OXFORD

Seeing is not stereotyping: the functional independence of categorization and stereotype activation

Tiffany A. Ito,[1] and Silvia Tomelleri[2]

[1]Department of Psychology and Neuroscience, University of Colorado, Boulder, CO 80309-0345, USA and
[2]Department of Developmental Psychology, University of Padova, Padua, 35131, Italy

Correspondence should be addressed to Tiffany A. Ito, Department of Psychology and Neuroscience, University of Colorado Boulder, 345 UCB, Boulder, CO 80309-0345, USA. E-mail: tiffany.ito@colorado.edu

Abstract

Social categorization has been viewed as necessarily resulting in stereotyping, yet extant research suggests the two processes are differentially sensitive to task manipulations. Here, we simultaneously test the degree to which race perception and stereotyping are conditionally automatic. Participants performed a sequential priming task while either explicitly attending to the race of face primes or directing attention away from their semantic nature. We find a dissociation between the perceptual encoding of race and subsequent activation of associated stereotypes, with race perception occurring in both task conditions, but implicit stereotyping occurring only when attention is directed to the race of the face primes. These results support a clear conceptual distinction between categorization and stereotyping and show that the encoding of racial category need not result in stereotype activation.

Key words: implicit stereotyping; racial categorization

Categorization and stereotyping have been treated as inextricably linked, facilitating the assumption that mere exposure to category exemplars activates stereotypical associations (Bargh, 1984, 1999; Devine, 1989; Dovidio *et al.*, 1986; Stangor *et al.*, 1992). However, despite assumptions that stereotyping follows from categorization in many theoretical accounts, empirical research reveals dissociations between the two processes with respect to automaticity, suggesting a more nuanced relation between the two. On the one hand, the activation and application of stereotypic associations are widely considered to be conditionally automatic (e.g. Blair and Banaji, 1996; Gilbert and Hixon, 1991; Jones and Fazio, 2010; Macrae *et al.*, 1997; Wittenbrink *et al.*, 2001), meaning that they depend to some degree on the existence of other enabling circumstances (Bargh, 1999) . In this way, stereotypes never operate in a way that is completely unaffected by other factors. One illustration of this was provided

by Macrae et al. who showed that implicit stereotyping depended on the way a group exemplar was processed, with greater gender stereotype activation when female face primes were processed in a way that facilitated semantic encoding (whether the picture depicted sometime animate or inanimate) as compared with more superficial encoding (detecting the presence of a dot in the picture). Similar results were obtained by Jones and Fazio (2010) for implicit racial stereotyping, with Black face primes facilitating response to guns only when the primes were processed in terms of their race (see also, Wheeler and Fiske, 2005).

In contrast, studies examining categorization, or the cognitive grouping of individuals with others perceived to be similar (Turner, 1987), more often conclude that it occurs unconditionally for dimensions such as race and gender. That is, grouping individuals in terms of apparent race and gender is assumed to

occur to the same degree across different processing goals (Ito and Urland, 2003, 2005; Kubota and Ito, 2007, 2016; Martin and Macrae, 2007; Mason *et al.*, 2006; Tomelleri and Castelli, 2012). Ito and Urland (2005) provide an example, showing that neural sensitivity to racial group membership occurs even when participants are focused on more individuating information or on non-semantic visual information. Stereotyping was not measured in these studies, but such manipulations can decrease stereotype application (Macrae *et al.*, 1997; Wheeler and Fiske, 2005), implying a dissociation between categorization and stereotyping. Similarly, in Gilbert and Hixon (1991), cognitively busy subjects generated fewer stereotype-consistent responses in a word completion task as compared with not busy subjects, but recall of the race of the research assistant administering the task was equally high in both conditions. This suggests that target race was categorized to the same degree in both conditions even though stereotype application differed.

More recent models of person perception recognize the potential for categorization and stereotyping to dynamically influence each other, and also for separate higher order processes to modulate both categorization and stereotyping (Freeman and Ambady, 2011). If one assumes that higher order processes can independently affect categorization and stereotyping, such models would allow for the possibility that factors like task goals could differentially impact categorization and stereotyping, creating dissociations between the two. Unfortunately, past studies are limited by not simultaneously and independently measuring categorization and stereotyping. As a consequence, there are no demonstrations of which we are aware assessing the potential for the same higher order goal to have different effects on categorization and stereotyping within the same participants and at the same time. Gilbert and Hixon (1991) did show a within subject dissociation between stereotyping (word completion) and categorization (recall of the assistant's race), but the measures were separated in time, with categorization assessed at the end of study, rather than simultaneous with the measure of stereotyping.

The absence of simultaneous assessment of categorization and stereotyping is likely due to their close conceptual and temporal relations, and the potential for each to occur quickly and without conscious awareness (e.g. Brewer, 1988; Fiske and Neuberg, 1990; Freeman and Ambady, 2011; Kunda and Thagard, 1996), making their independent assessment challenging. Fortunately, some of the signature features of event-related brain potentials (ERPs), including high temporal resolution and sensitivity to processes that occur outside of conscious awareness (Ito and Bartholow, 2009) make them well-suited to assess processes related to categorization at the same time that stereotype activation is assessed behaviorally. Capitalizing on this in the present study, we recorded ERPs to Black and White faces that served as primes in a sequential priming task in which participants' imperative task was to classify pictures of subsequently presented guns and insects (Judd *et al.*, 2004). We facilitated attention to prime race during one block of trials by simply asking participants to categorize each individual in terms of race. In contrast, we directed attention away from the semantic nature of the primes by having participants attend to a face-irrelevant physical feature in another block of trials using a task that has been shown to decrease implicit stereotype activation (Macrae *et al.*, 1997; Wheeler and Fiske, 2005). By varying the task performed with respect to the face primes, we could simultaneously assess the conditional automaticity of race processing—as reflected in ERP responses to the face primes—and the conditional automaticity of implicit stereotyping—via response latencies and accuracy to the imperative stimuli.

Given past research on the conditional activation of implicit stereotypes, we expected Black primes to facilitate categorization of stereotype-congruent guns in the race categorization condition, but not in the feature detection condition. Of critical interest are neural responses to faces of different races under these different conditions. To assess this, we capitalized on a large body of past research showing that ERPs are sensitive to racial category membership (Dickter and Bartholow, 2007; Ito and Urland, 2003, 2005; Kubota and Ito, 2007, 2016; Willadsen-Jensen and Ito, 2006, 2008, 2015). In these past studies, two components associated with selective attention differentiate between Black and White faces and have been used as evidence of racial categorization: P200s are consistently larger to Black than White faces while N200s are larger to White than Black faces among non-Black, largely White perceivers.[1] To the degree that social categorization reflects the grouping of individuals perceived to be similar (Turner, 1987), neural activity differentiating between members of different racial groups indicates that people with apparently different racial backgrounds are perceived dissimilarly. In this way, modulations in the amplitude of the P200 and N200 response can be taken as evidence that exemplars from different racial categories have to some degree been grouped into different categories. Of interest, here, is whether this occurs in both task conditions, or mirrors the behavioral results and only occurs when participants explicitly attend to race. If physiognomic cues associated with race are encoded automatically, regardless of task, the P200 and N200 should differentiate between Black and White faces in both task conditions. However, if sensitivity to social category information is also conditionally automatic, then Black/White differentiation in the P200 and N200 may be absent in the non-semantic task condition.

Materials and methods

Participants

Seventy-two University of Colorado undergraduates participated in exchange for course credit or monetary compensation. Data from seven participants were discarded because of a high number of incorrect responses, suggesting lack of attention. Four of these participants had high error rates both to the imperative stimulus and to questions about the face prime (explained shortly); the remaining three participants had high error rates to only the questions about the primes (<20% mean accuracy in all cases). This left 65 participants in the analyses. Demographic data were available for 61 of these, with 29 reporting their gender as female and 32 as male ($M_{age} = 21.54$ years). Forty-five reported their race as White, two as Black, five as Asian, five as Hispanic, one as Indian, one as multi-racial, and two as 'other' with no additional elaboration. Results do not change when Black participants are omitted from analyses so we opted for including all possible data. All participants performed the priming task, with 30 doing so while ERPs were also recorded (within this subset of participants, 17 were female, and racial background was reported as 23 White, two Hispanic,

1 Given that none of the participants who provided ERP data were Black, our review focuses on results obtained with non-Black participants. For clarity, we note that these ERP components are sensitive to ingroup/outgroup racial status. Specifically, whereas White participants show larger P200s to Blacks and larger N200s to Whites, Black participants show the inverse: larger P200s to Whites and larger N200s to Blacks (Dickter and Bartholow, 2007; see also Willadsen-Jensen and Ito, 2008, for similar results with White and Asian participants).

three Asian, one multiracial, one 'other' with no further elaboration).[2]

Stimuli

Primes consisted of yearbook-style color pictures of 20 Black and 20 White males. Ratings from an independent sample indicated that faces were clearly perceived as Black or White and equal in mean attractiveness. A white dot (3 mm diameter) was placed on a random subset of 10 Black and 10 White pictures in a random position that did not obscure a critical face feature (see Macrae *et al.*, 1997). A second subset of 20 total pictures with dots was also created; each participant saw only one set in order to counterbalance between participants the pictures with and without dots.

Targets consisted of eight color pictures each of handguns and insects. These categories were chosen because both share a negative valence while only handguns are stereotypically associated with Blacks (Judd *et al.*, 2004).

Procedure

All participants completed two different blocks of 320 trials each, once performing the race categorization task and once performing the feature (dot) detection task. Trial structure was identical in both blocks, consisting of the presentation of a Black or White face prime followed by either a gun or insect. Participants were told to indicate as quickly and accurately as possible whether the photo showed a gun or insect via button press, with responses labeled *gun* and *insect*. The two blocks differed in how participants were told to process the face primes. In the *race categorization task*, participants were told to mentally categorize the face as Black or White, then prepare to respond to the gun or insect. In the *feature detection task*, they were told to mentally determine whether the initial stimulus had a small dot on it.[3] Each trial began with the presentation of a fixation point for 500 ms followed by a face prime for 155 ms. A blank screen then occurred for 93 ms, followed by the target, which was visible until participants made their response (prime-target SOA = 248 ms, see Judd *et al.*, 2004).

In order to verify that participants processed face primes differently in the two conditions, a task-relevant question appeared immediately after the gun/insect response was made. In the race categorization task, these questions asked participants if the prime was either Black or White (*Black?* or *White?*). In the feature detection task, the questions asked whether a dot was present (*Dot?* or *No Dot?*). Each task-relevant question appeared on a random 50% of trials (i.e. in the race categorization task, participants were asked if the face was White on 50% of the trials and if it was Black on the other 50% of the trials); the correct answer to each type of question was yes on 50% of the trials. Answers were made via button press, with responses labeled *yes* and *no*. See Supplementary Materials for full analysis of the responses to the prime questions. ERP and behavioral analyses assessing primary hypotheses reported in the main text are based on trials for which the prime-related questions were answered correctly. A randomly determined inter-trial interval between 1500 and 2000 ms occurred after the prime-related response.

2   More behavioral than ERP data were collected to increase our ability to detect what we expected to be smaller effects in the behavioral (cf. Macrae *et al.*, 1997; Wheeler and Fiske, 2005) than ERP measures (cf. Ito and Urland, 2003).

3   Block order was inadvertently fixed for participants providing only behavioral data, with the dot detection task occurring first. Task order was counterbalanced across participants who provided ERP data. For those participants, task order did not qualify the Task × Prime Race × Target effects of interest.

Psychophysiological data collection and reduction

ERP data were recorded from 28 tin electrodes embedded in a stretch-lycra cap (Electro-Cap International, Eaton, OH, USA), positioned according to the 10–20 international system (Jasper, 1958). The ground electrode was mounted in the cap on the midline between the frontal pole and Fz. Tin electrodes were also placed over the left and right mastoid, with scalp data referenced online to the left mastoid. Electrodes on the supra- and sub-orbit of the left eye assessed vertical eye movements. Electrodes on the outer canthi of the left and right eyes assessed horizontal eye movements. Electrode impedances were below 10 kΩ. Electrode gel (Eaton, OH, USA) was used as the conducting medium for all electrodes. ERP recordings were amplified with a gain of 500 by NeuroScan Synamps (Sterling, VA, USA), with a bandpass of .15–30 Hz, and digitized at 1000 Hz.

Off-line, data were re-referenced to a computed average of the left and right mastoids. We then applied a regression procedure to correct for eyeblink artifact (Semlitsch *et al.*, 1986). Epochs were next created starting 100 ms before prime onset and continuing for 1000 ms after prime onset, then baseline corrected to the mean voltage of the pre-stimulus period. Each trial was then visually inspected for remaining blink or muscle artifact. If artifact was detected in any channel, the entire trial was removed from analyses.

Separate average waveforms were then created for each participant as a function of prime race, whether the picture had a dot on it, and what task was being performed with respect to the face primes in that block, resulting in eight separate averages for each participant. The average number of trials retained per subject per condition was 33.83 (s.d. = 6.18). Analyses focused on the P200 and N200, two components that past research repeatedly demonstrates are sensitive to social category membership, specifically showing differences in responses to Black and White faces (Correll *et al.*, 2006; Ito and Urland, 2003; 2005; Kubota and Ito, 2007; Willadsen-Jensen and Ito, 2006; 2015). Local peak component amplitudes were scored by locating the maximal positive deflection between 120 and 190 ms (P200)[4] and the maximal negative deflection between 190 and 290 ms (N200). These latency windows were determined based on visual inspection of peak latencies in the grand averages and past studies (Correll *et al.*, 2006; Ito and Urland, 2003; 2005; Kubota and Ito, 2007; Willadsen-Jensen and Ito, 2006; 2015).

Results

We first examined behavioral responses for evidence of task effects on implicit stereotyping. Priming effects in response latency as compared with accuracy are most likely when a task lacks a tight response deadline (as was the case in this study). However, some studies either lacking a response deadline or with a long deadline do find similar patterns of racial bias in both response latencies and accuracy (e.g. Correll *et al.*, 2002; Judd *et al.*, 2004). In light of this, we report analyses on both response latencies and accuracy. We then examined ERP responses to see if comparable task modulations were observed during the encoding of the face primes, which would suggest that perception of physiognomic cues indicative of category membership occurs conditionally based on perceivers' goals, or if race perception occurs automatically and equally regardless

4   Differences in peak latency suggest the P200 we quantified is distinct from the face-specific N170. In these data, the P200 had a peak latency of ~144 ms whereas the N170 visible over lateral temporal regions had a later peak latency of ~167 ms.
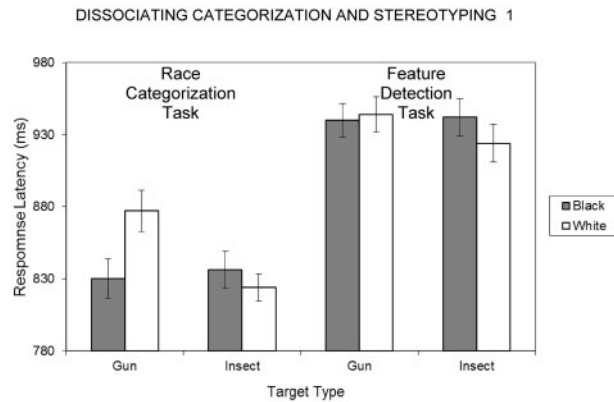
DISSOCIATING CATEGORIZATION AND STEREOTYPING 1

Fig. 1. Mean response latencies (in ms) as a function of prime race and target object separately for the race categorization and feature detection tasks. Error bars show ±1 SEM, based on within-subject error variance (Cousineau, 2005).

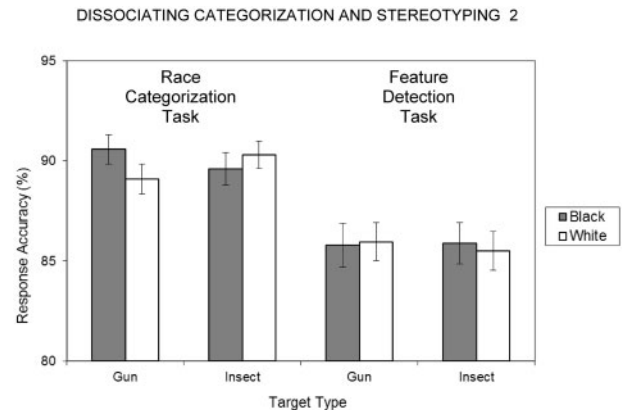DISSOCIATING CATEGORIZATION AND STEREOTYPING 2

Fig. 2. Mean accuracy as a function of prime race and target object separately for the race categorization and feature detection tasks. Error bars show ±1 SEM, based on within-subject error variance (Cousineau, 2005).

of task. Preliminary analyses revealed similar behavioral effects among ERP and non-ERP participants, so data from all participants were combined in the behavioral analyses.

Behavioral responses

Response latencies. A 2 (Task: race categorization, feature detection) × 2 (Dot Presence: yes, no) × 2 (Prime Race: Black, White) × 2 (Target: gun, insect) within subjects ANOVA was performed on trials on which the gun or insect was correctly classified, after dropping trials on which participants' responses exceeded ±3 s.d. relative to their mean. Response latencies on individual trials were log-transformed, then averaged within condition. All analyses were performed on the log-transformed values but raw response latencies are reported, here, for ease of interpretation. The predicted Task × Prime Race × Target Type interaction indicative of changes in implicit stereotyping as a function of task was significant, $F(1,64) = 5.00$, $P = .029$, $\eta_p^2 = .072$. To test for task effects on implicit stereotyping, we examined separate Prime Race × Target ANOVAs within each task condition (see Figure 1). As expected, this interaction was significant in the race categorization condition, $F(1,64) = 19.49$, $P < .0001$, $\eta_p^2 = .233$. Implicit stereotyping was seen in faster responses to guns after Black ($M = 830.00$ ms, s.d. $= 263.29$) than White primes ($M = 877.76$ ms, s.d. $= 329.58$), $F(1,64) = 16.54$, $P < .0001$, $\eta_p^2 = .205$. There was a marginally significant effect in the opposite direction for insects, with a trend toward faster responses following White ($M = 824.36$ ms, s.d. $= 287.67$) than Black primes ($M = 836.75$ ms, s.d. $= 295.74$), $F(1,64) = 3.34$, $P = .072$, $\eta_p^2 = .050$.

The Prime Race × Target interaction was also significant in the feature detection task, $F(1,64) = 5.14$, $P = .027$, $\eta_p^2 = .074$. However, there was no evidence that responses to guns were facilitated by Black ($M = 940.16$, s.d. $= 36.48$) as compared with White primes ($M = 944.80$, s.d. $= 37.83$), $F(1,64) = 0.98$, $P = .325$, $\eta_p^2 = .015$. However, replicating the trend in the race categorization task, the simple effect of race was significant for insects in the feature detection task, with faster responses following White ($M = 924.05$, s.d. $= 35.7$) than Black primes to insects ($M = 942.74$, s.d. $= 35.95$), $F(1,64) = 4.04$, $P = .049$, $\eta_p^2 = .059$.[5]

Accuracy. Response accuracy within each condition revealed the same pattern as response latencies, with implicit stereotyping only in the race categorization condition. The expected three-way Prime Race × Target × Task interaction was significant, $F(1,64) = 4.39$, $P = .040$, $\eta_p^2 = .064$. Separate Prime Race × Target ANOVAs within each task condition revealed the expected significant interaction in the race categorization condition, $F(1,64) = 4.98$, $P = .029$, $\eta_p^2 = .072$. As expected, participants more accurately categorized guns following Black ($M = 90.58\%$, s.d. $= .14$) than White primes ($M = 89.10\%$, s.d. $= .16$), $F(1,64) = 4.21$, $P = .044$, $\eta_p^2 = .062$ (see Figure 2). There was no difference in accuracy to insects as a function of prime race in the race categorization task ($F(1,64) = 1.33$, $P = .253$, $\eta_p^2 = .020$), and no Prime Race × Target Type interaction in the dot detection task ($F(1,64) = 0.64$, $P = .426$, $\eta_p^2 = .010$).[6]

ERPs

In past research with non-Black, largely White participants, P200s are consistently larger to Black than White faces whereas N200s are larger to White than Black faces (Dickter and Bartholow, 2007; Ito and Urland, 2003, 2005; Kubota and Ito, 2007; Willadsen-Jensen and Ito, 2006, 2008). Our interest, here, is whether this occurs in both task conditions, or whether the lack of implicit stereotyping in the feature detection task occurs because it decreases attention to race. This was tested in separate 2 (Task) × 2 (Dot Presence) × 2 (Prime Race) ANOVAs on P200 and N200 amplitude, with all factors within-subjects. Since race-sensitive P200 and N200 effects are typically largest at midline scalp sites (e.g. Ito and Urland, 2003), an additional three-level within-subjects factor of scalp site (Fz, Cz, Pz) was also included in both models.

---

5   There were a number of other significant effects in the omnibus response latency analysis but none were of theoretical interest. We list them for completeness: (i) Task main effect, $F(1,64) = 43.23$, $P < .0001$, $\eta_p^2 = .403$, (ii) Dot Presence main effect, $F(1,64) = 21.82$, $P < .0001$, $\eta_p^2 = .254$, (iii) Prime Race × Target Type interaction, $F(1,64) = 25.13$, $P < .0001$, $\eta_p^2 = .282$, (iv) Task × Prime Race interaction, $F(1,64) = 5.04$,

$P = .028$, $\eta_p^2 = .073$, (v) Task × Target Type × Dot Presence interaction, $F(1,64) = 12.27$, $P = .001$, $\eta_p^2 = .161$. These showed faster responses in the race categorization than feature detection task and following primes without dots in some conditions (in the race categorization condition when the targets were guns and in the dot detection task when the targets were insects), or were subsumed by the predicted Task × Prime Race × Target Type interaction reported in the main text.

6   Other effects not of theoretical interest in the omnibus accuracy analysis: (i) Task main effect, $F(1,64) = 4.42$, $P = .039$, $\eta_p^2 = .065$, (ii) Dot Presence main effect, $F(1,64) = 5.42$, $P = .023$, $\eta_p^2 = .078$, and the (iii) Task × Dot Presence interaction, $F(1,64) = 4.84$, $P = .031$, $\eta_p^2 = .070$. These showed greater accuracy when targets were preceded by primes with dots and in the race categorization task, but only when the primes lacked dots.
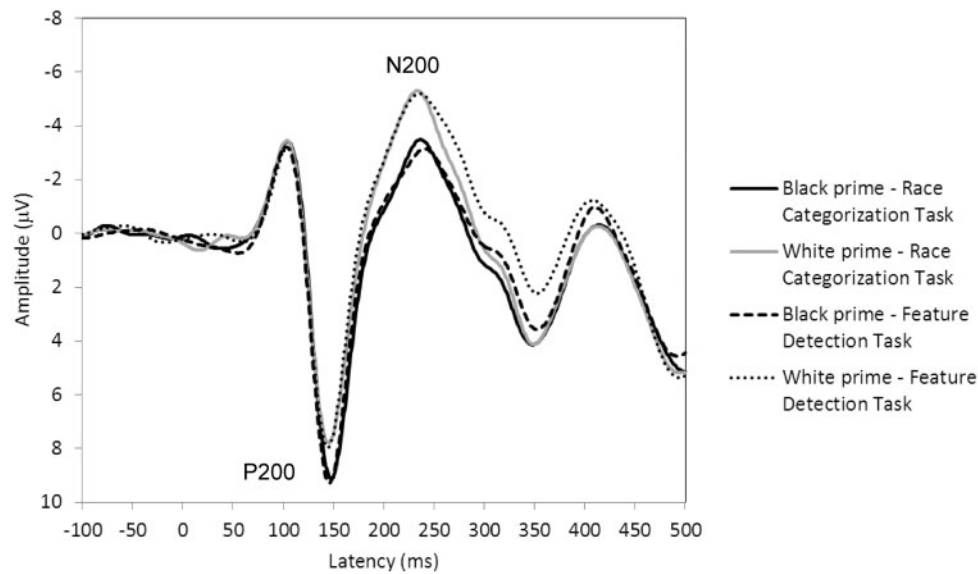
DISSOCIATING CATEGORIZATION AND STEREOTYPING 3

Fig. 3. Grand average ERP responses to Black and White face primes in the race categorization and feature detection tasks. Shown at electrode Cz, where the P200 and N200 were the largest.

Both analyses revealed main effects of Prime Race that replicated past research, as shown in Figure 3. P200s were larger to Black ($M = 9.32\,\mu V$, s.d. $= 3.44$) than White primes ($M = 8.25\,\mu V$, s.d. $= 3.47$), $F(1,29) = 39.92$, $P < .0001$, $\eta_p^2 = .579$. N200s were larger to White ($M = -6.29\,\mu V$, s.d. $= 2.91$) than Black primes ($M = -4.72\,\mu V$, s.d. $= 2.85$), $F(1,29) = 45.12$, $P < .0001$, $\eta_p^2 = .609$. Of critical importance, neither effect was moderated by task.[7]

Discussion

Results of the present study show that task goals have distinct effects on social categorization and stereotyping. Using measures that allow us to separately but simultaneously assess neural responses to faces of different races and the stereotypical associations activated by those faces, we find an equally high sensitivity to race when participants are explicitly categorizing faces in terms of race and when they are directed to perform a face-irrelevant feature detection task. In contrast, implicit

---

7 The following additional effects were significant in the omnibus analyses of the P200 and N200. In the P200 analysis, there was a significant Scalp Site main effect that took a quadratic form, with the largest P200s at Cz ($M = 9.63\,\mu V$, s.d. $= 4.10$) as compared to Fz ($M = 8.29\,\mu V$, s.d. $= 3.88$) and Pz ($M = 8.43\,\mu V$, s.d. $= 2.64$), $F(1,29) = 34.37$, $P < .0001$, $\eta_p^2 = .542$. The Prime Race main effect was also moderated by Scalp Site ($F(1,29) = 7.93$, $P = .014$, $\eta_p^2 = .215$). The simple effect of race, with larger P200s to Blacks than Whites, was significant at Fz, Cz, and Pz (Fs$(1,29) = 32.09$, 44.99, and 25.02, all ps $< .0001$, η_p^2s $= .525$, $.608$, $.463$, respectively), but the interaction indicates the magnitude of the race effect was smaller at Pz ($M = 0.78\,\mu V$, s.d. $= .85$) as compared to Fz ($M = 1.17\,\mu V$, s.d. $= 1.13$, $F(1,29) = 5.66$, $P = .024$, $\eta_p^2 = .163$) and Cz ($M = 1.25\,\mu V$, s.d. $= 1.02$, $F(1,29) = 23.11$, $P < .0001$, $\eta_p^2 = .443$). In the N200 analysis, there was also a main effect of scalp site that took a quadratic form, with the largest N200s at Cz ($M = -6.08\,\mu V$, s.d. $= 3.03$) as compared to Fz ($M = -5.59\,\mu V$, s.d. $= 2.97$) and Pz ($M = -4.83\,\mu V$, s.d. $= 2.96$), $F(1,29) = 23.88$, $P < .0001$, $\eta_p^2 = .452$. The main effect of Dot Presence was also significant, with larger N200s to pictures without dots ($M = -5.82\,\mu V$, s.d. $= 2.95$) than with dots ($M = -5.19\,\mu V$, s.d. $= 2.77$), $F(1,29) = 10.19$, $P < .003$, $\eta_p^2 = .250$.

stereotyping occurred only when participants focused on the racial category distinction. Other studies have demonstrated the effects of task instructions on implicit stereotyping (e.g. Jones and Fazio, 2010; Macrae *et al.*, 1997), but none included a separate measure of the way in which racial information was encoded. The present study, therefore, provides a unique demonstration of an on-line dissociation between the perceptual encoding of category information and the activation of category based semantic content, suggesting that knowing someone's race does not inevitably mean the relevant stereotype will be activated and/or applied.

This conclusion returns us to the issue that opened this article of how categorization and stereotyping are related. One possibility is that higher levels of input, in this case due to task goals, affect the categorization-stereotyping link. Lacking any top-down modulation signaling the need to inhibit categorization or stereotyping, mere exposure to category exemplars may directly trigger the activation of relevant stereotypes, as shown in the race categorization task where implicit stereotyping followed the racial modulation of neural responses. However, when other goals are activated, such as the goal to attend to a face-irrelevant dot stimulus, the direction of attention away from the semantic nature of the face stimuli may have allowed participants to still process its physical features to some degree (as reflected in racial modulation of neural responses) yet weakened the connection between such race processing and stereotypes associated with the group. While our design provides no direct evidence of the modulatory effects of task on the categorization-stereotyping relation, this account is consistent with models of person perception that highlight the role of so-called top-down forms of modulation (Freeman and Ambady, 2011) and also with other studies demonstrating the moderating role of task goals or contextual features on aspects of categorization and stereotyping (e.g. Krosch and Amodio, 2014; Macrae *et al.*, 1997; Senholzi and Ito, 2013; Van Bavel *et al.*, 2008; Wheeler and Fiske, 2005).

An important issue we were not able to address is how the feature detection task might affect the categorization-stereotyping link. Macrae *et al.* (1997) suggest the task may operate by

interfering with semantic appraisal of the individual. A cognitive load effect is also possible (cf. Gilbert and Hixon, 1991). Moreover, other manipulations may have similar modulatory effects via different mechanisms. Individuation, for instance, can similarly diminish stereotype activation but in a way that does not prevent semantic appraisal of the person under consideration. We thus expect that higher order modulation of categorization and stereotyping can occur via a number of different specific mechanisms.

We chose these particular measures because they have been used frequently to assess categorization (e.g. Ito and Urland, 2003) and stereotyping (e.g. Judd *et al.*, 2004). They also each have properties recommending their use for their respective purposes (e.g. ERPs have a fast temporal resolution and are sensitive to implicit processes; sequential priming tasks can assess the implicit activation of racial stereotypes). However, since the measures of categorization and stereotyping are different from each other, we cannot rule out the possibility that the nature of the tasks themselves contributed to the divergence in the effects obtained on them. The most problematic possibility is that ERPs are simply not sensitive to task manipulations and so would be incapable of detecting any task-related modulations of categorization. This seems unlikely as ERPs to faces of different races have shown sensitivity to task manipulations in past studies (e.g. Cunningham *et al.*, 2012; Senholzi and Ito, 2013) but task differences are nevertheless relevant to consider.

One unpredicted aspect of the results is the longer and less accurate responses in the feature detection than race categorization task (see notes 5 and 6). This could reflect a practice effect. While we varied task order for all the participants who had ERPs recorded, we inadvertently fixed the order for the participants who only contributed behavioral data, with the feature detection task performed first by all of the these participants. It seems likely that performance would get faster and more accurate over the course of the study. Thus, a practice effect combined with more participants doing the feature detection task first could explain why latencies were on average longer and accuracy was lower in the task. Another possibility is that the feature detection task took longer to perform than the race categorization task, regardless of the order in which it was completed, perhaps because dot location was not fixed, so performing the task could require scanning a large portion of the image on each trial. Neither difference is necessarily problematic for the issue of whether categorization and stereotyping are dissociable. Note that (i) task order was varied systematically for the ERP participants and (ii) if the feature detection task was more difficult it might be expected to interfere with the encoding of race-related information, thereby decreasing the likelihood of seeing race effects in the P200 and N200 in the feature detection task. Thus, if there was a task order effect, it was distributed equally across ERP participants, and if task difficulty had an effect, we would expect it to diminish neural sensitivity to race in the feature detection task, yet sensitivity to race occurred equally in both task conditions.

We nevertheless wanted to ensure that the lack of implicit stereotyping effects on behavior in the full sample in the feature detection task condition was due to a change in attentional focus, and not merely differences in task proficiency or attentional load. To assess this, we collected behavioral responses from an additional 60 participants with identical procedures save for systematically varying task order and increasing the SOA between the prime and target from 248 to 400 ms. We reasoned that if the feature detection task was inherently more difficult, giving participants longer to perform it before requiring attention to the target stimuli should reduce this load.

Slower responses in the feature detection task were eliminated in these Supplementary Data, but we still replicate the task moderating effects on implicit stereotyping. That is, participants were faster and more accurate to identify guns following Black than White primes in the race categorization task, whereas this difference was absent in the feature detection task. Full information on the Supplementary Data can be found in the Supplementary Materials. While we think these data are useful for ruling out the possibility that the modulation in implicit stereotyping obtained in the main study was an artifact of either a practice effect or differences in task difficulty, we stress again that the critical assessment of how perceivers' goals affected behavioral *vs* ERP responses should not have been affected.

In sum, results support a view of categorization and stereotyping as closely intertwined yet conceptually and functionally distinct. While this conclusion may seem to contradict theorizing that categorization necessarily leads to stereotyping, a careful reading of many of these theories indicates a more nuanced relation. For instance, Allport's statement that 'The category infuses all that it contains with the same ideational and emotional favor' (1954/1979, p. 21) seemingly implies that to categorize is to stereotype yet he later noted that while categorization and stereotyping are part of the same complex mental process, they are not identical; stereotyping only occurs if and when relevant category-based content is 'freighted' with the category (pp. 191–192). More recent treatments similarly conceptualize automatic categorization as distinguishable from the activation of category-based affect and cognitions (Brewer, 1988; Fiske and Neuberg, 1990; Freeman and Ambady, 2011; Kunda and Thagard, 1996), and the dissociation is also consistent with evidence of distinct brain systems subserving basic visual processes and the activation of semantic associations (e.g. Haxby *et al.*, 2000). Such theorizing and results from the present study highlight the importance of understanding at a mechanistic level how stereotype application can be prevented even when category information is encoded, which will ultimately benefit the design of maximally effective bias-reducing interventions.

Funding

Supplementary data

Supplementary data are available at *SCAN* online.

Conflict of interest. None declared.

Acknowledgements

References

Allport, G.W. (1954/1979). *The Nature of Prejudice*. Reading, MA: Perseus Books.

Bargh, J.A. (1984). Automatic and conscious processing of social information. In: Wyer, R.S., Jr., Srull, T.K., editors. *Handbook of Social Cognition*, Vol. **3**, Hillsdale, NJ: Erlbaum, pp. 1–43.

Bargh, J.A. (1999). The cognitive monster. The case against the controllability of automatic stereotype effects. In: Chaiken, S., Trope, Y., editors. *Dual-Process Theories in Social Psychology*. New York: The Guilford Press, pp. 361–382.

Blair, I.V., Banaji, M.R. (1996). Automatic and controlled processes in stereotype priming. *Journal of Personality and Social Psychology* **70**, 1142–63.

Brewer, M.C. (1988). A dual process model of impression formation. In: Wyer, R., Scrull, T., editors. *Advances in Social Cognition*, Vol. **1**, Hillsdale, NJ: Erlbaum, pp. 1–36.

Correll, J., Park, B., Judd, C.M., Wittenbrink, B. (2002). The police officer's dilemma: using ethnicity to disambiguate potentially threatening individuals. *Journal of Personality and Social Psychology* **83**, 1314–29.

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**, 120–8.

Cousineau, D. (2005). Confidence intervals in within-subjects designs: a simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology* **1**, 42–5.

Cunningham, W.A., Van Bavel, J.J., Arbuckle, N.A., Packer, D.J., Waggoner, A.S. (2012). Rapid social perception is flexible. Approach and avoidance motivational states shape P100 responses to other-race faces. *Frontiers in Human Neuroscience* **6**, 1–7.

Devine, P.G. (1989). Stereotypes and prejudice: their automatic and controlled components. *Journal of Personality and Social Psychology* **56**, 5–18.

Dickter, C.L., Bartholow, B.D. (2007). Event-related brain potential evidence of ingroup and outgroup attention biases. *Social, Cognitive, and Affective Neuroscience* **2**, 189–98.

Dovidio, J.F., Evans, N., Tyler, R.B. (1986). Racial stereotypes: the contents of their cognitive representations. *Journal of Experimental Social Psychology* **22**, 22–37.

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**, 1–73.

Freeman, J.B., Ambady, N., (2011). A dynamic interactive theory of person construal. *Psychological Review* **118**, 247–79.

Gilbert, D.T., Hixon, J.G. (1991). The trouble of thinking: activation and application of stereotypic beliefs. *Journal of Personality and Social Psychology* **60**, 509–17.

Haxby, J.V., Hoffman, E.Z., Gobbini, M.I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences* **4**, 223–33.

Ito, T.A., Bartholow, B.D. (2009). The neural correlates of race. *Trends in Cognitive Sciences* **13**, 524–31.

Ito, T.A., Urland, G.R. (2003). Race and gender on the brain: electrocortical measures of attention to race and gender of multiply categorizable individuals. *Journal of Personality and Social Psychology* **85**, 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**, 21–36.

Jasper, H.H. (1958). The ten twenty electrode system of the international federation. *Electroencephalography and Clinical Neurophysiology* **10**, 371–5.

Jones, C.R., Fazio, R.H. (2010). Person categorization and automatic racial stereotyping effects on weapon identification. *Personality and Social Psychology Bulletin* **36**, 1073–85.

Judd, C.M., Blair, I.V., Chapleau, K.M. (2004). Automatic stereotypes versus automatic prejudice: Sorting out the possibilities in the Payne (2001) weapon paradigm. *Journal of Experimental Social Psychology* **40**, 75–81.

Krosch, A.K., Amodio, D.M. (2014). Economic scarcity alters the perception of race. *Proceedings of the National Academy of Science of the United States of America* **111**, 9079–84.

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**, 738–52.

Kubota, J.T., Ito, T.A. (2016). Rapid race perception despite individuation and accuracy goals. *Social Neuroscience*. doi:10.1080/17470919.2016.1182585.

Kunda, Z., Thagard, P. (1996). Forming impressions from stereotypes, traits, and behaviors: a parallel-constraint-satisfaction theory. *Psychological Review* **103**, 284–308.

Macrae, C.N., Bodenhausen, G.V., Milne, A.B., Thorn, T.M.J., Castelli, L. (1997). On the activation of social stereotypes: the moderating role of processing objectives. *Journal of Experimental Social Psychology* **33**, 471–89.

Martin, D., Macrae, C.N. (2007). A face with a cue: exploring the inevitability of person categorization. *European Journal of Social Psychology* **37**, 806–16.

Mason, M.F., Cloutier, J., Macrae, C.N. (2006). On construing others: category and stereotype activation from facial cues. *Social Cognition* **24**, 540–62.

Payne, B.K. (2001). Prejudice and perception: the role of automatic and controlled processes in misperceiving a weapon. *Journal of Personality and Social Psychology* **81**, 181–92.

Semlitsch, H.V., Anderer, P., Schuster, P., Presslich, O. (1986). A solution for reliable and valid reduction of ocular artifacts, applied to the P300 ERP. *Psychophysiology* **23**, 695–703.

Senholzi, K.B., Ito, T.A. (2013). Structural face encoding: how task affects the N170's sensitivity to race. *Social Cognitive and Affective Neuroscience* **8**, 937–42. doi:10.1093/scan/nss091.

Stangor, C., Lynch, L., Duan, C., Glas, B. (1992). Categorization of individual on the basis of multiple social features. *Journal of Personality and Social Psychology* **62**, 207–18. doi:10.1037/0022-3514.62.2.207.

Tomelleri, S., Castelli, L. (2012). On the nature of gender categorization: pervasive but flexible. *Social Psychology* **43**, 14–27.

Turner, J.C. (1987). *Rediscovering the Social Group: A Self-Categorization Theory*. New York, NY, Basil Blackwell.

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* **11**, 1131–9.

Wheeler, M.E., Fiske, S.T. (2005). Controlling racial prejudice: social-cognitive goals affect amygdala and stereotype activation. *Psychological Science* **16**, 56–63.

Willadsen-Jensen, E.C., Ito, T.A. (2006). Ambiguity and the time-course of racial perception. *Social Cognition* **24**, 580–606.

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 Interpersonal Relations* **11**, 182–200.

Willadsen-Jensen, E.C., Ito, T.A. (2015). The effect of context on responses to racially-ambiguous faces: changes in perception and evaluation. *Social Cognitive and Affective Neuroscience* **10**(7),885–92. doi:10.1093/scan/nsu134.

Wittenbrink, B., Judd, C.M., Park, B. (2001). Spontaneous prejudice in context: variability in automatically activated attitudes. *Journal of Personality and Social Psychology* **81**, 815–27.