

Reconstructing Perceived and Retrieved Faces from Activity Patterns in Lateral Parietal Cortex

 Hongmi Lee¹ and  Brice A. Kuhl²

¹Department of Psychology, New York University, New York, New York 10003, and ²Department of Psychology, University of Oregon, Eugene, Oregon 97403

Recent findings suggest that the contents of memory encoding and retrieval can be decoded from the angular gyrus (ANG), a subregion of posterior lateral parietal cortex. However, typical decoding approaches provide little insight into the nature of ANG content representations. Here, we tested whether complex, multidimensional stimuli (faces) could be reconstructed from ANG by predicting underlying face components from fMRI activity patterns in humans. Using an approach inspired by computer vision methods for face recognition, we applied principal component analysis to a large set of face images to generate eigenfaces. We then modeled relationships between eigenface values and patterns of fMRI activity. Activity patterns evoked by individual faces were then used to generate predicted eigenface values, which could be transformed into reconstructions of individual faces. We show that visually perceived faces were reliably reconstructed from activity patterns in occipitotemporal cortex and several lateral parietal subregions, including ANG. Subjective assessment of reconstructed faces revealed specific sources of information (e.g., affect and skin color) that were successfully reconstructed in ANG. Strikingly, we also found that a model trained on ANG activity patterns during face perception was able to successfully reconstruct an independent set of face images that were held in memory. Together, these findings provide compelling evidence that ANG forms complex, stimulus-specific representations that are reflected in activity patterns evoked during perception and remembering.

Key words: angular gyrus; face perception; MVPA; reconstruction; retrieval; working memory

Significance Statement

Neuroimaging studies have consistently implicated lateral parietal cortex in episodic remembering, but the functional contributions of lateral parietal cortex to memory remain a topic of debate. Here, we used an innovative form of fMRI pattern analysis to test whether lateral parietal cortex actively represents the contents of memory. Using a large set of human face images, we first extracted latent face components (eigenfaces). We then used machine learning algorithms to predict face components from fMRI activity patterns and, ultimately, to reconstruct images of individual faces. We show that activity patterns in a subregion of lateral parietal cortex, the angular gyrus, supported successful reconstruction of perceived and remembered faces, confirming a role for this region in actively representing remembered content.

Introduction

Human neuroimaging studies of episodic memory have consistently implicated the lateral parietal cortex in memory retrieval (Wagner et al., 2005; Cabeza et al., 2008). In particular, the angular gyrus (ANG), a ventral subregion of lateral parietal cortex, exhibits increased activation when details of an experience are

successfully retrieved (Hutchinson et al., 2009; Spaniol et al., 2009) and ANG activity scales with the subjective vividness of retrieved memories (Kuhl and Chun, 2014). Because ANG involvement in memory retrieval generalizes across multiple modalities and stimulus types (Shannon and Buckner, 2004; Vilberg and Rugg, 2008), it has been argued that ANG plays a content-general role in memory retrieval (Cabeza et al., 2008). In other words, ANG activity may reflect whether information has been successfully retrieved without actively representing that information. However, this perspective has been challenged by recent evidence that activity patterns in ANG carry information about “what” is being remembered (Kuhl et al., 2013; Kuhl and Chun, 2014; Bird et al., 2015; St-Laurent et al., 2015). These examples of pattern-based fMRI evidence for mnemonic content representations in ANG have relied on two main analysis methods: pattern classification (Norman et al., 2006) and pattern similarity

Received Nov. 29, 2015; revised April 19, 2016; accepted April 21, 2016.

Author contributions: H.L. and B.A.K. designed research; H.L. performed research; H.L. analyzed data; H.L. and B.A.K. wrote the paper.

This work was supported by New York University Whitehead Fellowship to B.A.K. and Korea Foundation for Advanced Studies Doctoral Study Abroad Fellowship to H.L.

The authors declare no competing financial interests.

Correspondence should be addressed to Dr. Brice A. Kuhl, University of Oregon, 1585 East 13th Avenue, Eugene, OR 97403. E-mail: bkuhl@uoregon.edu.

DOI:10.1523/JNEUROSCI.4286-15.2016

Copyright © 2016 the authors 0270-6474/16/366069-14\$15.00/0

(Kriegeskorte et al., 2008). Pattern classification analyses attempt to read out a categorical label associated with a stimulus by mapping labels to activity patterns, with above-chance classification accuracy taken as evidence for content representations. Pattern similarity analyses, on the other hand, involve correlating activity patterns elicited by various stimuli, with higher correlations for “matching” than “nonmatching” stimuli taken as evidence for content-specific representations. Although these approaches have established that different types of content are associated with distinctive activation patterns in ANG, they provide very limited insight into the nature of the underlying content representations (Naselaris and Kay, 2015).

Here, we tested for mnemonic content representations in ANG using an innovative and more transparent form of pattern-based fMRI. Rather than testing for classification of a single stimulus dimension (e.g., face vs scene) or a correlation between stimulus activity patterns (match > nonmatch), we used a complex, multidimensional stimulus class (faces) and sought to map a wide range of continuously varying feature values to activity patterns within ANG. Specifically, we used a data-driven approach, originally developed for computer-based face recognition, to extract latent components from a large set of face images (*eigenfaces*) (Turk and Pentland, 1991) and then tested whether these face components could be predicted, in a cross-validated manner, by ANG activity patterns. Critically, these predictions could be converted, via an inverse transformation, to native “face space” (Cowen et al., 2014), allowing for the fMRI-based face information to be plainly viewed in the form of a reconstructed face image. Thus, in contrast to typical classification-/similarity-based analyses, this approach involves a mapping between multiple stimulus dimensions and voxel activity patterns and allows for more direct assessment and better characterization of the sources of information that contribute to ANG content representations.

Across two fMRI experiments, we first tested whether activity patterns in ANG enabled successful reconstructions of perceived (visually presented) faces. We tested the quality of perception-based reconstructions in two ways: (1) by comparing the similarity of predicted versus actual eigenface scores, which yields an objective measure of reconstruction accuracy; and (2) by measuring whether specific face dimensions (gender, affect, skin color, etc.) were subjectively apparent in the reconstructed faces. Next, using a model trained on perception data, we tested whether ANG activity patterns allowed for reliable reconstructions of faces held in memory using a retro-cue working memory paradigm (Harrison and Tong, 2009). Critically, because we used a distinct set of faces in the perception and working memory phases, this approach provided a stringent test of whether internally generated memory representations could be reconstructed by predicting the underlying feature representations from ANG activity patterns. For comparison, we also assessed perception- and memory-based reconstructions in face-sensitive areas of occipitotemporal cortex (OTC) and other subregions of lateral parietal cortex.

Materials and Methods

To reconstruct perceived and remembered face images from fMRI activity patterns, we ran two fMRI experiments each consisting of separate perception and memory phases. The two experiments were highly similar, with the critical difference being that in Experiment 2 we reduced variance in low-level visual properties of the face images. Additionally, we ran two behavioral studies using independent samples of subjects. In the first behavioral study, subjects evaluated reconstructed and original

face images on various dimensions so that we could characterize the subjective accuracy of reconstructed faces. In the second behavioral study, we used a change-detection working memory paradigm to test whether the confusability of individual face memories could be predicted by the similarity of eigenface scores.

Experiment 1

Participants

Twelve healthy subjects (eight female, age 19–28 years) completed 13 experimental sessions. Two sessions from an additional subject were excluded due to excessive sleepiness and movement during scanning. One additional session from 1 of the 12 included subjects was excluded due to excessive motion. All subjects were right-handed and reported normal or corrected-to-normal vision. Informed consent was obtained in accordance with procedures approved by the New York University Institutional Review Board.

Stimuli

A total of 1012 face images were selected from a variety of online sources, including publicly available face image databases (e.g., FEI Face Database; <http://fei.edu.br/~cet/facedatabase.html>; the Color FERET Database) (Phillips et al., 1998). Half of the selected faces were female. Both female and male faces varied in terms of age, ethnicity, expression, etc., with no deliberate face “categories” other than gender. All faces were forward-facing so that the eyes and mouth were visible in every image. The images were cropped and resized to 179 × 251 pixels, and the centers of the eyes and mouth were manually aligned across images. Of the 1012 faces, 36 were selected as test faces for the perception phase and 16 were selected as faces for the memory phase. These test/memory faces were pseudo-randomly selected to include a range of ethnicities and facial expressions. The specific faces used as test faces for the perception phase and for the memory phase were fixed (no counterbalancing) across subjects. The remaining 960 faces served as a “pool” of faces from which a different, random subset served as the perception-phase training faces for each subject.

A small percentage of scene trials were included in the perception phase to select face-preferring voxels. A total of 120 scene images were collected from freely available sources (e.g., <http://cvcl.mit.edu/MM/sceneCategories.html>) (Konkle et al., 2010). The scene images included both indoor and outdoor scenes from a variety of categories, such as mountains, playground, living room, etc. All scene images were cropped and resized to be the same size as the face images. The total number of face and scene images presented in the perception phase varied across subjects, depending on the number of sessions and scanning runs that were completed.

Experimental design and procedures

Before beginning the main experiment, subjects were given instructions and practiced the tasks. Subjects were also familiarized with all of the faces that would be used in the memory phase to facilitate vivid remembering. Familiarization consisted of subjects viewing the faces one at a time, for 2 s each. Subjects completed two familiarization blocks, each of which lasted 5 min and 52 s. Thus, the familiarization phase took ~12 min in total. Within each block, each face was presented 10 times with a randomized order. The lag between presentations of the same face was not controlled. No responses were required during the familiarization phase, but subjects were instructed to try to remember the visual details of each face. Inside the scanner, subjects first completed the memory phase and then the perception phase.

Perception phase. Subjects completed seven to nine fMRI runs of the perception phase, each of which consisted of 58 trials and lasted 7 min and 58 s. Every trial started with a face or a scene image (size: 9° × 12°) centrally presented over a black background screen for 2 s, followed by a 6 s fixation cross (Fig. 1A). Each scanning run contained 44 faces presented once each, 4 faces presented twice each (test faces) and 6 scene images. Face/scene images that appeared in a given run were never repeated across runs. During the perception phase, subjects performed a repetition detection task wherein they indicated on each trial whether the image was “new” (first presentation) or “old” (second presentation).

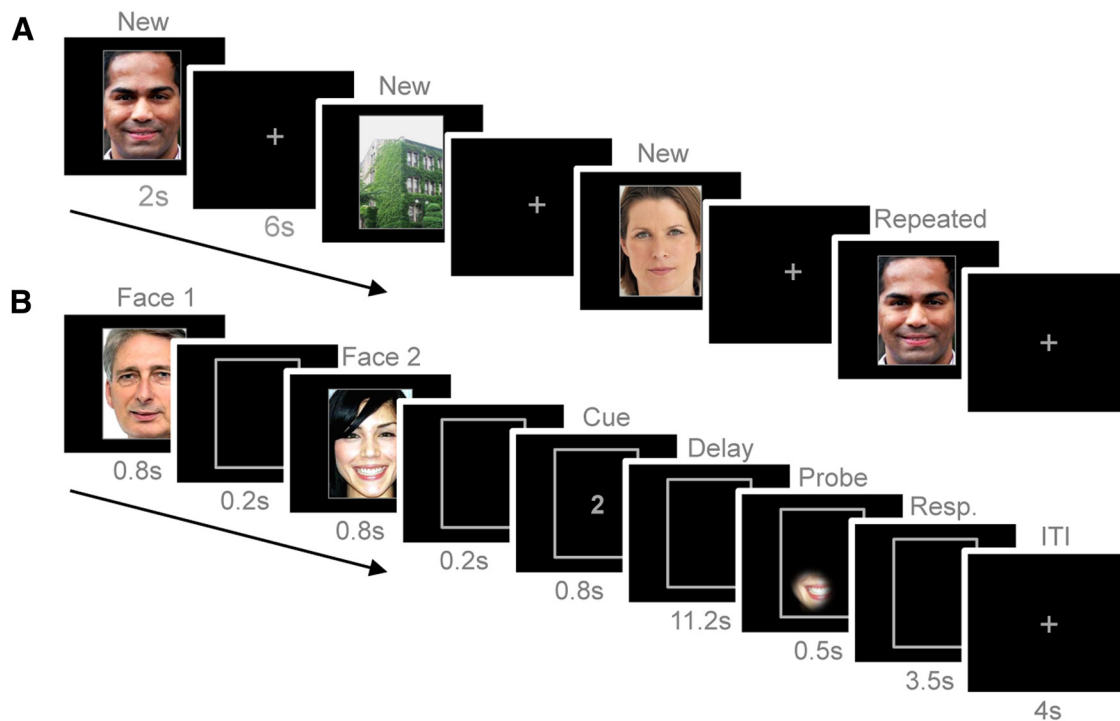


Figure 1. Task design of Experiment 1. **A**, In the perception phase, subjects viewed images of faces or scenes one at a time and judged whether each image was “old” (repeated within the run) or “new” (novel). The perception phase was similar in Experiment 2, with the main exception being that scene trials were not included. **B**, In the memory phase, subjects viewed two faces in a row, followed by a number cue indicating which of the two faces to maintain during the delay period. Following the delay period, a probe (partial face image) was presented and subjects indicated whether the probe matched the cued face or not.

Responses were made on a button box and were recorded if they were made within 7 s of the stimulus onset. Half of the training faces and half of the test faces in each run were female. The trial order was pseudo-randomized for each run with the constraints that test faces did not appear consecutively and there were at least three trials between repetitions of a particular test face. On average, subjects observed 413.5 faces per session (including both training and test images; range: 336–432) and 51.7 scenes (range: 42–54).

Memory phase. For the memory phase, we used a retro-cue working memory paradigm (Harrison and Tong, 2009), as shown in Figure 1B. Each trial started with brief, sequential presentations of two sample faces (800 ms) at the center of the display with a 200 ms gap in between. The image size was $9^\circ \times 12^\circ$, and the background color of the display was black. A thin gray rectangular frame surrounded the images and remained on the display throughout the trial to indicate the location/size of the face images while they were not present; 200 ms after the presentation of the second face, the number 1 or 2 appeared at the center of the empty gray frame for 800 ms, indicating which of the two faces should be maintained in memory (1 = remember first face, 2 = remember second face). Subjects were instructed to imagine the cued face as vividly as possible throughout the 11.2 s delay period that followed the cue. A memory probe was then presented (500 ms), which consisted of a face image seen through a small circular aperture. The diameter of the aperture was $\sim 2.5^\circ$, and the edge of the aperture was smoothed with a Gaussian filter ($\sigma = 1^\circ$). The location of the aperture was pseudo-randomly determined with constraints that the center of the aperture was on one of the quadrants of the image with equal probability within the session, and its distance to the center of the image did not exceed $\sim 2^\circ$ on both the horizontal and vertical dimensions. The probe face was the same as the cued face on one-half of the trials (“match” condition), and for the other one-half of trials, it was randomly selected from the set of faces used in the memory phase, excluding the sample faces from the current trial (“nonmatch” condition). Subjects had 4 s, from the probe onset, to indicate whether the probe matched the cued memory target or not by making a response on a button box. A 4 s fixation period followed the response window. In total, each memory phase trial lasted 22 s.

Subjects completed four memory phase scanning runs, each of which consisted of 16 trials and lasted 6 min and 6 s. For each trial, one of eight pairs of face images was used as the sample faces. The face pairs were fixed across subjects. Each face pair was repeated eight times throughout the session. The match/nonmatch condition, the presentation order of the two faces, and the memory retrieval cue (1 or 2) were fully counterbalanced within each pair. Thus, each face image served as the memory target four times. The order of trials was pseudo-randomized for each session with constraints that a face pair was not used as the set of sample faces in consecutive trials and all face pairs appeared at least once in a scanning run.

Both phases of the experiment were run in MATLAB using the Psychophysics Toolbox (Brainard, 1997). Visual stimuli were projected onto a screen at the end of the scanner bore and viewed through a mirror mounted in the head coil. Subjects made responses with an MRI-compatible button box using index and middle fingers. Response buttons were counterbalanced across subjects. Every scanning run started and ended with an additional 10 s and 4 s fixation period, respectively. Upon the end of each scanning run, subjects were given feedback on their performance (% accurate responses) on the screen and were allowed to have a short break.

fMRI data acquisition

fMRI scanning was conducted at the Center for Brain Imaging at New York University on the 3T Siemens Allegra head-only scanner. Functional data were collected using a head coil (NM011; NOVA Medical) for transmitting and an eight-channel phased array surface coil (NMSC071; NOVA Medical) for receiving. We obtained 34 oblique-coronal slices using a T2*-weighted gradient EPI sequence (TR = 2 s; TE = 30 ms; flip angle = 82° ; grid size 88×72 ; voxel size $2.5 \times 2.5 \times 2.5$ mm). The slices were oriented semiperpendicular to the calcarine sulcus and covered an area approximately between the occipital pole and the postcentral sulcus. A total of 239 volumes were collected for a perception run and 183 volumes for a memory run. We additionally collected proton density images with the same slice prescriptions as the functional images to improve functional-to-anatomical image coregistration. Whole-brain

high-resolution anatomical images were collected using a T1-weighted protocol (grid size 256×256 ; 176 slices; voxel size $1 \times 1 \times 1$ mm). In some cases, the high-resolution anatomical image was collected during a separate fMRI session. The Siemens Auto Align protocol was used to improve registration across sessions.

Experiment 2

Participants

Eleven right-handed subjects (nine female, 19–31 years) with normal or corrected-to-normal vision completed the experiment. Three additional subjects were excluded due to excessive motion. Informed consent was obtained in accordance with procedures approved by the New York University Institutional Review Board.

Stimuli

A total of 1184 face images were selected and prepared in the same way as in Experiment 1. Of these, 36 faces were selected as test faces for the perception phase and 24 were selected as faces for the memory phase. These faces were fixed (no counterbalancing) across subjects. The remaining 1124 faces served as a “pool” of faces for the perception phase, with a different, random subset of these faces serving as the perception-phase training faces for each subject. To reduce variability of low-level visual properties of the face images, image brightness and contrast were manually adjusted for individual faces. The SD of the distribution consisting of each image’s average pixel intensities was reduced by $\sim 10\%$ in Experiment 2 compared with Experiment 1, which was a statistically significant decrease ($p = 0.0007$ from a randomization test with 10,000 iterations).

An additional set of grayscale images of faces (with or without background scenes), scenes (corridors or houses), and objects (cars or guitars) were used for an independent localizer scan that took place on Day 1 of the experiment. There were 144 images for each image subcategory.

Experimental design and procedures

Experiment 2 consisted of two scanning sessions held on separate days. On Day 1, subjects practiced the memory task and were familiarized with the to-be-remembered faces. For the familiarization phase, subjects completed self-paced rounds during which they pressed a key to proceed to the next face. Each face image appeared once within each self-paced familiarization block. Eight subjects completed two self-paced familiarization blocks, and three subjects completed one block due to time constraints. Additionally, during the collection of high-resolution anatomical images, subjects again viewed each face, with a fixed timing of 2.5 s, for a total of 8 repetitions per face. Subjects also completed a functional localizer scan on Day 1, which was used to select face-prefering voxels. The second session (Day 2) was held within 3 d of the first session, and subjects completed the main experiment (memory phase followed by perception phase) as in Experiment 1. Before beginning the main experiment on Day 2, subjects practiced the tasks and were familiarized with the to-be-remembered faces once again ($1.5 \text{ s} \times 8$ repetitions per face). All procedures for the memory and perception phases were identical across experiments unless otherwise described below.

Localizer. The functional localizer scan was a block-design experiment using three categories of images: faces, scenes, and objects. Each category consisted of two subcategories (faces with and without backgrounds, corridors and houses, and guitars and cars), and each subcategory was repeated six times. Thus, there were 36 stimulus blocks. We also included 12 baseline (fixation) blocks. The order of blocks was randomized in a way that minimized the predictability of categories. Twelve unique images were presented per block. The duration of each image was 500 ms. All images were presented on a phase-scrambled noise pattern of size $17^\circ \times 17^\circ$. The task was an oddball detection task where subjects made a button press with their index finger whenever only a scrambled image appeared without an intact picture. The oddball trial randomly occurred in half of the stimulus blocks. The localizer scan began and ended with two additional baseline blocks (no images) and lasted 5 min and 12 s.

Perception phase. Subjects completed seven or eight perception-phase runs, each of which consisted of 52 trials and lasted 7 min and

10 s. No scene trials were included. Faces were presented on a gray background. As in Experiment 1, each run contained 4 test faces and each test face was presented twice. On average, subjects observed 362.2 faces (range: 336–384).

Memory phase. Subjects completed six memory-phase runs, each of which lasted 6 min 38 s. We used 24 faces, and the faces were randomly rearranged into 12 pairs for each subject. Stimuli were presented on a gray background, and a black rectangular frame was displayed on the screen while face images were not present. Subjects had 5 s to indicate, via button press, whether the memory probe matched the cued face or not. When the probe image was not the cued face (nonmatch trials), it was the “other” face from the sample period [this contrasts with Experiment 1, where nonmatch probes always excluded the uncued sample face]. Forcing subjects to select between the cued versus uncued face was intended to increase the difficulty of the task and subjects’ vigilance. A 5 s fixation cross followed the probe. Each trial lasted a total of 24 s.

fMRI data acquisition

A total of 156 functional volumes were collected for a localizer scan on Day 1. On Day 2, 215 volumes were obtained in each perception run, and 199 volumes were obtained in each memory run. For one subject, the localizer scan was run on Day 2 due to an equipment failure on Day 1. Whole-brain high-resolution anatomical images were collected on Day 1 or in a separate experimental session.

fMRI analysis

Preprocessing

Preprocessing of the functional data was conducted using FSL 5.0.5 (FMRIB Software Library, <http://www.fmrib.ox.ac.uk/fsl>) and custom scripts. The first five volumes from each scanning run were discarded. Images were corrected for head motion using MCFLIRT (Jenkinson et al., 2002). Scanning runs where subjects moved more than one voxel (2.5 mm) were excluded from further analyses; this resulted in the exclusion of one scanning run from one subject from Experiment 1. Motion-corrected images were smoothed with a Gaussian kernel with 1.7 mm SD (~ 4 mm FWHM). For the reconstruction of faces retrieved from memory, functional images were additionally detrended, high-pass filtered (cutoff = 0.01 Hz), and z-scored across time within each scanning run. High-resolution anatomical images were brain extracted using BET (Smith, 2002). The brain-extracted images were first coregistered to the proton density image of each session and then to the functional images using FLIRT (Jenkinson and Smith, 2001; Jenkinson et al., 2002).

ROIs

We generated a total of five ROIs: four anatomically defined subregions of posterior lateral parietal cortex (PPC) and a functionally defined set of face-prefering areas within the OTC (see Fig. 3A). All ROIs were generated in a subject-specific manner. Anatomical ROIs were first defined using FreeSurfer’s cortical parcellation scheme (<http://surfer.nmr.mgh.harvard.edu>). The subregions of PPC consisted of the bilateral ANG, supramarginal gyrus (SMG), intraparietal sulcus (IPS), and superior parietal lobule (SPL) as defined in FreeSurfer’s Destrieux atlas (Destrieux et al., 2010), except that our SMG ROI was a combination of SMG and the Jensen sulcus. The anatomical PPC ROIs were coregistered to the functional images and further masked by subject-specific whole-brain masks generated from functional images to exclude areas where signal dropout occurred. The number of voxels included in the anatomical masks varied across subjects and sessions (1605–2584 in ANG; 1386–3125 in SMG; 1248–2045 in IPS; 1726–2886 in SPL; 12,139–16,141 in OTC). We additionally generated unilateral ROIs for each PPC subregion to test for hemispheric differences in reconstruction accuracies. Across the perception and working memory phases, the only significant difference was that working memory reconstruction accuracy was moderately higher in left SPL than right SPL ($p = 0.04$); however, this difference was likely influenced by the fact that left SPL contained ~ 200 more voxels than right SPL. We therefore report results from the bilateral ROIs only.

To further increase the specificity of localization within ventral parietal cortex, we also generated finer-grained parietal subregions. As shown in Figure 7A, we extracted seven bilateral ventral parietal ROIs from the

17 network cortical parcellations based on resting-state functional connectivity described by Yeo et al. (2011). We included networks 7, 8, 12, 13, 15, 16, and 17. It should be noted that these network-based ROIs were smaller than the main anatomical PPC ROIs and varied considerably in size (mean number of voxels for each network: 568.8 in 7, 237.6 in 8, 795.5 in 12, 679.4 in 13, 309.5 in 15, 868.0 in 16, and 288.2 in 17). All ROIs were coregistered to the functional images and masked by subject-specific whole-brain masks.

The OTC ROI was generated by first combining anatomical ROIs corresponding to the bilateral occipital lobe and ventral temporal cortex (i.e., occipital pole, inferior occipital gyrus and sulcus, middle occipital gyrus, superior occipital gyrus, cuneus, lingual gyrus, fusiform gyrus, parahippocampal gyrus, calcarine sulcus, anterior and posterior transverse collateral sulcus, middle occipital sulcus and lunatus sulcus, superior occipital sulcus and transverse occipital sulcus, anterior occipital sulcus and preoccipital notch, lateral and medial occipitotemporal sulcus, parieto-occipital sulcus). We then selected the 2500 most face-preferring voxels within this anatomical OTC mask based on face-specific activation estimated from a GLM analysis using SPM8 (<http://www.fil.ion.ucl.ac.uk/spm>). In Experiment 1, face-specific activation was estimated from the perception phase data. We generated a design matrix with two main regressors: face trials and scene trials, each convolved with the hemodynamic response function. The resulting parameter estimates of the two conditions were contrasted to produce the t statistical map, and we selected the 2500 voxels with the highest t values. In Experiment 2, we used the independent localizer scan from Day 1 to select voxels. We generated a hemodynamic response function-convolved boxcar regressor for each stimulus subcategory and contrasted the two face subcategories against the scene and object subcategories. The resulting t map was coregistered to the functional images of the main experiment scanned on Day 2. Again, the 2500 voxels with the highest t values were selected. In both experiments, temporal derivatives of the regressors were also included in the design matrix. Six motion parameters and constant regressors for each scanning run were entered as regressors of no interest.

Face reconstruction analysis: perception

The face reconstruction analyses were implemented as previously described by Cowen et al. (2014) (Fig. 2). We first extracted eigenfaces by running a principal component analysis (PCA) on the entire pool of training faces (960 faces in Experiment 1; 1124 faces in Experiment 2). For each face image, the pixel intensities between 0 and 255 from all three color channels (i.e., red, green, blue) were normalized to have values between 0 and 1, and were then vectorized to produce a single vector of length 134,787 (175 pixels in x direction \times 251 pixels in y direction \times 3 color channels) representing the image. The PCA produced $N - 1$ total components, or eigenfaces, where N is the number of training faces. The eigenface scores of any face image Y (m faces \times n components) can be defined using the following formula:

$$Y = (F - F_{\text{mean}})C$$

where F (m faces \times k pixel values) is the face image vector, F_{mean} is the matrix of the same size as F consisting of the average image of all training faces, and C (k pixel values \times n components) is the eigenface vector.

To map eigenface scores to fMRI activity patterns, we first estimated the fMRI activation patterns evoked by each face by running a GLM analysis using SPM8. We concatenated all perception phase scanning runs and generated a design matrix, including each trial as a separate regressor. All trial regressors were convolved with the hemodynamic response function. Six motion parameters and constant regressors representing each scanning run were entered as regressors of no interest. The data were high-pass filtered at 0.01 Hz. One-sample t tests against a contrast value of zero were applied to the resulting parameter estimates to obtain a t value for each voxel. The t values were extracted from all voxels within an ROI to produce a vector of voxel activation (activity pattern) for each trial. We additionally regressed out effects of no interest, including stimulus type (i.e., faces vs scenes) and repetition condition (i.e., first vs second presentation) from the trial-by-trial activation patterns.

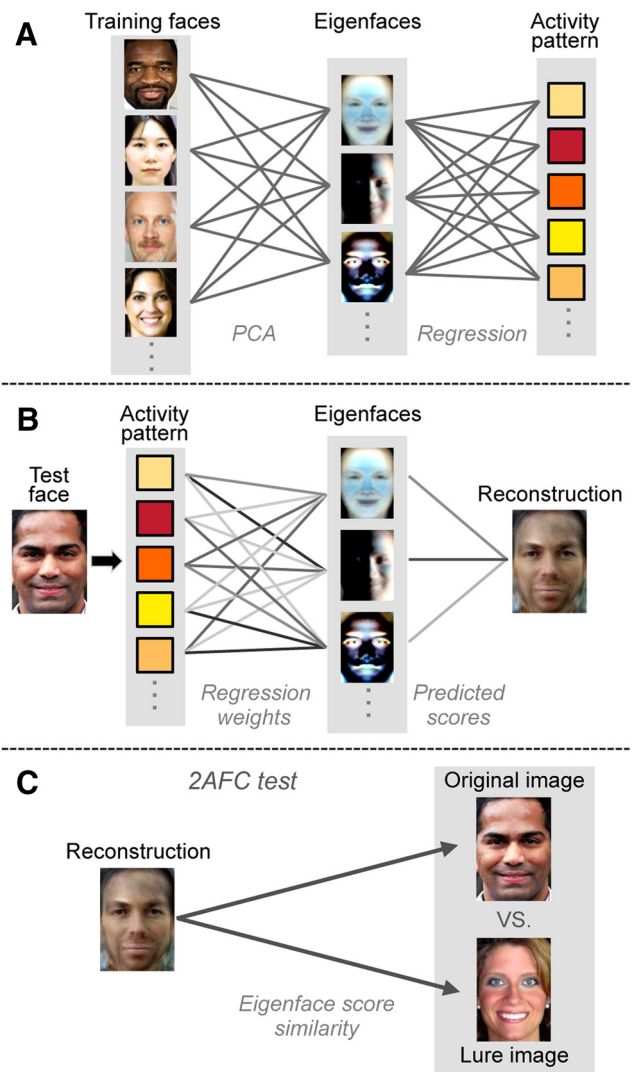


Figure 2. Schematic of the face reconstruction analysis. **A**, Model training. First, PCA was run on a large set of training face images to extract a relatively small number of eigenfaces, so that any face image could be efficiently expressed as a weighted sum of the eigenfaces. A regularized linear regression was next applied to estimate a model that could predict the eigenface scores of the training faces from the voxel responses they evoked. **B**, Model testing. Voxel responses evoked by an independent test face were measured, and the eigenface scores of the test face were predicted by applying the weights estimated from the trained regression model to the measured responses. Finally, a reconstruction was generated by taking the linear combination of the eigenfaces using the predicted scores. **C**, Assessing reconstruction accuracy. The accuracy of a reconstruction was evaluated with 2AFC tests, which compared the Euclidean distance between the eigenface scores of the reconstructed face and those of its original image or other test images (“lures”). A reconstruction was counted as accurate when it was more similar (lower Euclidean distance) to the original image than to the lure image.

Reconstructions were generated using an N -fold cross-validation, where N equals the number of scanning runs. For each iteration, the activation patterns of the four test faces from a “held-out” scanning run were averaged over repetitions, resulting in a single activation pattern for each of the four test faces in that run. However, to maximize power, only the 4 test faces from a given run were held out of the training; all of the nontest face trials in the run and all of the face trials in the other runs of the perception phase were used as training patterns to derive the relationship between eigenface scores and brain activation. The mapping between eigenface scores and brain activation is described by the following linear model:

$$Y_{\text{train}} = X_{\text{train}}W$$

where Y_{train} (m faces $\times n$ components) represents the eigenface scores of the original training faces, X_{train} (m faces $\times p$ voxels) represents the activation patterns of all training trials within an ROI, and W (p voxels $\times n$ components) represents the weights linking each voxel's BOLD signal intensity to eigenface scores. We used ridge regression with a penalization parameter (λ) of 1 to estimate the weight matrix \hat{W} with the following formula:

$$\hat{W} = (X_{\text{train}}^T X_{\text{train}} + \lambda I)^{-1} X_{\text{train}}^T Y_{\text{train}}$$

where I is the identity matrix. Given the estimated weights, we predicted the eigenface scores of test faces Y_{test} (q faces $\times n$ components) from the BOLD activity patterns of test faces X_{test} (q faces $\times p$ voxels) as follows:

$$Y_{\text{test}} = X_{\text{test}} \hat{W}$$

Both the eigenface scores of training faces and the brain activity patterns were de-meaned before being entered into the regression analysis. The mean eigenface scores of the training faces were then added back to the predicted scores. Finally, reconstructions of the test face image vectors F_{recon} (q faces $\times k$ pixel values) were generated via the following formula:

$$F_{\text{recon}} = Y_{\text{test}} C^T + F_{\text{mean}}$$

The accuracy of reconstruction for each face was determined by two-alternative forced choice (2AFC) tests, which assessed whether a reconstructed face was more similar to (1) the original face from which the reconstruction was generated, or (2) another randomly selected original image ("lure"). Similarity was measured by the Euclidean distance between predicted versus actual eigenface scores. Thus, for a given face, if the Euclidean distance between predicted and actual eigenface scores was lower than between predicted and lure eigenface scores, that reconstruction would be scored as "accurate" in the 2AFC test. Each reconstruction was tested against all other test faces. Thus, $N \times (N - 1)$ comparisons were made per ROI per session, where N is the total number of test faces. The proportion of trials for which the reconstruction was counted as correct represented the reconstruction accuracy of the ROI for the corresponding session (chance level = 50%).

Face reconstruction analysis: memory

To generate reconstructions of faces held in memory, we used the same approach that was used for reconstructing test faces during the perception phase, with a few minor differences. First, the regression model was trained on the entire set of perception phase data and tested on the "held out" memory phase data. However, rather than using t statistic maps for the activation patterns, here we used preprocessed raw fMRI data. We used raw activity patterns instead of t statistic maps derived from a GLM because it is difficult, with a GLM, to reliably separate activity evoked by the sample stimuli versus activity that is selective to the delay period. Thus, using the raw data represented a more conservative approach for measuring "true" delay-period activity. To create a single activity pattern for each face trial of the perception phase, volumes corresponding to 6–10 s poststimulus onset were averaged (the delay accounted for the hemodynamic response lag). To create a single activation pattern for each trial of the memory phase, volumes corresponding to 4–10 s post-cue onset were averaged. We averaged volumes across a longer time period for the memory phase data because we expected delay-period maintenance of the cued face to evoke more sustained processing than the perception phase. We regressed out the effects of conditions other than the identity of faces in both perception (i.e., stimulus types and repetition) and memory trials (i.e., repetition, match/nonmatch condition, and accuracy). Finally, activation patterns from the memory phase were averaged across the four repeated trials (i.e., the four trials on which a particular face was cued) to produce a single activation pattern per remembered face.

As with perception-phase data, reconstruction accuracy was determined via a 2AFC test where the reconstruction (predicted eigenface scores) was compared with the original image and a lure image. However, for the memory-phase reconstructions, the lure image for each 2AFC test trial was always the uncued face from the same trial (the "other face"

from the sample period). Because the reconstructed face was directly compared with the cued versus uncued faces, above-chance reconstruction accuracy could not be explained by persistence of visual responses evoked by the sample images; rather, successful reconstruction required top-down memory for the cued face. Again, we used the Euclidean distance between predicted and actual eigenface scores as a measure of similarity. The reconstruction was considered "accurate" if it was more similar to the target than to the lure image (chance level = 50%). The comparison was made for every face in the memory phase, and the proportion of accurate reconstructions represented the accuracy of the ROI.

Statistical tests

Unless otherwise noted, tests of statistical significance, including for reconstruction accuracies, were based on randomization (permutation) tests with 10,000 iterations. Permutation tests were used instead of parametric tests to minimize assumptions about the data (namely, that data were normally distributed or that chance accuracy is 50%). The specific approach we used involved a rigorous bootstrapping method that allows for group-level statistical tests (Stelzer et al., 2013). To derive chance-level accuracies for the face reconstructions, we randomly shuffled the face identity labels for reconstructed and original images and then computed similarity measures (Euclidean distance) between images. We computed this "null" result 100 times per session and averaged results over sessions from a single subject so that each subject had 100 chance accuracies. Next, we randomly selected one of the null accuracies per subject and averaged them over subjects to produce a group-level value representing mean chance accuracy. We repeated the group-level sampling 10,000 times to generate a null distribution of mean group-level accuracies. The p value was defined as the proportion of mean accuracies from the null distribution that were equal to or greater than the actual mean accuracy. In randomization tests that compared different conditions, we shuffled the labels for the conditions of interest to produce the null distribution of the difference between the conditions. In cases where there were more than two conditions to be compared, we generated the null distribution of F statistics by running ANOVAs based on the shuffled labels. The proportion of values from the null distribution that were equal to or more extreme than the actual difference or the F statistic was reported as the p value. Because we had a priori predictions concerning ANG (Kuhl and Chun, 2014), we do not apply corrections for multiple comparisons when analyses included multiple ROIs. Thus, all p values are uncorrected.

Subjective ratings of face images

To better characterize the information reflected in reconstructed face images, we conducted a separate behavioral experiment in which independent groups of subjects rated the original and reconstructed face images according to the following five dimensions: brightness of the skin, gender, emotional expression, dominance, and trustworthiness. The latter two dimensions (dominance and trustworthiness) were included based on evidence that they are major traits that explain >80% of variance in face-based social judgments (Todorov et al., 2008). Reconstructions were generated from ANG and OTC (based on 300 eigenfaces), using the perception-phase data from Experiments 1 and 2. Ratings based on the original faces were then compared with the ratings based on the reconstructions.

All subjects were recruited online via Amazon's Mechanical Turk (Mason and Suri, 2011) using the psiTurk (McDonnell et al., 2014) system. A total of 158 subjects were provided with an informed consent form approved by the University of Oregon Institutional Review Board, and rated the original and reconstructed face images from Experiment 1. A total of 99 subjects were provided with an informed consent form approved by the New York University Institutional Review Board, and rated the images from Experiment 2. For Experiment 1, 6 subjects rated the original test faces, 77 subjects rated the reconstructions generated from ANG, and 75 subjects rated the reconstructions generated from OTC. For Experiment 2, 6 subjects rated the original test faces, 46 subjects rated the reconstructions generated from ANG, and 47 subjects rated the reconstructions generated from OTC. Thus, independent samples of subjects rated the reconstructed and original faces.

The experiment was run in web browsers using JavaScript. Each subject rated the whole set of 36 original test faces or 28–36 reconstructed faces generated from a single session of a single fMRI subject. Reconstructions from a given fMRI session were rated by 4–6 Mechanical Turk subjects. The number of reconstructed faces depended on the number of scanning runs an fMRI subject completed within a session. The order of image presentation was randomized for each subject. In each trial, a face image and five slider bars representing the five dimensions appeared on the screen. Opposing adjectives (Dark-skinned vs Light-skinned, Feminine vs Masculine, Happy vs Unhappy, Dominant vs Submissive, and Trustworthy vs Untrustworthy) were presented on the left and right end of each slider bar. Sliders were initially presented at the middle of the slider bars. Subjects were asked to click on the sliders, using a 0–10 scale, to indicate how well each adjective described the face. Subjects received a pop-up message when they made responses too quickly (<2 s to rate all five dimensions) or did not make any changes on the sliders. These messages thus encouraged subjects to thoughtfully engage in the task. Seventeen subjects (6.6% of all subjects) received a warning that they responded too quickly. Seven subjects (2.7% of all subjects) received a warning that they failed to move the sliders. The average number of warnings (regardless of the message) per subject was 0.097.

The ratings for each dimension for each original face image were averaged across the six Mechanical Turk subjects. The ratings for the reconstructions from each ROI were first averaged within each fMRI session across Mechanical Turk subjects and then averaged across sessions (for cases where an fMRI subject completed more than one session of the main experiment). The mean ratings were finally averaged across fMRI subjects to generate a single score per dimension per reconstructed test face. The ratings for both the original and reconstructed images were *z*-scored across faces within experiments, and the ratings from the two experiments were then combined. The *z*-scored ratings of four test faces that were used in both Experiments 1 and 2 were averaged across experiments so that they were not “double-counted.” This resulted in a total of 68 test faces (32 unique to Experiment 1, 32 unique to Experiment 2, and 4 presented in both experiments) to be analyzed and a pair of scores (original vs reconstruction) per dimension per face and for each ROI.

Change detection experiment

To validate the use of eigenface scores in assessing remembered as well as perceived face representations, we tested whether eigenface score similarity could predict confusability between faces in a behavioral working memory task. We ran an independent experiment using a change detection paradigm (see Fig. 9A). On each trial, subjects studied a set of face images and were then tested, after a brief delay, as to whether a single probe face had changed or not. The probe, which appeared at one of the studied spatial locations, was either the same as or different from the face that had been studied at that location. To the extent that eigenface scores capture subjective face information, we expected that subjects should be less likely to detect face changes when the probe and studied faces were more similar in terms of eigenface scores. In other words, lower Euclidean distance between the studied and probe face was expected to result in greater confusability.

We recruited 33 subjects on Amazon’s Mechanical Turk using the psiTurk (McDonnell et al., 2014) system. Five additional subjects were excluded because they did not make enough incorrect responses when faces changed, leaving less than five analyzable trials. All subjects were provided an informed consent form approved by the University of Oregon Institutional Review Board. We used the same pool of training face images as in Experiment 1 (excluding the faces used in the memory phase). For each subject, a different set of 32 face images was randomly selected from the face pool and was repeatedly used throughout the experiment.

The experiment was run in web browsers using JavaScript. Each experimental block consisted of 24 trials. Each trial started with a fixation cross on a gray background. After 1.5 s, six randomly selected sample faces were presented for 2 s, followed by a 2 s delay. During the delay, the sample faces were masked by patterns composed of scrambled face parts. The faces and masks appeared at six fixed locations surrounding the fixation cross. The distance from the fixation to the center of an image

was $\sim 102\%$ of the height of the image. The average distance between the centers of neighboring images was the same as the height of the image. At the end of the delay, one of the six masks was replaced by a probe face. In half of the trials, the probe face was identical to the sample face that had appeared at that location (“match” condition). In the other half of the trials, the probe face was randomly selected from the set of 32 faces, excluding faces from the sample array on the current trial (“nonmatch” condition). The probe was equally likely to appear at any of the six locations in both conditions. The subjects’ task was to remember the sample faces throughout the delay and indicate whether the probe matched the sample (at that location) by pressing one of two keyboard buttons (“M” for match and “N” for nonmatch). The probe face and masks disappeared upon the subjects’ response, which initiated the next trial. Each subject completed a varied number of blocks (30 subjects completed 7 blocks; 1 subject completed 4 blocks; 2 subjects completed 3 blocks). The first block of the experiment served as a practice round and was excluded from analysis. Subjects were given feedback on their performance and were allowed a short break at the end of every block.

To obtain eigenface scores for each face in the experiment, PCA was applied to all 960 images from the training face pool. An eigenface score vector was constructed for each face, consisting of scores for the first 300 principal components. For each trial, dissimilarity between the sample and the probe face was defined as the Euclidean distance between the two eigenface score vectors, which was 0 in the match condition. Distance values in nonmatch trials were *z*-scored within subjects.

Results

Behavioral performance

Unless otherwise stated, for this and all following analyses, group-level results were obtained by averaging first across sessions within a subject and then across subjects. The mean sensitivities (d') for the repetition detection task performed during the perception phase were 2.32 and 2.28 ($SD = 0.69$ and 0.56) in Experiments 1 and 2, respectively. The average reaction times for the correct trials were 1472 ms and 1521 ms ($SD = 232$ and 532 ms), respectively. The average accuracies for the memory task in Experiments 1 and 2 were 93% and 91.9% ($SD = 5.4\%$ and 5.8%), with reaction times of the correct trials being 1682 ms and 1643 ms ($SD = 316$ ms and 403 ms), respectively. For both the perception phase and the memory phase, there was no difference between experiments in terms of accuracy or reaction time (p values >0.2).

Reconstruction of perceived faces

We first examined whether we could reconstruct perceived face images from ANG, other lateral parietal regions, and/or OTC (Fig. 3B). We used the first 300 eigenfaces, which explained 95.7% and 94.5% of the variance across images in Experiment 1 and Experiment 2, respectively. Example reconstructions from a single subject’s data are shown in Figure 4. Reconstruction accuracy was obtained from 2AFC tests, in which a reconstruction was counted as correct if its eigenface scores were more similar to those of the original image than to those of a lure image.

Consistent with our previous study (Cowen et al., 2014), robust reconstruction accuracies were achieved from face-prefering voxels in OTC in both Experiments 1 and 2 (66.7% and 58.0%, respectively), and these accuracies were well above chance (p values <0.0005). In both experiments, we were also able to reconstruct face images with above-chance accuracy from ANG and from all other parietal ROIs (ANG: 56.3%, SMG: 55.9%, IPS: 58.3%, SPL: 57.6% in Experiment 1; ANG: 55.4%, SMG: 57%, IPS: 56.9%, SPL: 57.7% in Experiment 2; p values <0.0005 ; for results from finer-grained parietal subregions, see Fig. 7B, left). No significant differences between experiments were found in any of the parietal ROIs (p values >0.4). For the

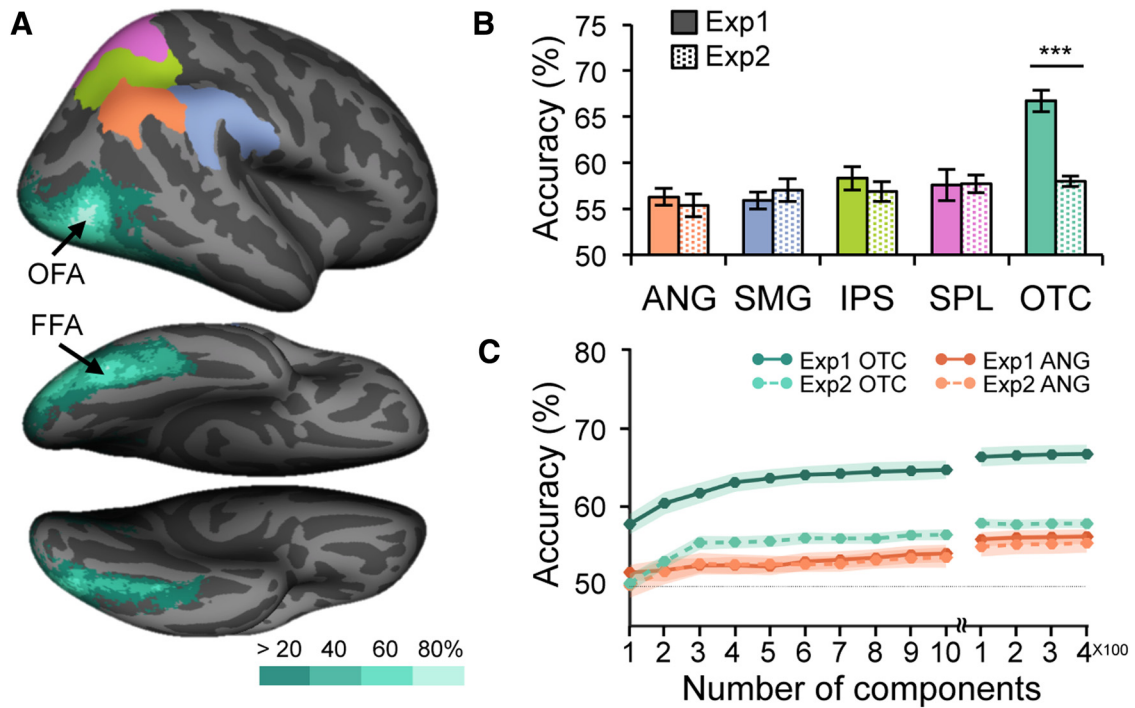


Figure 3. *A*, Anatomical and functional ROIs (orange represents ANG; blue represents SMG; yellow/green represents IPS; magenta represents SPL; turquoise represents probabilistic map of the 2500 most face-prefering voxels in OTC), visualized on the inflated surface of an averaged template brain supplied by FreeSurfer. Top, Right lateral view. Bottom, Ventral view. Turquoise color scale represents the percentage of sessions (computed from all sessions collapsed across experiments) where a voxel is included in the OTC functional ROI mask. Regions corresponding to the fusiform face area (FFA) and occipital face area (OFA) are included in the OTC functional ROI in most sessions. *B*, Mean reconstruction accuracies for each ROI in each experiment. Reconstructions were generated from the perception phase data using 300 eigenfaces. Error bars indicate SEM across subjects. *C*, Mean reconstruction accuracy for perceived faces in ANG and OTC as a function of the number of components. Shaded areas represent SEM across subjects. $***p < 0.001$.

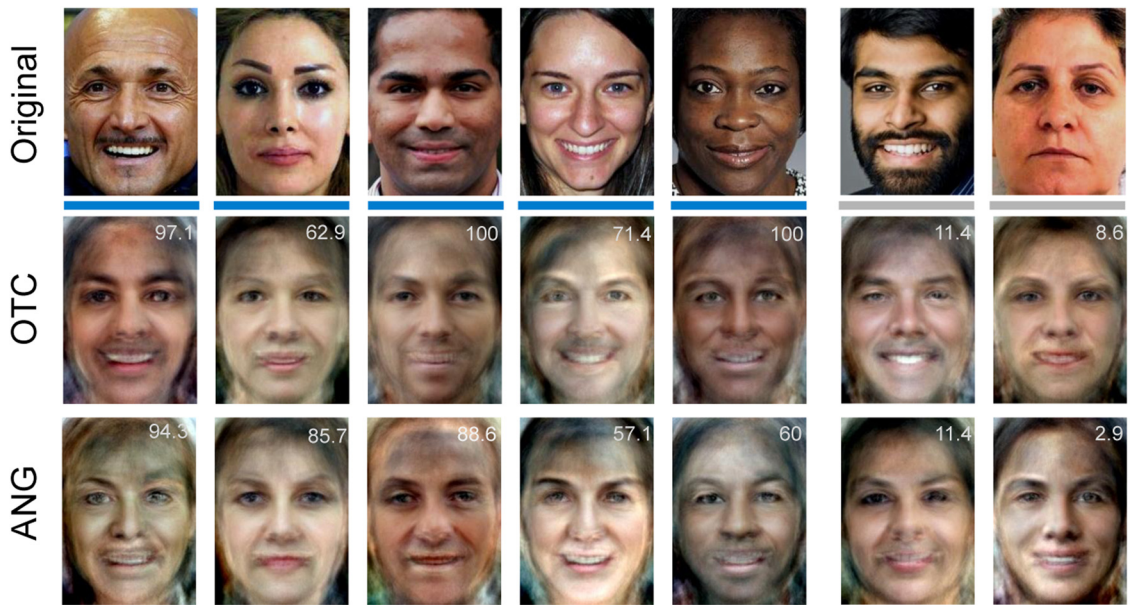


Figure 4. Example face images and their reconstructions from a single subject's perception phase data. Top row, Original test faces. Middle row, Reconstructions from OTC. Bottom row, Reconstructions from ANG. The five left-most columns (with blue bars above the reconstructions) represent examples of successful reconstructions ($>50\%$ 2AFC accuracy). The two right-most columns (with gray bars above the reconstructions) represent examples of unsuccessful reconstructions ($<50\%$ 2AFC accuracy). Gray numbers at the upper right-hand corners of each reconstruction indicate the mean 2AFC accuracy for each reconstructed image (for the selected subject). Reconstructions were generated using 300 eigenfaces.

OTC ROI, however, accuracy in Experiment 2 was substantially lower ($p = 0.0001$) than in Experiment 1 and was no longer higher than accuracy in the parietal ROIs ($p = 0.374$). While direct comparison across the two experiments requires caution because there were several minor procedural differences, it is very

likely that the drop in reconstruction accuracy in OTC in Experiment 2 was due to the reduced variability across stimuli in low-level visual information, such as overall luminance. In contrast, parietal regions were apparently not sensitive to these changes in low-level visual properties. We also confirmed that, for both ex-

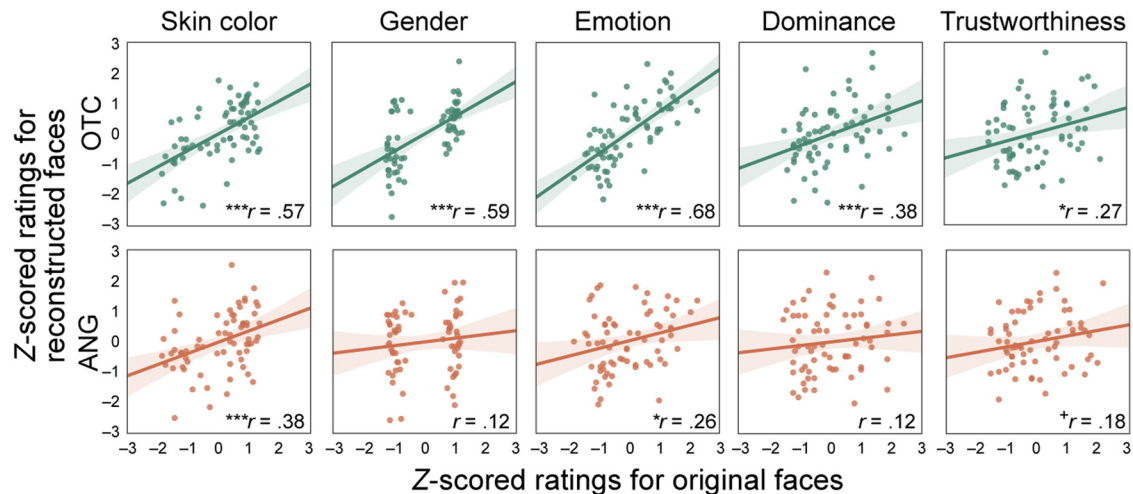


Figure 5. Correlations between z-scored subjective ratings for original faces and those for reconstructions from OTC (top row) and ANG (bottom row). Reconstructions were generated with the perception phase data using 300 eigenfaces. Each dot indicates a test face. Solid lines indicate fitted regression lines. Shaded areas represent 95% bootstrapping confidence intervals. $^+p < 0.1$. $*p < 0.05$. $***p < 0.001$.

periments, reconstruction accuracies remained significant for each ROI if accuracy was determined by comparing the raw pixel intensities of the original versus reconstructed images instead of the eigenface scores (p values < 0.0005).

We next compared reconstruction accuracies as a function of the number of eigenface components that were included in the models. Specifically, we generated reconstructions based on 1–10, 100, 200, 300, and 400 components for each of our primary ROIs: ANG and OTC. We started with the first component, which captured the most variance, and gradually added later components (Fig. 3C). In both ROIs, the first two components were sufficient to produce above-chance accuracy in both experiments (p values < 0.05). Importantly, although a few early components carried the most salient visual information (e.g., overall brightness or the direction of lighting) and thus were sufficient to produce above-chance accuracies, the accuracies obtained with 300 components were significantly higher than those obtained with only 10 components in both experiments and in both ROIs (p values < 0.005). Thus, while adding additional components yielded diminishing returns, as would be expected, relatively late components still contributed to reconstruction accuracy.

Finally, we tested whether reconstructions from OTC were correlated with those from PPC. For each reconstructed face, we correlated the predicted eigenface scores from OTC with the predicted eigenface scores in each PPC ROI. OTC reconstructions positively correlated with reconstructions from each PPC subregion (p values < 0.0005 , compared with randomization baseline; see Fig. 8, left). Moreover, OTC-PPC correlations significantly differed across the PPC ROIs ($p = 0.0005$), with OTC reconstructions being most similar to reconstructions from ANG.

Subjective ratings of reconstructed facial attributes

The above-chance reconstruction accuracies in ANG and OTC demonstrate that these regions carried enough information to distinguish individual faces. However, these reconstruction accuracies based on Euclidean distance of eigenface scores do not necessarily indicate that reconstructions were subjectively compelling. Moreover, reconstruction accuracy alone does not provide information about the specific face dimensions that may have been successfully reconstructed. To address these issues, we

had an independent set of human subjects rate the reconstructed faces on the following face dimensions: skin color, gender, emotional expression, dominance, and trustworthiness. The Pearson's correlation between the ratings for the reconstructions and the original face images served as a quantitative measure of how well each face dimension was reconstructed. The statistical significance of the correlations was determined with randomization tests by randomly pairing the reconstructions and original faces and computing r 10,000 times for each dimension in each ROI. The proportion of r values in the resulting null distribution that were equal to or greater than the actual correlation coefficient was used as the p value of the test.

In all five dimensions we tested, we found positive cross-image correlations between the average z-scored ratings for the original test faces and their reconstructions generated from OTC with 300 components (Fig. 5; p values < 0.05). In ANG, correlations were significantly positive for emotion (happy vs unhappy) ($r = 0.26$, $p = 0.015$) and skin color ($r = 0.38$, $p = 0.0008$), and marginally positive for trustworthiness ($r = 0.18$, $p = 0.069$). These results indicate that “high-level” face information was subjectively evident in the reconstructions generated from ANG and OTC.

To make it clearer that we were reconstructing more than low-level visual information, we also regressed out the skin color ratings (dark-skinned or light-skinned) from the other four dimensions (separately for the original images and reconstructions) and computed correlations with the residuals. Although the skin color dimension was related to the race or ethnicity of the individual, it was also tightly related to the overall luminance of the image. Skin color ratings also had weak to moderate correlations with other dimensions when we correlated ratings from the original images (r values = 0.22, 0.13, 0.24, and 0.11 for gender, emotion, dominance, and trustworthiness, respectively). In OTC, the correlations between the reconstructions and original images were still significantly positive for the remaining four dimensions after removing the effect of skin color (r values = 0.51, 0.67, 0.32, and 0.31 for gender, emotion, dominance, and trustworthiness, respectively; p values < 0.01). In ANG, the correlations remained positive for emotion ($r = 0.26$, $p = 0.016$) and marginally positive for trustworthiness ($r = 0.16$, $p = 0.093$).

Reconstruction of faces retrieved from memory

The preceding analyses indicate that individual face images can be reconstructed from ANG activity patterns when the face images are visually present. Next, we extended our method to test whether reconstructions could be generated for faces retrieved from memory in the absence of visual input. We estimated the relationship between eigenface scores and fMRI activity patterns with the perception phase data using 300 eigenfaces. We then applied the resulting model to the patterns obtained during the delay period of trials from the memory phase (4–10 s from the cue onset). The reconstruction of a face held in memory was considered accurate in the 2AFC test if the predicted eigenface scores were more similar to those of the cued face image than to the lure image (i.e., the noncued face from the same trial). Because we specifically compared the reconstructed image to the cued versus lure images, above-chance reconstruction accuracy could not reflect “carryover” activity from the sample images. Rather, above-chance reconstruction accuracy required top-down, selective maintenance of the cued face.

Combining data from Experiments 1 and 2, we found above-chance reconstruction accuracy in ANG (mean 54.3%, $p = 0.0021$; Fig. 6A). Considering the experiments separately, reconstruction accuracy in ANG was marginally above chance in Experiment 1 (mean 53.6%, $p = 0.06$) and significantly above chance in Experiment 2 (mean 54.9%, $p = 0.0097$), with no significant difference between the experiments ($p = 0.769$). Reconstructions were not significantly above chance in OTC or other PPC ROIs (p values >0.1). Examples of memory-based reconstructions from a single subject’s ANG are shown in Figure 6C. We also obtained virtually identical results if ANG reconstruction accuracy was determined based on raw pixel intensities of the images instead of the eigenface scores (accuracy across experiments: $p = 0.005$; Experiment 1: $p = 0.027$; Experiment 2: $p = 0.043$).

Among the finer-grained ventral PPC subregions from Yeo et al. (2011), reconstruction accuracies were above chance in most of the posterior networks that at least partially overlapped with the anatomically defined ANG ROI (networks 12, 13, 15, and 17; p values <0.05 ; see Fig. 7B, right). The localization of these memory-based reconstructions overlaps with an area of lateral parietal cortex previously associated with vivid remembering and stimulus-specific memory representations (Kuhl and Chun, 2014).

For qualitative assessment of the time course of memory-based reconstructions over the delay period, we also created reconstructions based on sliding time windows with a duration of 3 TRs (i.e., 3 brain volumes). As shown in Figure 6B, reconstruction accuracy in ANG was numerically highest toward the latter portion of the delay period. In OTC, recon-

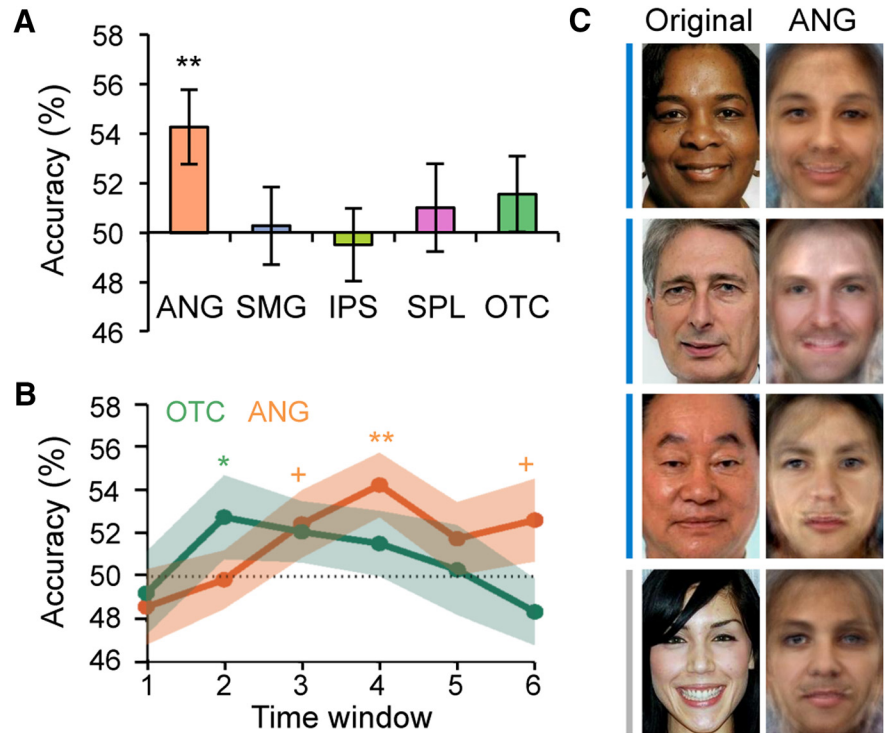


Figure 6. *A*, Mean memory-based reconstruction accuracies for each ROI, combined across experiments. Reconstructions were generated with patterns averaged across 3–5 TRs from the cue onset using 300 eigenfaces. Error bars indicate SEM across subjects. *B*, Mean reconstruction accuracies in ANG and OTC over the course of the delay period. The first time window started with the onset of the sample faces and the last time window stopped 2 TRs before the probe onset (to avoid “contamination” by the partial-face probe image). Reconstructions for each time window were generated using fMRI activation patterns averaged within the window. The average accuracy in ANG was significantly above chance at the fourth time window (mean 54.3%, $p = 0.0021$) and marginally above chance at the third (mean 52.4%, $p = 0.095$) and the sixth time window (mean 52.6%, $p = 0.065$). Shaded areas represent SEM across subjects. Dotted line indicates chance level. $+p < 0.1$. $*p < 0.05$. $**p < 0.01$. *C*, Example target faces from the memory phase (left column) and corresponding reconstructions from ANG from a representative single subject’s data (right column). The first three rows (with blue bars on the left) represent successful (more similar to the target than to the lure image) reconstructions, and the last row (with a gray bar on the left) represents an unsuccessful (more similar to the lure than to the target image) example for comparison. Reconstructions were generated with patterns averaged across 3–5 TRs from the cue onset using 300 eigenfaces.

struction accuracy was above chance early in the delay period (second time bin: mean 52.8%, $p = 0.032$), but accuracy decreased to chance levels afterward. In addition, correlations between predicted eigenface scores from OTC and ANG were significantly positive over each time bin during working memory trials (p values <0.002 ; significance determined via randomization test), but the correlations significantly increased over the course of the trial (last three time bins vs first three time bins: $p = 0.029$; Fig. 8, right). Although these results potentially suggest an interaction between ANG and OTC that increased as top-down memory representations were established, caution is warranted given the weak overall reconstruction accuracy from OTC. Of critical relevance, these findings provide evidence for stimulus-specific reconstructions of memory-based representations from ANG.

Validation of using eigenface score similarity

All of the above analyses use eigenface values as a means for mapping perceived or remembered face images to fMRI activity patterns. However, to further confirm the validity of using eigenfaces to study memory representations, we tested, in a separate behavioral study, whether behavioral measures of memory-based confusability between individual faces can be predicted by eigenface scores (see Fig. 9A).

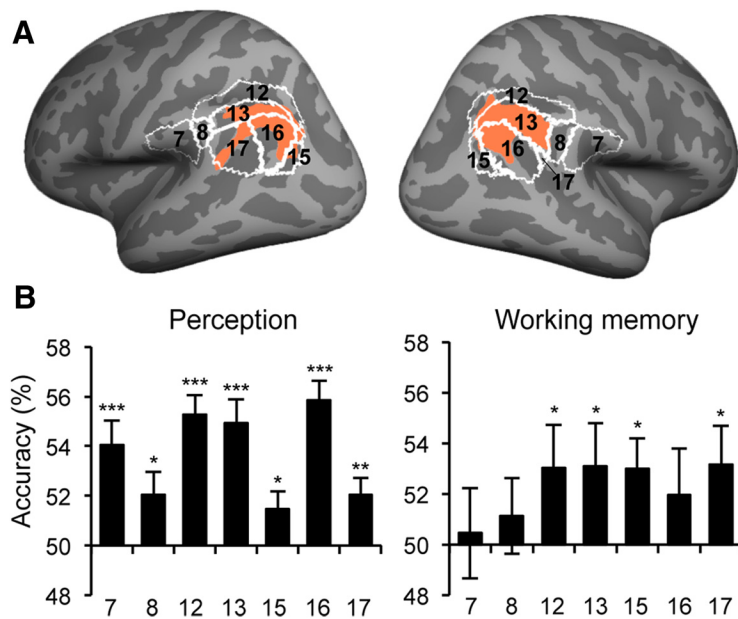


Figure 7. Reconstruction accuracy in ventral PPC subregions defined from independent intrinsic functional connectivity analyses (Yeo et al., 2011). **A**, ROI visualization on the lateral surface of both hemispheres of the FreeSurfer template brain. White lines indicate the boundaries of each subregion from Yeo et al. (2011). Black numbers indicate the names of the networks from which each subregion is derived. Orange areas represent anatomically defined ANG based on FreeSurfer parcellation. **B**, Mean accuracies for the perception-based (left) and working memory-based (right) reconstructions. All reconstructions were generated using 300 eigenface components. Memory-based reconstructions were generated from activation patterns averaged across 3–5 TRs from the cue onset. **A**, **B**, Means and SEMs were computed from the data combined across experiments. Error bars indicate SEM across subjects. Significance is against chance. * $p < 0.05$. ** $p < 0.01$. *** $p < 0.001$.

We first removed the 1.8% of trials which had a reaction time that deviated by more than 3 SDs from the mean of each subject. Mean accuracy on match trials was 68.6% (SD 13.9%) and mean accuracy on nonmatch (change) trials was 75.2% (SD 12.9%). Accuracy did not significantly differ across the six probe locations ($p = 0.059$). Of critical interest was whether the failure to detect changes on nonmatch trials was predicted by the similarity between the studied face (sample) and probe. Figure 9B shows examples of sample and probe face pairs with high, medium, and low eigenface similarity, selected from actual nonmatch trials. We first ran a logistic regression analysis using all nonmatch trials collapsed across subjects. The numbers of correct (match) and incorrect (nonmatch) responses were controlled to be identical within each subject by randomly selecting the same number of trials from each condition. On average, 31 nonmatch trials were included in the analysis per subject (range: 10–62 trials). Indeed, we found that greater eigenface score similarity (i.e., lower z -scored Euclidean distance) was associated with a higher probability of making incorrect match responses (see Fig. 9C; $\beta = -0.17$, odds ratio = 0.84, $p = 0.007$). This effect remained significant when we added the probe location as a fixed effect and subject and a subject-by-location interaction as random effects (likelihood ratio test against the null model; $\chi^2_{(1)} = 8.996$, $p = 0.003$). A simple comparison of the mean distances also confirmed that distances were lower for incorrect match responses (mean -0.15 , SD 0.31) than for correct change detection (mean 0.05, SD 0.08) without controlling for the number of trials ($p = 0.0005$). Thus, if two faces were more similar in terms of their eigenface scores, subjects were more likely to confuse them as the same face in a memory task. These results suggest that eigenface scores can be a valid measure of subjective similarity and can predict behavioral memory performance even though they are

derived from raw pixel intensities in a completely data-driven way.

Discussion

The current study tested for content representations in lateral parietal cortex during perception and memory. We used a data-driven method for extracting information from face images (eigenfaces) (Turk and Pentland, 1991) and then modeled relationships between eigenface scores and fMRI activity patterns, with the goal of predicting a face's eigenface scores from the fMRI activity patterns that it evoked. Because predicted eigenface scores can be transformed into native face space, this allowed us to reconstruct images of individual faces by linearly combining the predicted components (Cowen et al., 2014). Across two fMRI studies, we show that reconstructions were successfully generated from activity patterns within the angular gyrus when faces were visually present or maintained in memory, providing compelling evidence for stimulus-specific content representations in this subregion of lateral parietal cortex.

Nature of ANG content representations

Our findings, along with a handful of recent pattern-based fMRI studies, provide evidence that activity patterns in ANG reflect the contents of mnemonic processing (Kuhl and Chun, 2014; Bird et al., 2015; St-Laurent et al., 2015). To our knowledge, however, our study represents the first targeted effort to reconstruct mnemonic content from ANG, and our methodological approach affords several important insights into the nature of ANG content representations.

First, in contrast to more typical pattern classification or pattern similarity analyses, our analyses were based on relationships between neural activity patterns and latent components underlying stimuli (i.e., eigenfaces). This detail is important because it indicates that the stimulus-specific representations we observed in ANG cannot be attributed to subjects generating verbal labels or other stimulus-specific “tags.” Instead, reconstructions could only be “built” by combining predicted face “parts,” and these parts (eigenfaces) were derived from an entirely independent set of faces. Likewise, because we did not impose an explicit categorical task structure on the stimulus set (e.g., discrimination of happy vs sad faces), the content representations we observed in ANG are not easily explained in terms of adaptive coding of a behaviorally relevant dimension (Toth and Assad, 2002).

Another unique aspect of our approach is that reconstructed content could be assessed via objective and subjective measures. For our subjective test of reconstruction accuracy, we selected previously validated face dimensions (Todorov et al., 2008) and asked independent groups of subjects to rate the original and reconstructed faces along these dimensions. We found that emotional expression and skin color, and to a lesser extent, trustworthiness, were successfully reflected in reconstructions generated from ANG. These findings could either reflect low-level representations in ANG that give rise to these higher-level forms of face information or that ANG directly codes for higher-level face in-

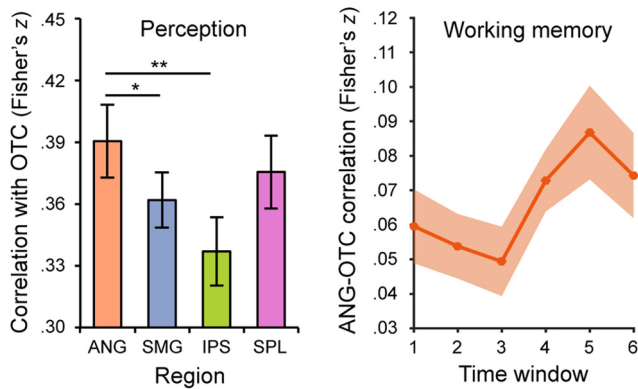


Figure 8. Correlations between predicted eigenface scores from OTC and PPC. For each reconstructed face, Pearson's correlation was computed between the 300 predicted eigenface scores from each region. The correlation coefficients were transformed to Fisher's z , and averaged across faces. Left, Correlations between OTC and each of the anatomically defined PPC subregions during the perception phase (ANG: 0.39, SMG: 0.36, IPS: 0.34, SPL: 0.38). OTC-ANG correlations were significantly higher than OTC-SMG or OTC-IPS correlations ($p = 0.037$ and $p = 0.0017$, respectively). OTC-SMG and OTC-SPL correlations were also higher than OTC-IPS correlation ($p = 0.042$ and $p = 0.0008$, respectively). Error bars indicate SEM across subjects. $*p < 0.05$. $**p < 0.01$. Right, Correlations between ANG and OTC during the working memory phase at each 3-TR-long time window over the course of the delay period. The first time window started with the onset of the sample faces. Correlations were significantly stronger during the last three time windows than the first three time windows ($p = 0.029$). Shaded areas represent SEM across subjects. In both panels, means and SEM were computed from the data combined across experiments.

formation that may be semantic, conceptual, or affective in nature. Consistent with the latter possibility, evidence from other fMRI studies indicates that ANG, along with the neighboring superior temporal sulcus, contributes to social evaluation of faces (Allison et al., 2000) and that ANG is a core region in semantic/conceptual processing (Binder and Desai, 2011). It is also notable that ANG reconstruction accuracies did not vary across Experiments 1 and 2, despite substantial differences in variance of low-level visual information (which contrasted sharply with OTC). Thus, although the reconstructions we generated were visual in format, it is possible that the underlying representations in ANG were not visual at all. This possibility reflects the fact that our method will reconstruct nonperceptual information so long as it covaries with eigenface scores. In a similar vein, establishing a mapping between ANG activity patterns and eigenface scores does not indicate that ANG face representations are based on eigenfaces. Rather, we only capitalize on the fact that eigenfaces (1) capture a large amount of face information and (2) can be correlated with patterns of brain activity. That said, our independent behavioral study confirmed that memory-based confusability between individual faces can be predicted by the similarity of corresponding eigenface scores, validating eigenfaces as a useful means for studying face memory.

Lateral parietal involvement in perception and memory

The role of lateral parietal cortex in memory has been a topic of increasing interest largely because of consistent observations that univariate ANG activation increases during successful, compared with unsuccessful, retrieval of event details, so-called "retrieval success effects" (Wagner et al., 2005; Rugg and Vilberg, 2013). In contrast, ANG tends to show little task-evoked activation during memory encoding/perception (Daselaar et al., 2009), suggesting that ANG contributions to memory are restricted to retrieval. Indeed, it has been argued that ANG tracks internally oriented thoughts/memories and

therefore does not code for externally presented information. However, there is recent evidence that ANG activity patterns reflect encoded content (Xue et al., 2013) and that content-sensitive ANG activity patterns evoked during encoding are reinstated at retrieval (Kuhl and Chun, 2014; Bird et al., 2015; St-Laurent et al., 2015). Thus, it is notable that we observed robust reconstructions generated from ANG activity patterns during perception and that a model trained on ANG perception-based patterns successfully transferred to ANG memory-based patterns. Thus, our findings strongly argue against a selective role for ANG in memory retrieval. Instead, although it is clear that univariate activity in ANG varies according to encoding versus retrieval processes (Vannini et al., 2011; Huijbers et al., 2013), patterns of activity in ANG "ride on top" of these univariate changes and reflect what is being perceived or remembered.

Our a priori interest in ANG was motivated by prior univariate (Wagner et al., 2005; Cabeza et al., 2008; Vilberg and Rugg, 2008) and pattern-based (Kuhl and Chun, 2014) fMRI studies that have related this region to memory. Notably, in a previous study in which we introduced the current face reconstruction method (Cowen et al., 2014), we also found that, outside of visual cortex, ANG was among the regions that carried the most information about perceived faces. There are, however, several examples of mnemonic content representations in other subregions of lateral parietal cortex. For example, successful decoding (Christophel et al., 2012; Christophel and Haynes, 2014; Bettencourt and Xu, 2016) and reconstruction (Sprague et al., 2014; Ester et al., 2015) of stimuli held in working memory has been observed in more dorsal aspects of lateral parietal cortex, including the intraparietal sulcus and superior parietal lobule. The "discrepancy" in localization across these studies could reflect several factors. For example, we used face stimuli, whereas other working memory studies have used fractals (Christophel et al., 2012), grating orientations (Ester et al., 2015; Bettencourt and Xu, 2016), or spatial locations (Sprague et al., 2014). Faces may recruit different and/or more diverse sources of information (perceptual, semantic, affective, social, etc.) and it has been proposed that ANG plays a role in integrating different kinds of face information (Joassin et al., 2011). Indeed, it has been argued, more broadly, that ANG contributes to memory by integrating event features into bound representations (Shimamura, 2011; Wagner et al., 2015). Thus, it may be that ANG involvement in the present study was at least partially related to the diversity of information that faces combine.

It is also notable that, although the memory task we used was nominally a "working memory" task, we deliberately encouraged contributions from long-term memory by prefamiliarizing subjects to the to-be-remembered faces. This was intended to increase the fidelity of the face representations subjects held in memory (i.e., that long-term memory might "boost" the quality of working memory representations). However, given that ANG has most typically been associated with long-term memory retrieval (Vilberg and Rugg, 2008; Hutchinson et al., 2009), it is possible that this factor explains why we observed localization within ANG, whereas other examples of working memory decoding/reconstruction have implicated regions outside of ANG. Notably, the localization of working memory reconstructions to ANG contrasted with our observation of reliable perception-based reconstructions across multiple parietal subregions.

One of the important remaining questions concerning content representations in ANG is as follows: why does ANG represent mnemonic content? One possibility is that lateral parietal

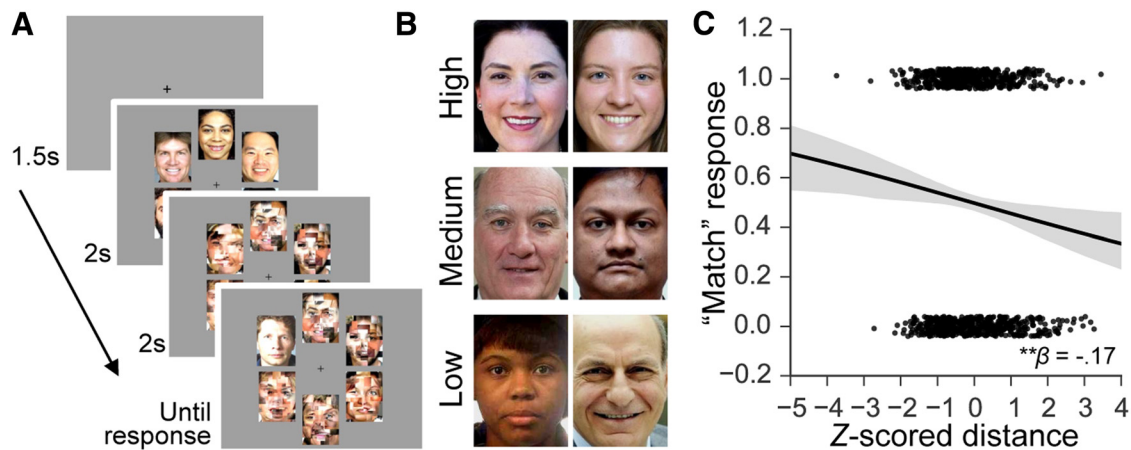


Figure 9. Behavioral change-detection experiment. **A**, Trial structure of the experiment. Subjects studied a set of six sample face images, followed by masks generated from scrambled face images. The subjects' task was to remember all six faces during the following brief delay. After the delay, a single probe face was presented at one of the six locations, and subjects indicated whether or not the probe face matched the sample face that had been studied at that location. **B**, Example pairs of sample and probe faces with high (top row), medium (middle row), and low (bottom row) eigenface score similarity. The z-scored Euclidean distances between the eigenface scores were -2.23 , 0.01 , and 2.72 , respectively. **C**, Logistic regression testing for the relationship between the probability of making a "match" response on nonmatch trials and eigenface score similarity between the sample and probe faces. The numbers of correct and incorrect trials were controlled to be identical within each subject. Each dot indicates a single trial in the nonmatch condition, and trials from all subjects are shown. A value of 1 on the y-axis represents a "match" response, and a value of 0 represents a "nonmatch" response. Smaller values on the x-axis represent higher similarity (lower Euclidean distance) between the faces. Solid line indicates the fitted logistic regression line. The shaded area represents 95% bootstrapping confidence intervals. $**p < 0.01$.

regions help align retrieved information with behavioral goals (Kuhl et al., 2013). Whereas prefrontal cortex may hold abstract representations of top-down goals (Miller and Cohen, 2001), and OTC may serve as an initial site where memories are reactivated (Danker and Anderson, 2010), lateral parietal cortex may function as an interface between prefrontal cortex and OTC, recoding or maintaining representations in-line with behavioral goals. From this perspective, it may be wondered why we observed such modest memory-based reconstruction effects in OTC. Notably, our working memory task included both target and distractor faces. Recently, it has been shown that working memory representations in lateral parietal cortex are significantly more robust to distractors than are representations in OTC (Bettencourt and Xu, 2016). Although speculative, it is possible that the inclusion of distractor faces in the present study depressed memory-based reconstructions in OTC to a greater extent than in ANG.

One important caveat when considering the functional contributions of ANG to memory is that damage to lateral parietal cortex does not produce profound impairments in objective memory accuracy but rather reduces memory confidence (Berryhill, 2012). However, a more recent study that tested paired associate learning shortly after onset of parietal damage found robust memory deficits, and deficits were most pronounced when damage included ANG (Ben-Zvi et al., 2015). Moreover, repeated stimulation of ANG has also been shown to modulate paired associate learning (Wang et al., 2014), although this may reflect downstream effects of stimulation given the functional and structural connectivity between angular gyrus and the hippocampus (Uddin et al., 2010). Thus, while several findings point to functional contributions of lateral parietal cortex to memory, additional investigation will be required to better understand the relative contributions of lateral parietal regions to subjective and objective aspects of successful remembering.

In conclusion, we used an innovative methodological approach that allows for perceived and remembered faces to be reconstructed from the neural activity patterns they evoke. We used this method to test for content representations within the angular gyrus and to gain insight into the nature of these representations. Our findings

uniquely implicate angular gyrus in representing high-level face information across perception and memory.

References

- Allison T, Puce A, McCarthy G (2000) Social perception from visual cues: role of the STS region. *Trends Cogn Sci* 4:267–278. [CrossRef Medline](#)
- Ben-Zvi S, Soroker N, Levy DA (2015) Parietal lesion effects on cued recall following pair associate learning. *Neuropsychologia* 73:176–194. [CrossRef Medline](#)
- Berryhill ME (2012) Insights from neuropsychology: pinpointing the role of the posterior parietal cortex in episodic and working memory. *Front Integr Neurosci* 6:31. [CrossRef Medline](#)
- Bettencourt KC, Xu Y (2016) Decoding the content of visual short-term memory under distraction in occipital and parietal areas. *Nat Neurosci* 19:150–157. [CrossRef Medline](#)
- Binder JR, Desai RH (2011) The neurobiology of semantic memory. *Trends Cogn Sci* 15:527–536. [CrossRef Medline](#)
- Bird CM, Keidel JL, Ing LP, Horner AJ, Burgess N (2015) Consolidation of complex events via reinstatement in posterior cingulate cortex. *J Neurosci* 35:14426–14434. [CrossRef Medline](#)
- Brainard DH (1997) The psychophysics toolbox. *Spat Vis* 10:433–436. [CrossRef Medline](#)
- Cabeza R, Ciaramelli E, Olson IR, Moscovitch M (2008) The parietal cortex and episodic memory: an attentional account. *Nat Rev Neurosci* 9:613–625. [CrossRef Medline](#)
- Christophel TB, Haynes JD (2014) Decoding complex flow-field patterns in visual working memory. *Neuroimage* 91:43–51. [CrossRef Medline](#)
- Christophel TB, Hebart MN, Haynes JD (2012) Decoding the contents of visual short-term memory from human visual and parietal cortex. *J Neurosci* 32:12983–12989. [CrossRef Medline](#)
- Cowen AS, Chun MM, Kuhl BA (2014) Neural portraits of perception: reconstructing face images from evoked brain activity. *Neuroimage* 94:12–22. [CrossRef Medline](#)
- Danker JF, Anderson JR (2010) The ghosts of brain states past: remembering reactivates the brain regions engaged during encoding. *Psychol Bull* 136:87–102. [CrossRef Medline](#)
- Daselaar SM, Prince SE, Dennis NA, Hayes SM, Kim H, Cabeza R (2009) Posterior midline and ventral parietal activity is associated with retrieval success and encoding failure. *Front Hum Neurosci* 3:13. [CrossRef Medline](#)
- Destrieux C, Fischl B, Dale A, Halgren E (2010) Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *Neuroimage* 53:1–15. [CrossRef Medline](#)

- Ester EF, Sprague TC, Serences JT (2015) Parietal and frontal cortex encode stimulus-specific mnemonic representations during visual working memory. *Neuron* 87:893–905. [CrossRef Medline](#)
- Harrison SA, Tong F (2009) Decoding reveals the contents of visual working memory in early visual areas. *Nature* 458:632–635. [CrossRef Medline](#)
- Huijbers W, Schultz AP, Vannini P, McLaren DG, Wigman SE, Ward AM, Hedden T, Sperling RA (2013) The encoding/retrieval flip: interactions between memory performance and memory stage and relationship to intrinsic cortical networks. *J Cogn Neurosci* 25:1163–1179. [CrossRef Medline](#)
- Hutchinson JB, Uncapher MR, Wagner AD (2009) Posterior parietal cortex and episodic retrieval: convergent and divergent effects of attention and memory. *Learn Mem* 16:343–356. [CrossRef Medline](#)
- Jenkinson M, Smith S (2001) A global optimisation method for robust affine registration of brain images. *Med Image Anal* 5:143–156. [CrossRef Medline](#)
- Jenkinson M, Bannister P, Brady M, Smith S (2002) Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage* 17:825–841. [CrossRef Medline](#)
- Joassin F, Pesenti M, Maurage P, Verreckett E, Bruyer R, Campanella S (2011) Cross-modal interactions between human faces and voices involved in person recognition. *Cortex* 47:367–376. [CrossRef Medline](#)
- Konkle T, Brady TF, Alvarez GA, Oliva A (2010) Scene memory is more detailed than you think: the role of categories in visual long-term memory. *Psychol Sci* 21:1551–1556. [CrossRef Medline](#)
- Kriegeskorte N, Mur M, Bandettini P (2008) Representational similarity analysis: connecting the branches of systems neuroscience. *Front Syst Neurosci* 2:4. [CrossRef Medline](#)
- Kuhl BA, Chun MM (2014) Successful remembering elicits event-specific activity patterns in lateral parietal cortex. *J Neurosci* 34:8051–8060. [CrossRef Medline](#)
- Kuhl BA, Johnson MK, Chun MM (2013) Dissociable neural mechanisms for goal-directed versus incidental memory reactivation. *J Neurosci* 33:16099–16109. [CrossRef Medline](#)
- Mason W, Suri S (2011) Conducting behavioral research on Amazon's Mechanical Turk. *Behav Res Methods* 44:1–23. [CrossRef Medline](#)
- McDonnell J, Martin J, Markant D, Coenen A, Rich A, Gureckis T (2014) psiTurk (Version 2.1.2)[Software]. New York: New York University. <https://github.com/NYUCCL/psiTurk>.
- Miller EK, Cohen JD (2001) An integrative theory of prefrontal cortex function. *Annu Rev Neurosci* 24:167–202. [CrossRef Medline](#)
- Naselaris T, Kay KN (2015) Resolving ambiguities of MVPA using explicit models of representation. *Trends Cogn Sci* 19:551–554. [CrossRef Medline](#)
- Norman KA, Polyn SM, Detre GJ, Haxby JV (2006) Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends Cogn Sci* 10:424–430. [CrossRef Medline](#)
- Phillips PJ, Wechsler H, Huang J, Rauss PJ (1998) The FERET database and evaluation procedure for face-recognition algorithms. *Image Vis Comput* 16:295–306. [CrossRef](#)
- Rugg MD, Vilberg KL (2013) Brain networks underlying episodic memory retrieval. *Curr Opin Neurobiol* 23:255–260. [CrossRef Medline](#)
- Shannon BJ, Buckner RL (2004) Functional-anatomic correlates of memory retrieval that suggest nontraditional processing roles for multiple distinct regions within posterior parietal cortex. *J Neurosci* 24:10084–10092. [CrossRef Medline](#)
- Shimamura AP (2011) Episodic retrieval and the cortical binding of relational activity. *Cogn Affect Behav Neurosci* 11:277–291. [CrossRef Medline](#)
- Smith SM (2002) Fast robust automated brain extraction. *Hum Brain Mapp* 17:143–155. [CrossRef Medline](#)
- Spaniol J, Davidson PS, Kim AS, Han H, Moscovitch M, Grady CL (2009) Event-related fMRI studies of episodic encoding and retrieval: meta-analyses using activation likelihood estimation. *Neuropsychologia* 47:1765–1779. [CrossRef Medline](#)
- Sprague TC, Ester EF, Serences JT (2014) Reconstructions of information in visual spatial working memory degrade with memory load. *Curr Biol* 24:2174–2180. [CrossRef Medline](#)
- Stelzer J, Chen Y, Turner R (2013) Statistical inference and multiple testing correction in classification-based multi-voxel pattern analysis (MVPA): random permutations and cluster size control. *Neuroimage* 65:69–82. [CrossRef Medline](#)
- St-Laurent M, Abdi H, Buchsbaum BR (2015) Distributed patterns of reactivation predict vividness of recollection. *J Cogn Neurosci* 27:2000–2018. [CrossRef Medline](#)
- Todorov A, Said CP, Engell AD, Oosterhof NN (2008) Understanding evaluation of faces on social dimensions. *Trends Cogn Sci* 12:455–460. [CrossRef Medline](#)
- Toth LJ, Assad JA (2002) Dynamic coding of behaviourally relevant stimuli in parietal cortex. *Nature* 415:165–168. [CrossRef Medline](#)
- Turk M, Pentland A (1991) Eigenfaces for recognition. *J Cogn Neurosci* 3:71–86. [CrossRef Medline](#)
- Uddin LQ, Supekar K, Amin H, Rykhlevskaia E, Nguyen DA, Greicius MD, Menon V (2010) Dissociable connectivity within human angular gyrus and intraparietal sulcus: evidence from functional and structural connectivity. *Cereb Cortex* 20:2636–2646. [CrossRef Medline](#)
- Vannini P, O'Brien J, O'Keefe K, Pihlajamäki M, Laviolette P, Sperling RA (2011) What goes down must come up: role of the posteromedial cortices in encoding and retrieval. *Cereb Cortex* 21:22–34. [CrossRef Medline](#)
- Vilberg KL, Rugg MD (2008) Memory retrieval and the parietal cortex: a review of evidence from a dual-process perspective. *Neuropsychologia* 46:1787–1799. [CrossRef Medline](#)
- Wagner AD, Shannon BJ, Kahn I, Buckner RL (2005) Parietal lobe contributions to episodic memory retrieval. *Trends Cogn Sci* 9:445–453. [CrossRef Medline](#)
- Wagner IC, van Buuren M, Kroes MC, Gutteling TP, van der Linden M, Morris RG, Fernández G (2015) Schematic memory components converge within angular gyrus during retrieval. *ELife* 4:e09668. [CrossRef Medline](#)
- Wang JX, Rogers LM, Gross EZ, Ryals AJ, Dokucu ME, Brandstatt KL, Hermler MS, Voss JL (2014) Targeted enhancement of cortical-hippocampal brain networks and associative memory. *Science* 345:1054–1057. [CrossRef Medline](#)
- Xue G, Dong Q, Chen C, Lu ZL, Mumford JA, Poldrack RA (2013) Complementary role of frontoparietal activity and cortical pattern similarity in successful episodic memory encoding. *Cereb Cortex* 23:1562–1571. [CrossRef Medline](#)
- Yeo BT, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, Hollinshead M, Roffman JL, Smoller JW, Zöllei L, Polimeni JR, Fischl B, Liu H, Buckner RL (2011) The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J Neurophysiol* 106:1125–1165. [CrossRef Medline](#)