Representations of individuals in ventral temporal cortex defined by faces and biographies

Sara C. Verosky [a,*,1], Alexander Todorov [b], Nicholas B. Turk-Browne[b]

[a] Department of Psychology, Harvard University, Cambridge, MA 02138, USA
[b] Department of Psychology, Princeton University, Princeton, NJ 08540, USA

ABSTRACT

The fusiform gyrus responds more strongly to faces than to other categories of objects. This response could reflect either categorical detection of faces or recognition of particular facial identities. Recent fMRI studies have attempted to address the question of what information is encoded in these regions, but have reported mixed results. We tested whether the creation of richer identity representations via training on visual and social information, and the use of an adaptation design, would reveal more robust representations of these identities in ventral temporal cortex. Examining the patterns of activation across voxels in bilateral fusiform gyri, we identified unique patterns for particular identities. Attaching distinctive biographical information to identities did not increase the strength of these representations, but did produce a grouping effect: faces associated with the same amount of biographical information were represented more similarly to each other. These results are consistent with the possibility that identity exemplars are represented in posterior visual areas best known for their role in representing categorical information, and suggest that these areas may be sensitive to some forms of non-visual information, including from the social domain.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

Multiple regions in ventral temporal cortex, including the lateral fusiform gyrus (FG), respond more strongly to faces than to other categories of objects (Kanwisher, McDermott, & Chun, 1997; McCarthy, Puce, Gore, & Allison, 1997; Tsao, Moeller, & Freiwald, 2008). However, since individual faces convey both category-level and exemplar-level information, these regions could be involved in the categorical detection of faces or the differentiation between specific face exemplars. Neuroimaging techniques such as adaptation and multi-voxel pattern analysis (MVPA) have been used to address the question of what these regions represent.

Univariate analyses of the FG have revealed adaptation for repeated versus novel facial identities (Rotshtein, Henson, Treves, Driver, & Dolan, 2005; Verosky & Turk-Browne, 2012). In such studies, the FG response measured by functional magnetic resonance imaging (fMRI) is attenuated when two faces depicting the same identity are shown sequentially, despite changes in physical properties such as viewpoint. Although multivariate analyses have revealed that distributed patterns of activation in ventral temporal cortex convey information about object categories (Haxby et al.,

2001), finding patterns associated with specific identities has proven more difficult. Two published studies have found evidence for identity-specific patterns in the right anterior temporal lobe (ATL; Kriegeskorte, Formisano, Sorger, & Goebel, 2007; Nestor, Plaut, & Behrmann, 2011), but only one of them additionally found evidence for such patterns in the FG (Nestor et al., 2011). In the present study, we replicate and extend this novel finding.

While previous studies have tended to define identity in terms of visual features, identity representations in everyday life are much more complex. For instance, mere acquaintances see each other multiple times, know each other's names, and often know detailed social information about each other (e.g., group membership). Becoming familiar with someone in one or more of these ways alters face processing in the brain: familiar faces engage a broad network of regions more strongly than unfamiliar faces (Gobbini & Haxby, 2006, 2007; Todorov, Gobbini, Evans, & Haxby, 2007; Cloutier, Kelley, & Heatherton, 2011). Many of these regions are involved in processing emotional and social information, and such higher-level knowledge may influence how identity is represented in perceptual brain regions.

Behaviorally, face recognition is superior for familiar versus unfamiliar faces across changes in visual properties (such as lighting, viewpoint, context, and emotional expression; Johnston & Edmonds, 2009), indicating that experience produces more invariant facial identity representations. Research in other domains further supports

---

* Corresponding author. Tel.: +1 617 384 5875.
E-mail address: sverosky@fas.harvard.edu (S.C. Verosky).
[1] Department of Psychology, Harvard University, Northwest Science Building 280.14, 52 Oxford Street, Cambridge, MA 02138, USA.

this notion. In the developmental literature, for example, teaching infants to individuate versus categorize monkey faces with verbal labels prevents a later decline in the ability to discriminate between those faces resulting from perceptual narrowing (Scott & Monesson, 2009). In the social psychological literature, warning subjects about the cross-race recognition deficit—that face recognition is worse for faces from other races—can reduce or eliminate the effect, suggesting that the deficit partly results from categorization of out-group members (Hugenberg, Miller, & Claypool, 2007). Together, these studies suggest that learning to individuate others through everyday interactions, verbal labels, or by increasing motivation, improves face processing.

Here we test whether richer experience with a set of identities produces neural representations that can be better identified in ventral temporal cortex. Prior to scanning, subjects completed a behavioral training task in which they saw multiple identities from different viewpoints and learned to associate each identity with a unique name. For half of the identities, subjects also learned fictional biographical information about that person. During a subsequent fMRI session, subjects were presented with sequential pairs of faces followed by pairs of names. For each pair of faces, the two faces depicted either the same identity or different identities. In univariate analyses, we measured whether there was an attenuated response to pairs of faces where the identity was the same versus different (fMRI adaptation or repetition suppression; Grill-Spector, Henson, & Martin, 2006; Turk-Browne, Scholl, & Chun, 2008) in bilateral FG regions of interest (ROIs). In multivariate analyses, we created an activation map for each identity and correlated these maps across runs within each ROI. We tested whether the activation maps for a given identity were more highly correlated with themselves than with those for other identities, which would provide evidence of distributed identity representations.

Additionally, we tested how learning biographical information influences identity representations in FG. For univariate analyses, we examined whether biographical information influences identity adaptation. For multivariate analyses, we considered two possibilities. First, biographical information might enhance the strength of representations for specific identities, such that the pattern correlation for the same identity across runs should be greater if it was versus was not associated with biographical information. Second, biographical information might affect identity representations more generally by adding a new dimension of information, in which case a given identity should

show greater pattern correlations with other identities that were associated with the same amount of biographical information.

2. Materials and methods

2.1. Subjects

Twenty-two subjects participated in the study for monetary compensation (5 males; mean age=22.2 years, range=18–37 years). Informed consent was obtained for a protocol approved by the Princeton University Institutional Review Board. All subjects were right-handed, had normal or corrected-to-normal vision, and reported no history of neurological or psychological disorders. The data from two subjects were excluded from analysis due to excessive head motion (greater than 3 mm in one or more runs), resulting in a final sample of 20 subjects.

2.2. Stimuli

Color photographs of 16 female facial identities from the Karolinska Directed Emotional Faces set (Lundqvist, Flykt, & Ohman, 1998) were used as stimuli. For each identity, there were three photographs: one facing forward and two with the head turned 30–40° to the left and right. The individuals displayed neutral facial expressions and eye gaze matched head direction.

2.3. Procedure

2.3.1. Training task

The day before the scan or earlier in the day of the scan, subjects completed a training task in which they learned a name for each face (Fig. 1). Faces were randomly assigned one of the twenty most popular female baby names for 1989 (www.ssa.gov/oact/babynames). This year was chosen to roughly match subjects' birth year, to ensure that they were familiar with the names. In addition, eight of the 16 faces were randomly paired with short biographical vignettes (mean words=186, SD=9). Four vignettes were primarily negative in tone and four were primarily positive (Appendix A in Supplementary material). We were not interested in valence effects due to limited power for examining such differences, but rather included valenced information to make the vignettes more memorable. Training was conducted by cycling between two phases, learning and test, until subjects reached criterion.

During the learning phase, subjects saw 16 identities, presented one at a time in a random order. For each identity, a name appeared in the center of the screen. When the subject pressed the space bar, the task forwarded to the next screen. For identities paired with biographical information, a paragraph of the associated vignette appeared in the center of the screen. Each vignette was broken into two paragraphs, shown on alternating blocks of the training task. Subjects were instructed to read the biographical information carefully and to try to form an impression of the person based on it (Hamilton, Katz, & Leirer, 1980). Subjects were not asked to recall the biographical information during the training or main tasks, but they were given a surprise memory
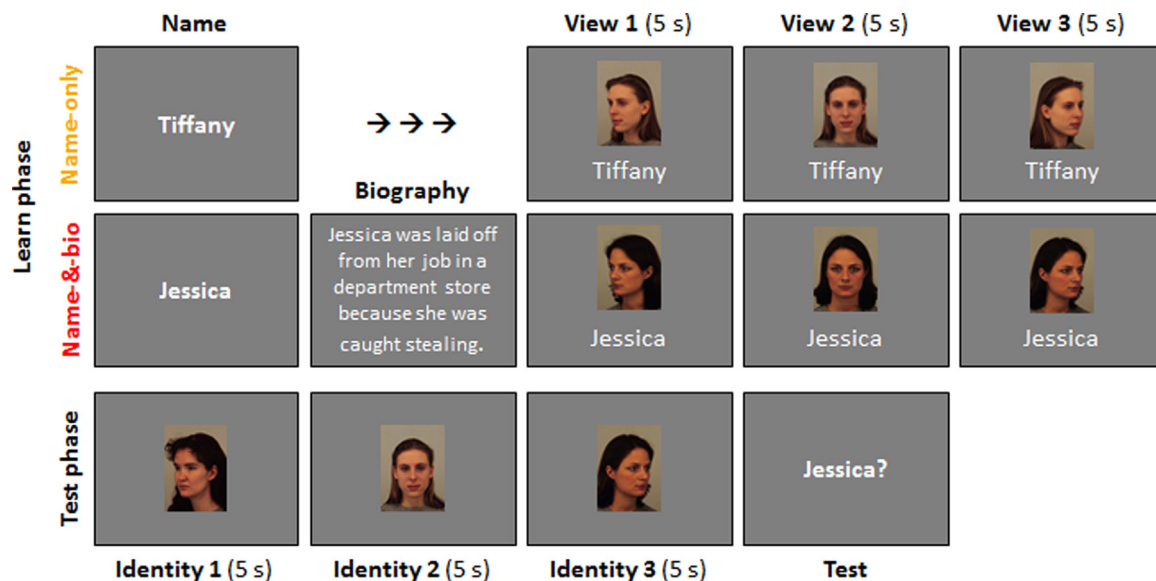
Fig. 1. Training task. On each trial of the learning phase, subjects were shown a name followed by three views of that facial identity. For faces associated with biographies, the biographical information appeared immediately after the name but before the faces. On each trial of the test phase, subjects were shown views of three different facial identities followed by a name and chose which face matched the name.

test at the end of the study as a manipulation check. When the participant pressed the space bar again, a sequence of three images appeared in the center of the screen. This screen immediately followed the name screen for identities without biographical information. The first image showed the person facing to the left, the next image facing forward, and the last facing right. Each image appeared for 5 s, with the name underneath. Because the biographical information was presented prior to the faces, subjects viewed the faces with and without biographical information for an equal amount of time.

During the test phase, subjects were tested on which name belonged to each face. On each trial, a sequence of three faces appeared in the center of the screen. Each face was shown for 5 s, and after the last face, a name appeared. Subjects indicated the serial position of the face that belonged with the name using the number keys 1–3. Subjects received feedback about whether their response on that trial was correct, as well as an overall accuracy score after the last trial. Each cycle of the test phase contained 48 trials, such that every view of each face was the correct answer for one trial.

To ensure that subjects saw both paragraphs of each vignette, they completed a minimum of two learning-test training cycles. After the second cycle, training stopped when subjects achieved 95+% accuracy in the preceding test phase. As a refresher, subjects completed an additional learning phase with all identities immediately before the scan. For identities paired with biographical information, the entire vignette was shown rather than just one paragraph.

2.3.2. Main task

During the scan session, subjects viewed stimuli on a projection screen at the back of the scanner bore via an angled mirror attached to the head coil. On each trial, subjects were presented with a sequence of two faces in the center of the screen (Fig. 2A). The faces subtended $5.0 \times 6.8°$, and appeared for 300 ms each separated by a 400 ms inter-stimulus interval. Two names appeared immediately afterwards, located above and below fixation. One name belonged to the second face and the other was a foil that did not belong to either face (vertical position randomized). Subjects were instructed that the identity of the two faces in each pair could either remain the same or change, but regardless, their task was to indicate which name belonged to the second face. They pressed a button with their right index finger to select the name above fixation and with their right middle finger for the name below fixation. The two names appeared for 2.5 s, regardless of when subjects responded. There was a randomly selected jittered inter-trial interval of 1.0 or 2.5 s, resulting in trial onset asynchronies of 4.5 or 6.0 s. To ensure familiarity with this task, subjects completed practice runs at the end of the behavioral training session until they reached a criterion of greater than 90% correct on a single run.

The fMRI response for each identity was assessed using *identity-repeat* trials, where the same identity was presented twice. As a baseline for adaptation, *identity-change* trials were also included, where two different identities were presented once. To increase the generalizability of findings across changes in stimuli, we included two types of runs: *view-repeat* runs, where the viewpoint of the faces in each pair was the same, and *view-change* runs, where there was a change in viewpoint from the first to the second face. Both run types had an identical 2 (social knowledge factor: name-and-biography, name-only) × 2 (identity repetition factor: identity-repeat, identity-change) design. For the view-repeat runs, there were three viewpoint conditions: the two faces in each trial faced to the left (left–left), right (right–right), or forward (center–center). For the view-change runs, the changes in view were limited to less than 45°, resulting in four conditions: left–center, right–center, center–left, and center–right. To match the frequencies of viewpoints in the view-repeat runs to the view-change runs, there were twice as many center–center trials as left–left and right–right trials. For both

view-change and view-repeat runs, there was an equal number of identity-repeat and identity-change trials, and the frequency of each viewpoint and identity was counterbalanced within each condition. Collapsing across viewpoint, each identity appeared in the identity-repeat condition 16 times. Although subjects completed eight functional runs, the frequency of identities was equated across pairs of runs and these pairs were later concatenated to produce four runs. Subjects alternated between pairs of view-repeat and view-change runs, with the starting run type counterbalanced across subjects. For instance, a given subject might have completed two view-repeat runs, two view-change runs, two view-repeat runs, and two view-change runs. Each run contained 64 trials and lasted 364.5 s.

2.3.3. Localizer task

After the main task, subjects completed two runs of a functional localizer task to identify face-selective subregions of FG. Each run contained color photographs of faces (a different set than used in the main task) and scenes in separate blocks. Subjects indicated whether faces were male or female and whether scenes were indoor or outdoor. Each block contained 12 images, presented for 500 ms each followed by a 1000-ms ISI, for a duration of 18 s (followed by 12 s of fixation). Images were presented in the center of the screen, and subtended $5 \times 5°$. Each run contained six face blocks and six scene blocks, shown in alternating order (starting category was counterbalanced across subjects), and lasted 369 s (including 9 s of discarded acquisitions).

2.3.4. Memory test

After completing the scan session, subjects were given a surprise test of memory for the biographical information presented with the faces. For each identity, subjects were shown the image of the person facing forward and were asked to indicate whether they had learned any biographical information about the person using the number keys 1 ("yes") and 2 ("no"). If they indicated that they had learned biographical information, they were asked to further indicate whether it was primarily negative or positive in tone using the number keys 3 ("negative") and 4 ("positive"). Each trial began with 1 s of fixation and each image remained onscreen until response. The 16 identities were shown in a random order.

2.4. fMRI acquisition

Data were collected on a Siemens 3 T Allegra head-only scanner using a Nova Medical head coil. Functional images were acquired using a T2*-weighted EPI sequence, with a 1500 ms TR, 28 ms TE, 64° flip angle, and a 64×64 matrix. Twenty-six interleaved oblique axial slices parallel to the anterior commissure/posterior commissure line covered the whole brain (3.5×3.5 mm inplane, 5 mm thick). A T1-weighted high-resolution anatomical image was acquired with an MPRAGE sequence for registration purposes. In all but three subjects, a second T1-weighted image coplanar with the functional slices was acquired to assist with registration.

2.5. fMRI analysis

2.5.1. Preprocessing

Analyses were conducted with FEAT in FSL (www.fmrib.ox.ac.uk/fsl). The first six volumes of each functional run were discarded to allow for T1 equilibration.
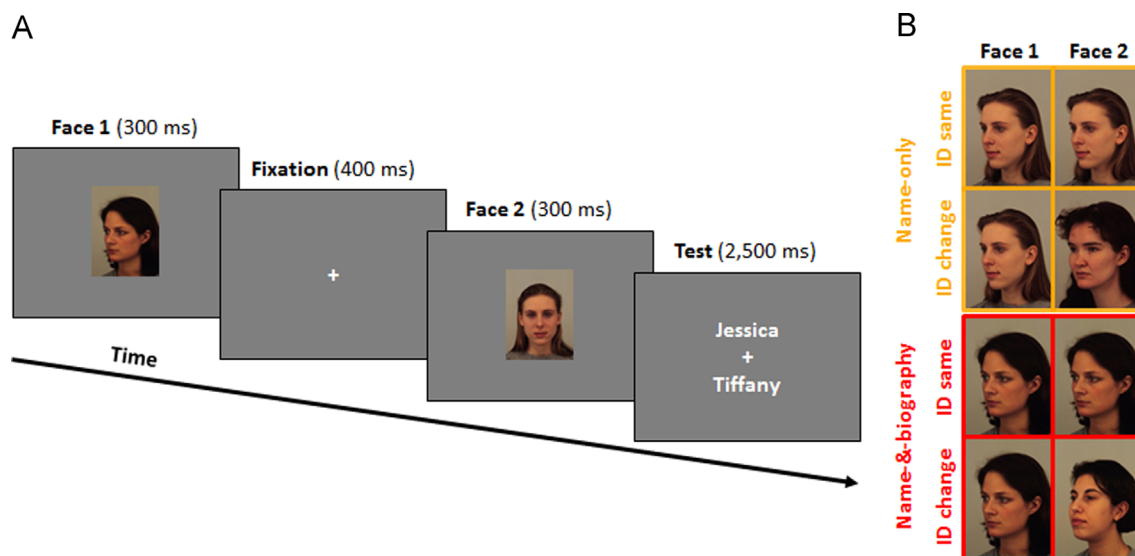
Fig. 2. Main task. (A) On each trial, subjects were shown a sequence of two faces followed by two names and chose which name matched the second face. (B) The first and second face in the main task depicted either the same person (identity-repeat) or different people (identity-change).

Data in each run were corrected for slice-acquisition time and head motion, spatially smoothed (5 mm FWHM), de-trended, and high-pass filtered (128 s period cutoff). There were eight total runs of the main task, four each of the view-repeat and view-change run types. Because the frequency of identities was equated across pairs of runs, the data from pairs of each run type were concatenated to produce four new runs (two view-repeat runs and two view-change runs). The concatenated functional runs were registered to the coplanar anatomical scan and then to the high-resolution anatomical scan. For subjects without a coplanar anatomical scan, functional runs were registered directly to the high-resolution anatomical scan. Data were normalized into Montreal Neurological Institute (MNI) space and interpolated to 2 mm isotropic voxels.

2.5.2. General linear modeling

Because of the rapid presentation of the two images in each trial and the sluggishness of the blood oxygen level-dependent (BOLD) response, we examined the combined response to pairs of images (e.g., Verosky & Turk-Browne, 2012). There were four identity-repeat trials for each identity in each concatenated run, which were used to estimate the voxelwise response to each identity in a general linear model (GLM). The model included a separate regressor for each of the 16 identities, including only identity-repeat trials. We also included a regressor for missed and incorrect responses, and two regressors for the identity-change condition—one for pairs of name-and-biography identities and the other for pairs of name-only identities (trials with different identities always contained two identities from the same social knowledge condition). Regressors were convolved with a double-gamma canonical hemodynamic response function (HRF). Temporal derivatives were included for each regressor to account for variability in response latency. Motion correction parameters were included as covariates of no interest.

2.5.3. Regions of interest

Localizer data were preprocessed using the same parameters as the main task. For each run, boxcar functions lasting 18 s were placed at that onset of face and scene blocks, convolved with an HRF, and entered into a GLM. The resulting parameter estimates were combined across the two localizer runs within-subject using a fixed-effects GLM. Parameter estimates for face and scene blocks were contrasted, and the peak face-selective voxel in an anatomically restricted region of the FG was identified (e.g., Verosky & Turk-Browne, 2012). For the sake of completeness, we also defined ROIs in ATL and performed the same analyses below (see Supplementary material).

Univariate and pattern analyses were conducted in ROIs in the bilateral FG. ROIs were created by centering 20-mm-spheres on the peak face-selective voxel in right
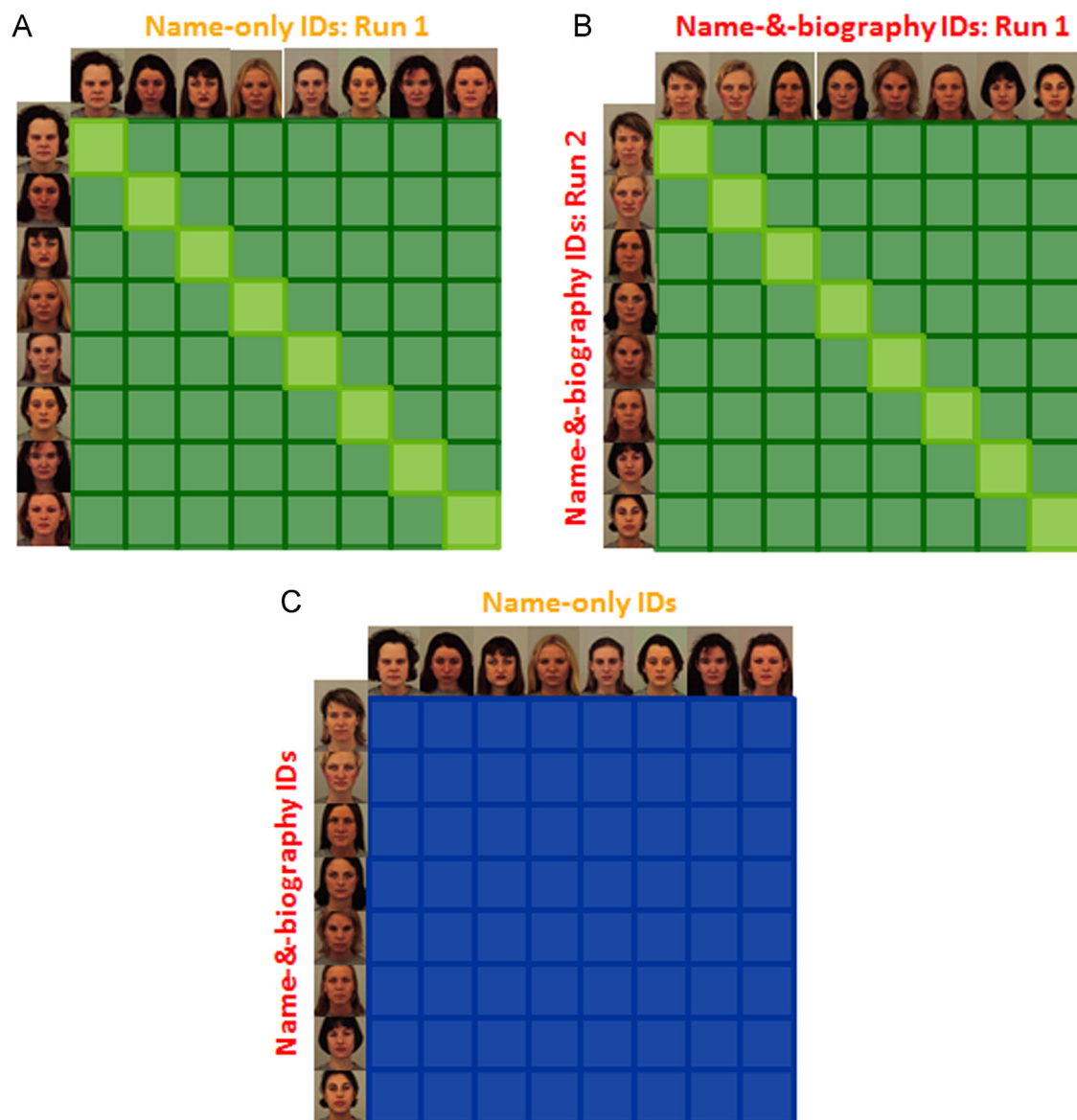
Fig. 3. Pattern analyses. An activation map was created for each identity, and these activation patterns were correlated across runs. The colors of the cells in the correlation matrices indicate assignment to condition. To test for identity-specific representations, pattern correlations were calculated for (A) name-only identities and (B) name-and-biography identities, and the resulting diagonal within-identity correlations (light green) were compared to off-diagonal within-group correlations (dark green). (C) To test for group-level representations, activation maps for name-only identities were correlated with name-and-biography identities across runs and the resulting between-group correlations (blue) were compared to within-group correlations (dark green). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

and left FG. Spheres of this size were chosen to include the fusiform face area, but also surrounding cortex. Previous studies suggest that the broader ventral temporal cortex carries information about object category (Haxby et al., 2001), as well as viewpoint-independent representations of object exemplars (Pourtois, Schwartz, Seghier, Lazeyras, & Vuilleumier, 2005; Pourtois, Schwartz, Spiridon, Martuzzi, & Vuilleumier, 2009). Using spheres of this size is also consistent with a recent finding that clusters of voxels with identity-specific patterns in FG extend beyond face-selective regions (Nestor et al., 2011).

2.5.4. Adaptation analyses

Adaptation was assessed by comparing BOLD responses for identity-change versus identity-repeat trials (e.g., Rotshtein et al., 2005; Verosky & Turk-Browne, 2012; Fig. 2B). Before extracting percent BOLD signal change, each ROI was masked with the contrast of faces > scenes from the independent localizer task at a threshold of $p = .05$, to avoid including voxels from white matter and outside the brain in the analysis. For each ROI, signal change was extracted for each condition and entered into a 2 (view-repeat, view-change) × 2 (within-identity, within-group) × 2 (name-and-biography, name-only) repeated-measures ANOVA.

2.5.5. Pattern analyses

The analyses above examined identity information within individual voxels. To test for identity information distributed in patterns of activation across voxels, we created a voxel map for each identity and correlated these maps across concatenated runs. With regard to the similarity metric, we chose to use a correlation analysis because it is a common and well-accepted approach (see Haxby et al., 2001) and because it has proven useful in previous studies from our group (Schapiro, Kustner, & Turk-Browne, 2012). Moreover, correlation values are easy to interpret, and this method does not require assumptions about normalization, preprocessing, or classification parameters, nor the extensive stimulus sets or numerous repetitions of the same stimulus that are required for other forms of RSA or MVPA. For each subject, voxel maps were created by contrasting activation for the identity-repeat trials of each identity with activation for the identity-change trials; this baseline controls for correlations driven by the mere presence of any face. The map for each identity was correlated with maps for all other identities in the same social knowledge group, producing an 8 × 8 correlation matrix for the name-and-biography identities and another one for the name-only identities (Fig. 3A and B). Correlations were computed separately for view-repeat and view-change runs, resulting in four total matrices.

For each matrix, correlations along the diagonal reflected the correlation of each identity with itself across runs (*within-identity*), while the correlations off the diagonal reflected the correlations for different identities in the same social knowledge group across runs (*within-group*). If specific identities have unique voxel patterns, then the within-identity correlations should be greater than within-group correlations. We tested this hypothesis with a permutation test. For each subject, we shuffled the columns of each correlation matrix in every possible way (8! = 40,320) and calculated the mean of the diagonal in each permutated matrix. We combined data across social knowledge groups by taking the mean of the name-and-biography and name-only diagonals for each permutation, and we combined across subjects by taking the group mean. For this comparison and subsequent comparisons, we computed the probability of the true observed group mean by dividing the number of permutations with means higher than the true mean by the total number of permutations. For the test of identity, effects with $p < .05$ indicate the presence of reliable patterns for specific identities.

Creating separate correlation matrices for name-and-biography and name-only identities allowed us to examine identity representations separately for each social knowledge group. To test whether learning biographical information produced more unique identity representations, we compared the strength of within-identity correlations between the two groups using a similar approach to above: For each subject, we shuffled the columns of each correlation matrix in every possible way. For each permutation, we subtracted the mean of the diagonal (within-identity) correlations for the name-only group from the mean of the diagonal correlations for the name-and-biography group. We then combined data across subjects by calculating the mean of this difference. We had a directional hypothesis that social knowledge would increase the strength of identity correlations, but we also wanted to consider the possibility that learning social knowledge might decrease the need for reliance on visual representations and thus reduce correlations: $p < .025$ means that identity correlations were stronger for name-and-biography than name-only identities, and $p > .975$ means the reverse.

In addition to examining patterns for specific identities, we also examined whether learning biographical information changed the nature of identity representations in a more general manner. Biographical details may serve as a feature in the neural representation of identity, and the presence of information along this dimension for name-and-biography but not name-only identities may have resulted in higher correlations within versus between social knowledge groups (Fig. 3C). Such a grouping effect would suggest that biographical information is represented in a categorical manner in FG (at least as measured by fMRI).

To examine group-level correlations, we contrasted the mean correlation of different identities from the same group (within-group) and the mean correlation of identities from different groups (*between-group*). For each subject and run type, a new 16 × 16 correlation matrix was generated for all identities. The matrix columns were shuffled 40,320 times (matching the identity analyses), given the intractability of the 16! permutations. For each permutation, we calculated the mean within-group correlation (name-and-biography with other name-and-biography, name-only with other name-only) and between-group correlation (name-and-biography with name-only). We then subtracted the between-group correlation from the within-group correlation and averaged this difference across subjects. Effects with $p < .05$ indicated that the patterns conveyed whether an identity had been linked to social knowledge.

For all the permutation tests, we examined whether correlations differed by run type by subtracting the mean correlation of one run type from the mean correlation of the other run type for each permutation. Similarly, we combined across run types by taking the mean of the view-repeat and view-change correlations. For the test of run type, we did not have a directional prediction: $p < .025$ means that correlations were stronger for view-repeat than view-change runs, and $p > .975$ means the reverse.

For inclusion in the pattern analyses, we selected voxels in each ROI that showed differential activity across identities in half of the data (the 'training' set). We ran a one-way ANOVA over the 16 identities in the first view-repeat and view-change run and selected voxels with $p < .05$ (one-tailed). This feature selection was not biased because it was done blind to the other half of the data (the 'test' set). That is, the correlation approach we used can be thought of as a simple pattern classifier, in which the similarity of each test pattern was computed relative to the prototype training patterns to which feature selection had been applied. Regardless, we repeated the analyses without feature selection (including all voxels in each ROI) and obtained similar results (Supplementary material). Because permutation tests are not conventional, we also repeated all analyses using standard parametric tests (Supplementary material).

To examine our ability to classify patterns, we compared pairs of correlation values and designated each comparison as correct if the means were in the expected direction. For the identity-level correlations, we compared the correlation of each identity with itself across runs to the correlation of that identity with the other seven identities in the same social knowledge group across runs. If the within-identity correlation was greater than the within-group correlation, the comparison was assigned a value of 1; otherwise it was assigned a value of 0. Since there were eight identities in each social knowledge group, this resulted in 56 comparisons for each social knowledge group for each run type. For the group-level correlations, we compared the correlation of each identity with the seven other identities in the same social knowledge group to the correlation of that identity with the eight identities in the other social knowledge. All correlations were calculated across concatenated runs. Since there were 7 within-group correlations and 16 between-group correlations for each identity, this resulted in 896 comparisons for each social knowledge group for each run type. After performing all the comparisons, we calculated the mean accuracy for each subject across social knowledge groups and run types, and compared the resulting values to chance (50%) using one-sample t-tests across subjects.

3. Results

3.1. Behavior

3.1.1. Training

In the first cycle of training, test accuracy was significantly above chance ($M = 69.58\%$, $SD = 13.91\%$, chance = 33.33%; $t(19) = 11.76$, $p < .001$). In the second cycle, test accuracy improved further ($M = 90.73\%$; $SD = 10.22\%$; vs. chance, $t(19) = 25.27$, $p < .001$). Nine subjects reached criterion during this cycle. In the third cycle, test accuracy was at ceiling ($M = 96.78\%$, $SD = 3.00\%$; vs. chance, $t(10) = 70.53$, $p < .001$), and nine more subjects reached criterion. In the fourth cycle, the remaining two subjects reached criterion ($M = 98.96\%$, $SD = 1.04\%$). Across all cycles, accuracy did not differ for name-and-biography ($M = 85.07\%$, $SD = 7.83\%$) and name-only identities ($M = 85.00\%$, $SD = 7.89$; $t < 1$), and neither did response time (RT; name-and-biography: $M = 1322$ ms, $SD = 431$ ms; name-only identities: $M = 1266$ ms, $SD = 363$; $t(19) = 1.19$, $p = .25$). Accuracy and RT did not differ in any individual test phase either (all $ps > .12$).

3.1.2. Main task

Accuracy on the main task was near ceiling ($M = 97.80\%$, $SD = 2.11\%$). There was a marginally significant main effect of identity, such that subjects were slightly more accurate for the identity-repeat ($M = 97.99\%$; $SD = 1.98\%$) than for the identity-change trials ($M = 97.62\%$, $SD = 2.31\%$; $F(1,19) = 3.61$, $p = .07$,

$\eta_p^2 = .16$). None of the other effects reached significance (biographical information x identity: $F(1,19)=3.43$, $p=.08$, $\eta_p^2=.15$; biographical information x run type: $F(1, 19)=1.03$, $p=.32$, $\eta_p^2=.05$; $F < 1$ in all other cases).

In RTs, the main effect of identity was not significant ($F < 1$) but there was a marginally significant interaction between identity and run type ($F(1, 19)=4.10$, $p=.06$, $\eta_p^2=.18$). For view-repeat runs, subjects responded similarly to identity-repeat ($M=733$ ms, $SD=96$ ms) and identity-change trials ($M=727$, $SD=101$; $t < 1$), but for view-change runs, subjects were marginally significantly faster for identity-repeat ($M=711$, $SD=95$) than identity-change trials ($M=727$, $SD=105$; $t(19)=1.92$, $p=.07$). No other effects reached significance (biographical information: $F(1,19)=1.87$, $p=.19$, $\eta_p^2=.09$; run type: $F(1,19)=1.46$, $p=.24$, $\eta_p^2=.07$; biographical information x identity: $F(1,19)=3.45$, $p=.08$, $\eta_p^2=.15$; biographical information x run type: $F < 1$; biographical information x run type x identity: $F(1,19)=1.77$, $p=.20$, $\eta_p^2=.09$). This analysis was based on correct trials only, but including RTs from incorrect trials did not change the pattern of results.

3.1.3. Memory test

Subjects accurately identified which faces were associated with biographical information ($M=97.81\%$, $SD=5.08\%$; vs. chance=50%: $t(19)=86.12$, $p < .001$), and whether that information was negative or positive ($M=94.37\%$, SD=8.58%; vs. chance=50%: $t(19)=49.20$, $p < .001$), despite never having been tested on this information during training.

3.2. fMRI

3.2.1. Adaptation analyses

To maximize power, we based our initial analysis on data from both the view-repeat and view-change runs, with run type included as a factor in the model. Within the right FG ROI, there was a significant effect of identity ($F(1,19)=23.48$, $p < .001$, $\eta_p^2=.55$), defined as greater signal change for identity-change ($M=.40$, $SD=.20$) than identity-repeat trials ($M=.35$, $SD=.17$). Neither the main effect of biographical information ($F(1,19)=1.43$, $p=.25$, $\eta_p^2=.07$) nor the interaction between biographical information and identity change reached significance ($F(1,19)=2.79$, $p=.11$, $\eta_p^2=.13$), indicating that biographical information did not influence overall signal change or the degree of identity adaptation. The main effect of run type and other interactions did not reach significance either ($F(1,19)=1.61$, $p=.22$, $\eta_p^2=.08$ for the interaction between biographical information and run type, $F < 1$ in all other cases). Within the left FG ROI, there was numerically greater signal change for the identity-change ($M=.33$, $SD=.25$) than identity-repeat trials ($M=.30$, $SD=.24$), however the main effect of identity did not reach significance ($F(1,19)=2.03$, $p=.17$, $\eta_p^2=.10$). No other effects in this region reached significance either ($F(1,19)=1.86$, $p=.19$, $\eta_p^2=.09$ for the main effect of run type, $F(1,19)=1.67$, $p=.21$, $\eta_p^2=.08$ for the interaction between biographical information and run type, $F < 1$ in all other cases).

Because the view-change runs provided a more stringent test of the presence of identity information than the view-repeat runs, we also examined identity adaptation for these runs alone (Fig. 4). Within the right FG ROI, the main effect of identity was again significant ($F(1,19)=11.12$, $p=.003$, $\eta_p^2=.37$), with greater signal change for the identity-change ($M=.39$. $SD=.17$) than identity-repeat trials ($M=.35$, $SD=.17$). Neither the main effect of biographical information ($F(1,19)=1.93$, $p=.18$, $\eta_p^2=.09$) nor the interaction between biographical information and identity change reached significance ($F(1,19)=1.82$, $p=.19$, $\eta_p^2=.09$). Within the left FG ROI, none of the effects reached significance ($F(1,19)=1.04$,
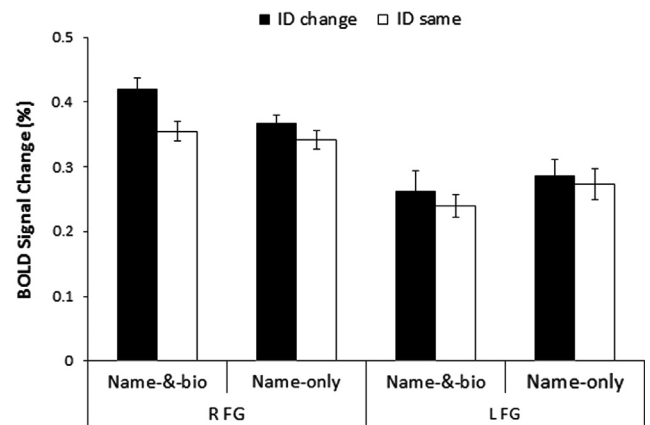

Fig. 4. Identity adaptation. Percent BOLD signal change for identity-change (black) and identity-repeat (white) trials for name-and-biography and name-only identities in the right and left FG ROIs for the view-change runs. Error bars depict ± 1 s.e.m.

$p=.32$, $\eta_p^2=.05$ for the main effect of biographical information, $F < 1$ for the other effects).

3.2.2. Pattern analyses

To create identity maps, average activation for the identity-repeat trials of a given identity was contrasted with a baseline of activation for identity-change trials. Removing this same baseline from all identities controlled for activation related to generic face processing, and as such, reduced overall correlation values. For instance, the mean correlation of maps belonging to different identities was higher before versus after removing the baseline ($r=.455$ vs. $r=.026$). In essence, the correlations remaining after removing the baseline reflect the *change* in correlation from the generic pattern correlation for any two faces.

As with the adaptation analyses, we based our initial analysis on data from both the view-repeat and view-change runs. To test for identity-specific patterns, we compared within-identity and within-group pattern correlations. We found that within-identity correlations were significantly greater than within-group correlations in both the right (permutation $p=.03$) and left ($p=.03$) FG ROIs, indicating the presence of identity representations. The amount of identity-specific information was not qualified by run type in either ROI (right: $p=.59$; left: $p=.71$).

To quantify the size of the identity effect, we designated each within-identity versus within-group comparison as correct or incorrect before collapsing across identities. Identity classification was significantly greater than chance in the right FG ROI ($M=52.81\%$, $SD=5.22\%$, $t(19)=2.41$, $p=.03$), but not the left FG ROI ($M=51.92\%$, $SD=6.88\%$, $t(19)=1.25$, $p=.23$), suggesting that patterns in the right FG ROI were more reliable. Classification accuracy did not differ by run type in either ROI (right: $t < 1$; left: $t(19)=1.23$, $p=.23$). Although classification accuracy was low, it is comparable to accuracy seen in other tasks that involve discriminating between perceptually similar stimuli. For instance, accuracy for discriminating between different emotional expressions in the STS was 52.7% in a recent paper (Said, Moore, Norman, Haxby, & Todorov, 2010).

Learning biographical information about an individual might produce a more distinctive identity representation, resulting in stronger pattern correlations for the specific identities that were associated with vignettes. To test this, we compared the magnitude of within-identity correlations for the name-and-biography versus name-only identities. However, these correlations did not differ in either the right (permutation $p=.86$) or left ($p=.34$) FG ROIs, and this was not qualified by run type (right: $p=.82$; left:

$p=.87$). Classification accuracy for the name-and-biography versus the name-only identities did not differ in the right or left FG ROIs either ($t<1$ in both cases), and this was not qualified by run type (right: $t(19)=1.15$, $p=.27$; left: $t<1$).

Although social knowledge did not increase the magnitude of identity correlations, it may have nevertheless altered the nature of the underlying representations in some way. Specifically, insofar as the availability of biographical information is represented categorically, identities should be represented more similarly to other identities in the same social knowledge group. In support of this possibility, within-group correlations were greater than between-group correlations in the right FG ROI (permutation $p=.004$) and marginally greater in the left FG ROI ($p=.05$). In the right FG ROI, this effect of group interacted with run type ($p=.98$; note: $p>.975$ is a significant effect in one direction): within-group correlations were significantly greater than between-group correlations for view-change ($p=.0009$) but not view-repeat runs ($p=.30$). In the left FG ROI, the effect of group did not interact with run type ($p=.32$). In line with these results, when within-group versus between-group comparisons were designated as correct or incorrect before collapsing across identities, group classification accuracy was significantly greater than chance in the right FG ROI ($M=51.93\%$, $SD=3.26\%$, $t(19)=2.65$, $p=.02$), but not the left FG ROI ($M=50.64\%$, $SD=3.77\%$, $t<1$). Classification accuracy did not interact with run type in either the right or left FG ROI however ($t<1$ in both cases).

Conducting the pattern analyses on the view-change runs alone yielded similar, although slightly weaker, results (Fig. 5). In particular, the within-identity correlations went from being significantly greater than the within-group correlations to marginally significantly greater in both the right (permutation $p=.06$) and left FG ROIs ($p=.05$). Identity classification accuracy in the right FG ROI remained significant ($M=53.04\%$, $SD=6.48\%$, $t(19)=2.09$, $p=.05$), while it again failed to reach significance in the left FG ROI ($M=53.22\%$, $SD=8.39\%$, $t(19)=1.71$, $p=.10$). Meanwhile, as in the analysis on both run types, the magnitude of within-identity correlations did not differ for the name-and-biography versus name-only identities in either ROI (right: $p=.52$; left: $p=.14$). Similarly, identity classification accuracy for the name-and-biography identities versus name-only identities did not differ either FG ROI ($t<1$ in both cases). Finally, as mentioned in the previous paragraph in regard to the interaction between group and run type, the within-group correlations were significantly

higher than the between-group correlations in the right FG ROI ($p=.0009$); however, the difference went from being marginally significant to not significant in the left FG ROI ($p=.20$). Group classification accuracy in the right FG ROI was marginally significant ($M=52.99\%$, $SD=7.58\%$, $t(19)=1.76$, $p=.09$), but was again non-significant in the left FG ROI ($M=55.45\%$, $SD=6.81\%$, $t<1$).

4. Discussion

We found evidence for individual-level information in the ventral temporal cortex, a region of the brain best known for its role in processing visual categories. In univariate analyses, there was fMRI adaptation when an identity was repeated versus changed in the right FG. In multivariate analyses, activation maps in bilateral FG for a given identity were more highly correlated with maps of the same identity versus different identities across runs. These two complementary techniques reveal the presence of facial identity information in the ventral temporal cortex at different spatial scales: fMRI adaptation indicates that identity-specific information exists within single voxels, while MVPA indicates that this information is also conveyed by distributed spatial patterns of activity across multiple voxels.

4.1. Identity representations

We succeeded in finding identity-level patterns in ventral temporal cortex, despite the fact that this has proven difficult in the past (Kriegeskorte et al., 2007). One difference between our study and previous studies is that we created rich identity representations through training. There are two unique and potentially important aspects of our training: First, subjects learned invariant identity representations by associating multiple viewpoints of each face with a name. Interestingly, the other study to find facial identity representations in posterior temporal cortex (Nestor et al., 2011) also employed multiple face exemplars for each identity—in their case, with different emotional expressions. Viewing faces over changes in visual properties may help create more invariant and robust representations (Li & DiCarlo, 2008).

The second unique aspect of our training was the use of many different identities. Previous studies have used two identities (Kriegeskorte et al., 2007) or four identities (Nestor et al., 2011), while we used 16 identities. Increasing the number of identities may have increased the difficulty of discriminating between them. Because our training task did not end until subjects had learned to accurately discriminate identities, subjects may have acquired more differentiated representations via perceptual learning (Goldstone, 1998). The fact that we found identity-level information in ventral temporal cortex suggests that our training may have helped. Nevertheless, future work will be needed to directly compare identity representations for trained versus novel identities.

Beyond training, the manner in which we presented faces may have increased the sensitivity of our study. Previous studies presented one face per trial (Kriegeskorte et al., 2007; Nestor et al., 2011). In contrast, we presented pairs of faces in an fMRI adaptation design. Since pattern analyses were based on trials where identities repeated, it is possible that identity adaptation on these trials produced more distinct patterns. According to the sharpening model (Wiggs & Martin, 1998; Grill-Spector et al., 2006), adaptation occurs via local competition between selective and non-selective neurons, resulting in sparser and more selective neural activity for repeated objects. Trials containing both a trial-novel identity and its repetition may have evoked a more identity-specific response than a trial-novel identity would have evoked on its own. Other models of adaptation, such as fatigue (e.g., McMahon & Olson, 2007), make the opposite
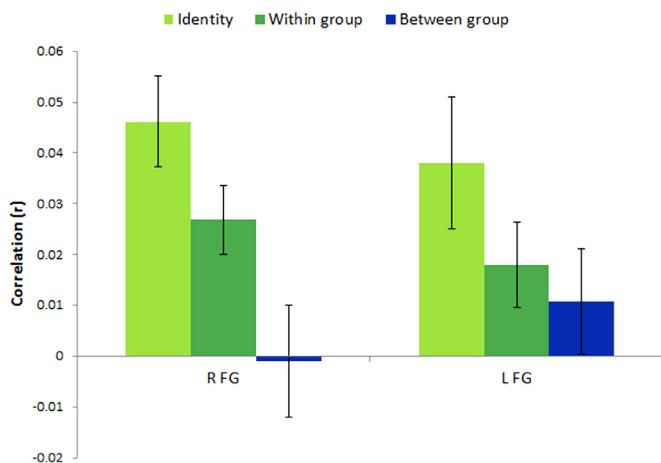
Fig. 5. Pattern correlations. Mean Pearson correlation coefficients for one identity with itself across runs (light green), for two different identities that were either both name-and-biography or name-only identities (dark green), and for one name-and-biography identity with one name-only identity (blue) for the view-change runs. Error bars depict ± 1 s.e.m. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

prediction, in which case our adaptation design may have been detrimental to finding identity representations.

4.2. Anterior temporal lobe

We found identity-level information in ventral temporal cortex, but this does not mean that the information originated there. Previous work has found identity-specific patterns in the right ATL (Kriegeskorte et al., 2007; Nestor et al., 2011), and has suggested that identity information may propagate from there back to ventral temporal cortex. Although ventral temporal cortex was our central focus, we considered the role of ATL by applying the same univariate and multivariate analyses to this region (see Supplementary material). For the univariate analyses, we found identity *enhancement* – the opposite of adaptation (Turk-Browne et al., 2008) – in bilateral ATL. Adaptation and enhancement have both been observed in the same regions of the brain with fMRI, and enhancement has been theorized to be tied to visual prediction (de Gardelle, Waszczuk, Egner, & Summerfield, 2013). Although it is not clear why the FG shows one pattern of activity while the ATL shows another, the univariate results do indicate that the ATL is sensitive to identity. For the multivariate analyses, we found identity-specific patterns in the right ATL but not the left ATL. The finding of unique patterns for specific identities in the right ATL replicates previous findings (Kriegeskorte et al., 2007; Nestor et al., 2011). This finding is also consistent with evidence that the right hemisphere is critical for face processing (Gazzaniga & Smylie, 1983; De Renzi, Perani, Carlesimo, Silveri, & Fazio, 1994; Le Grand, Mondloch, Maurer, & Brent, 2003), and may even be the source of face representations in the left hemisphere (Verosky & Turk-Browne, 2012).

4.3. Face selectivity

We investigated whether identity- and group-level information was carried in face-selective voxels in the vicinity of the Fusiform Face Area (FFA) by varying the size of the spherical ROIs and choosing voxels based on their degree of face selectivity in the localizer task (see Supplementary material). For the univariate analyses, we observed consistent identity adaptation in the right FG, but more variable adaptation in the left FG. Importantly, the identity adaptation in both regions persisted when we examined runs where there was a change in viewpoint between the first and second face in a trial, indicating that the observed adaptation was not due to low-level visual properties of the images being repeated. Thus, while there is viewpoint-independent identity information present bilaterally, the information in the right FG appears to be easier to detect. To put this finding in context, previous studies have reported mixed results: from failure to find viewpoint-independent identity information (e.g., Andrews and Ewbank. (2004)), to finding it only outside of face-selective regions (Pourtois et al., 2005, 2009) or within face-selective regions (Ewbank & Andrews, 2008; Verosky & Turk-Browne, 2012). While it is not clear what enabled us to find viewpoint-independent identity adaptation where others have failed, two possible explanations are the training task where subjects were familiarized with the identities, and the main task which required a focus on identity.

The multivariate analyses reported in the main text based feature selection on univariate sensitivity to identity. In the Supplementary material, we report multivariate analyses that instead based feature selection on face selectivity from the localizer task. In these analyses, there was identity-level information in the left FG, but not the right FG. There was also evidence for group-level information in bilateral FG, although this effect was more consistent in the right FG. While the lack of a multivariate identity effect in face-selective voxels from the right FG may come as a surprise, it is worth keeping two things in

mind. First, another study that found identity-specific patterns in ventral temporal cortex reported that the clusters showing this effect partially overlapped with the FFA, but were not face selective overall (Nestor et al., 2011). Second, the fact that there was an identity effect in the face-selective voxels in the left but not the right FG does not mean that the identity effect in the left FG is, in general, more robust. Instead, it may be more useful to think about these results in terms of face selectivity: it is possible that face-selective voxels in the right versus left hemisphere, or voxels with differing degrees of face-selectivity, carry different types of information.

In summary, univariate and multivariate analyses both indicate the presence of identity information within and outside of face-selective subregions of the FG. Future work will be needed to examine the types of identity information carried in these subregions.

4.4. Person knowledge

We had initially entertained the possibility that biographical information might increase the magnitude of identity adaptation and increase the strength of pattern correlations for specific identities, but this was not found to be the case. However, it did group identity representations around a common pattern: identity representations correlated more strongly with other identity representations that were associated with the same amount of biographical information. Consistent with the finding that even passive viewing of familiar faces engages regions throughout the brain more strongly than viewing of unfamiliar faces (Gobbini & Haxby, 2006, 2007; Todorov et al., 2007; Cloutier et al., 2011), we observed this common pattern despite the fact that the name-matching task used during scanning did not require accessing biographical information.

While the presence of associated biographical information may account for the grouping effect, there are also other potential explanations. For instance, this effect could be due to changes in attention or to spontaneous categorization of faces into groups. The biographies included valenced information, and it is possible that simply viewing the faces associated with these biographies could have increased arousal. In addition, the fact that faces in each group were always paired with other faces from the same social knowledge condition may have served as a cue to group the faces together. Attention to faces (Wojciulik, Kanwisher, & Driver, 1998) and experimentally created group membership (Van Bavel, Packer, & Cunningham, 2011) have both been found to modulate activity in the FG, indicating that this region is sensitive to top-down influence. However, while we do not know whether the group effect is due to biographical information per se, it is worth noting that we did not observe an effect of biographical information in the univariate analyses, which we might have if the effect was related to attention or arousal. Thus, our data suggest that the FG is sensitive to the social context associated with faces, even when that context is not relevant to the task at hand. Future work could investigate how the content of the biographical information, such as learning that faces belong to a particular social group (e.g., political party, profession, etc.), alters face representations.

It has often been assumed that representations in ventral temporal cortex are unimodally visual, but a growing body of work suggests that this part of the brain may be sensitive to information from other sensory modalities (von Kriegstein, Kleinschmidt, Sterzer, & Giraud, 2005; Mahon, Anzellotti, Schwarzbach, Zampini, & Caramazza, 2009; Van den Hurk, Gentile, & Jansma, 2011). For example, animal sounds presented auditorily elicit category-specific responses in ventral temporal cortex of congenitally blind adults, suggesting that visual experience is not necessary for this categorical organization (Mahon et al., 2009). In addition, the FFA responds more strongly to familiar versus unfamiliar voices during a speaker recognition task, further demonstrating that ventral temporal cortex can be co-opted for

processing sound (Von Kriegstein et al., 2005). Finally, the FG also responds differently to words describing categories of objects (e.g., pets, sports, etc.) and people, indicating a potential sensitivity to semantic content (Van den Hurk et al., 2011). Identities in our study were randomly assigned to be paired with biographical information or not, such that faces in each social knowledge group were not intrinsically more similar to each other. The finding that these faces were nevertheless represented more similarly in ventral temporal cortex suggests that this region is sensitive to information beyond mere appearance, including perhaps social information.

4.5. Conclusions

In sum, we obtained converging evidence for identity representations in bilateral FG from adaptation and pattern analyses. These findings highlight several potential avenues for future research, including investigating factors that might contribute to finding identity representations. Such factors include: familiarization with identities during training, exposure to identities across multiple viewpoints, associating names with identities, using a reasonably large number and range of identities, and employing cover tasks that focus attention on identity-level information. Subjects were also trained to associate biographical information with half of the identities. Although learning this information neither increased the magnitude of identity adaptation nor increased the strength of identity-specific pattern similarity, it did cluster identities with similar amounts of biographical information, suggesting that FG may also be sensitive to social context.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.neuropsychologia.2013.07.006.

References

Andrews, T. J., & Ewbank., M. P. (2004). Distinct representations for facial identity and changeable aspects of faces in the human temporal lobe. *NeuroImage*, *23*, 905–913.

Cloutier, J., Kelley, W. M., & Heatherton, T. F. (2011). The influence of perceptual and knowledge-based familiarity on the neural substrates of face perception. *Social Neuroscience*, *6*, 63–75.

De Renzi, E., Perani, D., Carlesimo, G. A., Silveri, M. C., & Fazio, F. (1994). Prosopagnosia can be associated with damage confined to the right hemisphere—an MRI and PET study and a review of the literature. *Neuropsychologia*, *32*, 893–902.

de Gardelle, V., Waszczuk, M., Egner, T., & Summerfield, C. (2013). Concurrent repetition enhancement and suppression responses in extrastriate visual cortex. *Cerebral Cortex*, 23, 2235–2244.

Ewbank, M. P., & Andrews, T. J. (2008). Differential sensitivity for viewpoint between familiar and unfamiliar faces in human visual cortex. *NeuroImage*, *40*, 1857–1870.

Gazzaniga, M. S., & Smylie, C. S. (1983). Facial recognition and brain asymmetries: clues to underlying mechanisms. *Annals of Neurology*, *13*, 536–540.

Gobbini, M. I., & Haxby, J. V. (2006). Neural response to the visual familiarity of faces. *Brain Research Bulletin*, *71*, 76–82.

Gobbini, M. I., & Haxby, J. V. (2007). Neural systems for recognition of familiar faces. *Neuropsychologia*, *45*, 32–41.

Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology*, *49*, 585–612.

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends in Cognitive Science*, *10*, 14–23.

Hamilton, D. L., Katz, L. B., & Leirer, V. O. (1980). Cognitive representation of personality impressions: organization processes in first impression formation. *Journal of Personality and Social Psychology*, *39*, 1050–1063.

Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, *28*, 2425–2430.

Hugenberg, K., Miller, J., & Claypool, H. M. (2007). Categorization and individuation in the cross-race recognition deficit: toward a solution to an insidious problem. *Journal of Experimental Social Psychology*, *43*, 334–340.

Johnston, R. A., & Edmonds, A. J. (2009). Familiar and unfamiliar face recognition: a review. *Memory*, *17*, 577–596.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, *17*, 4302–4311.

Kriegeskorte, N., Formisano, E., Sorger, B., & Goebel, R. (2007). Individual faces elicit distinct response patterns in human anterior temporal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *104*, 20600–20605.

von Kriegstein, K., Kleinschmidt, A., Sterzer, P., & Giraud, A. (2005). Interaction of face and voice areas during speaker recognition. *Journal of Cognitive Neuroscience*, *17*, 367–376.

Le Grand, R., Mondloch, C. J., Maurer, D., & Brent, H. P. (2003). Expert face processing requires visual input to the right hemisphere during infancy. *Nature Neuroscience*, *6*, 1108–1112.

Li, N., & DiCarlo, J. J. (2008). Unsupervised natural experience rapidly alters invariant object representations in visual cortex. *Science*, *321*, 1502–1507.

Lundqvist, D., Flykt, A., & Ohman, A. (1998). *The Karolinska Directed Emotional Faces (KDEF)*. Stockholm: Department of Neurosciences Karolinska Hospital.

McCarthy, G., Puce, A., Gore, J. C., & Allison, T. (1997). Face-specific processing in the human fusiform gyrus. *Journal of Cognitive Neuroscience*, *9*, 605–610.

McMahon, D. B., & Olson, C. R. (2007). Repetition suppression in monkey inferotemporal cortex: relation to behavioral priming. *Journal of Neurophysiology*, *97*, 3532–3543.

Mahon, B. Z., Anzellotti, S., Schwarzbach, J., Zampini, M., & Caramazza, A. (2009). Category-specific organization in the human brain does not require visual experience. *Neuron*, *63*, 397–405.

Nestor, A., Plaut, D. C., & Behrmann, M. (2011). Unraveling the distributed neural code of facial identity through spatiotemporal pattern analysis. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 9998–10003.

Pourtois, G., Schwartz, S., Seghier, M. L., Lazeyras, F., & Vuilleumier, P. (2005). Portraits or people? Distinct representations of face identity in the human visual cortex. *Journal of Cognitive Neuroscience*, *17*, 1043–1057.

Pourtois, G., Schwartz, S., Spiridon, M., Martuzzi, R., & Vuilleumier, P. (2009). Object representations for multiple visual categories overlap in lateral occipital and medial fusiform cortex. *Cerebral Cortex*, *19*, 1806–1819.

Rotshtein, P., Henson, R. N. A., Treves, A., Driver, J., & Dolan, R. J. (2005). Morphing Marilyn into Maggie dissociates physical and identity face representations in the brain. *Nature Neuroscience*, *8*, 107–113.

Said, C. P., Moore, C. D., Norman, K. A., Haxby, J. V., & Todorov, A. (2010). Graded representations of emotional expressions in the left superior temporal sulcus. *Frontiers in Systems Neuroscience*, *4*, 6.

Schapiro, A. C., Kustner, L. V., & Turk-Browne, N. B. (2012). Shaping of object representations in the human medial temporal lobe based on temporal regularities. *Current Biology*, *22*, 1622–1627.

Scott, L. S., & Monesson, A. (2009). The origin of biases in face perception. *Psychological Science*, *20*, 676–680.

Todorov, A., Gobbini, M. I., Evans, K. K., & Haxby, J. V. (2007). Spontaneous retrieval of affective person knowledge in face perception. *Neuropsychologia*, *45*, 163–173.

Tsao, D. Y., Moeller, S., & Freiwald, W. A. (2008). Comparing face patch systems in macaques and humans. *Proceedings of the National Academy of Sciences of the United States of America*, *105*, 19514–19519.

Turk-Browne, N. B., Scholl, B. J., & Chun, M. M. (2008). Babies and brains: Habituation in infant cognition and functional neuroimaging. *Frontiers in Human Neuroscience*, *2*, 16.

Van Bavel, J. J., Packer, D. J., & Cunningham, W. A. (2011). Modulation of the fusiform face following minimal exposure to motivationally relevant faces: evidence of in-group enhancement (not out-group disregard). *Journal of Cognitive Neuroscience*, *23*, 3343–3354.

Van den Hurk, J., Gentile, F., & Jansma, B. M. (2011). *What's behind a face: person context coding in fusiform face area as revealed by multivoxel pattern analysis*. *Cerebral Cortex*, 212893–2899.

Verosky, S. C., & Turk-Browne, N. B. (2012). Representations of facial identity in the left hemisphere require right hemisphere processing. *Journal of Cognitive Neuroscience*, *24*, 1006–1017.

Wiggs, C. L., & Martin, A. (1998). Properties and mechanisms of perceptual priming. *Current Opinion in Neurobiology*, *8*, 227–233.

Wojciulik, E., Kanwisher, N., & Driver, J. (1998). Modulation of activity in the fusiform face area by covert attention: an fMRI study. *Journal of Neurophysiology*, *79*, 1574–1579.