
- Hugenberg, K., Young, S.J., Bernstein, M.J., Sacco, D.F. (2010). The categorization-individuation model: an integrative account of the other-race recognition deficit. *Psychological Review*, 117, 1168–87.

- Ito, T.A. (2013). Imaging the pictures in our heads: using ERPs to inform our understanding of social categorization. In: Derks, B., Scheepers, D., Ellemers, N., editors. *Neuroscience of Prejudice and Intergroup Relations*. London, UK: Psychology Press, pp. 25–44.
- Ito, T.A., Bartholow, B.D. (2009). The neural correlates of race. *Trends in Cognitive Sciences*, 13(12), 524–31.
- 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(4), 616–26.
- Ito, T.A., Urland, G.R. (2005). The influence of processing objectives on the perception of faces: an ERP study of race and gender perception. *Cognitive, Affective and Behavioral Neuroscience*, 5(1), 21–36.
- Jetten, J., Spears, R., Manstead, A.S.R. (1999). Group distinctiveness and intergroup discrimination. In: Ellemers, N., Spears, R., Doosje, B., editors. *Social Identity: Context, Commitment, Content*. Oxford, England: Blackwell Science Ltd, pp. 107–26.
- Jetten, J., Spears, R., Postmes, T. (2004). Intergroup distinctiveness and differentiation: a meta-analytic integration. *Journal of Personality and Social Psychology*, 86(6), 862–79.
- 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(5), 738–52.
- Lalonde, R.N. (2002). Testing the social identity-intergroup differentiation hypothesis: ‘We’re not American eh!’. *British Journal of Social Psychology*, 41(4), 611–30.
- Lucas, H.D., Chiao, Y., Paller, K. (2011). Why some faces won’t be remembered: brain potentials illuminate successful versus unsuccessful encoding for same-race and other-race faces. *Frontiers in Human Neuroscience*, 5, 20.
- Luck, S.J. (2005). *An Introduction to the Event-Related Potential Technique*. Cambridge, MA: The MIT Press.
- Ma, Y., Yang, S., Han, S. (2011). Attitudes influence implicit racial face categorization in a perceptual task. *Group Processes and Intergroup Relations*, 14, 887–99.
- Macrae, C.N., Bodenhausen, G.V. (2000). Social cognition: thinking categorically about others. *Annual Review of Psychology*, 51, 93–120.
- Macrae, C.N., Cloutier, J. (2009). A matter of design: priming context and person perception. *Journal of Experimental Social Psychology*, 45(4), 1012–15.
- Maner, J., Miller, S.L., Moss, J., Leo, J., Plant, E.A. (2012). Selective responses to threat: the roles of race and gender in decisions to shoot. *Journal of Personality and Social Psychology*, 103, 70–83.
- Miller, S.L., Maner, J.K., Becker, D.V. (2010). Self-protective biases in group categorization: threat cues shape the psychological boundary between “us” and “them”. *Journal of Personality and Social Psychology*, 99, 62–77.
- Otten, S., Mummendey, A., Blanz, M. (1996). Intergroup discrimination in positive and negative outcome allocations: the impact of stimulus valence, relative group status, and relative group size. *Personality and Social Psychology Bulletin*, 22, 568–81.
- Phalet, K., Ter Wal, J., editors. (2004). *Moslim in Nederland [Muslim in the Netherlands]*. Den Haag: Sociaal en Cultureel Planbureau.
- Preacher, K.J., Hayes, A.F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, 40, 879–91.
- Quinn, K.A., Rosenthal, H.E.S. (2012). Categorizing others and the self: how social memory structures guide social perception and behavior. *Learning and Motivation*, 43(4), 247–58.
- Rudman, L.A., Feinberg, J., Fairchild, K. (2002). Minority members’ implicit attitudes: automatic ingroup bias as a function of group status. *Social Cognition*, 20(4), 294–320.
- Tajfel, H., Billig, M.G., Bundy, R.P., Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology*, 1(2), 149–78.
- Tajfel, H., Turner, J.C. (1986). The social identity theory of intergroup behavior. In: Worchel, S., Austin, W.G., editors. *The Psychology of Intergroup Relations*. Chicago, IL: Nelson-Hall, pp. 7–24.
- Taylor, S., Fiske, S., Etcoff, N., Ruderman, A. (1978). Categorical and contextual bases of person memory and stereotyping. *Journal of Personality and Social Psychology*, 36, 778–93.
- Turner, J.C., Hogg, M.A., Oakes, P.J., Reicher, S.D., Wetherell, M.S. (1987). *Rediscovering the Social Group: A Self-Categorization Theory*. New York, NY: Basil Blackwell.
- Turner, J.C., Oakes, P.J., Haslam, S.A., McGarty, C. (1994). Self and collective: cognition and social context. *Personality and Social Psychology Bulletin*, 20(5), 454–63.
- Van Bavel, J.J., Cunningham, W.A. (2010). A social neuroscience approach to self and social categorisation: a new look at an old issue. *European Review of Social Psychology*, 21, 237–84.
- Van Bavel, J.J., Packer, D.J., Cunningham, W.A. (2008). The neural substrates of in-group bias: a functional magnetic resonance imaging investigation. *Psychological Science*, 19(11), 1131–9.
- Wenzel, M., Mummendey, A. (1996). Positive-negative asymmetry of social discrimination: a normative analysis of differential evaluations of in-group and out-group on positive and negative attributes. *British Journal of Social Psychology*, 35(4), 493–507.
- Wheeler, M.E., Fiske, S.T. (2005). Controlling racial prejudice: social-cognitive goals affect amygdala and stereotype activation. *Psychological Science*, 16(1), 56–63.
- Willadsen-Jensen, E.C., Ito, T.A. (2008). A foot in both worlds: Asian Americans’ perceptions of Asian, White, and racially ambiguous faces. *Group Processes and Intergroup Relations*, 11(2), 182–200.
- Zyerbyt, V., Dumont, M., Wigboldus, D., Gordijn, E. (2003). I feel for us: the impact of categorization and identification on emotions and action tendencies. *British Journal of Social Psychology*, 42(4), 533–49.