



## OPEN ACCESS

## EDITED BY

Eduar Herrera,  
ICESI University, Colombia

## REVIEWED BY

Ye Li,  
Intuitive Surgical, Inc., United States  
Jungmi Choi,  
Human Anti-Aging Standards Research  
Institute, Republic of Korea

## \*CORRESPONDENCE

Adrian Nestor  
✉ adrian.nestor@utoronto.ca

RECEIVED 11 December 2024

ACCEPTED 17 February 2025

PUBLISHED 05 March 2025

## CITATION

Shoura M, Liang YZ, Sama MA, De A and  
Nestor A (2025) Revealing the neural  
representations underlying other-race face  
perception.  
*Front. Hum. Neurosci.* 19:1543840.  
doi: 10.3389/fnhum.2025.1543840

## COPYRIGHT

© 2025 Shoura, Liang, Sama, De and Nestor.  
This is an open-access article distributed  
under the terms of the [Creative Commons  
Attribution License \(CC BY\)](#). The use,  
distribution or reproduction in other forums is  
permitted, provided the original author(s) and  
the copyright owner(s) are credited and that  
the original publication in this journal is cited,  
in accordance with accepted academic  
practice. No use, distribution or reproduction  
is permitted which does not comply with  
these terms.

# Revealing the neural representations underlying other-race face perception

Moaz Shoura, Yong Z. Liang, Marco A. Sama, Arijit De and  
Adrian Nestor\*

Department of Psychology at Scarborough, University of Toronto, Toronto, ON, Canada

The other-race effect (ORE) refers to poorer recognition for faces of other races than one's own. This study investigates the neural and representational basis of ORE in East Asian and White participants using behavioral measures, neural decoding, and image reconstruction based on electroencephalography (EEG) data. Our investigation identifies a reliable neural counterpart of ORE, with reduced decoding accuracy for other-race faces, and it relates this result to higher density of other-race face representations in face space. Then, we characterize the temporal dynamics and the prominence of ORE for individual variability at the neural level. Importantly, we use a data-driven image reconstruction approach to reveal visual biases underlying other-race face perception, including a tendency to perceive other-race faces as more typical, younger, and more expressive. These findings provide neural evidence for a classical account of ORE invoking face space compression for other-race faces. Further, they indicate that ORE involves not only reduced identity information but also broader, systematic distortions in visual representation with considerable cognitive and social implications.

## KEYWORDS

EEG, other-race effect, image reconstruction, visual bias, face space

## Introduction

The other-race effect (ORE), poorer recognition for faces of a different race than one's own, has been consistently observed across an array of cultures and races (Hugenberg et al., 2010; Malpass and Kravitz, 1969). Given the scope of its real-world impact, ranging from failures of eye-witness testimony (Wells and Olson, 2001) and identity verification (Sporer et al., 2022) to difficulties with social interactions (McKone et al., 2023), ORE features prominently in the study of visual cognition (for reviews see Ficco et al., 2023; Kubota et al., 2012; Rossion and Michel, 2011). Hence, much has been learned about the relative contribution of perceptual, memory and social factors to its emergence as well as about the neural systems which they engage (Herzmann et al., 2018; Hughes et al., 2019; Natu et al., 2011; Wang et al., 2023; Zhou et al., 2020). Yet, much less is known about the neural representations underlying ORE, their visual content and intrinsic biases.

Accordingly, here, we aim to uncover visual representations involved in other-race (OR) versus same-race (SR) face perception with the aid of neural decoding and image reconstruction applied to electroencephalography (EEG) data (Nemrodov et al., 2018; Nestor et al., 2020). Specifically, we investigate the representational structure and the visual content responsible for ORE in East Asian and White participants. Further, we examine the prominence of ORE in neural face processing and its dynamics.

A seminal framework in the study of ORE is that of psychological face space, a multidimensional construct in which faces are represented as points and their pairwise distances correspond to their perceived similarity (Valentine, 1991). In this space, OR faces

are often described as separate from SR ones (Jaquet et al., 2008) and more densely clustered (Byatt and Rhodes, 2004; Papesh and Goldinger, 2010). Such clustering accounts for the diminished discriminability of OR faces in experimental settings as well as for the real-world visual phenomenology (e.g., “to the uninitiated American, all Asiatics look alike, while to the Asiatic all white men look alike”; Feingold, 1914). The representational homogeneity of OR faces also finds ground at the neural level, where it elicits repetition suppression, as reported across multiple neuroimaging modalities (Hughes et al., 2019; Zhou et al., 2020). Specifically, viewing pairs of different-identity OR faces may elicit a reduction in neural response comparable to that induced by same-identity face pairs, suggesting that OR faces *look alike* even to the relevant neural population (Reggev et al., 2020; Vizioli et al., 2010).

The density of OR clusters highlights a reduction in facial identity information for OR representations and a tendency to collapse them to a prototypical face. However, the optimization of face space dimensions for SR representations (Valentine, 1991) also suggests that OR faces are encoded by suboptimal visual features corresponding to these dimensions (Dahl et al., 2014; O’Toole et al., 1991). Accordingly, a search for ORE-relevant features has led to extensive debates regarding the mis/use of facial shape and surface information (Balas et al., 2011; Balas and Nelson, 2010; Michel et al., 2013; Zhou et al., 2021), and the differential reliance on holistic/featural information (Harrison et al., 2014; Michel et al., 2006; Tanaka et al., 2004; Zhao et al., 2014). Yet, the misrepresentation of OR faces, beyond a reduction in identity information, remains to be elucidated. The present work addresses this challenge by investigating and leveraging the neural counterpart of ORE.

Neural sensitivity to ORE has previously been captured by event-related potentials (ERPs; Tüttenberg and Wiese, 2023), such as enhanced amplitude of the N170 component for OR faces (Senholzi and Ito, 2013; Walker et al., 2008; Wiese and Schweinberger, 2018); however, other studies reported no difference (Caldara et al., 2004; Tanaka and Pierce, 2009). Further, functional magnetic resonance imaging (fMRI) studies have reported higher activity in the fusiform face area (FFA) for SR than OR faces (Feng et al., 2011; Golby et al., 2001), though this difference may be more prominent for unfamiliar faces (Kim et al., 2006) and/or reflect a more general benefit for peer perception (Dai and Scherf, 2023). Relevantly here, multivariate fMRI analyses have revealed different spatial activation patterns in the ventral cortex for OR versus SR faces (Contreras et al., 2013; Natu et al., 2011; Wang et al., 2023) and have suggested different dynamics (Zhou et al., 2020).

Thus, while prior work highlights important aspects of OR perception, such as face-space clustering (Byatt and Rhodes, 2004; Papesh and Goldinger, 2010) and misrepresentation (Balas and Nelson, 2010; Zhou et al., 2021), their neural basis remains largely unexplored. Similarly, the full extent and the nature of OR face representations require broader investigation. Here, we attempt to shed light on these topics by: (i) relating neural decoding to behavioral ORE estimates; (ii) characterizing the temporal dynamics of OR versus SR perception, and (iii) retrieving the neural representations underlying ORE. To anticipate, our results show that: (i) ORE can be reliably related to differences in the neural decoding of OR versus SR faces; (ii) the neural dynamics of ORE evince an extensive time course, and (iii) OR face representations exhibit a typicality bias as well as age and expressiveness biases.

## Materials and methods

### Participants

A total of 43 adults (age range: 18–30 years; 28 females) from the University of Toronto community volunteered for the EEG experiment. After excluding three participants (2 East Asian, 1 White) due to technical difficulties with the EEG recordings, 20 identified themselves as East Asian and 20 as White. A majority (60%) of the remaining East Asian participants were international students from a Han Chinese background; in contrast, all White participants were locals (from Toronto, Ontario). All participants were right-handed, with normal or corrected-to-normal vision, and no history of neurocognitive disorders.

Prior work (Estudillo, 2021; McKone et al., 2012) on ORE with East Asian and White participants has estimated a medium effect size in both populations (Cohen’s  $d \geq 0.62$ ). A power analysis (JASP 0.17.1; [jasp-stats.org](http://jasp-stats.org)) for an effect size of 0.60 (80% power; 5% Type I error rate) indicated that a sample size of 19 participants per group is needed. Thus, our final sample size should allow detecting ORE reliably.

Further, to validate behaviorally our EEG-based reconstruction results, we recruited online (Prolific; [www.prolific.co](http://www.prolific.co)) 50 additional adults (18–35 years; 23 female), referred to below as *validators*. Two were removed because of face recognition scores below the normative range (see below) and two due to failing reliability checks (i.e., negative correlation with themselves across repeated trials). Of the remaining validators, 23 were East Asian, born in East Asian countries, and 23 White, born in North America or Europe.

All participants/validators provided informed consent and received monetary compensation. All procedures were approved by the Research Ethics Board at the University of Toronto.

### Stimuli

For the EEG experiment we selected 30 images of East Asian males and 30 of White males of similar age, with frontal view, frontal gaze and neutral expression, from the Chicago Face Database (Ma et al., 2015). Images were standardized by: (1) aligning the position of the eyes and nose; (2) cropping to display only internal facial features; (3) normalizing with the same mean/contrast values in each CIEL\*a\*b\* channel; and (4) resizing to  $98 \times 75$  pixels. Resulting stimuli subtended a visual angle of  $4^\circ \times 2.6^\circ$ .

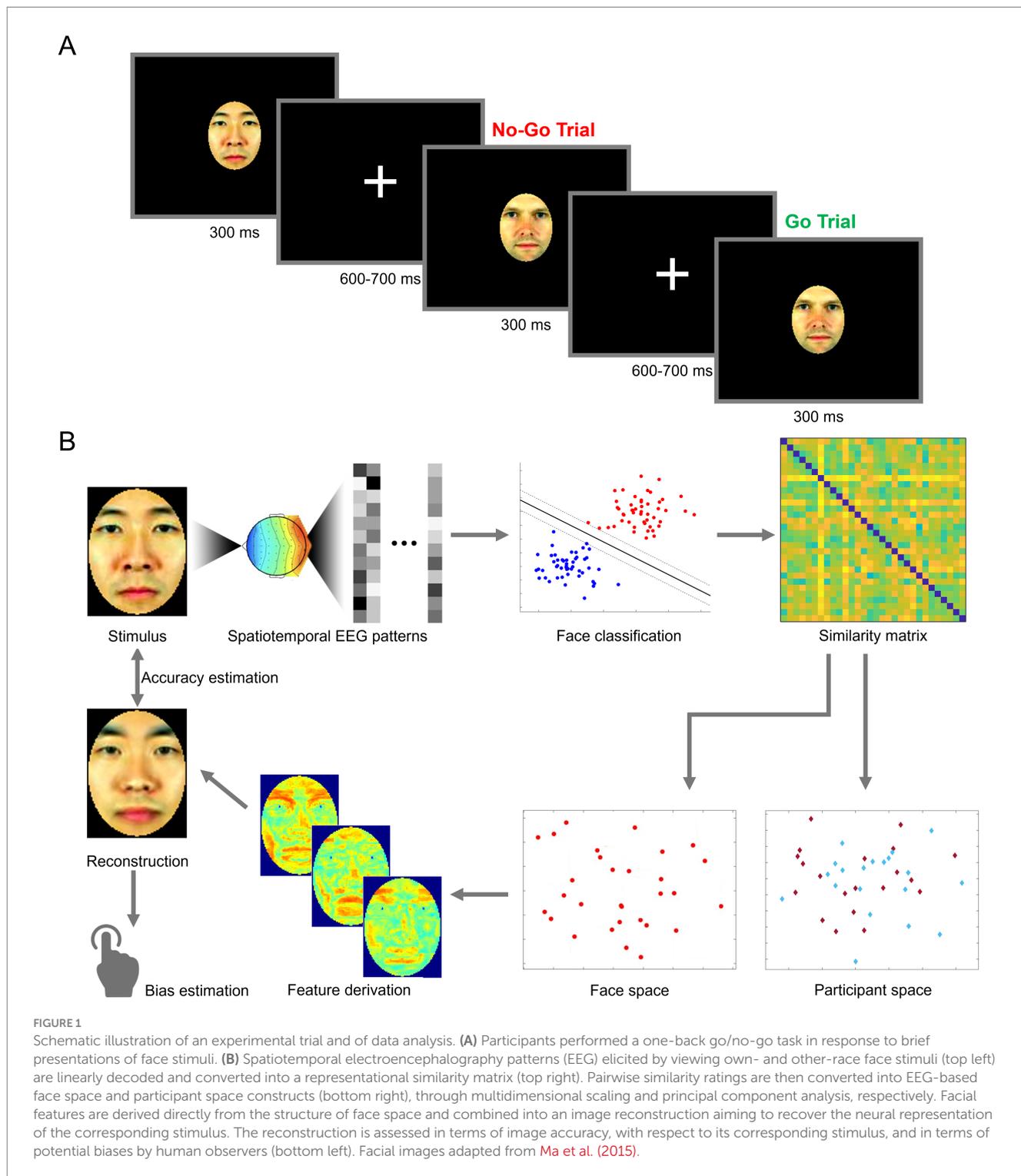
### Experimental procedures

First, participants completed two versions of the Cambridge Face Memory Test (CFMT) with Chinese (McKone et al., 2012) and White (Duchaine and Nakayama, 2006) face stimuli. This served to assess their face processing proficiency relative to the normative range (i.e., mean  $\pm$  2SD, as determined by prior work; Bowles et al., 2009; McKone et al., 2012) and, also, to behaviorally quantify ORE. Both tests have excellent psychometric properties (e.g., Cronbach’s  $\alpha$  of 0.89–0.90; McKone et al., 2012; Wilmer et al., 2010) and can be used to estimate individual ORE via the traditional subtraction method

(i.e., CFMT score for SR minus that for OR faces; [Estudillo, 2021](#); [Wan et al., 2015](#)).

EEG testing was conducted across two 2.5-h sessions on different days. Participants performed a go/no-go task by pressing a key upon the repetition of the same image on two consecutive trials. Data collection was divided across 32 blocks spanning the two sessions. Within any block, each stimulus was displayed four times, and a

random 10% of trials featured a repetition (for a total of 264 trials/block). Stimuli were presented in pseudorandom order (e.g., to avoid repeated “go” trials). Each stimulus was displayed for 300 ms, followed by a variable 600–700 ms inter-stimulus interval during which a center fixation cross replaced the stimulus (see [Figure 1A](#)). Each session began with a practice block. Stimulus presentation and data collection relied on MATLAB and Psychtoolbox 3.0.8 ([Brainard, 1997](#);



Pelli, 1997). Both sessions were identical in structure, with stimuli presented on a 60 Hz, 1920 × 1080p display, from a distance of 80 cm.

## Analyses

### Behavioral data analysis

Own-race CFMT scores were inspected to confirm that participants exhibit normal face recognition abilities. A two-way ANOVA (2 CFMT versions: East Asian, White; 2 participant groups: East Asian, White) was followed by within-group comparisons (paired *t*-tests) to confirm and compare ORE size.

To add confidence in significant findings and to aid the interpretation of statistically nonsignificant results, we employed Bayesian hypothesis testing (JASP 0.17.1) to report  $BF_{10}$  values, providing weight in favor of the alternative ( $BF_{10} > 1$ ) or null ( $BF_{10} < 1$ ) hypotheses. The analysis relied on standard default distributions for unspecified priors (e.g., uniform distribution for ANOVA analogs and Cauchy's distribution for *t*-test analogs). We interpret results using established benchmarks (BF values in favor of alternative hypotheses: 1–3 anecdotal, 3–10 substantial, 10–30 strong, 30–100 very strong, > 100 extreme; and in favor of null hypothesis: 0.33–1 anecdotal, 0.10–0.33 substantial, 0.01–0.03 very strong, < 0.01 extreme; Wagenmakers et al., 2011).

### EEG acquisition and preprocessing

EEG signals were recorded with a 64-channel Biosemi ActiveTwo system (Biosemi B.V.). Electrodes were configured according to the international 10/20 system. Offsets were maintained under a threshold of 40 mV. The signal was low-pass filtered using a fifth order sinc filter with a half-power cutoff at 204.8 Hz and, then, digitized at 512 Hz with 24 bits of resolution. All data were digitally filtered offline (zero-phase 24 dB/octave Butterworth filter) with a bandpass of 0.1–40 Hz and segmented into epochs extending from –100 to 900 ms. Epochs underwent direct current (DC) removal, linear detrending and baseline correction. Noisy electrodes were interpolated if necessary (no more than two electrodes per participant) and epochs were re-referenced to the average reference. Preprocessing relied on Letswave 6 (Mouraux and Iannetti, 2008) and Infomax ICA (Delorme et al., 2007) for artifact removal.

Relatively few trials contained false alarms as participants performed the task at ceiling (mean sensitivity  $d' = 2.73$ ). “Go” trials and those containing artifacts or false alarms were excluded, resulting in an average of 98.1% remaining trials across participants for further analysis.

### Pattern classification analyses

We selected 12 electrodes positioned over occipitotemporal (OT) regions (P5, P7, P9, PO3, PO7, O1 on the left, and P6, P8, P10, PO4, PO8, O2 on the right). Their choice was based on their value in decoding facial information in the temporal domain (e.g., as previously determined by multivariate channel selection; Nemrodov et al., 2020) – for univariate results, see [Supplementary materials and Supplementary Figure S1](#).

Consistent with this choice, decoding analyses relying on all electrodes yielded lower, though still significant, levels of classification accuracy compared to those relying on all electrodes (see [Supplementary Figure S2](#)).

With respect to the temporal information included in the analysis, two different approaches were employed. First, for temporally-cumulative analyses (Nemrodov et al., 2018; Roberts et al., 2019), we selected a large 50–650 ms time window and, then, concatenated ERP amplitude values across electrodes and time points resulting in 3684-feature patterns (i.e., 12 OT electrodes by 307 time points) for each trial. The length of this windows aimed to capture both early and late stages of face processing (Ghuman et al., 2014; Nemrodov et al., 2016; Roberts et al., 2019). Patterns corresponding to repetitions of the same stimulus within a block were averaged and normalized via z-scoring separately for each feature (i.e., for each combination of time point and OT channel). This yielded, for each participant, a total of 32 observations, one per block, for each stimulus.

Second, for time-resolved analyses (Dobs et al., 2019; Nemrodov et al., 2016), we considered smaller ~10 ms windows (i.e., five ~1.95 ms adjacent time bins). Decoding performance was estimated at each position, for a total of 508 intervals, by sliding this window one bin at a time between –100 and 700 ms.

Pairwise classification was then computed for each participant and stimulus pair, for a total of 1770 pairs across face identity and race. To this end, we used linear SVM ( $c = 1$ ) with leave-one-block-out cross-validation (across 32 blocks from both experimental sessions). The procedure yielded decoding estimates (i.e., classification accuracy) for each pair of faces. These estimates were then separately assessed for: (1) SR faces (e.g., White faces for White participants), (2) OR faces (e.g., White faces for East Asian participants), and (3) between-race faces (i.e., White versus East Asian faces for both participant groups). Chance-level accuracy was derived by randomly shuffling stimulus labels 1,000 times and recomputing classification estimates. Average accuracy across participants was compared to permutation-based chance both for temporally-cumulative analyses (one-sample two-tailed *t*-tests, Bonferroni-corrected across comparisons) and time-resolved analyses (Wilcoxon signed-rank test, FDR-corrected across time points).

### EEG-based face space and participant space

To visualize the representational space underlying race perception, we derived an EEG-based face space for each group of participants. To this end, decoding estimates derived from temporally-cumulative analyses for all pairs of faces were averaged across participants in each group and organized into a dissimilarity matrix. Then, metric multidimensional scaling (MDS) was applied to each matrix to derive a face space construct for each participant group (i.e., 40-dimension spaces accounting for at least 90% variance). We anticipated that faces in this space would be clearly separated by race and, also, that OR faces would cluster more tightly than SR ones. To this aim, we computed and compared average pairwise Euclidean distances between each pair of SR faces, each pair of OR faces and each pair of between-race faces separately for each participant group and corresponding space.

A complementary analysis sought to estimate a participant space and account for its structure. To this end, principal component analysis (PCA) was applied to EEG-based decoding accuracy vectors across all participants. Specifically, for each participant we considered a decoding performance vector consisting of the pairwise classification accuracies across all stimuli (i.e., a 1770-element vector yielded by all pairs of 60 stimuli). These vectors were then concatenated across all participants (i.e., into a 1770 × 40 matrix).

Following the application of PCA to this matrix, the first two components of the resulting space were probed for their relationship with behavioral metrics: we assessed the contribution of participant race, facial recognition ability (as captured by SR CFMT scores) and ORE (captured by the difference between SR and OR CFMT scores; Estudillo, 2021; Wan et al., 2015) to these components. To this end, behaviorally derived scores were correlated with PC scores for each dimension.

### Image reconstruction procedure

The reconstruction procedure was applied to neural data to recover and to visualize internal representation associated with each face, as well as to probe for potential visual misrepresentations associated with ORE. Our reconstruction approach follows earlier efforts (Nemrodov et al., 2018; Roberts et al., 2019), which leverage the face space framework to extract visual features and recover neural representations by harnessing the rich spatiotemporal information of EEG patterns (see Figure 1B). This approach involves a sequence of steps, adapted from prior work (for review see Nestor et al., 2020).

First, estimates of pairwise decoding accuracy for all faces of the same race were placed into a dissimilarity matrix and averaged across participants, separately for each group. To reconstruct a target face, this was removed from the data and the structure of the resulting matrix was used to estimate the representation of the target while avoiding dependency.

Second, metric multidimensional scaling (MDS) was applied to the dissimilarity matrix to derive a face space. A total of 20 dimensions were retained (accounting for over 90% variance separately for each group and stimulus race).

Third, for each dimension of face space, corresponding coefficients were normalized via *z*-scoring and a classification image (CI) was computed through reverse correlation (see Murray, 2011, for review). Specifically, we computed the average of all non-target images weighted by their coefficients on the dimension considered.

Fourth, significant CI information was assessed using a pixelwise test. To this end, images were shuffled relative to their coefficients on a given dimension and a new CI was derived (for a total of 1,000 permutations). Then, the value of each pixel in the original CI was compared to those at the same location in the permutation-based CIs. Significant pixels were identified, separately for each CIE  $L^*a^*b^*$  color channel (permutation test, FDR-corrected across pixels,  $q < 0.1$ ). Informative CIs/dimensions were determined based on the presence of at least 10 significant pixels in at least one color channel. All other CIs were eliminated from the procedure ensuring that only significant facial information was included in the reconstruction.

Fifth, the target face was projected into the existing face space. This was computed by Procrustes alignment of a space *X* containing all faces, including the target, to the space *Y* used for feature derivation (and not including the target). Then, the target was projected from *X* to *Y* using the alignment coefficients. This procedure ensures that the features of the target are not used in its reconstruction, thus avoiding dependency.

Finally, each significant CI was weighted by its coordinates in the corresponding dimension. Then, the resulting images were linearly combined across dimensions, effectively generating an image analog of a given location in face space.

The approach above was applied separately for each face (i.e., by treating each face as a reconstruction target). By considering decoding

patterns for SR and OR faces, we aimed to recover both types of visual representations separately in East Asian and White participants.

Further, image reconstructions were also computed based on a theoretical observer. To this end, the approach above was applied to L2 pixelwise distances across stimuli (instead of neural-based dissimilarity estimates.) Since stimuli were matched in terms of contrast, pose, gaze, cropping, etc., this similarity space, though based on a low-level metric, should still reflect relative information about the visual appearance of different individuals. This analysis yielded comparable levels of reconstruction accuracy for the two races (90.57 and 92.18% for East Asian and White faces, respectively; Mann–Whitney U test across stimulus faces,  $p = 0.221$ ,  $BF_{10} = 0.468$ ). Importantly, this ensured that any potential differences in neural-based reconstruction could not be explained away by obvious differences in the image properties of the two stimulus races.

### Image reconstruction evaluation

The accuracy of the image reconstruction for any particular face was assessed by determining the proportion of instances on which a reconstructed image was closer to its target stimulus than to any other face of the same race via a pixelwise L2 distance. To be clear, 100% reconstruction does not indicate a perfect replication of the corresponding stimulus, but only that the reconstruction was more similar to this stimulus than to any other stimulus of the same race. Then, accuracy was compared against chance (i.e., a one-sample test against 50%) and against each other for reconstructions of different stimulus races using a bootstrap test (10,000 iterations).

To rule out the possibility that any differences between OR and SR reconstructions simply reflect differences in image quality (e.g., due to image blur, pixel noise, spatial distortions) we computed estimates of each reconstructed image via two complementary metrics. Specifically, we appealed to a common reference-based metric, the structural similarity index (SSIM; Wang et al., 2004), as well as to a reference-free metric that approximates perceptual judgements, the blind/referenceless image spatial quality evaluator (BRISQUE; Mittal et al., 2012). OR and SR estimates were then compared to each other using a Wilcoxon signed-rank test separately for each metric and stimulus race.

Last, to assess local, low-level pictorial differences of SR versus OR face reconstructions, we subtracted corresponding images generated from the two groups (i.e., a reconstruction of a given face based on the data of White participants from a reconstruction of the same facial identity based on data from East Asian participants). Then, the outcomes were assessed with a pixelwise permutation-based test. Specifically, we randomly shuffled the labels of the participants across groups (i.e., East Asian, White), recomputed average dissimilarity matrices, reconstructed face images for each group and subtracted corresponding images generated from the two groups. The initial image differences were then compared with their permutation-based counterparts to identify pixels yielding values different from chance (two-tailed pixelwise permutation test, 1,000 permutations; FDR-corrected across pixels, separately for each color channel). This analysis was conducted for every facial identity as well as for the average of all facial identities of the same race.

### Behavioral evaluation of reconstruction results

A different group of participants (*validators*) evaluated and compared the reconstruction results. First, validators completed the

two versions of CFMT, with East Asian (McKone et al., 2012) and White face stimuli (Duchaine and Nakayama, 2006), to assess face processing abilities and ORE.

Then, they viewed 120 image reconstructions (i.e., 30 East Asian faces, reconstructed twice, from each group of EEG participants, and 30 White faces, also reconstructed twice). Since these stimuli are reconstructions of percepts elicited by stimuli in our EEG experiment, their appearance is standardized the same manner (e.g., with respect to size).

On each trial, validators viewed pairs of reconstructions of the same face identity (i.e., derived from East Asian versus White participant data) and selected the face which appeared younger, more expressive, or more typical of its race. Faces were displayed until a response was recorded, with self-paced breaks between blocks to minimize fatigue. Each pair of reconstructions was presented twice, by swapping the left/right position of the reconstructions on the screen. The experiment comprised 6 blocks (3 facial attributes  $\times$  2 stimulus races) of 60 trials (30 reconstructions of a given race presented twice). Block and trial order were randomized. Stimulus presentation and data collection relied on PsychoPy (Peirce et al., 2019). Validators completed testing during a 45-min online session.

After data collection, selection rates were averaged across trials separately by each stimulus race and facial trait for each validator. Selection rates were coded so that values above 50% indicate that reconstructions from OR participants (i.e., White face images reconstructed from East Asian participants, or vice versa) appear younger, more expressive, or more typical of their race. Conversely, a score below 50% indicate that validators judge reconstructions from their own racial group in this manner. Selection scores were then compared to chance (one-sample *t*-tests across validators against 50%, Bonferroni-corrected) and with each other (three-way mixed ANOVA; two validator groups: East Asian and White, two stimulus races, and three facial traits: age, expressiveness, typicality). Last, trait-specific selection rates, averaged across validators, were correlated with each other across all facial identities of the same race (e.g., age and expressiveness rates for East Asian faces) to examine whether judgments capture same/different underlying bias(es) across traits.

## Results

### Behavioral performance

Face recognition abilities were assessed using the CFMT with Chinese (McKone et al., 2012) and White (Duchaine and Nakayama, 2006) face stimuli across all participants. An analysis of recognition performance (two-way ANOVA; 2 test versions: Chinese, White  $\times$  2 participant groups: East Asian, White) revealed an effect of test version ( $F(1,38) = 10.42$ ,  $p = 0.003$ ,  $\eta_p^2 = 0.22$ ,  $BF_{10} = 2.01$ ) and participant race ( $F(1,38) = 10.90$ ,  $p = 0.002$ ,  $\eta_p^2 = 0.22$ ,  $BF_{10} = 17.85$ ), along with an interaction ( $F(1,38) = 38.1$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.50$ ,  $BF_{10} = 35.69$ ). *Post hoc* tests revealed that, as expected, White participants outperformed East Asian participants on the CFMT-White ( $t(38) = 5.43$ ,  $p < 0.001$ ,  $d = 1.72$ ); however, the two participant groups performed comparably on the CFMT-Chinese ( $t(38) = 0.66$ ,  $p > 0.999$ ). The latter result may be due to participant background, as ORE tends to be diminished in cities with a diverse multicultural

population (Zhou et al., 2022) –all White participants were locals from a highly diverse, multicultural city (see Participants).

More importantly, ORE was confirmed by comparing SR to OR CFMT scores both for East Asian (two-tailed paired *t*-test,  $t(19) = 5.78$ ,  $p < 0.001$ ,  $d = 1.29$ ,  $BF_{10} = 3178.03$ ) and White participants ( $t(19) = 2.53$ ,  $p = 0.02$ ,  $d = 0.57$ ,  $BF_{10} = 5.60$ )—see Figure 2A. Overall, these results confirm the presence of ORE in both groups, provide evidence for the robustness of our behavioral measures and motivate our investigation into their relationship with neural-based effects below.

### Neural decoding of SR and OR faces

Pairwise face decoding relying on temporally-cumulative decoding (Nemrodov et al., 2018; Roberts et al., 2019) revealed above-chance classification accuracy within and across stimulus race for both participant groups (one-sample *t*-tests across participants against permutation-based chance, all  $p$ 's  $< 0.001$ , all  $d$ s  $> 2.21$ , all  $BF_{10}$  values  $> 3.71 \times 10^6$ ) – see Figure 2B. An assessment of classification accuracy (two-way ANOVA; 2 stimulus races  $\times$  2 participant groups) revealed no main effects but a significant interaction ( $F(1,38) = 26.98$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.42$ ,  $BF_{10} = 36.1$ ). Subsequent tests indicated that the decoding performance for East Asian faces was comparable in the two participant groups ( $t(38) = 0.22$ ,  $p = 0.83$ ). In contrast, White faces were marginally better decoded by White participants compared to East Asian participants ( $t(38) = 2.55$ ,  $p = 0.058$ ). We note that these results mirror the pattern of behavioral results described above, which we ascribe to participant background.

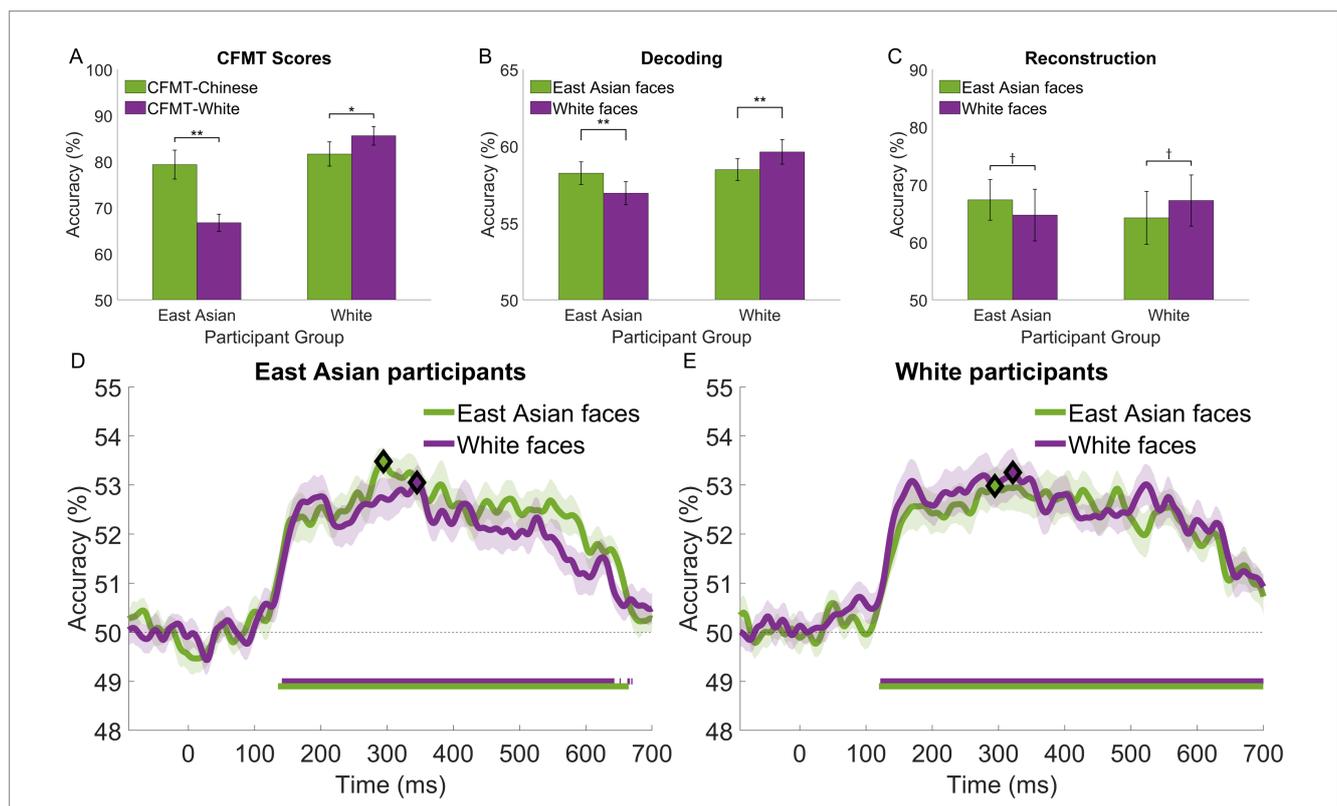
Importantly, a comparison of SR and OR face decoding revealed a significant advantage for the former both for East Asian ( $t(19) = 3.29$ ,  $p = 0.004$ ,  $d = 0.74$ ,  $BF_{10} = 22.62$ ) and White participants ( $t(19) = 4.51$ ,  $p < 0.001$ ,  $d = 1.0$ ,  $BF_{10} = 245.31$ ). Arguably, these results provide evidence for a neural-based ORE counterpart.

Further, between-race face decoding yielded higher decoding than within-race decoding both for East Asian participants (relative to White faces:  $t(19) = 8.41$ ,  $p < 0.001$ ,  $d = 1.88$ ,  $BF_{10} = 3.71 \times 10^5$  and East Asian faces:  $t(19) = 3.58$ ,  $p = 0.002$ ,  $d = 0.8$ ,  $BF_{10} = 40.62$ ) and for White participants (relative to White faces  $t(19) = 5.36$ ,  $p < 0.001$ ,  $d = 1.2$ ,  $BF_{10} = 1405.80$  and East Asian faces:  $t(19) = 6.85$ ,  $p < 0.001$ ,  $d = 1.53$ ,  $BF_{10} = 2.40 \times 10^4$ ).

We note that the present results are based on data from two experimental sessions for all participants. Separate session-specific analyses only yielded a significant SR advantage for the first but not the second session, potentially indicative of perceptual learning effects (see Supplementary Figure S3).

### Neural dynamics

To evaluate the temporal profile of face processing, decoding was conducted similarly over successive ~10-ms intervals (instead of a large window spanning most of the trial). For both participant groups and stimulus races, we note above-chance decoding (Wilcoxon signed-rank test against permutation-based chance; FDR-corrected,  $q < 0.05$ ) starting around 130 ms, peaking between 300–350 ms and tapering off after 600 ms – see Figures 2D,E for East Asian and White participants, respectively.



**FIGURE 2** Performance associated with own- and other-race face perception for East Asian and White participants: (A) face recognition accuracy and other-race effects estimated with CFMT-Chinese and CFMT-White; (B) temporally-cumulative pairwise stimulus decoding by stimulus race (based on a 50–650 ms interval × 12 electrode patterns); (C) accuracy of EEG-based image reconstructions; (D,E) time-resolved face decoding (based on a 10 ms sliding window × 12 electrode patterns) for East Asian and White participants, respectively. Participants exhibit a systematic own-race advantage, which reaches significance (two-tailed *t*-tests across stimulus race) for (A,B). The time course of face decoding shows extensive intervals of above-chance decoding (Wilcoxon signed-rank test against permutation-based chance; FDR-corrected across time points,  $q < 0.05$ ; horizontal bars at the bottom of each plot) and higher estimates of own-race face decoding (though not significant after FRD correction). Diamonds mark the peaks of decoding accuracy in (D,E). Error bars for (A–C) and shaded areas for (D), (E) indicate  $\pm 1$  SE ( $^{\dagger} p < 0.10$ ,  $^* p < 0.05$ ,  $^{**} p < .01$ ).

Interestingly, for both groups, SR faces supported higher decoding accuracy across most time points after 130 ms, though not significantly after multiple-comparison correction (Wilcoxon signed-rank tests across stimulus race,  $q < 0.05$ ). This suggests that the SR decoding advantage found by temporally-cumulative analysis is not due to any specific, restricted time window. Rather, it likely relies on aggregating complementary information over an extended interval. Accordingly, our results below capitalize on temporally-cumulative results rather than on temporally-restricted ones (e.g., such as those corresponding to decoding peaks).

### Neural-based face space

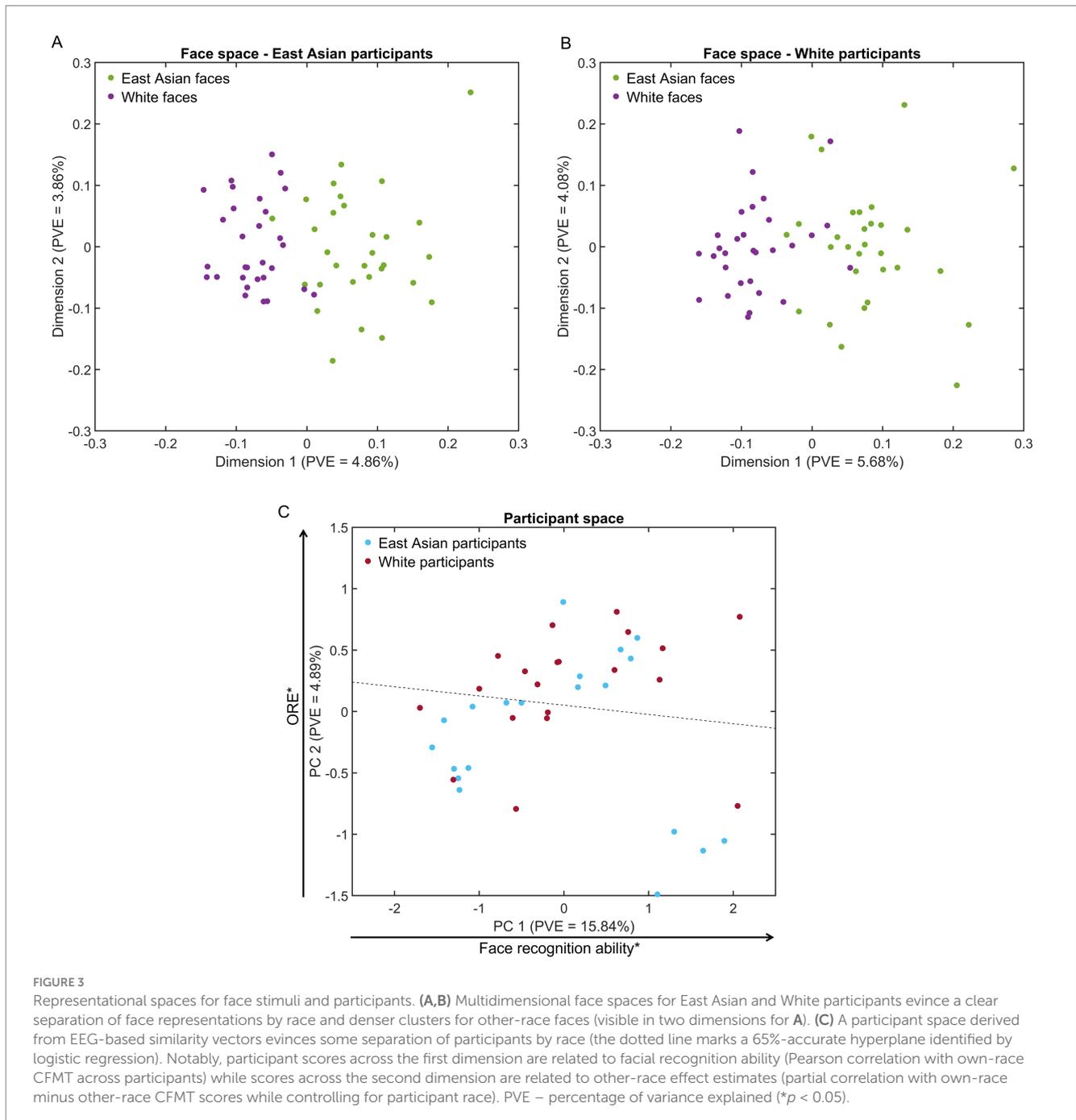
Accuracy estimates of pairwise face decoding, averaged across participants, were converted via multidimensional scaling (MDS) into a face space separately for each group (i.e., 40-dimension spaces accounting for at least 90% variance).

The results evince a clear separation of faces by race (see Figures 3A,B). Additionally, we note a higher density for OR than SR faces, consistent with face space theory (Valentine, 1991) – in two dimensions this latter result is more clearly apparent for East Asian participants (Figure 3A). To assess this difference more rigorously,

we computed pairwise face distances, using a Euclidean metric, across all dimensions. A comparison of such distances indicate that SR distances are, on average, 2.78% larger than OR ones for East Asian participants and 1.44% larger for White participants (Mann–Whitney *U* test, both  $p$ 's  $< 0.001$ ,  $BF_{10}$  values  $> 42.76$ ) – see also Supplementary Figure S4 for heatmaps of all pairwise distances. Similarly, between-race distances are, on average, 2.36 and 4.34% larger than SR ones for East Asian and White participants (both  $p$ 's  $< 0.001$ ,  $BF_{10}$  values  $> 8303.45$ ).

To confirm the robustness of the results above, we considered the potential impact of stimulus outliers. To this end, we identified stimuli whose distance from the center of their corresponding face group was larger than 2.5 SDs in face space, separately for each participant and stimulus race. This procedure identified one East Asian stimulus outlier for East Asian participants and one outlier per stimulus race for White participants. A comparison of SR, OR and between-race distances after outlier removal replicated the pattern of results above (all  $p$ 's  $< 0.001$ , all  $BF_{10}$  values  $> 82.05$ ).

These results reflect differences in decoding accuracy, as reported above, both within and across groups. Specifically, SR faces are better decoded than OR ones, leading to larger pairwise distances in face space for the former. At the same time, between-race faces are better decoded than within-race faces leading to



even larger distances. Interestingly though, we note the prominence of race-based separation (i.e., along the first dimension of face space). These results, based on neural data, complement face space accounts of race informed by behavioral (Byatt and Rhodes, 2004; Papesh and Goldinger, 2010) and computational work (O'Toole et al., 1991; Caldara and Abdi, 2006; Wang et al., 2023).

To be clear, not all dimensions are likely to be (equally) informative regarding visual face representations and, thus, the estimates above can only serve as a coarse indication of ORE. Hence, a more careful examination of representational content and its sensitivity to ORE has to consider dimension/feature selection (e.g., as implemented by image reconstruction below).

## Relationship between neural decoding and behavioral performance

Behavioral and neural-based ORE estimates were computed for each participant via subtraction (i.e., SR-OR scores; Estudillo, 2021; Wan et al., 2015) from CFMT and decoding scores, respectively. These estimates were related to each other across participants from both groups (Pearson correlation,  $r(38) = 0.68$ ,  $p < 0.001$ ,  $BF_{10} = 1.30 \times 10^4$ ). To assess whether this correlation reflected within-group differences, and not just a categorical difference across participant race, we conducted a partial correlation while controlling for participant race. This analysis yielded similar results ( $r(38) = 0.41$ ,  $p = 0.009$ ,  $BF_{10} = 5.53$ ). The

outcome highlights the relationship between behavioral and neural-based ORE.

Further, an exploratory analysis of EEG-based participant space, estimated via PCA from pairwise decoding results across participants, revealed some separation between the two participant groups, primarily along the second dimension. To visualize and quantify this separation, participant race was classified in PC space via logistic regression (i.e., across vectors of PCA coefficients corresponding to each participant). The classification reached an accuracy of 65% – see [Figure 3C](#) for the separating hyperplane.

To gain more insight into the structure of this space, as reflected by its first two dimensions, we assessed its relationship with estimates of facial recognition: own-race CFMT scores and ORE scores (i.e., SR – OR CFMT scores). We found that PC1 significantly correlated with SR CFMT scores ( $r(38) = 0.35, p = 0.029, BF_{10} = 1.98$ ), while PC2 evinced a significant correlation with ORE scores ( $r(38) = 0.37, p = 0.02, BF_{10} = 2.70$ ). Given that the stimuli and procedures used in our EEG experiment are considerably different from those used by the CFMT, these results demonstrate how race-related effects in face perception generalize across different stimulus sets and experimental procedures as well as across different data types. At the same time though, we acknowledge the need to confirm such correlational results using larger participant samples.

Thus, a neural-based participant space appears to be structured primarily by face recognition ability and ORE. Our present findings bolster the prominence of ORE at the neural level and provides ground for our investigation into the neural representations underlying ORE.

## Image reconstruction of SR and OR faces

EEG-based image reconstruction ([Nemrodov et al., 2018](#); [Nestor et al., 2020](#)) recovered the visual content of SR and OR face representations. Above-chance reconstruction accuracy was found for both participant groups and stimulus races (two-tailed bootstrap test; all  $p$ 's < 0.001). Also, a trend for higher SR relative to OR accuracy ([Figure 2C](#)) was found for both East Asian (two-tailed bootstrap test;  $p = 0.062$ ) and White participants ( $p = 0.065$ ). We note that theoretical observer reconstructions supported equivalent levels of accuracy for the two races (see Image reconstruction procedure). Thus, any differences based on neural results are likely to stem from differences in the perceptual representations of our participants rather than, simply, from those in the low-level image properties of our stimuli.

Next, we considered that lower OR reconstruction accuracy may reflect poorer image quality relative to its SR counterpart (e.g., due to loss of high-frequency spatial information or to spatial distortions introduced by the reconstruction method). Accordingly, we evaluated image quality via a common reference-based metric (SSIM) of each reconstruction relative to its corresponding stimulus as well as via a reference-free metric (BRISQUE). A comparison of OR to SR reconstructions revealed no significant difference by either metric for faces of either race (Wilcoxon signed-rank test; all  $p$ 's > 0.42, all  $BF_{10}$  values < 0.329, except for East Asian face reconstructions assessed with BRISQUE,  $p = 0.073, BF_{10} = 1.73$ ). Hence, we conclude that any differences between OR and SR reconstructions do not reflect mere image quality differences (e.g., image blur).

Further, to assess local, low-level pictorial differences between SR and OR reconstructions, a pixelwise permutation test was conducted, separately for each facial identity and color channel (two-tailed test, FDR corrected across pixels,  $q < 0.05$ ). This revealed differences for all channels ([Figure 4A](#)) across multiple facial areas (e.g., around the eyes, eyebrows, nose). The analysis was repeated for image reconstructions averaged across all facial identities of each race (e.g., all Asian face reconstructions based on data from East Asian participants vs. those based on data from White participants) – see [Figure 4B](#). However, the results are not immediately interpretable in terms of a systematic bias and somewhat less informative for White face reconstructions.

These results suggest that any divergence between SR versus OR image reconstructions may not be fully captured by low-level visual estimates. Accordingly, next, we asked a separate group of participants to assess the degree of race typicality for each of four sets of reconstructions (i.e., East Asian and White facial image reconstructions recovered from both groups of participants). In addition, we evaluated the possibility that OR visual representations are biased with respect to other facial traits, such as age ([Dehon and Brédart, 2001](#); [Mousavi et al., 2018](#)) and expressiveness ([Jiang et al., 2023](#); [Yan et al., 2016](#)), and asked validators to also judge these traits.

## Behavioral evaluation of image reconstructions

For each facial identity, validators viewed and compared its corresponding reconstructions derived from East Asian and White participants. Specifically, on separate trials, they selected the image in a pair that appeared younger, more expressive and more typical for its own race.

Average scores above 50% were noted for all three traits and both stimulus races (one-sample two-tailed  $t$ -tests, all  $p$ 's < 0.001, all  $d$ s > 0.58, Bonferroni-corrected, all  $BF_{10}$  values > 109). These results indicate that, overall, OR faces are perceived as younger, more expressive, and more typical of their race ([Figure 4C](#)).

An assessment of selection scores (three-way mixed ANOVA; two validator groups: East Asian and White, two stimulus races, three facial traits) found no main effect of group ( $F(1,44) = 0.46, p = 0.50, BF_{10} = 0.232$ ) or trait ( $F(2,88) = 1.84, p = 0.17, BF_{10} = 0.178$ ) but a main effect of stimulus race (higher scores for East Asian than White faces:  $F(1,44) = 43.23, p < 0.001, \eta_p^2 = 0.50, BF_{10} = 3.54 \times 10^5$ ) as well as an interaction between stimulus race and trait ( $F(2,88) = 7.39, p = 0.001, \eta_p^2 = 0.14, BF_{10} = 6.21 \times 10^4$ ). Subsequent tests indicated that age and typicality scores were significantly more pronounced for East Asian than White face stimuli ( $t(44) = 6.13, p < 0.001, d = 1.30$  and  $t(44) = 5.87, p < 0.001, d = 1.25$ , respectively) but not expressiveness ( $t(44) = 1.92, p = 0.23$ ).

Next, trait-specific selection rates, averaged across participants, were correlated with each other across all facial identities of the same race (i.e., six correlations for each pair of traits and each stimulus race). All correlations were significant (all  $p$ 's < 0.001, all  $BF_{10}$  values > 28.36, except for the correlation between expression and typicality of East Asian stimuli,  $p = 0.004, BF_{10} = 12.91$ ; Bonferroni-corrected). To assess whether results were entirely driven by typicality, next, we computed partial correlations across age and expressiveness while controlling for typicality. Again, correlations reached significance for both stimulus races (both  $p$ 's < 0.001). Last, for completeness, these

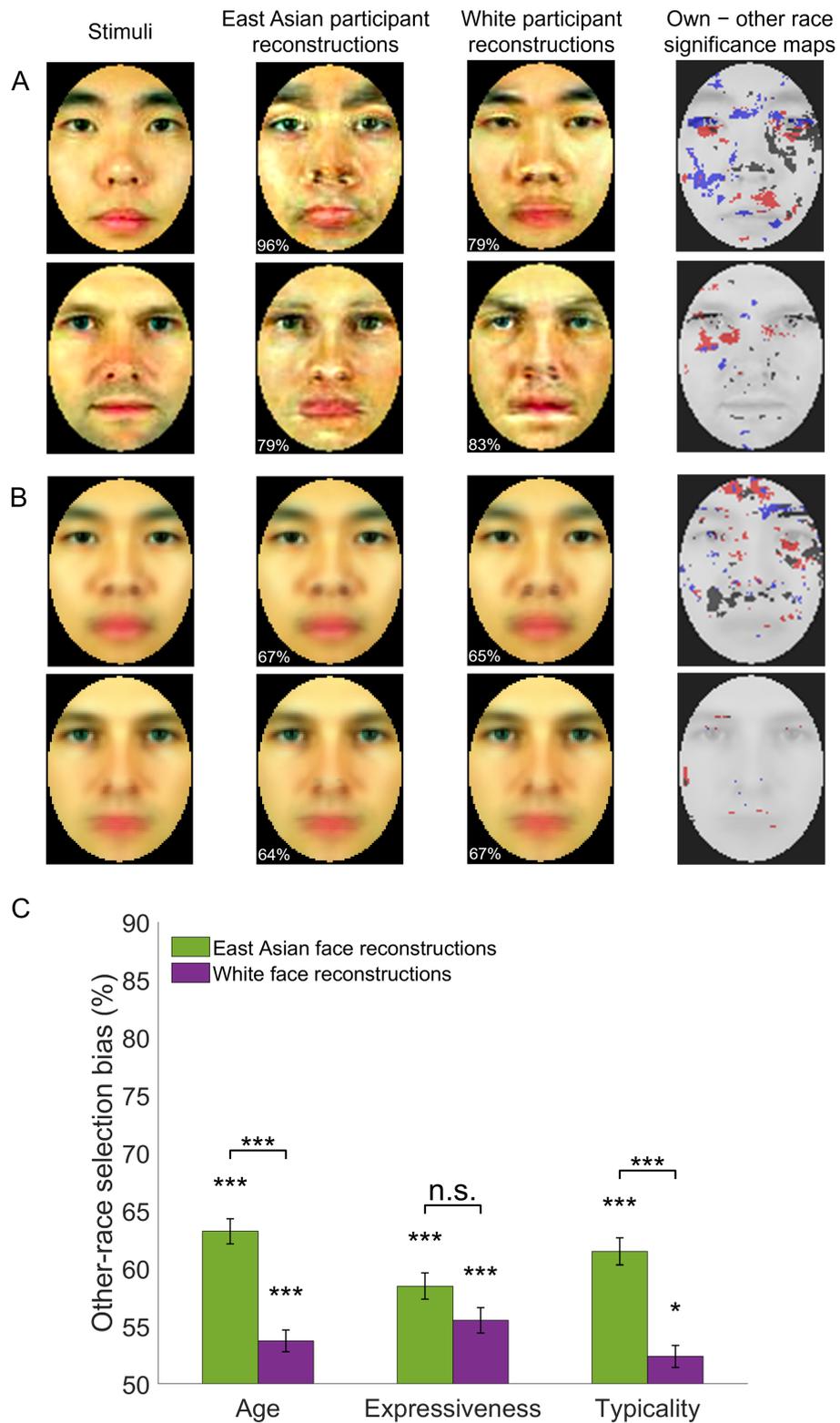


FIGURE 4

Own-race and other-race facial image reconstructions – examples and assessment. (A) Examples of individual face image reconstructions are shown for East Asian and White participants along with significance maps of the difference between own- versus other-race reconstructions (permutation-based pixelwise test, FDR-corrected,  $q < 0.05$ ; dark/red/blue pixels mark areas yielding significantly different values in the luminance, red-green and yellow-blue CIEL\*a\*b\* color channels). Numbers in the bottom left corner of each reconstruction indicate accuracy relative to the stimuli in the left column. (B) Averages of all reconstructions of the same race along with corresponding accuracy estimates and significance maps. (C) Other-race reconstructions are judged as younger, more expressive and more typical for their own race relative to own-race reconstructions of the same facial identities. Selection biases are significantly above

(Continued)

FIGURE 4 (Continued)

chance (two-tailed  $t$ -tests against 50% chance in a 2-alternative forced choice task) and more pronounced for East Asian stimuli (paired  $t$ -tests for each facial trait). Error bars indicate  $\pm 1$  SE (\* $p < 0.05$ , \*\*\* $p < 0.001$ ). Facial images adapted from Ma et al. (2015).

results were replicated after conducting the analysis separately for each group of validators, East Asian and White (both  $p$ 's  $< 0.001$ ).

Overall, these results indicate that OR faces, regardless of participant race, are perceived as younger, more expressive, and more typical of their race. Last, these biases appear to reflect overlapping but not identical visual sources.

## Discussion

The present work investigates OR face perception with respect to its neural basis, its processing dynamics and its representational content. By relating neural decoding and image reconstruction results to behavioral performance in East Asian and White participants this investigation leads to several notable findings.

First, neural decoding, relying on temporally-cumulative occipitotemporal signals, mirrors the ORE evinced by behavioral performance. Specifically, between-race face decoding is more accurate than within-race decoding, consistent with the separability of neural patterns for OR from SR faces (Natu et al., 2011; Wang et al., 2023). More importantly, within-race face decoding is more accurate for SR than OR faces in both East Asian and White participants. This provides a decoding-based neural counterpart of ORE at the group level (i.e., across participant race), complementing results based on repetition suppression (Hughes et al., 2019; Vizioli et al., 2010; Zhou et al., 2020). Yet, behaviorally, ORE is also known to exhibit a gradient across individuals from the same group (e.g., Megreya et al., 2011). Hence, here, we also examine ORE's prominence in neural processing at the individual level. This investigation points to ORE and face recognition ability as main components of variability in representational similarity across participants. Thus, our findings speak to the neural basis of ORE, both across and within groups, while also serving as a platform for the rest of our investigation.

Second, the temporal profile of decoding exhibits an extensive window of significance, between around 130–600 ms, consistent with prior work on the neural dynamics of facial identity perception (Ambrus et al., 2019; Dobs et al., 2019; Kovács et al., 2023; Nemrodov et al., 2018). Overall, this profile is similar for SR and OR faces as no specific time points, in isolation, yield a significant difference. However, an SR decoding advantage is, at least, numerically apparent during most of the 130–600 ms interval (Figures 2D,E). This suggests the presence of complementary information over time, which accrues into an overall SR advantage, captured by temporally-cumulative analysis – to be clear, the latter analysis concatenates, rather than averages, temporal information, thus allowing decoding to exploit temporal patterns over this extended window. While an ORE univariate effect for the N170 component was apparent for East Asian participants (see [Supplementary materials](#) and [Supplementary Figure S1](#)), our decoding results indicate that ORE is unlikely to be solely linked to a restricted temporal window. The interpretation above, of complementary information over time,

also agrees with evidence for multiple stages of neural processing associated with ORE across an extensive cortical network (Natu et al., 2011; Zhou et al., 2020).

Third, face spaces derived from EEG data evince meaningful, informative structure. Previous fMRI (Loffler et al., 2005; Nestor et al., 2016; Carlin and Kriegeskorte, 2017) and neurophysiology (Chang et al., 2017) studies have yielded evidence for a neural-based face space whose dimensions encode visual features relevant for identity recognition. Our results agree with this prior work, though important aspects of these spaces, such as specific nonlinearities (Carlin and Kriegeskorte, 2017; Yang et al., 2023), remain to be examined for their EEG-derived counterparts here. Our investigation goes, however, beyond the ability of such spaces to encode individual appearance. Specifically, we focus on the separability of faces by race and the higher representational density of OR versus SR.

Relevantly here, prior behavioral (Tanaka et al., 2013; Papesh and Goldinger, 2010) and computational work (O'Toole et al., 1991) has pointed to OR compression in face space. For instance, Caldara and Abdi (2006) trained an autoassociative network only with White or only with Asian face images. In both cases, the authors found higher density for the left-out race in the face space projections of the network. Namely, they found larger pairwise distances, on average, for SR faces. Our results provide a neural counterpart of these simulations based on pairwise face discriminability. Overall, such findings suggest that visual experience with one race amplifies SR face dissimilarity by optimizing face space structure for SR face encoding and recognition (Valentine, 1991).

Fourth, we demonstrate the richness of visual information present in face space by successfully deploying its dimensions and associated features in image reconstruction. Accordingly, face representations, recovered through reconstruction, reveal visual differences between SR and OR faces which do not reflect mere differences in image quality. Specifically, image reconstructions assessed by a separate group of validators, revealed that OR faces are represented as more typical for their race. This agrees with OR compression in face space, as noted above, as well as with ORE phenomenology (i.e., “they all look the same”; Ackerman et al., 2006; Feingold, 1914; Laurence et al., 2016). Notably though, we also find largely separate biases in the perception of age and expressiveness, which we address next.

Little is known about OR biases in age perception, with a handful of prior studies yielding conflicting results (Dehon and Brédart, 2001; Mousavi et al., 2018; Zhao and Bentin, 2008). Recent behavioral work suggests that OR faces are represented as younger in both East Asian and White participants (Shoura et al., 2024). Further, an *illusion* of Asian youthfulness is suggested by inter-ethnic differences in both skin physiology (e.g., more collagen) and skeletal structure (Shirakabe et al., 2003; Vashi et al., 2016) relative to White faces. Therefore, we reasoned that OR faces may be perceived as younger than SR ones, at least by White participants. Our results support the hypothesis above as well as a reciprocal bias, with Asians also perceiving White faces as younger. Thus, the present findings speak to a general OR bias in age estimation and, critically, they reveal the visual representations

supporting prior reports of such biases (Dehon and Brédart, 2001; Mousavi et al., 2018).

Regarding expressiveness, recent work has found poorer performance in OR expression recognition (Jiang et al., 2023; Yan et al., 2016), possibly due to cultural differences in the representation of emotional expressions (Chen et al., 2024; Jack et al., 2012). While our stimuli display neutral expressions, emotion can be perceived even in neutral faces as a function of facial structure (Neth and Martinez, 2009; Said et al., 2009) and person knowledge (Suess et al., 2015). Hence, we reasoned that biases in expression recognition may also extend to the degree of expressiveness perceived in OR neutral faces. The results bore out this hypothesis and, again, they revealed the representations underlying this bias. Further work though will be needed to uncover the precise nature of this bias (e.g., as driven by valence, arousal).

From a practical standpoint, we note that perceptual biases have critical implications for eyewitness testimony. For instance, a typicality bias, whereby OR faces are perceived as less distinctive, may exacerbate misidentifications in legal settings (Wells and Olson, 2001; Meissner and Brigham, 2001), even when confidence is high (Dodson and Dobolyi, 2016). Additionally, biases in expressiveness can influence judgments about emotional intent, introducing another layer of distortion in face recognition (Jiang et al., 2023). Prior research has evaluated training programs emphasizing individuation strategies, which may help mitigate such perceptual biases, enhance cross-racial identification and improve the reliability of eyewitness testimony (Young et al., 2012; Lebrecht et al., 2009).

While our examination of potential biases was driven by specific hypotheses, current results may be further queried for other facial trait biases (e.g., in attractiveness, competence). Methodologically, this opportunity illustrates the benefit of data-driven approaches aimed at recovering internal representations (Chen et al., 2024; Nestor et al., 2016, 2020; Zhan et al., 2019), including their ability to uncover new perceptual biases. Further, theoretically, it showcases how encoding OR faces in a suboptimal face space (Dahl et al., 2014; O'Toole et al., 1991), crafted for SR recognition, may lead to an array of representational distortions impacting multiple facial traits. In turn, practically, such an array of biases can shed new light on an ORE-induced decrement in the quality of social interactions (McKone et al., 2023), beyond difficulties with person identification.

Regarding the nature of the information underlying ORE, a qualitative advantage for SR versus OR emerged early, around 130 ms, consistent with the grounding of ORE in perception (Megreya et al., 2011). However, this advantage, exploited by our temporally-cumulative analyses, extended up to ~600 ms. Thus, these results do not allow linking ORE solely to an early, restricted temporal window associated with automatic perceptual processing. Prior work has argued for the contribution of memory (Herzmann et al., 2022; Tanaka et al., 2004; Yaros et al., 2019; Zhao et al., 2014) and socio-cognitive factors (Hugenberg et al., 2010; Schwartz et al., 2023) to ORE. Accordingly, we cannot rule out that such factors may have contributed, in addition to perceptual ones, to the visual representations assessed here.

Furthermore, we note that the main source of ORE may well vary across different groups. For instance, it may be driven by perceptual experience for Asian-White groups but by social-motivational factors across Black-White groups (Wan et al., 2015). Accordingly, investigating the nature and extent of visual biases across other groups (e.g., Black-White) could be very informative in this respect. Further,

evaluating memory-based representations (Chang et al., 2017; Zhan et al., 2019) and their ORE biases could help clarify the interplay between perception and memory.

To conclude, the present work integrates measures of behavioral performance, neural decoding and image reconstruction to yield new insights into the representational basis of ORE and its dynamics. Our findings reveal multiple biases in OR face perception with significant theoretical, methodological and practical implications. More generally, they open new avenues for exploring racial biases in face recognition and pave the way to studying visual misrepresentations via image reconstruction.

## Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: open science framework: <https://osf.io/u3rsh/>.

## Ethics statement

The studies involving humans were approved by Research Ethics Board at the University of Toronto. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

## Author contributions

MoS: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. YL: Data curation, Formal analysis, Investigation, Writing – original draft. MaS: Data curation, Formal analysis, Investigation, Writing – original draft. AD: Data curation, Investigation, Writing – original draft. AN: Conceptualization, Formal analysis, Funding acquisition, Methodology, Resources, Supervision, Validation, Writing – original draft, Writing – review & editing.

## Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC).

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations,

## References

- Ackerman, J. M., Shapiro, J. R., Neuberg, S. L., Kenrick, D. T., Becker, D. V., Griskevicius, V., et al. (2006). They all look the same to me (unless they're angry): from out-group homogeneity to out-group heterogeneity. *Psychol. Sci.* 17, 836–840. doi: 10.1111/j.1467-9280.2006.01790.x
- Ambrus, G. G., Kaiser, D., Cichy, R. M., and Kovács, G. (2019). The neural dynamics of familiar face recognition. *Cereb. Cortex* 29, 4775–4684. doi: 10.1093/cercor/bhz
- Balas, B., and Nelson, C. A. (2010). The role of face shape and pigmentation in other-race face perception: an electrophysiological study. *Neuropsychologia* 48, 498–506. doi: 10.1016/j.neuropsychologia.2009.10.007
- Balas, B., Westerlund, A., Hung, K., and Nelson, C. A. (2011). Shape, color and the other-race effect in the infant brain. *Dev. Sci.* 14, 892–900. doi: 10.1111/j.1467-7687.2011.01039.x
- Bowles, D. C., McKone, E., Dawel, A., Duchaine, B., Palermo, R., Schmalz, L., et al. (2009). Diagnosing prosopagnosia: effects of ageing, sex, and participant-stimulus ethnic match on the Cambridge face memory test and Cambridge face perception test. *Cogn. Neuropsychol.* 26, 423–455. doi: 10.1080/02643290903343149
- Brainard, D. H. (1997). The psychophysics toolbox. *Spat. Vis.* 10, 433–436. doi: 10.1163/156856897x00357
- Byatt, G., and Rhodes, G. (2004). Identification of own-race and other-race faces: implications for the representation of race in face space. *Psychon. Bull. Rev.* 11, 735–741. doi: 10.3758/BF03196628
- Caldara, R., and Abdi, H. (2006). Simulating the 'other-race' effect with autoassociative neural networks: further evidence in favor of the face-space model. *Perception* 35, 659–670. doi: 10.1068/p5360
- Caldara, R., Rossion, B., Bovet, P., and Hauert, C. A. (2004). Event-related potentials and time course of the 'other-race' face classification advantage. *Neuroreport* 15, 905–910. doi: 10.1097/00001756-200404090-00034
- Carlin, J. D., and Kriegeskorte, N. (2017). Adjudicating between face-coding models with individual-face fMRI responses. *PLoS Comput. Biol.* 13:e1005604. doi: 10.1371/journal.pcbi.1005604
- Chang, C. H., Nemrodov, D., Lee, A. C. H., and Nestor, A. (2017). Memory and perception-based facial image reconstruction. *Sci. Rep.* 7:6499. doi: 10.1038/s41598-017-06585-2
- Chen, C., Messinger, D. S., Chen, C., Yan, H., Duan, Y., Ince, R. A. A., et al. (2024). Cultural facial expressions dynamically convey emotion category and intensity information. *Curr. Biol.* 34, 213–223.e5. doi: 10.1016/j.cub.2023.12.001
- Contreras, J. M., Banaji, M. R., and Mitchell, J. P. (2013). Multivoxel patterns in fusiform face area differentiate faces by sex and race. *PLoS One* 8:e69684. doi: 10.1371/journal.pone.0069684
- Dahl, C. D., Chen, C. C., and Rasch, M. J. (2014). Own-race and own-species advantages in face perception: A computational view. *Sci. Rep.* 4:6654. doi: 10.1038/srep06654
- Dai, J., and Scherf, K. S. (2023). The privileged status of peer faces: subordinate-level neural representations of faces in emerging adults. *J. Cogn. Neurosci.* 35, 715–735. doi: 10.1162/jocn\_a\_01966
- Dehon, H., and Brédart, S. (2001). An 'other-race' effect in age estimation from faces. *Perception* 30, 1107–1113. doi: 10.1068/p3122
- Delorme, A., Sejnowski, T., and Makeig, S. (2007). Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. *NeuroImage* 34, 1443–1449. doi: 10.1016/j.neuroimage.2006.11.004
- Dobs, K., Isik, L., Pantazis, D., and Kanwisher, N. (2019). How face perception unfolds over time. *Nat. Commun.* 10:1258. doi: 10.1038/s41467-019-09239-1
- Dodson, C. S., and Dobolyi, D. G. (2016). Confidence and eyewitness identifications: the cross-race effect, decision time and accuracy. *Appl. Cogn. Psychol.* 30, 113–125. doi: 10.1002/acp.3178
- Duchaine, B., and Nakayama, K. (2006). The Cambridge face memory test: results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic participants. *Neuropsychologia* 44, 576–585. doi: 10.1016/j.neuropsychologia.2005.07.001
- Estudillo, A. J. (2021). Self-reported face recognition abilities for own-and other-race faces. *J. Crim. Psychol.* 11, 105–115. doi: 10.1108/jcp-06-2020-0025
- Feingold, G. A. (1914). The influence of environment on identification of persons and things. *J. Am. Inst. Crim. Law Criminol.* 5, 39–51. doi: 10.2307/1133283
- Feng, L., Liu, J., Wang, Z., Li, J., Li, L., Ge, L., et al. (2011). The other face of the other-race effect: an fMRI investigation of the other-race face categorization advantage. *Neuropsychologia* 49, 3739–3749. doi: 10.1016/j.neuropsychologia.2011.09.031
- Ficco, L., Müller, V. I., Kaufmann, J. M., and Schweinberger, S. R. (2023). Socio-cognitive, expertise-based and appearance-based accounts of the other-'race' effect in face perception: A label-based systematic review of neuroimaging results. *Br. J. Psychol.* 114, 45–69. doi: 10.1111/bjop.12595
- Ghuman, A. S., Brunet, N. M., Li, Y., Konecky, R. O., Pyles, J. A., Walls, S. A., et al. (2014). Dynamic encoding of face information in the human fusiform gyrus. *Nat. Commun.* 5:5672. doi: 10.1038/ncomms5672
- Golby, A. J., Gabrieli, J. D. E., Chiao, J. Y., and Eberhardt, J. L. (2001). Differential responses in the fusiform region to same-race and other-race faces. *Nat. Neurosci.* 4, 845–850. doi: 10.1038/90565
- Harrison, S. A., Gauthier, I., Hayward, W. G., and Richler, J. J. (2014). Other-race effects manifest in overall performance, not qualitative processing style. *Vis. Cogn.* 22, 843–864. doi: 10.1080/13506285.2014.918912
- Herzmann, G., Minor, G., and Curran, T. (2018). Neural evidence for the contribution of holistic processing but not attention allocation to the other-race effect on face memory. *Cogn. Affect. Behav. Neurosci.* 18, 1015–1033. doi: 10.3758/s13415-018-0619-z
- Herzmann, G., Ogle, O., and Curran, T. (2022). More elaborate processing of own-race faces and less elaborate processing of other-race faces contribute to the other-race effect in face memory. *Br. J. Psychol.* 113, 1033–1055. doi: 10.1111/bjop.12581
- Hugenberg, K., Young, S. G., Bernstein, M. J., and Sacco, D. F. (2010). The categorization-individuation model: an integrative account of the other-race recognition deficit. *Psychol. Rev.* 117, 1168–1187. doi: 10.1037/a0020463
- Hughes, B. L., Camp, N. P., Gomez, J., Natu, V. S., Grill-Spector, K., and Eberhardt, J. L. (2019). Neural adaptation to faces reveals racial outgroup homogeneity effects in early perception. *Proc. Natl. Acad. Sci. USA* 116, 14532–14537. doi: 10.1073/pnas.1822084116
- Jack, R. E., Garrod, O. G. B., Yu, H., Caldara, R., and Schyns, P. G. (2012). Facial expressions of emotion are not culturally universal. *Proc. Natl. Acad. Sci. USA* 109, 7241–7244. doi: 10.1073/pnas.1200155109
- Jaquet, E., Rhodes, G., and Hayward, W. (2008). Race-contingent aftereffects suggest distinct perceptual norms for different race faces. *Vis. Cogn.* 16, 734–753. doi: 10.1080/13506280701350647
- Jiang, Z., Recio, G., Li, W., Zhu, P., He, J., and Sommer, W. (2023). The other-race effect in facial expression processing: behavioral and ERP evidence from a balanced cross-cultural study in women. *Int. J. Psychophysiol.* 183, 53–60. doi: 10.1016/j.ijpsycho.2022.11.009
- Kim, J. S., Yoon, H. W., Kim, B. S., Jeun, S. S., Jung, S. L., and Choe, B. Y. (2006). Racial distinction of the unknown facial identity recognition mechanism by event-related fMRI. *Neurosci. Lett.* 397, 279–284. doi: 10.1016/j.neulet.2005.12.061
- Kovács, G., Li, C., Ambrus, G. G., and Burton, A. M. (2023). The neural dynamics of familiarity-dependent face identity representation. *Psychophysiology* 60:e14304. doi: 10.1111/psyp.14304
- Kubota, J. T., Banaji, M. R., and Phelps, E. A. (2012). The neuroscience of race. *Nat. Neurosci.* 15, 940–948. doi: 10.1038/nn.3136
- Laurence, S., Zhou, X., and Mondloch, C. J. (2016). The flip side of the other-race coin: they all look different to me. *Br. J. Psychol.* 107, 374–388. doi: 10.1111/bjop.12147
- Lebrecht, S., Pierce, L. J., Tarr, M. J., and Tanaka, J. W. (2009). Perceptual other-race training reduces implicit racial bias. *PLoS One* 4:e4215. doi: 10.1371/journal.pone.0004215

- Loffler, G., Yourganov, G., Wilkinson, F., and Wilson, H. R. (2005). fMRI evidence for the neural representation of faces. *Nat. Neurosci.* 8, 1386–1391. doi: 10.1038/nn1538
- Ma, D. S., Correll, J., and Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behav. Res. Methods* 47, 1122–1135. doi: 10.3758/s13428-014-0532-5
- Malpass, R. S., and Kravitz, J. (1969). Recognition for faces of own and other race. *J. Pers. Soc. Psychol.* 13, 330–334. doi: 10.1037/h0028434
- McKone, E., Dawel, A., Robbins, R. A., Shou, Y., Chen, N., and Crookes, K. (2023). Why the other-race effect matters: poor recognition of other-race faces impacts everyday social interactions. *Br. J. Psychol.* 114, 230–252. doi: 10.1111/bjop.12508
- McKone, E., Stokes, S., Liu, J., Cohan, S., Fiorentini, C., Pidcock, M., et al. (2012). A robust method of measuring other-race and other-ethnicity effects: the Cambridge face memory test format. *PLoS One* 7:e47956. doi: 10.1371/journal.pone.0047956
- Megreya, A. M., White, D., and Burton, A. M. (2011). The other-race effect does not rely on memory: evidence from a matching task. *Q. J. Exp. Psychol.* 64, 1473–1483. doi: 10.1080/17470218.2011.575228
- Meissner, C. A., and Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychol. Public Policy Law* 7, 3–35. doi: 10.1037/1076-8971.7.1.3
- Michel, C., Rossion, B., Bühlhoff, I., Hayward, W. G., and Vuong, Q. C. (2013). The contribution of shape and surface information in the other-race face effect. *Vis. Cogn.* 21, 1202–1223. doi: 10.1080/13506285.2013.823141
- Michel, C., Rossion, B., Han, J., Chung, C.-S., and Caldara, R. (2006). Holistic processing is finely tuned for faces of one's own race. *Psychol. Sci.* 17, 608–615. doi: 10.1111/j.1467-9280.2006.01752.x
- Mittal, A., Moorthy, A. K., and Bovik, A. C. (2012). No-reference image quality assessment in the spatial domain. *IEEE Trans. Image Process.* 21, 4695–4708. doi: 10.1109/TIP.2012.2214050
- Mouraux, A., and Iannetti, G. D. (2008). Across-trial averaging of event-related EEG responses and beyond. *Magn. Reson. Imaging* 26, 1041–1054. doi: 10.1016/j.mri.2008.01.011
- Mousavi, S. M., Chen, M., and Oruc, I. (2018). Ethnicity and gender effects in the perception of age in faces. *J. Vis.* 18:1098. doi: 10.1167/18.10.1098
- Murray, R. F. (2011). Classification images: A review. *J. Vis.* 11:2. doi: 10.1167/11.5.2
- Natu, V., Raboy, D., and O'Toole, A. J. (2011). Neural correlates of own- and other-race face perception: spatial and temporal response differences. *NeuroImage* 54, 2547–2555. doi: 10.1016/j.neuroimage.2010.10.006
- Nemrodov, D., Ling, S., Nudnou, I., Roberts, T., Cant, J. S., Lee, A. C. H., et al. (2020). A multivariate investigation of visual word, face, and ensemble processing: perspectives from EEG-based decoding and feature selection. *Psychophysiology* 57:e13511. doi: 10.1111/psyp.13511
- Nemrodov, D., Niemeier, M., Mok, J. N. Y., and Nestor, A. (2016). The time course of individual face recognition: A pattern analysis of ERP signals. *NeuroImage* 132, 469–476. doi: 10.1016/j.neuroimage.2016.03.006
- Nemrodov, D., Niemeier, M., Patel, A., and Nestor, A. (2018). The neural dynamics of facial identity processing: insights from EEG-based pattern analysis and image reconstruction. *ENEURO* 5. doi: 10.1523/ENEURO.0358-17.2018
- Nestor, A., Lee, A. C. H., Plaut, D. C., and Behrmann, M. (2020). The face of image reconstruction: Progress, pitfalls, prospects. *Trends Cogn. Sci.* 24, 747–759. doi: 10.1016/j.tics.2020.06.006
- Nestor, A., Plaut, D. C., and Behrmann, M. (2016). Feature-based face representations and image reconstruction from behavioral and neural data. *Proc. Natl. Acad. Sci. USA* 113, 416–421. doi: 10.1073/pnas.1514551112
- Neth, D., and Martinez, A. M. (2009). Emotion perception in emotionless face images suggests a norm-based representation. *J. Vis.* 9, 5.1–5.511. doi: 10.1167/9.1.5
- O'Toole, A., Deffenbacher, K., Abdi, H., and Bartlett, J. (1991). Simulating the 'other-race effect' as a problem in perceptual learning. *Connect. Sci.* 3, 163–178. doi: 10.1080/09540099108946583
- Papesh, M. H., and Goldinger, S. D. (2010). A multidimensional scaling analysis of own- and cross-race face spaces. *Cognition* 116, 283–288. doi: 10.1016/j.cognition.2010.05.001
- Pearce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., et al. (2019). PsychoPy2: experiments in behavior made easy. *Behav. Res. Methods* 51, 195–203. doi: 10.3758/s13428-018-01193-y
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies\*. *Spat. Vis.* 10, 437–442. doi: 10.1163/156856897x00366
- Reggev, N., Brodie, K., Cikara, M., and Mitchell, J. P. (2020). Human face-selective cortex does not distinguish between members of a racial outgroup. *ENEURO* 7, ENEURO.0431-19.2020. doi: 10.1523/ENEURO.0431-19.2020
- Roberts, T., Cant, J. S., and Nestor, A. (2019). Elucidating the neural representation and the processing dynamics of face ensembles. *J. Neurosci.* 39, 7737–7747. doi: 10.1523/JNEUROSCI.0471-19.2019
- Rossion, B., and Michel, C. (2011). "An experience-based holistic account of the other-race face effect" in Oxford handbook of face perception. Eds. A. J. Calder, G. Rhodes, M. H. Johnson and J. V. Haxby (Oxford: Oxford University Press).
- Said, C. P., Sebe, N., and Todorov, A. (2009). Structural resemblance to emotional expressions predicts evaluation of emotionally neutral faces. *Emotion* 9, 260–264. doi: 10.1037/a0014681
- Schwartz, L., Cohen, M., Xu, S., Liu, J., and Yovel, G. (2023). The social-encoding benefit in face recognition is generalized to other-race faces. *Br. J. Psychol.* 114, 213–229. doi: 10.1111/bjop.12592
- Senholzi, K. B., and Ito, T. A. (2013). Structural face encoding: how task affects the N170's sensitivity to race. *Soc. Cogn. Affect. Neurosci.* 8, 937–942. doi: 10.1093/scan/nss091
- Shirakabe, Y., Suzuki, Y., and Lam, S. M. (2003). A new paradigm for the aging Asian face. *Aesth. Plast. Surg.* 27, 397–402. doi: 10.1007/s00266-003-2099-x
- Shoura, M., Bernhardt-Walther, D., and Nestor, A. (2024). Unraveling other-race face perception with GAN-based image reconstruction. *PsyArXiv*. [Preprint]. Available at: [https://osf.io/preprints/psyarxiv/qz7ms\\_v1](https://osf.io/preprints/psyarxiv/qz7ms_v1)
- Sporer, S. L., Gehrke, J., and McClure, K. A. (2022). Patrolling the Borders: experience with multiple out-group members is associated with the own-ethnicity effect. *J. Police Crim. Psychol.* 37, 433–446. doi: 10.1007/s11896-020-09387-x
- Suess, F., Rabovsky, M., and Abdel Rahman, R. (2015). Perceiving emotions in neutral faces: expression processing is biased by affective person knowledge. *Soc. Cogn. Affect. Neurosci.* 10, 531–536. doi: 10.1093/scan/nsu088
- Tanaka, J. W., Heptonstall, B., and Hagen, S. (2013). Perceptual expertise and the plasticity of other-race face recognition. *Vis. Cogn.* 21, 1183–1201. doi: 10.1080/13506285.2013.826315
- Tanaka, J. W., Kiefer, M., and Bukach, C. M. (2004). A holistic account of the own-race effect in face recognition: evidence from a cross-cultural study. *Cognition* 93, B1–B9. doi: 10.1016/j.cognition.2003.09.011
- Tanaka, J. W., and Pierce, L. J. (2009). The neural plasticity of other-race face recognition. *Cogn. Affect. Behav. Neurosci.* 9, 122–131. doi: 10.3758/CABN.9.1.122
- Tittenberg, S. C., and Wiese, H. (2023). Event-related brain potential correlates of the other-race effect: A review. *Br. J. Psychol.* 114, 24–44. doi: 10.1111/bjop.12591
- Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *Quart. J. Exp. Psychol. Sect. A* 43, 161–204. doi: 10.1080/1464079108400966
- Vashi, N. A., Castro, B., Maymone, M. D. E., and Kundu, R. V. (2016). Aging differences in ethnic skin. *J. Clin. Aesthet. Dermatol.* 9, 31–38
- Vizioli, L., Rousselet, G. A., and Caldara, R. (2010). Neural repetition suppression to identity is abolished by other-race faces. *Proc. Natl. Acad. Sci. USA* 107, 20081–20086. doi: 10.1073/pnas.1005751107
- Wagenmakers, E.-J., Wetzels, R., Borsboom, D., and van der Maas, H. L. J. (2011). Why psychologists must change the way they analyze their data: the case of psi: comment on Bem (2011). *J. Pers. Soc. Psychol.* 100, 426–432. doi: 10.1037/a0022790
- Walker, P. M., Silvert, L., Hewstone, M., and Nobre, A. C. (2008). Social contact and other-race face processing in the human brain. *Soc. Cogn. Affect. Neurosci.* 3, 16–25. doi: 10.1093/scan/nsm035
- Wan, L., Crookes, K., Reynolds, K. J., Irons, J. L., and McKone, E. (2015). A cultural setting where the other-race effect on face recognition has no social-motivational component and derives entirely from lifetime perceptual experience. *Cognition* 144, 91–115. doi: 10.1016/j.cognition.2015.07.011
- Wang, Z., Bovik, A. C., Sheikh, H. R., and Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* 13, 600–612. doi: 10.1109/TIP.2003.819861
- Wang, A., Sliwiska, M. W., Watson, D. M., Smith, S., and Andrews, T. J. (2023). Distinct patterns of neural response to faces from different races in humans and deep networks. *Soc. Cogn. Affect. Neurosci.* 18, 1–15. doi: 10.1093/scan/nsad059
- Wells, G. L., and Olson, E. A. (2001). The other-race effect in eyewitness identification: what do we do about it? *Psychol. Public Policy Law* 7, 230–246. doi: 10.1037/1076-8971.7.1.230
- Wiese, H., and Schweinberger, S. R. (2018). Inequality between biases in face memory: event-related potentials reveal dissociable neural correlates of own-race and own-gender biases. *Cortex* 101, 119–135. doi: 10.1016/j.cortex.2018.01.016
- Wilmer, J. B., Germine, L., Chabris, C. F., Chatterjee, G., Williams, M., Loken, E., et al. (2010). Human face recognition ability is specific and highly heritable. *Proc. Natl. Acad. Sci. USA* 107, 5238–5241. doi: 10.1073/pnas.0913053107
- Yan, X., Andrews, T. J., Jenkins, R., and Young, A. W. (2016). Cross-cultural differences and similarities underlying other-race effects for facial identity and expression. *Q. J. Exp. Psychol.* 69, 1247–1254. doi: 10.1080/17470218.2016.1146312
- Yang, C. T., Fu, H. L., Hsieh, C. J., Hsu, P. H., and Townsend, J. T. (2023). On the varying density of faces in face space: explanation of the Atypicality effect in terms of a Riemannian metric and potential application to the other race effect. doi: 10.20944/preprints202304.0140.v1

- Yaros, J. L., Salama, D. A., Delisle, D., Larson, M. S., Miranda, B. A., and Yassa, M. A. (2019). A memory computational basis for the other-race effect. *Sci. Rep.* 9:19399. doi: 10.1038/s41598-019-55350-0
- Young, S. G., Hugenberg, K., Bernstein, M. J., and Sacco, D. F. (2012). Perception and motivation in face recognition: A critical review of theories of the cross-race effect. *Personal. Soc. Psychol. Rev.* 16, 116–142. doi: 10.1177/1088868311418987
- Zhan, J., Garrod, O. G. B., van Rijsbergen, N., and Schyns, P. G. (2019). Modelling face memory reveals task-generalizable representations. *Nat. Hum. Behav.* 3, 817–826. doi: 10.1038/s41562-019-0625-3
- Zhao, L., and Bentin, S. (2008). Own- and other-race categorization of faces by race, gender, and age. *Psychon. Bull. Rev.* 15, 1093–1099. doi: 10.3758/PBR.15.6.1093
- Zhao, M., Hayward, W. G., and Bühlhoff, I. (2014). Holistic processing, contact, and the other-race effect in face recognition. *Vis. Res.* 105, 61–69. doi: 10.1016/j.visres.2014.09.006
- Zhou, Y., Gao, T., Zhang, T., Li, W., Wu, T., Han, X., et al. (2020). Neural dynamics of racial categorization predicts racial bias in face recognition and altruism. *Nat. Hum. Behav.* 4, 69–87. doi: 10.1038/s41562-019-0743-y
- Zhou, X., Itz, M. L., Vogt, S., Kaufmann, J. M., Schweinberger, S. R., and Mondloch, C. J. (2021). Similar use of shape and texture cues for own- and other-race faces during face learning and recognition. *Vis. Res.* 188, 32–41. doi: 10.1016/j.visres.2021.06.014
- Zhou, X., Mondloch, C. J., Chien, S. H. L., and Moulson, M. C. (2022). Multi-cultural cities reduce disadvantages in recognizing naturalistic images of other-race faces: evidence from a novel face learning task. *Sci. Rep.* 12:8950. doi: 10.1038/s41598-022-11550-9