

The Neural Basis of Contextual Influences on Face Categorization

Jonathan B. Freeman¹, Yina Ma¹, Maria Barth², Steven G. Young³, Shihui Han⁴ and Nalini Ambady²

¹Department of Psychological and Brain Sciences, Dartmouth College, Hanover, NH 03755, USA, ²Department of Psychology, Stanford University, Stanford, CA 94305, USA, ³School of Psychology, Fairleigh Dickinson University, Teaneck, NJ, 07666, USA and ⁴Department of Psychology, Peking University, Beijing, 100871, China

Address Correspondence to Jonathan Freeman, Dartmouth College, Department of Psychological and Brain Sciences, 6207 Moore Hall, Hanover, NH 03755, USA. Email: jon.freeman@dartmouth.edu

From only brief exposure to a face, individuals spontaneously categorize another's race. Recent behavioral evidence suggests that visual context may affect such categorizations. We used fMRI to examine the neural basis of contextual influences on the race categorization of faces. Participants categorized the race of faces that varied along a White-Asian morph continuum and were surrounded by American, neutral, or Chinese scene contexts. As expected, the context systematically influenced categorization responses and their efficiency (response times). Neuroimaging results indicated that the retrosplenial cortex (RSC) and orbitofrontal cortex (OFC) exhibited highly sensitive, graded responses to the compatibility of facial and contextual cues. These regions showed linearly increasing responses as a face became more White when in an American context, and linearly increasing responses as a face became more Asian when in a Chinese context. Further, RSC activity partially mediated the effect of this face-context compatibility on the efficiency of categorization responses. Together, the findings suggest a critical role of the RSC and OFC in driving contextual influences on face categorization, and highlight the impact of extraneous cues beyond the face in categorizing other people.

Keywords: context, face processing, fMRI, scene processing, social categorization

Introduction

On catching sight of another person, we rapidly glean a variety of information from his or her face. Among the possible perceptions that may be made, most important perhaps are social categories such as race or gender. Once perceived, these categories often alter subsequent person processing and lead to divergent cognitive, affective, and behavioral outcomes (Macrae and Bodenhausen 2000; Freeman and Ambady 2011). Although commonly presented as isolated stimuli in the laboratory, faces are rarely encountered in isolation in the real world. Instead, they are typically embedded in rich contexts, with cues inherent to the target person (e.g., bodily cues) contextualizing the face, as well as the larger scene and environment. For example, we might catch sight of another person dining in a crowded restaurant or strolling down the street.

A number of studies have shown that contextual cues can bear substantial effects on focal stimuli, such as faces or objects. For example, it has long been known that objects presented in a congruent relative to incongruent visual scene are processed more efficiently (Biederman et al. 1982). In face perception research, contextual effects have mainly been investigated with respect to facial emotion. The emotional characteristics of bodily or scene cues surrounding a face are often found to influence the perception of the face's emotion,

particularly when the emotion is ambiguous (e.g., Russell et al. 2003; Aviezer et al. 2008; Masuda et al. 2008; Righart and De Gelder 2008; Barrett and Kensinger 2010).

More recently, contextual effects have been examined with respect to a face's race. Rather than the race categorization process being guided by a mere "read out" of facial features, it appears to rapidly integrate a variety of information sources that include contextual cues. Contexts such as attire or hair cues that immediately surround a face and are stereotypically associated with particular race categories influence both the outcomes and the process of categorizing a face's race (MacLin and Malpass 2001; Freeman et al. 2011). Even facial cues that are not directly related to race can also serve as a "context" that constrains its perception, such as facial emotion or gender (Hugenberg and Bodenhausen 2004; Johnson et al. 2012). Social category encoding may also be dynamically shaped by high-level social contexts, such as one's culture or the assignment to novel groups (Chiao et al. 2008; Van Bavel et al. 2008). Such findings are consistent with the longstanding claim that race is socially constructed and understood through a variety of social factors (e.g., Castano et al. 2002; Eberhardt et al. 2003). Recent neural network models suggest that race categorization involves a competition process in which both facial and contextual cues dynamically weigh in, thereby mutually constraining one another until a stable categorization is achieved (Freeman and Ambady 2011). Thus, contextual cues may trigger expectations that exert a top-down influence on face processing early on, while perceptual processing is still ongoing. Such an account is consistent with more general evidence for early interactions between object and context processing and top-down facilitation of visual object recognition (for review, see Humphreys et al. 1997).

Functional magnetic resonance imaging (fMRI) studies over the past decade have consistently identified a network of regions that are involved in associating focal stimuli with the surrounding visual context (Bar 2004; Epstein et al. 2007; Epstein 2008). These regions are commonly activated by contextual associations, including the parahippocampal cortex (PHC), the retrosplenial cortex (RSC), and the orbitofrontal cortex (OFC). Bar (2004) has argued that, early in processing, coarse low-spatial-frequency information about an object is projected from early visual areas to the OFC. Tentative hypotheses about that object available in the OFC are then fed back to the temporal cortex to facilitate likely interpretations of the object. Beyond the hypotheses triggered by the visual object information itself, there are also influences triggered in top-down fashion by the PHC and RSC, which are thought to analyze contextual associations between the object and surrounding scene context. These regions influence the tentative

hypotheses activated by the object information, by facilitating those that are most likely to appear in the given scene context (thus utilizing top-down expectations).

From this perspective, the OFC is proposed to serve as an integrative center of various information sources (e.g., early visual object areas, PHC, and RSC) that continuously updates a representation of contextual information. It then uses this information to provide top-down modulation of the object recognition process (Bar 2004). As for the PHC and RSC, both have been implicated in processing spatial and location-based information (e.g., Henderson et al. 2008). The PHC, in particular, and the parahippocampal place area (PPA) it contains show selective responses to images of scenes and landmarks (Epstein and Kanwisher 1998). Although these 2 regions have often shown similar responding in fMRI studies of context processing, recent work has found evidence for their dissociable roles. The PHC, for example, tends to host more allocentric representations of the context, whereas the RSC hosts more egocentric representations (Epstein et al. 2007). A number of studies have suggested that the PHC may have a greater involvement in processing the context's perceptual details, and these then feed into a higher level representation of the context available in the RSC. Given this, representations in the PHC are thought to be more perceptual, whereas representations in the RSC are thought to be more semantic and integrated with higher order information. Thus, the RSC appears to host generic representations of the context that are not fettered to particular physical information (e.g., Epstein 2008).

Although the neural basis of contextual influences on object perception has been studied extensively, the neural basis of contextual influences on face perception has remained less clear. Given that the neural mechanisms underlying face and object perception are dissociable (Haxby et al. 2000; Kanwisher and Yovel 2006), it is possible that separate mechanisms may also mediate their ability to integrate contextual cues. In the present work, therefore, we sought to investigate the neural mechanisms underlying the influence of scene contexts on face perception. Previous research suggests that the encoding of social categories such as race involves inherently graded representations of category-specifying facial information (Blair et al. 2002; Locke et al. 2005; Freeman et al. 2008; Freeman and Ambady 2011), which is evident at the neural level in a number of regions, including the OFC (Ronquillo et al. 2007; Freeman et al. 2010). The current study was designed in such a way as to allow us to additionally assess whether such graded social category representations may be systematically altered by the visual context.

We were also interested here in how expectations that are specifically social may constrain processing of a focal stimulus, such as a face. Prior studies examining compatibility between scenes and objects have generally used full-fledged semantic violations (e.g., Gronau et al. 2008). In the social domain, however, expectations may be semantically associated with a given social category but be stereotype-based and therefore only suggestive of that category; thus, incompatibilities often do not wholesale violate semantic principles. For example, we may have the stereotypic expectation that a dancer on a ballet stage would be female, or that an individual on the streets of a Chinese city would be Asian (which, of course, as stereotypes are in fact readily violated). These social expectations are distinct and more subtle than, for instance, the expectation that a basketball (rather than a watermelon) would be shot into a

basketball hoop. These latter kinds of full-fledged semantic violations have long been used to examine contextual scene influences on object processing (e.g., Mudrik et al. 2010). Here, we sought to examine whether the influences of social category-based expectations triggered by a scene context are mediated through the more generic context network seen in prior work, including the RSC, PHC, and OFC, which may constrain face categorization.

We were also interested in whether this network of regions may be sensitive to a perceiver's cultural background, given that East Asian individuals more readily incorporate context information than Western individuals (Nisbett et al. 2001). Specifically, previous studies have found that East Asian individuals allocate more attention to the inter-relationships between a focal stimulus and its visual context than Western individuals, presumably due to East Asian collectivist tendencies (Masuda and Nisbett 2001; Nisbett et al. 2001). Thus, we sought to examine the neural basis of contextual influences across both an East Asian and Western culture.

In the present study, blood oxygenation level-dependent (BOLD) responses were measured with fMRI while American and Chinese participants categorized the race of faces varying along a White-Asian morph continuum and embedded in either American, neutral, or Chinese scene contexts. Behaviorally, we predicted that the scene context would influence face-categorization responses and their efficiency. Categorization responses should be biased to be congruent with the context, and response times should become increasingly fast as the compatibility between face and context increases. At the neural level, we predicted that the RSC, PHC, and OFC would be sensitive to this compatibility between face and context, and that one or several of these regions may partly account for the behavioral effects on face categorization.

Materials and Methods

Participants

Seventeen Caucasian American volunteers (8 males) were scanned at the Harvard Center for Brain Science (Cambridge, MA, USA) and 15 native Chinese volunteers (7 males) were scanned at the Beijing MRI Center for Brain Research. Excessive artifacts required us to drop 1 Chinese participant, leaving 14 for analysis. All participants were right-handed, had normal or corrected-to-normal vision, gave informed consent consistent with the local Institution Review Board, and were paid for participation.

Materials

We used FaceGen Modeler (Singular Inversions) to generate realistic faces that were morphed by race, based on anthropometric parameters of human population (Blanz and Vetter 1999). Face stimuli were comprised of 16 computer-generated face identities (8 males) that were morphed along a 9-point continuum, from Asian (morph -4) to White (morph 4). Morphing allowed us to precisely manipulate race-related cues and unconfound other perceptual information. A total of 90 scene context stimuli (30 American, 30 neutral, 30 Chinese) were obtained from public-domain websites. Examples included scenes such as a traditionally American restaurant (American context), a nondescript grassy field (neutral context), or a traditionally Chinese house (Chinese context). Each face was presented 3 times and placed in the center location of a scene stimulus (once in an American context, once in a neutral context, and once in a Chinese context). Sample stimuli are shown in Figure 1. Face-scene pairings were counterbalanced across participants, yielding a total of 432 trials per participant: 16 face identities \times 9 race levels \times 3 contexts. Importantly, this design allowed us

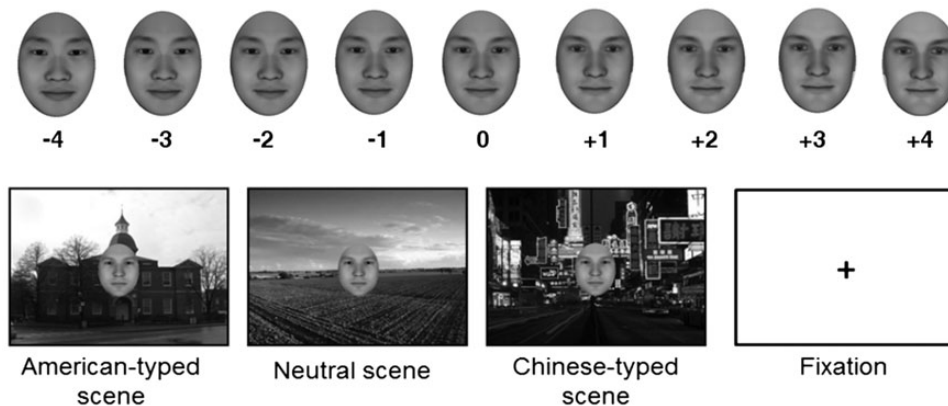


Figure 1. At the top is a sample morph continuum, from Asian (−4) to White (+4). On the bottom is a sample American-typed, neutral, and Chinese-typed scene context (with face stimulus at center), or fixation cross (baseline).

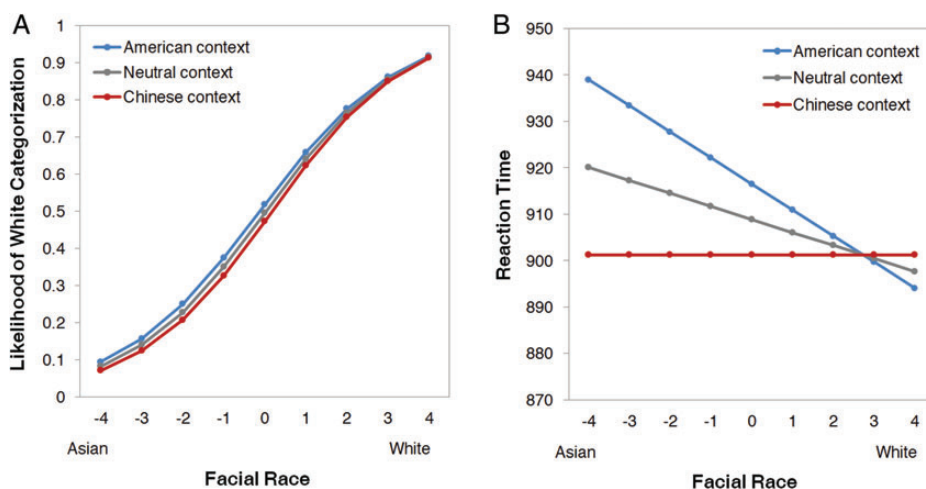


Figure 2. Predicted values for the likelihood of a White categorization (A) and categorization response time (B) are plotted as a function of facial race morph level, separately for American, neutral, and Chinese contexts. Both categorization responses and their efficiency were systematically influenced by the context.

to control for any potential differences in the visual properties of the American, neutral, and Chinese scene stimuli. This is because, across participants, each facial-morph level was paired with the same scene stimuli. Thus, comparisons of neural responses to various face-scene combinations could not reflect inherent differences in the scene stimuli, but instead reflect only the compatibility of the face and scene. These same face and scene stimuli were used in a previous behavioral study (Freeman et al. 2013).

Experimental Procedure

The 432 face stimuli and an additional 144 baseline trials (fixation cross) were viewed each for 2 s in 1 of 2 pseudorandomized orders, each sequenced in a manner as to maximize the efficiency of event-related BOLD signal estimation (Dale 1999). The entire task was divided into 2 functional runs. Participants were told that they would be presented with images of individuals in various settings, and that they would be asked to categorize the race of individuals' faces as quickly and accurately as possible using a button press.

fMRI Acquisition

The same scanner model was used across the 2 study sites (Siemens 3 T Tim Trio), and the imaging parameters were identical. Anatomical images were acquired using a T_1 -weighted protocol (256×256 matrix, 128 1.33-mm sagittal slices). Functional images were acquired using a single-shot gradient echo-planar imaging sequence (TR = 2000 ms,

TE = 30 ms). Thirty-two interleaved oblique-axial slices ($3 \times 3 \times 3$ mm voxels; slice gap = 0.54 mm) parallel to the AC-PC line were obtained. Analysis of the imaging data was conducted using BrainVoyagerQX (Brain Innovation, Maastricht, Netherlands). Functional imaging data preprocessing included 3D motion correction, slice timing correction (sinc interpolation), spatial smoothing using a 3D Gaussian filter (8-mm FWHM), and voxelwise linear detrending and high-pass filtering of frequencies (above 3 cycles per time-course). Structural and functional data of each participant were transformed to standard Talairach stereotaxic space.

fMRI Analysis

Original morph values of face stimuli (−4 = Asian, 4 White) were rescaled (0 = Asian, 1 = White): [−4, −3, −2, −1, 0, 1, 2, 3, 4] to [0, 0.125, 0.25, 0.375, 0.5, 0.625, 0.75, 0.875, 1]. Individual participants' BOLD signals were modeled in an event-related design using a general linear model (GLM) with 3 dichotomous predictors denoting a face in either an American, neutral, or Chinese context, and 3 parametric predictors denoting facial race (using rescaled morph values) separately for American, neutral, and Chinese contexts. The coding of the most Asian face as 0 and most White face as 1 was arbitrary. To further explore the fMRI effects, we constructed another GLM design matrix which included a total of 27 predictors used to separate out BOLD signals associated with each of the 9 levels of race \times 3 contexts. Conditions in all GLMs were modeled as boxcar functions (for parametric predictors, the amplitude of which was parametrically varied) and convolved with

a 2-gamma hemodynamic response function. First-level GLM analyses conducted on individual participants' fMRI signal were submitted to a second-level random effects analysis, treating participants as a random factor. Statistical maps depicting BOLD responses were overlaid onto an average image of all participants' corresponding structural data.

In all regression analyses involving behavioral data, we adopted a multilevel generalized estimating equation approach that can incorporate trial-by-trial data while accounting for the intracorrelations in repeated-measures designs (Zeger and Liang 1986). In the present study, this included intracorrelations associated with individual subjects and with individual stimulus tokens (face identities). We report unstandardized regression coefficients. To control for multiple statistical testing of voxels within the entire brain, we maintained a false positive detection rate of $P < 0.01$ by using a voxelwise threshold of $P < 0.001$ and a minimum cluster-size extent of 300 mm^3 . The minimum cluster-size threshold needed to maintain an experiment-wide α of 0.01 was empirically determined by a Monte Carlo simulation accounting for the spatial correlations between neighboring voxels (Forman et al. 1995).

Results

Participant's culture was initially included as a factor in analyses, but all effects were nonsignificant (P 's > 0.1). Thus, participant culture is discussed no further.

Behavioral

First, to examine the relationship between face race and scene context on race categorization, we regressed categorization responses (0 = Asian, 1 = White) onto facial-morph level (-0.5 most prototypically Asian, 0.5 most prototypically White), context (-0.5 = Chinese, 0 = neutral, 0.5 = American), and the interaction (using logistic regression). This revealed a significant main effect of facial morph, confirming our morphing manipulation. As a face became more White, the likelihood of a White categorization linearly increased, $B = 4.62$, $Z = 24.89$, $P < 0.0001$. Importantly, this also revealed a significant main effect of context, with an American context increasing the likelihood of a White categorization (simple slope, $P < 0.01$) and a Chinese context increasing the likelihood of an Asian categorization (simple slope, $P < 0.0001$), relative to a neutral context, $B = 0.18$, $Z = 4.31$, $P < 0.0001$ (Fig. 2B). The interaction was not significant, $B = -0.15$, $Z = 0.95$, $P = 0.34$. These results are consistent with previous behavioral work (Freeman et al. 2013).

We also regressed categorization response times onto the same variables (using normal regression). This revealed a significant main effect of facial morph, with response times quickened as a face became more White, $B = -22.46$, $Z = 2.20$, $P < 0.05$. There was also a significant main effect of context, with responses times overall quicker for faces paired with Chinese contexts, $B = 15.26$, $Z = 2.91$, $P < 0.01$. More importantly, there was a significant interaction, $B = -44.91$, $Z = 2.81$, $P < 0.01$. Depicted in Figure 2B, as a face became more White in an American context or became more Asian in a Chinese context, response times linearly decreased. Conversely, as a face became more White in a Chinese context or became more Asian in an American context, response times linearly increased. The simple slopes indicated that the interactive effect was more robust for Chinese contexts (simple slope, $P = 0.001$) than American contexts (simple slope, $P = 0.11$). Overall, race categorizations were highly sensitive to the compatibility between facial and contextual cues.

The behavioral results show that the scene context surrounding a face systematically influenced categorization responses as well as the efficiency of categorization. In our

neuroimaging analyses, we examined the neural basis of these contextual influences.

Neuroimaging

First, we sought to identify any regions that may have shown a parametric effect of facial race, regardless of context, including a linearly increasing or decreasing response as race became more White or Asian. To that end, we conducted a whole-brain analysis contrasting an aggregate of all facial race predictors (in American, neutral, and Chinese contexts) against baseline ($P < 0.01$, corrected). Due to the arbitrary coding of the most Asian face as 0 and most White face as 1, positive beta values indicate linearly increasing responses as a face became more White, and negative beta values indicate linearly increasing responses as a face became more Asian. Interestingly, this analysis revealed activation in a region of the RSC that was negatively associated with the facial race predictors (Tables 1 and 2), indicating linearly increasing responses as a face became more Asian, independent of context. Thus, this region was particularly sensitive overall to the presence of Asian-related cues.

More importantly, to identify which regions may have shown a parametric effect of facial race differentially by context (conceptually, a facial race \times context interaction), we conducted a whole-brain analysis of variance (ANOVA) that included a single within-subject factor comprising the 3 parametric predictors (facial race in American, neutral, and Chinese contexts). This revealed robust activity in the RSC and OFC, $P < 0.01$, corrected (Tables 1 and 2). As seen in Figure 3, whereas these regions exhibited positive beta values in American contexts (linear increases as a face became more White), they exhibited negative beta values in Chinese contexts (linear increases as a face became more Asian), with beta values hovering around 0 for faces in neutral contexts (indicating a lack of correlation with facial race).

To better specify the pattern of activation in these regions, we constructed another GLM design matrix that comprised a total of 27 predictors used to separate out BOLD signals associated with each of the 9 levels of race \times 3 contexts. Beta values were then extracted from all voxels within the RSC and OFC regions elicited by the whole-brain ANOVA. As such, these beta values reflect the mean BOLD response during the presentation of each 1 of the 27 specific face-context combinations (rather than reflecting a parametric correlation, as in the pre-

Table 1

Regions of activation elicited by the whole-brain contrast of 3 facial-race effects (across American, neutral, and Chinese contexts) $>$ baseline: FDR, $P < 0.01$, corrected

Region	x	y	z	Mean t	mm^3
Retrosplenial cortex	0	-68	0	-4.21	1992

Table 2

Regions of activation elicited by the whole-brain ANOVA testing differential activation by the 3 facial-race effects (across American, neutral, and Chinese contexts): $P < 0.01$, corrected

Region	x	y	z	Mean F	mm^3
Retrosplenial cortex	-8	-66	2	8.99	1409
Orbitofrontal cortex	9	36	0	8.87	1028
Lingual gyrus	11	-76	-2	8.39	703
Medial frontal gyrus	-5	-20	65	8.44	468

vious analysis). Analyses on these extracted beta values are for descriptive purposes and for better specifying the pattern of parametric correlations in these regions. Using the same coding scheme as in the behavioral regression analyses, these beta values were then regressed onto facial-morph level, context, and the interaction. In the RSC, the main effect of morph level was not significant, $B = 0.01$, $Z = 0.10$, $P = 0.92$. Interestingly, the main effect of context, however, was significant, with overall stronger responses for Chinese contexts,

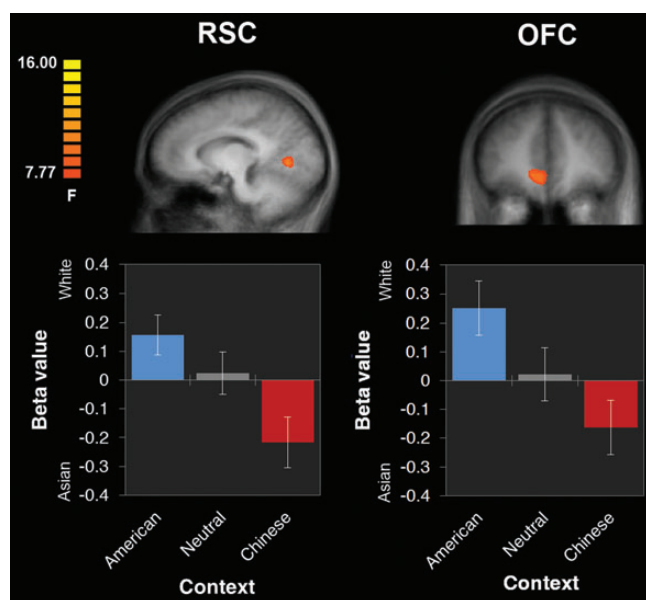


Figure 3. BOLD responses in the RSC and OFC elicited by a whole-brain analysis of variance, testing differential activation by 3 parametric predictors for facial race (arbitrarily coded such that the most Asian face = 0 and most White face = 1): 1 for American contexts, 1 for neutral contexts, and 1 for Chinese contexts. Due to the coding of the parametric predictors, positive beta values indicate increasing responses as a face became more White, and negative beta values indicate increasing responses as a face became more Asian. Thus, the RSC and OFC showed increasing responses as a face became more White in an American context, but increasing responses as a face became more Asian in a Chinese context.

$B = -0.11$, $Z = 2.41$, $P < 0.05$. More importantly, there was a significant interaction, $B = 0.42$, $Z = 3.69$, $P < 0.001$. As shown in Figure 4, the RSC showed linearly increasing responses as facial and contextual cues became more compatible. For American contexts, RSC activation increased as a face became more White (simple slope, $P < 0.0001$); for Chinese contexts, RSC activation increased as a face became more Asian (simple slope, $P < 0.01$). In the OFC, neither the main effects of morph level ($B = 0.07$, $Z = 1.22$, $P = 0.22$) nor context ($B = -0.02$, $Z = 0.38$, $P = 0.71$) were significant. More importantly, the interaction was significant, $B = 0.40$, $Z = 2.73$, $P < 0.01$. Shown in Figure 4, OFC activation increased as a face in an American context became more White (simple slope, $P < 0.01$) and as a face in a Chinese context became more Asian (simple slope, $P < 0.05$). Note that all OFC responses were negative; this is consistent with the role of ventral aspects of the medial prefrontal cortex in the default network, which is typically more active during baseline comparison conditions (e.g., Mason et al. 2007).

We also sought to test whether RSC or OFC responses may have influenced response times and had a role in driving the efficiency of face categorization. Separately regressing categorization response times onto the extracted RSC and OFC beta values (indexing mean activation to the specific face-context combinations) indicated that stronger RSC responses were associated with quicker response times, $B = -13.31$, $Z = 2.47$, $P = 0.01$. OFC responses, on the other hand, were not significantly associated with response times, $B = -7.89$, $Z = 1.57$, $P = 0.12$. This result suggests that the RSC may have played an important role in the contextual modulation of face-categorization response times.

To explore this possibility further, we constructed a standard mediational model to test whether RSC activation may have mediated the influence of the facial race \times context interaction on response times (Baron and Kenny 1986). However, given that the design included multilevel within-subject factors, we used hierarchical linear modeling and Monte Carlo simulations to estimate the indirect effect (Preacher and Selig 2012). Unstandardized regression coefficients and significance levels for paths in the model are depicted in Figure 5. As

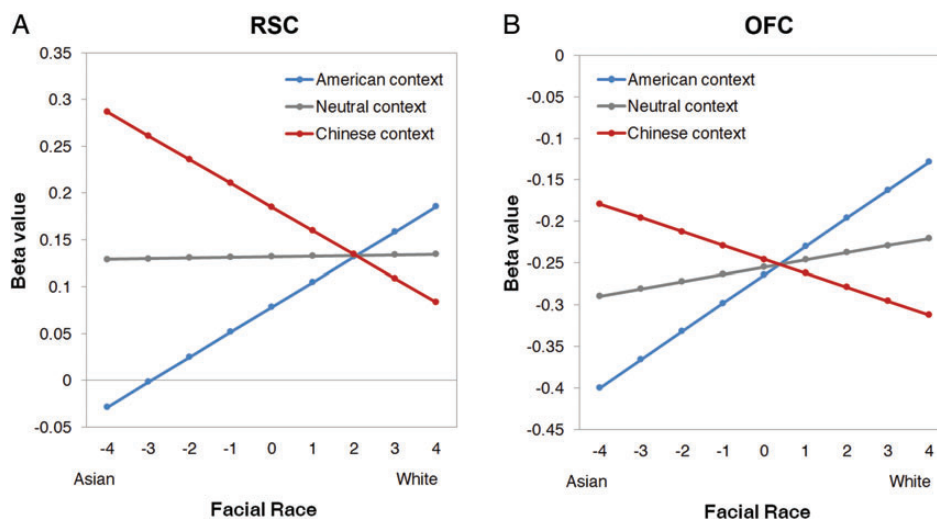


Figure 4. Predicted values for the magnitude of response (beta values) in the RSC and OFC are plotted as a function of each level of facial race, separately for American, neutral, and Chinese contexts.

expected given the previous analyses, the facial race \times context effect was a significant predictor of response times, $B = -43.33$, $SE = 12.85$, $P = 0.001$, as well as a significant predictor of RSC beta values, $B = 0.42$, $Z = 0.11$, $P < 0.001$. RSC beta values were also a significant predictor of response times (when including the interaction effect as well), $B = -11.24$, $SE = 5.44$, $P < 0.05$. Importantly, the inclusion of RSC beta values into the regression predicting response times from the facial race \times context interaction led the interaction effect to drop in size, $B = -38.65$, $SE = 13.99$, $P < 0.01$, suggesting partial mediation. Indeed, a Monte Carlo simulation estimated the 95% confidence interval of the indirect effect $[-10.72, -0.22]$, which excluded zero and thereby indicated the indirect effect was significant at the $P < 0.05$ level. Thus, RSC activation partially mediated the influence of face-context compatibility on categorization response times (see Fig. 5).

Discussion

The present study explored the influence of visual context on face categorization at the behavioral and neural levels. As predicted, the compatibility of facial and contextual cues strongly influenced the efficiency of race categorization. As a face became more White in an American context or more Asian in a Chinese context (relative to neutral), response times became correspondingly faster. As the cues became more incompatible (relative to neutral), response times became correspondingly slower. Thus, consistent with prior work (Freeman et al. 2013), race-relevant scene information was integrated into categorizations of facial race, leading to both facilitation (in cases of compatibility) and competition (in cases of incompatibility). More importantly, such contextual integration was reflected in activity in the RSC and OFC. These regions showed parametrically increasing responses as facial and contextual cues became more compatible and decreasing responses as the cues became more incompatible. Finally, RSC activity was found to partially mediate the effect of face-context compatibility on categorization response times, suggesting that the RSC had an important role in driving the contextual modulation of face categorization.

Over the past decade, the neural basis of contextual influences on object perception has been investigated thoroughly. These studies have pointed to important roles of the PHC, RSC, and OFC in mediating contextual associations. For instance, these regions show increased activation when participants are

presented with objects that contain strong relative to weak contextual associations (for review, see Bar 2004). The present results suggest that the RSC and OFC process contextual associations not only in object perception but also face perception, specifically in perceiving facial categories such as race. Thus, the present results extend our knowledge of these regions' participation in context processing, by demonstrating their more generalized involvement in multiple classes of perceptual stimuli (objects and faces).

The nature of the contextual associations that we show here are processed by the RSC and OFC is also noteworthy. Prior studies have shown these regions to be sensitive to wholesale semantic violations, such as cases where a focal stimulus is embedded in an extremely improbable context, for example, a watermelon being thrown into a basketball hoop (Gronau et al. 2008; Mudrik et al. 2010). Here, the relationships between focal face stimuli and scene contexts were never wholesale violated; they were only more or less congruent with semantically associated stereotypes: the typical co-occurrence between American environments and White individuals, and Chinese environments and Asian individuals. The present results suggest that the influences of social category-based expectations triggered by a scene context are mediated by a more general context network identified in prior work, which is able to systematically modulate face categorizations. This also suggests that social context effects may be implemented by relatively domain-general mechanisms.

A novel finding from this work is that the RSC and OFC exhibited highly sensitive, linear relationships with the compatibility of facial and contextual cues. These regions showed linearly increasing responses as facial and contextual cues become more compatible, and decreasing responses as they become more incompatible. Prior fMRI studies on the context network have generally used objects or scenes in isolation; those that have used them in tandem have used dichotomously congruent or incongruent conditions. For example, the RSC has been shown to exhibit stronger responses in the period prior to recognizing a target object when preceded by a congruent relative to incongruent prime (Eger et al. 2007). Such results are consistent with increased RSC activation with increases in the congruency of facial and contextual cues. In the present work, we were able to demonstrate the RSC and OFC's fine-grained, linear sensitivity to these cues' congruency by using continuous morphing where the congruency between face and context could be parametrically varied. This provides an important extension of previous work, by demonstrating the graded nature of these regions' representation of interacting facial and contextual cues.

The results of a recent behavioral study using the same stimuli as those of the present study may provide some clues as to the computations being conducted in the RSC and OFC. Participants' computer-mouse movements were measured in real time as they categorized the face-context pairs (starting at the bottom-center of the screen en route to "White" and "Asian" category responses at the 2 top corners). As the compatibility between facial and contextual cues increased, mouse trajectories exhibited an increasingly more direct, facilitated approach toward the selected category. As the cues became more incompatible, trajectories exhibited an increasing partial attraction toward the unselected category (on the opposite side of the screen) before settling onto the ultimately selected category, relative to neutral contexts (Freeman et al. 2013). These

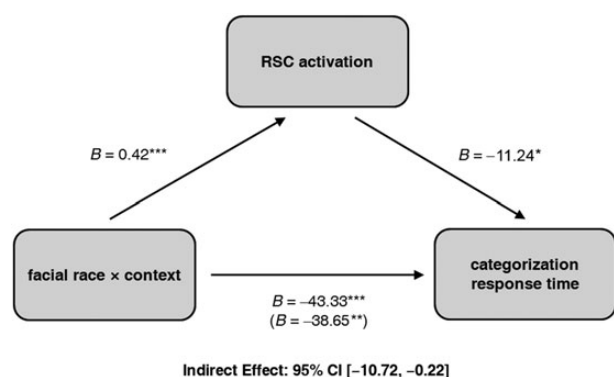


Figure 5. Paths in a mediation model, showing that RSC activity partially mediated the relationship between face-context compatibility and categorization response times.

results suggested, consistent with recent connectionist models, that a category competition process underlies face categorizations (e.g., White vs. Asian); during this process, both bottom-up facial cues and top-down expectations triggered by the scene context weigh in on the competition and mutually constrain one another until “compromising” over time (Freeman and Ambady 2011). Given that in the present study the RSC and OFC activity exhibited a similar graded, fine-grained response to the compatibility of facial and contextual cues, it is possible that the regions may be neural correlates of this competition process theorized to integrate facial and contextual information. Future work could utilize both methodologies together to specifically address this issue.

It is worth noting that the present study involved both American and native Chinese participants, and one might have expected the Chinese participants to utilize contextual cues to a greater extent than American participants (Nisbett et al. 2001). In contrast, cultural differences were not obtained in either behavioral or neural results. This can likely be explained by the results of the mouse-tracking study described above, which also involved American and Chinese participants. In that study, no cultural differences were found in terms of the magnitude of contextual influences, measured by ultimate categorization responses and mouse trajectories; instead, the differences were more subtle. Assessed through time-course analyses of the mouse trajectories, the researchers found that contextual influences began earlier in the categorization time-course for Chinese relative to American participants (Freeman et al. 2013). Thus, contextual cues were integrated into face categorizations slightly earlier in time in Chinese participants, but the 2 cultures did not differ in the overall degree to which contextual cues were utilized. It is therefore not surprising that, in the present study, we do not find overall differences in categorization responses or neural activity measured by fMRI (a methodology which is relatively time-insensitive). The present results therefore suggest that, similarly across both cultures, the RSC and OFC may play a role in driving contextual influences on face categorizations, perhaps in a culturally invariant manner (at least assessed by fMRI).

As discussed earlier, the network of regions involved in processing contextual associations has typically included the RSC, OFC, and the PHC (Bar 2004; Kveraga et al. 2007). In the present study, we found that the RSC and OFC, but not the PHC, were sensitive to the compatibility of a focal stimulus (the face) with contextual cues. This suggests an interesting dissociation between the PHC and the other regions in the context network. For example, the PHC is thought to be more involved in processing a context’s lower-level perceptual information, whereas the RSC more involved in processing its higher order semantic “gist” information (Epstein et al. 2007; Epstein 2008). It is not surprising, then, that the RSC but not the PHC represented the compatibility of facial and contextual cues here, given that such compatibility is based on semantic category information extracted from a face and a scene. If the compatibility between cues were lower-level in nature (e.g., the face and scene sharing similar overlapping visual information), it is possible that the PHC would play an important role in representing the compatibility. Future studies might directly compare these regions’ representations of face-context compatibilities that range from more perceptual to more semantic. Such work could advance our understanding of these regions’ roles in processing contextual associations more broadly.

In summary, our perceptions of other people are rarely made without context; we see others in specific environments. Although contextual effects have long been documented in perceiving emotion, recent behavioral work has also shown such effects in perceiving static categorical characteristics such as race as well. The present study characterized the neural basis of such contextual influences on face categorization. We found that the RSC and OFC exhibited a highly sensitive and graded relationship with the compatibility of race-related facial and contextual cues. Further, activity in the RSC partially mediated the effect of this compatibility on the efficiency of perceptual responses. Together, the findings demonstrate the involvement of the RSC and OFC in driving contextual influences on face categorization, and further highlight the role of extraneous factors beyond the face in perceiving other people.

Funding

This research was supported in part by National Science Foundation EAPSI grant (OISE-1107874) to J.B.F., National Institutes of Health grant (F31-MH092000) to J.B.F., National Science Foundation grant (BCS-0435547) to N.A., and National Natural Science Foundation of China grants (30910103901, 91024032) to S.H.

Notes

Conflict of Interest: None declared.

References

- Aviezer H, Hassin RR, Ryan J, Grady C, Susskind J, Anderson A, Movcovitch M, Bentin S. 2008. Angry, disgusted, or afraid? Studies of the malleability of emotion perception. *Psychol Sci.* 19:724–732.
- Bar M. 2004. Visual objects in context. *Nat Rev Neurosci.* 5:617–629.
- Baron RM, Kenny DA. 1986. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J Pers Soc Psychol.* 51:1173–1182.
- Barrett LF, Kensinger EA. 2010. Context is routinely encoded during emotion perception. *Psychol Sci.* 21:595–599.
- Biederman I, Mezzanotte RJ, Rabinowitz JC. 1982. Scene perception: detecting and judging objects undergoing relational violations. *Cogn Psychol.* 14:143–177.
- Blair IV, Judd CM, Sadler MS, Jenkins C. 2002. The role of afrocentric features in person perception: judging by features and categories. *J Pers Soc Psychol.* 83:5–25.
- Blanz V, Vetter T. 1999. A morphable model for the synthesis of 3D faces. In: SIGGRAPH’99. Los Angeles: ACM Press. p. 187–194.
- Castano E, Yzerbyt VY, Bourguignon D, Seron E. 2002. Who may enter? The impact of in-group identification on in-group/out-group categorization. *J Exp Soc Psychol.* 38:315–322.
- Chiao JY, Iidaka T, Gordon HL, Nogawa J, Bar M, Aminoff E, Sadato N, Ambady N. 2008. Cultural specificity in amygdala response to fear faces. *J Cogn Neurosci.* 20:2167–2174.
- Dale AM. 1999. Optimal experimental design for event-related fMRI. *Hum Brain Mapp.* 8:109–114.
- Eberhardt JL, Dasgupta N, Banasyzynski TL. 2003. Believing is seeing: the effects of racial labels and implicit beliefs on face perception. *Pers Soc Psychol Bull.* 29:360–370.
- Eger E, Henson RN, Driver J, Dolan RJ. 2007. Mechanisms of top-down facilitation in perception of visual objects in fMRI. *Cereb Cortex.* 17:2123–2133.
- Epstein R, Kanwisher N. 1998. A cortical representation of the local visual environment. *Nature.* 392:598–601.
- Epstein RA. 2008. Parahippocampal and retrosplenial contributions to human spatial navigation. *Trends Cogn Sci.* 12:388–396.

- Epstein RA, Parker WE, Feiler AM. 2007. Where am I now? Distinct roles for parahippocampal and retrosplenial cortices in place recognition. *J Neurosci*. 27:6141–6149.
- Forman SD, Cohen JD, Fitzgerald M, Eddy WF, Mintun MA, Noll DC. 1995. Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. *Magn Reson Med*. 33:636–647.
- Freeman JB, Ambady N. 2011. A dynamic interactive theory of person construal. *Psychol Rev*. 118:247–279.
- Freeman JB, Ambady N, Rule NO, Johnson KL. 2008. Will a category cue attract you? Motor output reveals dynamic competition across person construal. *J Exp Psychol Gen*. 137:673–690.
- Freeman JB, Ma Y, Han S, Ambady N. 2013. Influences of culture and visual context on real-time social categorization. *J Exp Soc Psychol*. 49:206–210.
- Freeman JB, Penner AM, Saperstein A, Scheutz M, Ambady N. 2011. Looking the part: social status cues shape race perception. *PLoS ONE*. 6:e25107.
- Freeman JB, Rule NO, Adams RB, Ambady N. 2010. The neural basis of categorical face perception: graded representations of face gender in fusiform and orbitofrontal cortices. *Cereb Cortex*. 20:1314–1322.
- Gronau N, Neta M, Bar M. 2008. Integrated contextual representation for objects' identities and their locations. *J Cogn Neurosci*. 20:371–388.
- Haxby JV, Hoffman EA, Gobbini MI. 2000. The distributed human neural system for face perception. *Trends Cogn Sci*. 4:223–233.
- Henderson JM, Larson CL, Zhu DC. 2008. Full scenes produce more activation than close-up scenes and scene-diagnostic objects in parahippocampal and retrosplenial cortex: an fMRI study. *Brain Cogn*. 66:40–49.
- Hugenberg K, Bodenhausen GV. 2004. Ambiguity in social categorization: the role of prejudice and facial affect in race categorization. *Psychol Sci*. 15:342–345.
- Humphreys GW, Riddoch MJ, Price CJ. 1997. Top-down processes in object identification: evidence from experimental psychology, neuropsychology, and functional anatomy. *Philos Trans R Soc B*. 352:1275–1282.
- Johnson KL, Freeman JB, Pauker K. 2012. Race is gendered: how covarying phenotypes and stereotypes bias Sex categorization. *J Pers Soc Psychol*. 102:116–131.
- Kanwisher N, Yovel G. 2006. The fusiform face area: a cortical region specialized for the perception of faces. *Philos Trans R Soc Lond B Biol Sci*. 361:2109–2128.
- Kveraga K, Boshyan J, Bar M. 2007. Magnocellular projections as the trigger of top-down facilitation in recognition. *J Neurosci*. 27:13232–13240.
- Locke V, Macrae CN, Eaton JL. 2005. Is person categorization modulated by exemplar typicality? *Soc Cogn*. 23:417–428.
- MacLin OH, Malpass RS. 2001. Racial categorization of faces: the ambiguous-race face effect. *Psychol Public Policy Law*. 7:98–18.
- Macrae CN, Bodenhausen GV. 2000. Social cognition: thinking categorically about others. *Annu Rev Psychol*. 51:93–120.
- Mason MF, Norton MI, Van Horn JD, Wegner DM, Grafton ST, Macrae CN. 2007. Wandering minds: the default network and stimulus-independent thought. *Science*. 315:393–395.
- Masuda T, Ellsworth PC, Mesquita B, Leu J, Tanida S, Van de Veerdonk E. 2008. Placing the face in context: cultural differences in the perception of facial emotion. *J Pers Soc Psychol*. 94:365–381.
- Masuda T, Nisbett RE. 2001. Attending holistically versus analytically: comparing the context sensitivity of Japanese and Americans. *J Pers Soc Psychol*. 81:922–934.
- Mudrik L, Lamy D, Deouell LY. 2010. ERP Evidence for context congruity effects during simultaneous object–scene processing. *Neuropsychologia*. 48:507–517.
- Nisbett RE, Peng K, Choi I, Norenzayan A. 2001. Culture and systems of thought: holistic versus analytic cognition. *Psychol Rev*. 108:291–310.
- Preacher KJ, Selig JP. 2012. Advantages of Monte Carlo confidence intervals for indirect effects. *Commun Methods Meas*. 6:77–98.
- Righart R, De Gelder B. 2008. Recognition of facial expressions is influenced by emotional scene gist. *Cogn Affect Behav Neurosci*. 8:264–272.
- Ronquillo J, Denson TF, Lickel B, Lu Z-L, Nandy A, Maddox KB. 2007. The effects of skin tone on race-related amygdala activity: an fMRI investigation. *Soc Cogn Affect Neurosci*. 2:39–44.
- Russell JA, Bachorowski J, Fernandez-Dols JM. 2003. Facial and vocal expressions of emotion. *Annu Rev Psychol*. 54:329–285.
- Van Bavel JJ, Packer DJ, Cunningham WA. 2008. The neural substrates of in-group bias: a functional magnetic resonance imaging investigation. *Psychol Sci*. 19:1131–1139.
- Zeger SL, Liang KY. 1986. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics*. 42:121–130.