

Dynamic representations of race: processing goals shape race decoding in the fusiform gyri

Christian Kaul,^{1,2} Kyle G. Ratner,¹ and Jay J. Van Bavel¹

¹Department of Psychology, New York University, 6 Washington Place, New York, NY 10003, USA and ²Center for Neural Science, New York University, 4 Washington Place, New York, NY 10003, USA

People perceive and evaluate others on the basis of social categories, such as race, gender and age. Initial processing of targets in terms of visually salient social categories is often characterized as inevitable. In the current study, we investigated the influence of processing goals on the representation of race in the visual processing stream. Participants were assigned to one of two mixed-race teams and categorized faces according to their group membership or skin color. To assess neural representations of race, we employed multivariate pattern analysis to examine neural activity related to the presentation of Black and White faces. As predicted, patterns of neural activity within the early visual cortex and fusiform gyri (FG) could decode the race of face stimuli above chance and were moderated by processing goals. Race decoding in early visual cortex was above chance in both categorization tasks and below chance in a prefrontal control region. More importantly, race decoding was greater in the FG during the group membership vs skin color categorization task. The results suggest that, ironically, explicit racial categorization can diminish the representation of race in the FG. These findings suggest that representations of race are dynamic, reflecting current processing goals.

Keywords: race; goals; top-down; multivariate pattern analysis; fusiform gyri; face perception

Social categorization may be an inevitable aspect of human life (Allport, 1954). Extensive research suggests that social category information—such as race, gender and age—is automatically encoded (Brewer, 1988; Fiske and Neuberg, 1990). This information can have a powerful influence on judgments and evaluations of others [see Devine (1989) and Fiske (1998) for a review]—even when the categories are irrelevant to the current task or context (Taylor *et al.*, 1978; Brewer, 1988; Devine, 1989; Ito and Urland, 2005). Race, in particular, affects perceptual processing within a few hundred milliseconds (Ito and Urland, 2003; Cunningham *et al.*, 2012) and appears to be difficult to ignore (Park and Rothbart, 1982; Hewstone *et al.*, 1991; Stangor *et al.*, 1992; Richeson and Shelton, 2003; Olsson *et al.*, 2005). As a consequence, several researchers have argued that encoding race may be unavoidable (e.g. Devine, 1989; Fiske *et al.*, 1999; Ito *et al.*, 2007). In the current research, we investigated the influence of processing goals on neural representations of race in the visual processing stream. Specifically, we examined the influence of shallow (i.e. determining a target's skin color) vs deep (i.e. determining a target's group membership) modes of social categorization on race decoding in early visual cortex and fusiform gyri (FG) using multivariate pattern analysis (MVPA)—an analytic technique that allows for the identification of race-based representations in the absence of mean-level differential activity between two racial categories.

Several studies have examined the processing of race in the human brain (for reviews, see Amodio and Lieberman, 2009; Cunningham and Van Bavel, 2009). The core and extended face network appear to play a critical role in discerning race from faces (see Eberhardt, 2005; Ito and

Bartholow, 2009; Macrae and Quadflieg, 2010; Kubota *et al.*, 2012 for reviews). Specifically, a face sensitive sub-region of the FG has been implicated in the perception and recognition of social categories (Van Bavel *et al.*, 2008, 2011), including race (Golby *et al.*, 2001; Lieberman *et al.*, 2005; Chiu *et al.*, 2011; Feng *et al.*, 2011). The FG is a key region involved in face processing (Sergent *et al.*, 1992; Puce *et al.*, 1995; Kanwisher *et al.*, 1997), especially aspects of the identity of a face (Ishai *et al.*, 1999; Grill-Spector *et al.*, 2004). A neuroimaging study by Golby *et al.* (2001) reported the first evidence that Black and White participants both show increased blood-oxygenation level-dependent (BOLD) activity in the FG, in response to seeing own-race vs other-race stimulus faces. Since that initial study, several others have found similar results (e.g. Lieberman *et al.*, 2005; Natu *et al.*, 2010). Moreover, in the study by Golby *et al.*, participants with the greatest relative activity in the FG to own-race faces also had superior recognition memory for own-race vs other-race faces, a behavioral phenomenon that has been termed 'the own-race bias' (Malpass and Kravitz, 1969). These results suggest that race is encoded during face perception.

Although there is now extensive research on the neural substrates of perceiving race, there is relatively little research on the influence of processing goals on these biases (but see Lieberman *et al.*, 2005; Wheeler and Fiske, 2005). There is reason to believe, however, that processing goals can influence early components of the perceptual (e.g. Amodio, 2010; Cunningham *et al.*, 2012) and evaluative (e.g. Cunningham *et al.*, 2008) systems (see Van Bavel *et al.*, 2012 for a recent review). Moreover, behavioral research suggests that the activation and application of racial stereotypes and evaluative biases depend on factors such as cognitive load (Gilbert and Hixon, 1991; Blair and Banaji, 1996; Sherman *et al.*, 2000) and processing goals (Mitchell *et al.*, 2003; Van Bavel and Cunningham, 2009). Consistent with this perspective, one neuroimaging study found that amygdala activity was higher to other-race (vs own-race) faces when people made categorical judgments about the faces. However, this pattern was reversed when people made individuating judgments about the faces (Wheeler and Fiske, 2005).

In a series of earlier studies, we investigated the influence of the social context on activity in FG (Van Bavel *et al.*, 2008, 2011). We

Received 26 March 2012; Accepted 24 November 2012

Advance Access publication 29 November 2012

The authors thank William Cunningham for generously funding the data collection phase of this research, and David Amodio, Tobias Brosch, Sharon David, Alumit Ishai, Lisa Kaggen, Dominic Packer, Chris Said, Jillian Swencionis, Jenny Xiao, Sophie Wharton and members of the NYU Social Perception and Evaluation Lab (@vanbavellab) for help with various aspects of this research.

This research was supported by a Feodor-Lynen-Award from the Alexander von Humboldt Foundation to C.K., a National Science Foundation Graduate Research Fellowship to K.G.R. and a Social Sciences and Humanities Research Council of Canada Award to J.V.B. Data were collected at the Queen's University MRI facility.

C.K. and K.G.R. contributed equally to this work.

Correspondence should be addressed to Jay J. Van Bavel, Department of Psychology, New York University, 6 Washington Place, New York, NY 10003, USA. E-mail: jay.vanbavel@nyu.edu

made race orthogonal to group membership by assigning people to mixed-race groups (see also Van Bavel and Cunningham, 2009). In contrast to other studies that have demonstrated preferential activation of the FG to own-race faces (Golby *et al.*, 2001; Feng *et al.*, 2011), we found no effects of race on FG activity (i.e. no difference in mean activity to Black *vs* White faces). Instead, we found greater activity in the FG to own-group *vs* other-group faces, regardless of race. These studies suggest that the intergroup context may shape the motivational relevance of different social categories and consequently alter social biases (Van Bavel and Cunningham, 2011, 2012).

More recently, we used MVPA to examine whether race is represented in patterns of activation within the FG and early visual cortex despite the fact that mean level BOLD activity reflects *group membership* rather than *race* (Ratner *et al.*, 2013). MVPA uses the information carried by fine-grained patterns of BOLD activity within different brain regions to 'decode' the presence or absence of different categories of stimuli or visual features (Haynes and Rees, 2006; Norman *et al.*, 2006; Mur *et al.*, 2009). Recent studies have used MVPA to compare the intensity of emotions perceived from face movements, body movements and vocal intonation (Peelen *et al.*, 2010), and visually salient social categories (Natu *et al.*, 2010; Chiu *et al.*, 2011; Kaul *et al.*, 2011). Studies have also shown that neural activity in early visual cortex (i.e. V1) can successfully decode several different colors better than other visual regions (Brouwer and Heeger, 2009).

In our previous MVPA study, we demonstrated that race was indeed represented in the FG even though overall BOLD activity in the FG was driven by own-group membership (Ratner *et al.*, 2013). During neuroimaging, participants in this experiment categorized each face in terms of their group membership (e.g. is this person a member of your own-group?). This task could not be completed using simple perceptual information because the two groups were *not* associated with different visual cues (e.g. different jersey colors). As such, participants had to recall the group membership of each target face. We speculated that race representation in the FG during this task might reflect the use of race as a cue for discerning group membership—a functional utilization of racial information for task completion. Specifically, we surmised that skin color and physiognomic information that co-vary with race could be used as retrieval cues in memory for associations learned about a particular face. Thus, by representing this individuating information during high-level face processing, participants could improve their performance on the categorization task.

In the present study, we examined the influence of processing goals on representations of race in the visual processing stream. We reasoned that representations of race in the visual processing stream might vary according to the relevance of different computational processes for task goals. For example, a goal to identify target faces should recruit the FG because this brain region plays a central role in processing the subordinate-level identity of stimuli (Gauthier *et al.*, 1997), especially faces (Ishai *et al.*, 1999; Grill-Spector *et al.*, 2004). On the other hand, a goal that requires identifying only the low-level visual features of a face (e.g. skin color categorization) should recruit brain regions associated with more basic aspects of visual processing, such as detecting color or contrast (Brouwer and Heeger, 2009). As such, we predicted that patterns of neural activity within the early visual cortex, but not a size-matched gray-matter control region (CTR), would decode the race of face stimuli above chance. More importantly, we predicted that rapidly categorizing a face based on superficial visual features (i.e. skin color) as opposed to more effortful, non-perceptual information (i.e. group membership) would recruit higher-order visual processing to a lesser extent, and would therefore ironically result in less race decoding in the FG.

METHODS

Participants

Data from 17 White participants (10 females; mean age = 25) were analyzed for this study using an existing dataset (Van Bavel *et al.*, 2008). [Five additional participants were not analyzed for reasons described in our previous article (Van Bavel *et al.*, 2008).] Participants reported no abnormal neurological history and had normal or corrected-to-normal vision. Participants gave written informed consent to participate in the study and were paid \$50 for completing the study.

Procedure

Group assignment

After arriving at the neuroimaging center, participants were informed that they were in a study exploring learning about groups, that they had been randomly assigned to the Leopards or Tigers, and that it was important for them to learn the members of their group and a competing group. Participants then spent 3 min memorizing the team membership of 24 faces presented simultaneously: 12 members of the Leopards and 12 members of the Tigers. There were six Black and six White males on each team. Faces were randomly assigned to a team and fully counterbalanced. During the second learning task, participants saw and categorized each face according to whether the face was affiliated with the Leopards or Tigers (see Van Bavel *et al.*, 2008 for more details).

Categorization task

Participants completed six runs of four blocks containing 12 trials for a total of 288 trials during fMRI. On each trial, participants categorized one of the 24 faces in one of two ways (Figure 1). On *skin color* trials, participants categorized each face according to skin color (Black or White). On *group membership* trials, participants categorized each face according to team membership (Leopard or Tiger). Team and race labels were counterbalanced (left *vs* right) within runs, creating four randomized blocks within each run. Each of the 24 faces was categorized twice in each run (once by team membership and once by race). Direction screens were presented before each block for 6 s to cue the categorization required for the following block of 12 trials. Each face appeared for 2 s, during which time participants responded with a button box in their right hand. To allow for estimation of the hemodynamic signal, fixation crosses appeared between names for 3, 4, 6 or 8 s (in pseudo-random order).

Neuroimaging parameters, acquisition, and preprocessing

All imaging was conducted with a Siemens 3T scanner. For whole-brain functional coverage of changes in BOLD activity, 32 axial slices (slice thickness = 3.5 mm, 0.5 mm skip) were prescribed parallel to the AC–PC line. Nearly isotropic functional images were acquired from inferior to superior using a single-shot gradient echo planar pulse sequence (TE = 25 ms, TR = 2000 ms, in-plane resolution = 3.5 × 3.5 mm, matrix size = 64 × 64 and FOV = 224 mm). Data were preprocessed using SPM8: all data were realigned to the first image and corrected for slow signal drift with a 128 s high-pass filter, but otherwise left untransformed. Following functional imaging, a high-resolution MPRAGE anatomical image (176 sagittal slices; TE = 2.15 ms; TR = 1760 ms; resolution = 1 × 1 × 1 mm) was collected. During analysis, this structural scan was co-registered to the functional data of each participant.

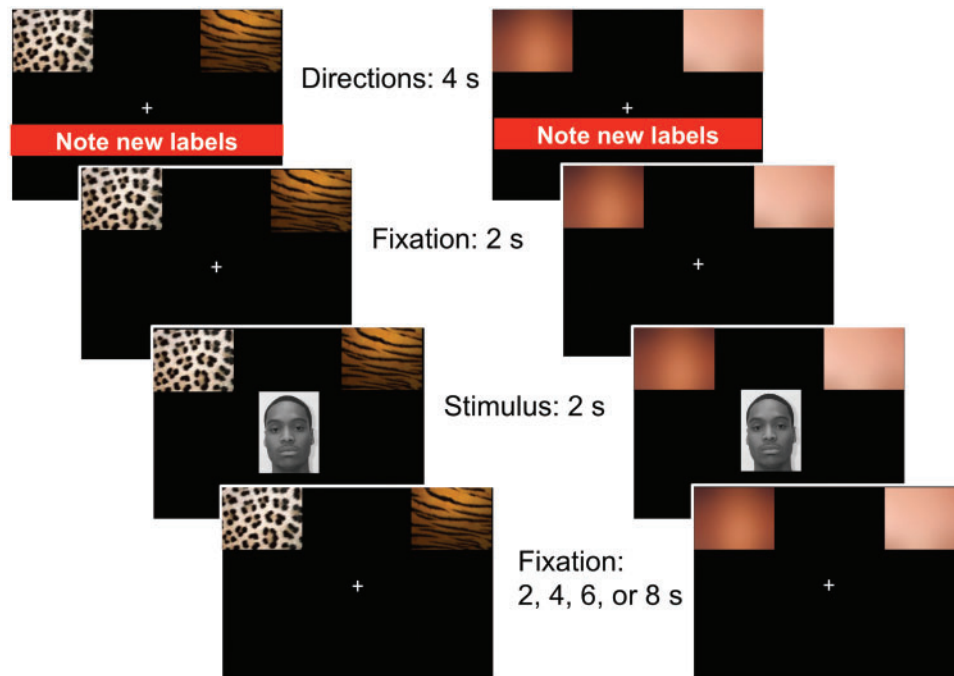


Fig. 1 Sample trials of the skin color and group membership categorization tasks. There were two skin color and two group membership categorization blocks in each of six runs. Each block started with a directions screen followed by 12 randomly presented faces. Faces presented within each block were separated by fixation crosses. After the completion of each block, directions for the next block appeared. Images are not drawn to scale.

Region of interest localization and analysis

FreeSurfer (<http://surfer.nmr.mgh.harvard.edu>) was used for segmentation and cortical inflation and flattening using each participant's structural image (previously co-registered to all functional data). We used SPM8 to perform a within-participant analysis, with a voxel-wise general linear model comprising 10 boxcar waveforms. The study used a 2 (group membership: in-group, out-group) \times 2 (race: Black, White) \times 2 (categorization task: group membership, skin color) experimental design and two additional regressors to model direction screens and fixation (rest). For each MRI time series, global changes in activity were removed by proportional scaling of each session.

To identify regions-of-interest (ROI) for each participant, we overlaid the contrast of *all faces vs rest* at a FWE-corrected significance level of $P < 0.01$ onto each participants' inflated brain surface. On the basis of this overlay, we identified the FG bilaterally and a size-matched CTR in the medial orbitofrontal cortex. Finally, each participant's V1 was anatomically defined using a probabilistic map resulting from an automated algorithm within FreeSurfer (Hinds *et al.*, 2008). Because each ROI was defined at the participant-level, the size of each ROI varied between participants. Note that, as a result of defining ROIs on the inflated brain surface, all ROI masked images contained gray-matter voxels only.

Univariate analyses

In a prior univariate analysis of this dataset (Van Bavel *et al.*, 2008), we reported that the mean BOLD signal in the FG did not significantly differ between Black and White faces when participants categorized faces according to *skin color* or *group membership*. In the present study, we replicated these analyses, but used the preprocessing steps and number of voxels that we used for MVPA. The data for the univariate analyses were spatially smoothed to maximize the

signal-to-noise ratio (Mikl *et al.*, 2008). Due to the possibility that spatial smoothing can remove fine-grained pattern information, we did not spatially smooth the data prior to the MVPA (Kriegeskorte *et al.*, 2006; Mur *et al.*, 2009, but see Kamitani and Sawahata, 2010; Op de Beeck, 2010). After smoothing the data, we conducted separate analyses on the BOLD responses in our ROIs (FG, V1 and CTR) to Black and White faces independently for the two categorization tasks (group membership and skin color).

Multivariate analysis

Unsmoothed, realigned fMRI data were analyzed using the MATLAB routines provided in the Princeton MVPA Toolbox (www.csmb.princeton.edu/mvpa). Using the toolbox, the time series from each voxel was de-trended to remove linear and quadratic trends, and z-scored to normalize the time series to have a mean of zero and a variance of one. Condition onsets were adjusted for the lag in hemodynamic response function by shifting all block-onset timings by three volumes (6 s). To determine classification accuracies, only classification with unseen and independent test data was considered using a leave-one-session-out cross-validation method (Mur *et al.*, 2009). In the actual classification step, we used a Gaussian Naïve Bayes classifier algorithm (Mitchell *et al.*, 2004) within the MVPA toolbox.

Classification accuracies were averaged across the five cross-validations for each ROI in each participant. Thus, for each participant this procedure yielded exactly one mean prediction accuracy per ROI (i.e. 17 observations per ROI). We used *t*-tests to test for significant differences in race decoding from chance (two categories = 50%) in each of the three ROIs: V1, FG, and CTR. Second, we tested for a significant difference against a second control-baseline as defined by the classification accuracy within the CTR using *t*-tests. We also used *t*-tests to examine lateralization effects in the FG.

RESULTS

Behavioral results

We compared participants' *reaction time* (the time in milliseconds between the presentation of a face and the response to it) and *accuracy* (the proportion of trials responded to correctly during the 2-s face presentation) during the skin color *vs* group membership blocks of the face categorization task. Trials on which the reaction time was ≤ 300 ms were excluded (Van Bavel *et al.*, 2008). Participants were faster, $F(16) = 152.59, P < 0.01$, and more accurate $F(16) = 12.06, P < 0.01$, when categorizing faces during the skin color (897 ms; 94%) *vs* the group membership (1242 ms; 83%) task. These results are consistent with the suggestion that the group membership categorization involved deeper processing.

Univariate results

Mean BOLD signal in all three ROIs did not significantly differ between Black and White faces ($P_s \geq 0.48$).

MVPA results

We expected to find task-dependent differences in race decoding when comparing neural responses during the *group membership* task (deep encoding) *vs* the *skin color* task (shallow encoding). We hypothesized that race would be decoded in the FG when participants engaged in deeper processing to determine the identity of faces (Grill-Spector *et al.*, 2004) during the group membership task. In contrast, we expected that race would not be decoded in the FG when participants

engaged in the superficial processing sufficient for the skin color task. Both tasks were analyzed separately and then compared for race decoding accuracies in our three ROIs (V1, FG and CTR). Figure 2 shows these regions on the inflated surface of each hemisphere of a sample participant.

Following our previous research (Ratner *et al.*, 2013), we expected that multivariate patterns of neural activity within the early visual cortex and FG could decode facial race above chance. As predicted, in the *group membership* task, race was decoded above chance in V1, 54.6%, $t(16) = 6.86, P < 0.01$ and FG, 57.8%, $t(16) = 9.30, P < 0.01$. The FG data were collapsed across hemispheres. However, separate analysis confirmed that the right and left FG both successfully predicted facial race at similar levels [right FG: 57.1%, $t(16) = 8.08, P < 0.01$; left FG: 55.6%, $t(16) = 6.96, P < 0.01$] and there were no differences when comparing the left and right FG, $t(16) = -1.27, n.s.$ Importantly, race could not be decoded in the CTR, 50.3%, $t(16) = 0.32, n.s.$ This is a direct replication of results from Ratner *et al.* (2013) with slightly superior decoding accuracies, which may be due to more accurate ROI-definition provided by the ability to co-register BOLD response to structural images. This suggests that the race decoding effects were not due to global BOLD activity. In other words, patterns of BOLD activity in V1 and FG could decode racial face when the task required deeper processing (i.e. recalling the identity of each face to determine group membership).

Given the superficial processing sufficient for the *skin color* task, we expected that race decoding would not be apparent in the FG. As predicted, in the skin color task, race was decoded above chance in V1, 56.2%, $t(16) = 7.52, P < 0.01$, but not FG, 50.1%, $t(16) = 0.03, n.s.$, or CTR, 51.0%, $t(16) = 0.85, n.s.$ (Figure 3, light gray bars). The absence of race decoding in the FG during the skin color task was consistent in both hemispheres, left FG = 50.1%, $t(16) = 0.06, n.s.$; right FG = 50.5%, $t(16) = 0.62, n.s.$ The lack of race decoding in FG during the skin color task is consistent with research showing that the FG is involved in processing the structure and identity of faces (Ishai *et al.*, 1999; Grill-Spector *et al.*, 2004), not low-level visual features.

To directly test whether multivariate patterns of neural activity within V1 and FG were moderated by processing goals, we compared decoding accuracies in FG and V1 during both tasks. Replicating the results of our previous research (Ratner *et al.*, 2013), during the *group*

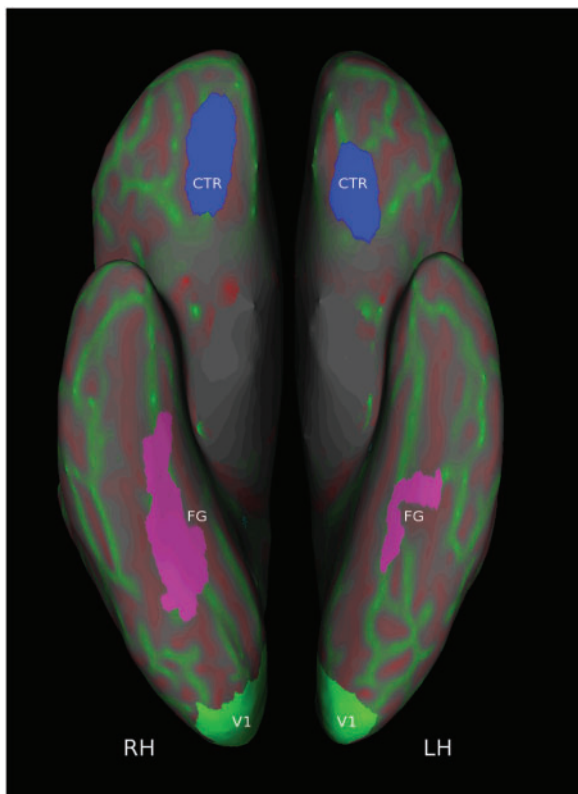


Fig. 2 An inflated surface map of the brain showing the location of the ROIs. The figure shows an inferior view of the right (RH) and left (LH) hemispheres of one participant reconstructed using Freesurfer. Red colors indicate sulci, while green colors designate gyri. Shown in green are bilateral estimations of early visual cortex (V1) as calculated by a probabilistic map based on each participant's individual structural image (Hinds *et al.*, 2008). Depicted in violet and blue are the locations of a face-sensitive region of the bilateral FG and a size-matched control region (CTR).

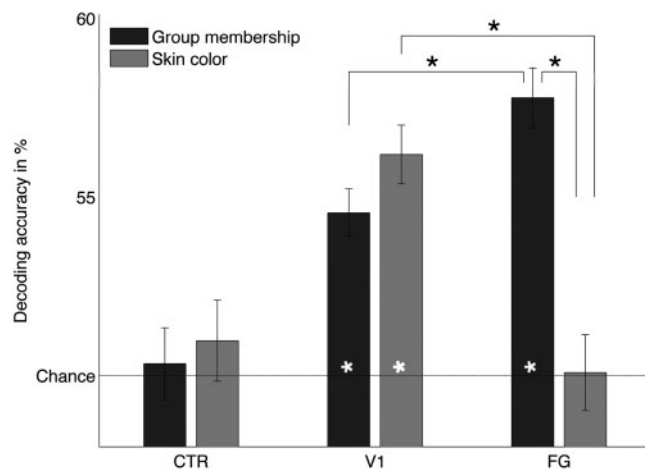


Fig. 3 Mean decoding accuracies for Black vs White faces across the two tasks shows the influence of processing goals on facial race pattern classification accuracy in the FG, early visual cortex (V1) and a size-matched (CTR). Face-race decoding during the group membership task (dark gray) was significantly ($*P < 0.01$) different from chance in V1 and FG. However, during the skin color task, face-race decoding was significantly different from chance in V1 but not in FG. Close to chance decoding accuracy in the CTR during both tasks confirmed the validity of the V1 and FG results.

membership task, race decoding was significantly higher in FG than V1, $t(16) = 3.01$, $P = 0.03$. In contrast, during the *skin color* task, race decoding was significantly greater in V1 than FG, $t(16) = 4.54$, $P < 0.01$. More importantly, race decoding in FG was higher during the group membership task than the skin color task, $t(16) = 5.68$, $P < 0.01$. These results suggest that, ironically, explicit racial categorization can diminish the representation of race in the FG. Moreover, race decoding in V1 was higher during the skin color task than the group membership task, $t(16) = 1.54$, n.s. There were no meaningful correlations between race decoding results across tasks, V1: $r = 0.16$, n.s.; FG: $r = -0.22$, n.s. Taken together, these results are consistent with our hypothesis that processing goals can moderate neural representations of race in the visual processing stream.

To further investigate the influence of processing goals on V1 and FG, we compared individual differences in the race decoding accuracies of these regions across participants. Interestingly, individual differences in race decoding in FG and V1 were correlated in the *skin color* task, $r = 0.51$, $P = 0.04$, but not the *group membership* task, $r = 0.20$, n.s. These results suggest that FG and V1 might be using similar information during the skin color task. Indeed, the skin color task does not require the structural categorization of faces, but mere attention to visually salient, low-level perceptual features inherent in the faces. It is important to note, however, that it might not be color per se that drives skin color categorization. For example, differences in facial features or other low-level visual features such as contrast (the difference between light and dark) may account for race decoding in early visual cortex.

Control analyses

We attempted to predict facial race in both studies from a frontal control area not known to contain neuronal populations selective for faces (Kaul et al., 2011; Ratner et al., 2013). Classification performance in the CTR did not differ from chance (Figure 3). We then defined the distribution of decoding results using this CTR as a baseline (instead of 50% chance). Testing against this alternate baseline, we replicated race decoding in FG, $t(16) = 5.69$, $P < 0.01$ and V1, $t(16) = 3.51$, $P = 0.02$, during the *group membership* task, and race decoding in V1, $t(16) = 3.72$, $P < 0.01$, but not FG, $t(16) = -0.57$, n.s., during the *skin color* task. These results make it unlikely that race decoding was due to confounds, such as a general increase in blood flow (Kaul et al., 2011).

DISCUSSION

The current research demonstrates the dynamic nature of representations of social categories, like race, in the human brain. We examined whether the representations of race in primary visual cortex (V1) and the FG were sensitive to different processing goals. To address this issue, we asked people to categorize faces according to *skin color* (shallow encoding) or *group membership* (deep encoding). We hypothesized that race would be represented in V1 during both tasks, but that race might only be represented in the FG when participants during the group membership task were required to process the identity of each face. Using MVPA—an analytic technique that allows for the identification of race-based representations in the absence of mean-level differential activity between two racial categories—we found that different processing goals did indeed elicit different *patterns* of activation in V1 and FG. As predicted, race decoding in V1 was above chance in both categorization tasks and race decoding was below chance in a prefrontal CTR. More importantly, race decoding was higher in the FG during the *group vs skin color* categorization task. The results suggest that, ironically, explicit racial categorization can diminish the representation of race in the FG.

These results extend existing work on the representation of social categories (Natu et al., 2010; Kaul et al., 2011; Ratner et al., 2013) by demonstrating that processing goals can moderate neural representations of race in the visual processing stream. Replicating our previous research (Ratner et al., 2013), we found evidence of race decoding in V1 and FG when participants categorized faces on the basis of their *group membership*. However, race decoding in FG was non-significant when participants categorized faces on the basis of their *skin color*. It might seem counterintuitive that categorizing a face based on skin color could *decrease* race decoding. In fact, a recent article suggests that attending to certain social category can *increase* decoding accuracy in the FG (Chiu et al., 2011). However, there is extensive research showing that the FG is involved in processing the *identity* of faces (Ishai et al., 1999; Grill-Spector et al., 2004), whereas early visual cortex is important for perceiving contrast and colors (Brouwer and Heeger, 2009). As such, the lack of race decoding in the FG during the skin color task likely reflects the fact that the task did not require processing the identity of each face. Taken together, our results suggest that mere attention to a social category like race does not necessarily enhance its representation or have uniform effects on different regions in the visual processing stream. Instead, the computational process of each brain region may dictate the relationship between processing goals and the representations of visual input.

Several aspects of our study corroborate the idea that face processing was shallow during the skin color task. For instance, our behavioral data indicate that participants were much faster and more accurate during the skin color task compared with the group membership task. In addition, decoding accuracies between the V1 and FG were highly correlated during the skin color task, but not during the group membership task. These correlations suggest that FG activity patterns reflect more basic visual information during the skin color categorization task and additional structural information during the group membership categorization task. In conclusion, FG decodes race less when participants are required to categorize targets according to superficial visual cues than when they need to process faces on a deeper level.

Theoretical and societal implications

From a cognitive neuroscience perspective, our work demonstrates how processing goals can influence *patterns* of neural activation in the visual processing stream (Van Bavel and Cunningham, 2011). We suggest that when processing goals only require simple categorization on visual dimensions (e.g. skin color), neural regions that represent low-level visual features (e.g. color, contrast) may be sufficient for visual categorization. However, when recalling and identifying an individual target are important, parts of the visual system that support subordinate-level recognition may also be recruited. Critically, we are not implying that stimuli that require shallow processing are only processed in early visual cortex. Acting on perceptions requires motor planning and execution, which likely involves regions in the dorsal visual stream, prefrontal cortex, and the motor cortex. Our point is simply that areas of the visual stream necessary for higher-order conceptual representation of interracial faces are sensitive to processing goals. More generally, our results suggest that the representation of race is dynamic, reflecting current processing goals that arise from the social environment (Van Bavel et al., 2013). This work adds to a growing body of research that shows that person perception and evaluation are characterized by a widely distributed network of brain regions that dynamically integrate a constellation of bottom-up (e.g. skin color) and top-down (e.g. processing goals) information (e.g. Cunningham et al., 2007; Freeman and Ambady, 2011).

The current research also has implications for the social psychological understanding of person perception. Influential continuum models of person perception contend that salient social category cues lead to less individuation and more categorical processing (Fiske and Neuberg, 1990; Fiske *et al.*, 1999). In our previous work (Ratner *et al.*, 2013), we speculated that social category cues, like race, can be used to serve individuating processing goals, such as narrowing the pool of possible memory traces during retrieval (see also Brewer, 1988). Whereas our earlier work merely alluded to this possibility, the current research provides a more stringent test. Our data suggest that categorical information can lead to individuation, which then can inform categorization decisions (e.g. determining whether this person is part of my team). Further research should continue to investigate this possibility.

This work may also have broader implications relevant to social action. Current conventional wisdom suggests that in an egalitarian society, citizens should be 'colorblind' to race (Jones, 1998; Schofield, 2007; Carbado and Harris, 2008; Plaut *et al.*, 2009). Although the pro-social intentions behind colorblind initiatives are laudable, recent work has indicated that such efforts do not always lead to favorable outcomes (Norton *et al.*, 2006; Apfelbaum *et al.*, 2008, 2010). Inline with this latter research, our work illustrates the complex role that racial cues play in the processing of visual information, particularly faces. We assessed race decoding in a context where race (i.e. skin color) was either explicitly the categorical cue or incidental to the task. Our work suggests that categorizing an individual based on skin color does not necessarily involve race representation in the FG, a higher-order visual area associated with face processing. Moreover, inconsistent with reasoning that follows from a colorblind perspective, race representation in the FG was actually stronger when people categorized faces based on a group membership dimension that was orthogonal to race. We posit that this finding is one case where racial information can be useful for individual identification. Thus, this work suggests a more nuanced understanding of how race operates in perception, which may inspire new approaches to reduce discrimination and intergroup conflict in society.

REFERENCES

- Allport, G.W. (1954). *The Nature of Prejudice*. Reading, MA: Addison Wesley.
- 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., Lieberman, M.D. (2009). Pictures in our heads: contributions of fMRI to the study of prejudice and stereotyping. In: Nelson, T., editor. *Handbook of Prejudice, Stereotyping, and Discrimination*. New York: Erlbaum Press, pp. 347–66.
- Apfelbaum, E.P., Pauker, K., Sommers, S.R., Ambady, N. (2010). In blind pursuit of racial equality? *Psychological Science*, 21, 1587–92.
- Apfelbaum, E.P., Sommers, S.R., Norton, M.I. (2008). Seeing race and seeming racist? Evaluating strategic colorblindness in social interaction. *Journal of Personality and Social Psychology*, 95, 918–32.
- Blair, I.V., Banaji, M.R. (1996). Automatic and controlled processes in stereotype priming. *Journal of Personality and Social Psychology*, 70(6), 1142–63.
- Brewer, M.B. (1988). A dual process model of impression formation. In: Srull, T.K., Wyer, R.S., editors. *Advances in Social Cognition*, Vol. 1, Hillsdale, NJ: Lawrence Erlbaum Associates, Inc, pp. 1–36.
- Brouwer, G.J., Heeger, D.J. (2009). Decoding and reconstructing color from responses in human visual cortex. *Journal of Neuroscience*, 29(44), 13992–4003.
- Carbado, D.W., Harris, C.I. (2008). The new racial preferences. *California Law Review*, 96, 1139–214.
- Chiu, Y.-C., Esterman, M., Rosen, H., Yantis, S. (2011). Decoding task-based attentional modulation during face categorization. *Journal of Cognitive Neuroscience*, 23, 1198–204.
- Cunningham, W.A., Van Bavel, J.J. (2009). A neural analysis of intergroup perception and evaluation. In: Berntson, G.G., Cacioppo, J.T., editors. *Handbook of Neuroscience for the Behavioral Sciences*. Hoboken, NJ: Wiley, pp. 975–84.
- Cunningham, W.A., Van Bavel, J.J., Ar buckle, N.L., 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, 140.
- Cunningham, W.A., Van Bavel, J.J., Johnsen, I.R. (2008). Affective flexibility: evaluative processing goals shape amygdala activity. *Psychological Science*, 19(2), 152–60.
- Devine, P.G. (1989). Stereotypes and prejudice: their automatic and controlled components. *Journal of Personality and Social Psychology*, 56(1), 5–18.
- Eberhardt, J.L. (2005). Imaging race. *American Psychologist*, 60, 181–90.
- Feng, L., Liu, J., Wang, Z., Li, J., Li, L., Ge, L., Tian, J., Lee, K. (2011). The other face of the other-race effect: an fMRI investigation of the other-race face categorization advantage. *Neuropsychologia*, 49(13), 3739–49.
- Fiske, S.T. (1998). Stereotyping, prejudice, and discrimination. In: Gilbert, D.T., Fiske, S.T., Lindzey, G., editors. *The Handbook of Social Psychology*, 4th edn. New York: McGrawHill, pp. 357–411.
- Fiske, S.T., Lin, M.H., Neuberg, S.L. (1999). The continuum model: ten years later. In: Trope, S.C.Y., editor. *Dual Process Theories in Social Psychology*. New York: Guilford, pp. 231–54.
- 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. In: Zanna, M.P., editor. *Advances in Experimental Social Psychology*, Vol. 23, New York: Academic Press, pp. 1–74.
- Freeman, J.B., Ambady, N. (2011). A dynamic interactive theory of person construal. *Psychological Review*, 118, 247–79.
- Gauthier, I., Anderson, A.W., Tarr, M.J., Skudlarski, P., Gore, J.C. (1997). Levels of categorization in visual recognition studied with functional MRI. *Current Biology*, 7, 645–51.
- Gilbert, D.T., Hixon, J.G. (1991). The trouble of thinking: activation and application of stereotypic beliefs. *Journal of Personality & Social Psychology*, 60(4), 509–17.
- Golby, A.J., Gabrieli, J.D.E., Chiao, J.Y., Eberhardt, J.L. (2001). Differential fusiform responses to same- and other-race faces. *Nature Neuroscience*, 4, 845–50.
- Grill-Spector, K., Knouf, N., Kanwisher, N. (2004). The fusiform face area subserves face perception, not generic within-category identification. *Nature Neuroscience*, 7, 555–62.
- Haynes, J.-D., Rees, G. (2006). Decoding mental states from brain activity in humans. *Nature Reviews Neuroscience*, 7, 523–34.
- Hewstone, M., Hantzi, A., Johnston, L. (1991). Social categorization and person memory: the pervasiveness of race as an organizing principle. *European Journal of Social Psychology*, 21, 517–28.
- Hinds, O.P., Rajendran, N., Polimeni, J.R., Augustinack, J.C., Wiggins, G., Wald, L.L. (2008). Accurate prediction of V1 location from cortical folds in a surface coordinate system. *Neuroimage*, 39(4), 1585–99.
- Ishai, A., Ungerleider, L.G., Martin, A., Schouten, J.L., Haxby, J.V. (1999). Distributed representation of objects in the human ventral visual pathway. *Proceedings of the National Academy of Sciences*, 96(16), 9379–84.
- 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 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, 21–36.
- Ito, T.A., Willadsen-Jensen, E.C., Correll, J. (2007). Social neuroscience and social perception: new perspectives on categorization, prejudice, and stereotyping. In: Harmon-Jones, E., Winkielman, P., editors. *Social Neuroscience: Integrating Biological and Psychological Explanations of Social Behavior*. New York: Guilford, pp. 401–21.
- Jones, J.M. (1998). Psychological knowledge and the new American dilemma of race. *Journal of Social Issues*, 54, 641–62.
- Kamitani, Y., Sawahata, Y. (2010). Spatial smoothing hurts localization but not information: pitfalls for brain mappers. *Neuroimage*, 49, 1949–52.
- Kanwisher, N., McDermott, J., Chun, M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for the perception of faces. *Journal of Neuroscience*, 17, 4302–11.
- Kaul, C., Rees, G., Ishai, A. (2011). The gender of face stimuli is represented in multiple regions in the human brain. *Frontiers in Human Neuroscience*, 4, 238.
- Kriegeskorte, N., Goebel, R., Bandettini, P. (2006). Information-based functional brain mapping. *Proceedings of the National Academy of Sciences*, 103(10), 3863–8.
- Kubota, J.T., Banaji, M.R., Phelps, E.A. (2012). The neuroscience of race. *Nature Neuroscience*, 15, 940–8.
- Kurzban, R., Tooby, J., Cosmides, L. (2001). Can race be erased? Coalitional computation and social categorization. *Proceedings of the National Academy of Sciences*, 98, 15387–92.
- Lieberman, M.D., Hariri, A., Jarcho, J.M., Eisenberger, N.L., Bookheimer, S.Y. (2005). An fMRI investigation of race-related amygdala activity in African-American and Caucasian-American individuals. *Nature Neuroscience*, 8, 720–2.
- Macrae, C.N., Quadflieg, S. (2010). Perceiving people. In: Gilbert, D.T., Fiske, S.T., Lindzey, G., editors. *The Handbook of Social Psychology* 5th edn. New York: McGraw-Hill.
- Malpass, R.S., Kravitz, J. (1969). Recognition for faces of own and other 'race'. *Journal of Personality and Social Psychology*, 13, 330–4.
- Mikl, M., Mareček, R., Hlustik, P.M.P., Drastich, A., Chlebus, P. (2008). Effects of spatial smoothing on fMRI group inferences. *Magnetic Resonance Imaging*, 26, 490–503.

- Mitchell, J.P., Nosek, B.N., Banaji, M.R. (2003). Contextual variations in implicit evaluation. *Journal of Experimental Psychology: General*, 132, 455–69.
- Mitchell, T.M., Hutchinson, R., Niculescu, R.S., Pereira, F., Wang, X. (2004). Learning to decode cognitive states from brain images. *Machine Learning*, 57, 145–75.
- Mur, M., Bandettini, P.A., Kriegeskorte, N. (2009). Revealing representational content with pattern-information fMRI—an introductory guide. *Social Cognitive and Affective Neuroscience*, 4(1), 101–9.
- Natu, V., Raboy, D., O'Toole, A.J. (2010). Neural correlates of own- and other-race face perception: spatial and temporal response differences. *Neuroimage*, 54(3), 2547–55.
- Norman, K.A., Polyn, S.M., Detre, G.J., Haxby, J.V. (2006). Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences*, 10, 424–30.
- Norton, M.I., Sommers, S.R., Apfelbaum, E.P., Pura, N., Ariely, D. (2006). Colorblindness and interracial interaction: playing the political correctness game. *Psychological Science*, 17, 949–53.
- Olsson, A., Ebert, J.P., Banaji, M.R., Phelps, E.A. (2005). The role of social groups in the persistence of learned fear. *Science*, 309, 785–7.
- Op de Beeck, H. (2010). Against hyperacuity in brain reading: spatial smoothing does not hurt multivariate fMRI analyses? *Neuroimage*, 49, 1943–8.
- Park, B., Rothbart, M. (1982). Perception of out-group homogeneity and levels of social categorization: memory for the subordinate attributes of in-group and out-group members. *Journal of Personality and Social Psychology*, 42, 1051–68.
- Peelen, M.V., Atkinson, A.P., Vuilleumier, P. (2010). Supramodal representations of perceived emotions in the human brain. *Journal of Neuroscience*, 30(30), 10127–34.
- Plaut, V.C., Thomas, K.M., Goren, M.J. (2009). Is multiculturalism or color blindness better for minorities? *Psychological Science*, 20, 444–6.
- Puce, A., Allison, T., Gore, J.C., McCarthy, G. (1995). Face-sensitive regions in human extrastriate cortex studied by functional MRI. *Journal of Neurophysiology*, 74(3), 1192–9.
- Ratner, K.G., Kaul, C., Van Bavel, J.J. (2013). Is race erased? Decoding race from patterns of neural activity when skin color is not diagnostic of group boundaries. *Social Cognitive and Affective Neuroscience*, 8(7), 750–5.
- Richeson, J.A., Shelton, J.N. (2003). When prejudice does not pay: effects of interracial contact on executive function. *Psychological Science*, 14(3), 287–90.
- Schofield, J.W. (2007). The colorblind perspective in school: causes and consequences. In: Banks, J.A., Banks, C.A.M., editors. *Multicultural Education: Issues and Perspectives*. New York, NY: Wiley, pp. 271–95.
- Sergent, J., Ohta, S., MacDonald, B. (1992). Functional neuroanatomy of face and object processing. A positron emission tomography study. *Brain*, 115, 15–36.
- Sherman, J.W., Macrae, C.N., Bodenhausen, G.V. (2000). Attention and stereotyping: cognitive constraints on the construction of meaningful social impressions. *European Review of Social Psychology*, 11, 145–75.
- Stangor, C., Lynch, L., Duan, C., Glass, B. (1992). Categorization of individuals on the basis of multiple social features. *Journal of Personality and Social Psychology*, 62, 207–81.
- Taylor, S.E., Fiske, S.T., Etoff, N.L., Ruderman, A. (1978). Categorical and contextual bases of person memory and stereotyping. *Journal of Personality and Social Psychology*, 36, 778–93.
- Van Bavel, J.J., Cunningham, W.A. (2009). Self-categorization with a novel mixed-race group moderates automatic social and racial biases. *Personality and Social Psychology Bulletin*, 35, 321–35.
- Van Bavel, J.J., Cunningham, W.A. (2011). A social neuroscience approach to self and social categorisation: a new look at an old issue. *European Review of Social Psychology*, 21(1), 237–84.
- Van Bavel, J.J., Cunningham, W.A. (2012). A social identity approach to person memory: group membership, collective identification, and social role shape attention and memory. *Personality and Social Psychology Bulletin*, 38(12), 1566–78.
- 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, 1131–9.
- Van Bavel, J.J., Packer, D.J., Cunningham, W.A. (2011). Modulation of the fusiform face area following minimal exposure to motivationally relevant faces: evidence of in-group enhancement (not out-group disregard). *Journal of Cognitive Neuroscience*, 23(11), 3343–54.
- Van Bavel, J.J., Xiao, Y.J., Cunningham, W.A. (2012). Evaluation as a dynamic process: moving beyond dual system models. *Social and Personality Psychology Compass*, 6(6), 438–54.
- Van Bavel, J.J., Xiao, Y.J., Hackel, L. (2013). Social identity shapes social perception and evaluation: using neuroimaging to look inside the social brain. In: Derk, B., Scheepers, D., Ellemers, N., editors. *The Neuroscience of Prejudice and Intergroup Relations*. New York: Psychology Press.
- Wheeler, M.E., Fiske, S.T. (2005). Controlling racial prejudice: social-cognitive goals affect amygdala and stereotype activation. *Psychological Science*, 16(1), 56–63.