Supporting Information

Nestor et al. 10.1073/pnas.1102433108

SI Text

Stimulus Preparation. *Faces.* Stimulus choice and construction were guided by two opposing goals. On the one hand, stimuli had to be as similar as possible with respect to a number of characteristics: high-level attributes (e.g., sex or age), low-level image descriptors (e.g., average luminance or contrast), and external feature properties (e.g., hair color or volume) to eliminate confounds with facial identity. On the other hand, individual faces needed to be as different from each other as possible to maximize the discriminability of the visually based activation patterns they elicit. To accommodate these different demands, we proceeded as follows.

First, we started with all front-view faces from the Face-Place 3.0 face database (www.face-place.org) and we narrowed down this dataset to young Caucasian adult male faces displaying a minimum of three basic emotional expressions (1) in addition to neutral expressions. This procedure ensures substantial within-identity image variability while preserving natural poses that are easy to interpret. In addition, we eliminated all faces that displayed facial hair, glasses, or other adornments, leaving us with a set of 128 faces (32 identities × 4 expressions).

Second, faces were normalized to the same size, subsampled to a lower resolution, and masked. More precisely, an oval mask was applied to all images to remove background and hair and also to reduce the dimensionality of the space (Fig. S6).

Third, we converted images to CIEL*a*b*, the color space that comes closest to that of human vision (2). Each image was normalized next with the same mean and contrast value separately for each of the three color channels: L* (corresponding to luminance), a* (corresponding to red:green), and b* (corresponding to yellow:blue).

Fourth, we computed pairwise similarity measures across all faces with a neutral expression. More specifically, we applied principal component analysis (PCA) to all faces and their mirror symmetric versions (3). We selected the projections on the first 40 principal components for each image and computed Mahalanobis distances between these lower-dimension patterns for each pair of neutral faces. A Mahalanobis metric was deployed given that it outperforms other types of metric with regard to both automatic face recognition (4) and modeling human face recognition (5). On the basis of these measurements, from all possible sets of four neutral faces we selected the set that minimized the average similarity score. We also ensured that each pair of faces within this set scored a similarity value below the average within the larger initial set.

Finally, we restored the original homogeneous background and applied the same hair feature to all four faces and their nonneutral versions (happy, sad, and disgusted). The resulting 16 images (Fig. 1) served as experimental stimuli for our individuation task.

A different set of faces was used for the functional localizers. ***Orthographic forms (OFs).*** Four five-letter pseudowords (Fig. S1) were presented in four different types of font (Arial Black, Comic Sans MS, Courier, and Lucida Handwriting). The pseudowords had the same syllable structure but were orthographically dissimilar in that they had no common letter in the same position. Moreover, they were composed of different sets of letters (with the exception of plang and greld that shared the letter "l" in different positions).

Subjects. Eight Caucasian young adults (five females, age range 18–22) from the Carnegie Mellon University community participated in the experiment. All subjects were right-handed and had normal or corrected-to-normal vision. None of the subjects had any history of neurological disorders. Two other subjects participated in the experiment; however, their data were excluded from analysis due to large head movements (more than a voxel) during at least one of three scanning sessions.

Informed consent was obtained from all subjects. The Institutional Review Board of Carnegie Mellon University approved all imaging and behavioral procedures.

Behavioral Procedures. Before scanning, subjects were presented with the 16-face stimuli described above and were trained to associate each facial identity with one of four buttons. None of the subjects were previously familiar with any of the faces presented nor were they given any biographical information with regard to them. Similarly, subjects were presented with the 16 OF stimuli and were trained to associate each individual OF with a button (face and OF responses were made using different hands randomly assigned to each category). Subjects practiced the task until accuracy reached ceiling (>98%). Training took place at least 1 d before each subject's first scanning session and was also briefly repeated before each scanning session.

During localizer scans, subjects performed a one-back task (same/different image). During the remaining functional scans, they performed the individuation task described above.

Stimuli were presented in the center of the screen against a black background and subtended a visual angle of 3.2° × 4.1°. Stimulus presentation and response recording relied on Matlab (Mathworks) and Psychtoolbox 3.0.8 (6, 7).

Experimental Design. Eight participants were each scanned for a total of 21 functional runs spread across three 1-h sessions. Of these, 17 runs used a slow event-related design whereas the rest used a block protocol suitable for functional localizers.

Localizer scans contained blocks of images grouped by category: faces, common objects, houses, words, and pseudofont strings. Each block consisted of back-to-back presentations of 15 stimuli for a total of 14 s (930 ms per stimulus). Stimulus blocks were separated by 10 s of fixation and were preceded by a 10-s fixation interval at the beginning of each run. No single stimulus was repeated within the course of a run. Each localizer scan contained 10 stimulus blocks, 2 for each stimulus category, and had a total duration of 250 s.

Runs with an individuation task used a slow event-related design with the following structure: a bright fixation cross was presented in the middle of the screen for 100 ms and then a stimulus appeared for 400 ms and was replaced by a lower-contrast fixation cross until the end of the event for 9.5 s. Each run contained a set of 32 such events following 10 s of fixation (for a total of 330 s). All face and OF stimuli described above were presented exactly once during each run. Stimuli were displayed in pseudorandom order to maximize uncertainty about stimulus identity (8) under the constraint that no more than three stimuli from the same category (face or OF) could be presented in a row.

Our decision to include OF stimuli along with faces was motivated by several different factors. First, inclusion of a different category was expected to reduce possible habituation/adaptation effects caused by prolonged exposure to the same small set of faces. Second, faces and OFs are perceptually highly dissimilar. Thus, although pattern discrimination for faces at the individual level is bound to be challenging for any type of method, discrimination of faces and OFs at the category level should be relatively easy and could serve as a robust benchmark for our classification method. Third and most important, the information

map obtained for face individuation could arguably be a generic individuation map, that is, not face specific but process specific. If so, we would expect other categories of objects with which we have extensive individuation experience, such as OFs, to produce similar information maps. Analysis of OF discriminability within the context of the same experiment provides us with a first test of this hypothesis. Finally, we opted for using pseudowords instead of actual words because they are unfamiliar (like faces) and minimize semantic processing while engaging similar mechanisms for OF processing (9, 10).

Functional scans were equally divided across three different sessions (seven scans per session) conducted on separate days. A structural scan was also performed at the beginning (or the end) of each session.

Imaging Parameters. Subjects were scanned in a Siemens Allegra 3T scanner with a single-channel head coil. Functional images were acquired with an echo-planar imaging sequence (TR 2 s, time to echo 31 ms, flip angle 79°, 2.5-mm isotropic voxels, field of view 240 × 240 mm^2, 27 oblique slices covering the ventral stream). An MP-RAGE sequence (1-mm^3 voxels; 192 slices of size 256 × 256 mm^2) was used for anatomical imaging.

Preprocessing. Functional scans were slice scan time corrected, motion corrected, coregistered to the same anatomical image, and normalized to percentage of signal change using AFNI (11). Functional localizer data were smoothed with a Gaussian kernel of 7.5 mm FWHM. No spatial smoothing was performed on the rest of the data to allow multivariate analysis to exploit high-frequency information (12).

Standard Univariate Analysis. After completion of preprocessing steps we discarded the first 5 vol of each run to allow the hemodynamics to achieve a steady state and to minimize transient effects of magnetic saturation. Next, we fitted each type of block with a boxcar predictor and convolved it with a gamma hemodynamic response function (13). A general linear model (14) was applied to estimate the coefficient of each predictor independently for each voxel. Statistical maps were computed by t tests of pairwise comparisons between different block types. Face-selective areas were detected using a face–object contrast. Correction for multiple comparisons was implemented by controlling the false discovery rate under the assumption of positive/no correlation (15).

Spatiotemporal Information-Based Brain Mapping. A manually drawn cortical mask was constructed for each subject's brain. Fig. S2*A* shows the corresponding group mask. Searchlight analysis was carried out by walking a sphere voxel-by-voxel across the entire volume of the mask, extracting the spatial–temporal patterns recorded at each location, and testing them for the presence of relevant information via multivariate analysis. More specifically, a sphere with a 5-voxel radius was centered on each voxel within the cortical mask and intersected with the mask to restrict analysis to cortical voxels. Activation values across this restricted set of voxels at three different time points (4, 6, and 8 s after stimulus onset) were extracted for each stimulus presentation and concatenated into a single pattern.

Our choice of a 5-voxel spatial radius was based on pilot data not included in the current analysis. In addition, to test the sensitivity of our results as a function of this parameter, we conducted identical analyses for searchlight radii of 4 and 6 voxels. We note that increasing the size of the searchlight may both benefit and hurt the mapping results and their interpretation. A larger searchlight augments the amount of potentially useful information but also increases the dimensionality of the patterns leading to more overfitting. Also, the larger the searchlight is, the less local the mapping results will be: Highly local information will be exploited by all searchlight masks that contain it over a larger area, thus leading to a more diffuse map—see Fig. S3 for an example. Our choice represents a compromise between searching for local information and exploiting a sufficient amount of spatial information.

The temporal size of the window was selected to capture the peak of the hemodynamic response function (HRF) (16). We note that a full-blown version of a spatial–temporal searchlight would have to walk a window in both space and time. Whereas this approach may provide a more detailed assessment of the temporal–spatial profile of information maps, such analysis comes at significant additional computational cost. As an alternative to this approach, we restrict our analysis to spatial mapping and keep the position of our temporal window fixed.

Next, to boost the signal-to-noise ratio (SNR) of our patterns, we averaged stimulus-specific patterns by stimulus identity. Thus, all patterns elicited during a functional run by images of the same individual, irrespective of the expression displayed, were combined into a single one. This procedure produced 17 different patterns, 1 per run, for each of four different facial identities. A similar procedure was used for OF stimuli.

To measure identity discriminability, we applied multiclass SVM classification using a one-against-one approach to speed up computations (17)—that is, each facial identity is compared with every other one at a time. Our particular choice of classifier is linear SVM with a trainable c term because it appears to perform better or, at least, equivalently to other classifiers tested on neuroimaging data (18, 19). Leave-one-run-out cross-validation was carried out for each pair of facial identities. At the same time, nested cross-validation within each training set was conducted to optimize the c term (allowed to range between 2^{-4} and 2^{10}) and minimize overfitting. Discriminability was next encoded using the sensitivity measure d' (20). Voxelwise averaging of these estimates across each of the six pairs of identities compared produced subject-specific information maps.

Because it appears that multivariate analysis is able exploit high-frequency spatial information (12, 21), we attempted to minimize the amount of distortion of the functional data and preserve this information. Thus, multivariate analysis was carried out on unsmoothed data in each subject's native space. However, for the purpose of group analysis all information maps were brought into Talairach space. Group information maps were obtained by averaging across subjects and statistical effects were computed using a one-sample t test against chance ($d' = 0$). Finally, multiple-comparison correction was implemented using FDR.

Whereas the analysis described above was designed to take advantage of information distributed across patterns of activation, it is possible that patterns per se contribute little, if anything, to the effects detected. In other words, it is possible that multivariate effects present in the data can be accounted for by univariate effects. To test this hypothesis, we carried out an analysis following the same procedure with the sole difference that patterns are averaged into single values previous to classification. This simplification renders the analysis comparable to a univariate t test.

Finally, we conducted a similar set of multivariate analyses to examine expression discrimination and OF individuation as well as category-level classification (faces versus OFs). More precisely, we computed discrimination performance among (*i*) four different expressions across changes in facial identity, (*ii*) four different OF identities across changes in font type, and (*iii*) two categories across variations in both identity and category-specific changes. In all other respects, the computation of the respective information maps follows the procedure described above.

With respect to category-level discrimination, we note two factors that need to be taken into account. First, the categories being discriminated, faces and OFs, are very dissimilar, both perceptually and conceptually. Second, the SL size was larger than that typically used, e.g., a 2-voxel radius (22), and represented a compromise between maintaining a local encoding con-

straint and maximizing the amount of spatial information as discussed above. Thus, it is possible that category information is somewhat more focal than we ended up finding (Fig. S2). Nevertheless, category differentiation was sufficiently dispersed to produce a rather diffuse information-based map.

Analyses were carried out in Matlab using the SVMLIB 2.88 library for pattern classification (23).

Note on the Use of Spatiotemporal Information in Pattern Classification. The use of spatiotemporal information for multivariate analyses (24) presents us with an interesting opportunity. The temporal properties of the BOLD signal (e.g., time to peak or time to rise) provide a rich source of information regarding the neural dynamics (25). However, both their interpretation in relationship with the actual neural dynamics and their estimation can be problematic (26)—although not more so than that of the ubiquitously used signal amplitude. Multivariate spatiotemporal analysis (24, 27) allows us to bypass the latter problem in that no estimation of temporal properties (or amplitudes for that matter) is required. Rather, the use of such information is implicit and, thus, eliminates the issue of model (mis)specification (26).

Furthermore, if selecting a single time point for the analysis, it is unclear which one encodes the most diagnostic spatial information. The HRF peak may lead, in certain cases, to the best decoding accuracy (28). However, the shape of the HRF, including the timing of the peak, can vary significantly among cortical areas and across subjects (29). Fortunately, the use of multiple time points allows for the possibility that diagnostic spatial information, whether corresponding to the response peak or not, can be present at different times in different areas or across different subjects. Our analysis above uses a combination of these two approaches by exploiting multiple time points while restricting their number to those likely to capture the HRF peak.

Overall, the advantages above make spatiotemporal analysis very appealing as long as the increase in pattern dimensionality introduced by this approach can be handled adequately (e.g., by the use of classifiers that scale well with dimensionality).

RFE Analysis. RFE serves three different but related goals: dimensionality reduction within the original feature space, optimization of classification performance, and feature ranking (30). The analysis proceeds as follows: (*i*) we train a linear SVM classifier on a given feature set, (*ii*) we compute a ranking criterion for all features, (*iii*) we eliminate the feature with the smallest rank, and (*iv*) we repeat until no features are left in the set.

One of the simplest feature ranking criteria for linear SVM, and the one we follow here, is based on maximizing the separating margin width of the classifier (30). More specifically, the algorithm eliminates within each iteration the feature with the smallest $c_i = w_i^2$, where w_i is the weight corresponding to feature i. This procedure has the effect of maintaining the largest possible margin width $W = \|w\|$ at each iteration step. The number of iteration steps corresponds, in this version of the algorithm, to the total number of features in the initial set. Whereas batch elimination provides an easy alternative to speeding up computations, it may lead to suboptimal estimates of performance and also compromise feature ranking. For this reason, we favored single-feature elimination in our analysis.

RFE analysis has been successfully used in the past to reduce the dimensionality of fMRI data (31) and to map voxel diagnosticity in category-level discrimination (32). Here, we use it both to map feature (voxel X time point) diagnosticity and to improve on the classification models derived for individuation.

The analysis was separately applied to each pair of facial identities. Diagnosticity rankings as well as performance estimates were computed by averaging across all six different pairs.

Contribution of Low-Level Image (Dis)Similarity to Individuation Performance. Low-level similarity was computed between any two images of different facial identities. More precisely, we applied an L_2 (Euclidian) metric (4) to estimate low-level image dissimilarity. To ensure the robustness of the results, the metric was applied in three different ways corresponding to different ways of extracting information: (*i*) to entire images using only the luminance channel, (*ii*) to cropped images using only luminance, and (*iii*) to cropped images using all color channels (Fig. S6).

The manipulations above are motivated, first, by the privileged role of internal features relative to external ones in face perception (33) and, second, by the contribution of color to face processing (34, 35). Thus, to deal with the first issue we used cropping to eliminate external features (e.g., hair and face outline) and to retain internal ones (e.g., eyes and mouth). To deal with the second, we combined similarity measures computed independently for each color channel. However, whereas color is known to be involved both in low-level (36) and in high-level face processing (35), the relative contribution of different color channels is still unclear (37). For this reason, all channels were given equal weight in computing the estimates corresponding to case *iii* above. Specifically, channel-specific estimates were z-scored across image pairs and then averaged to produce single values for each pair. Finally, the values obtained for all 16 image pairs corresponding to two different identities were averaged to produce a single score (Table S1).

The three types of measurement are in overall agreement with each other: identities 1 and 2 along with 2 and 4 are relatively similar to each other whereas 3 and 4 are the most dissimilar (identity numbers refer to the columns in Fig. 1).

Next, dissimilarity estimates were correlated with individuation performance across identity pairs separately for each ROI, experimental subject, and type of measurement. The resulting correlation coefficients were converted into normally distributed variables using Fisher's z transform, allowing us to conduct parametric tests on the results:

$$z = \frac{1}{2}\ln\frac{1+r}{1-r}.$$

Finally, average subject scores were compared against chance via one-group t statistics (Fig. S4).

Pairwise ROI Analysis. ROI-based classification patterns provide only a coarse and summary measure of the relevant information present in the ROIs, namely in the activation patterns they host. However, they can be useful in that they offer an estimate of common biases in misclassification.

To compare classification patterns produced by different ROIs we used both partial correlation and conditional mutual information while controlling for the pattern of correct (true) labels.

Correlation coefficients were converted to z scores using Fisher's z transform.

Conditional mutual information (38) was computed as follows:

$$I(C_1, C_2|T) = \sum_{\substack{C_1,C_2\in\{0,1\}\\ T\in\{0,1\}}} p(C_1, C_2, T)\log\left(\frac{p(C_1, C_2|T)}{p(C_1|T)p(C_2|T)}\right).$$

Here C_1 and C_2 are binary variables encoding the classification labels for two different regions and T is a binary variable encoding the true labels.

The two measures were separately computed for each pair of ROIs and averaged across face pairs and subjects.

1. Eckman EP (1972) *Nebraska Symposium on Motivation*, ed Cole CJ (Univ of Nebraska Press, Lincoln, NE), pp 207–283.
2. Brainard DH (2003) *The Science of Color*, ed Shevell SK (Optical Society of America, Washington, DC), pp 191–216.
3. Turk M, Pentland A (1991) Eigenfaces for recognition. *J Cogn Neurosci* 3:71–86.
4. Moon H, Phillips PJ (2001) Computational and performance aspects of PCA-based face-recognition algorithms. *Perception* 30:303–321.
5. Burton AM, Miller P, Bruce V, Hancock PJ, Henderson Z (2001) Human and automatic face recognition: A comparison across image formats. *Vision Res* 41:3185–3195.
6. Pelli DG (1997) The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spat Vis* 10:437–442.
7. Brainard DH (1997) The psychophysics toolbox. *Spat Vis* 10:433–436.
8. Wager TD, Nichols TE (2003) Optimization of experimental design in fMRI: A general framework using a genetic algorithm. *Neuroimage* 18:293–309.
9. Vigneau M, Jobard G, Mazoyer B, Tzourio-Mazoyer N (2005) Word and non-word reading: What role for the Visual Word Form Area? *Neuroimage* 27:694–705.
10. Polk TA, Farah MJ (2002) Functional MRI evidence for an abstract, not perceptual, word-form area. *J Exp Psychol Gen* 131:65–72.
11. Cox RW (1996) AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res* 29:162–173.
12. Swisher JD, et al. (2010) Multiscale pattern analysis of orientation-selective activity in the primary visual cortex. *J Neurosci* 30:325–330.
13. Boynton GM, Engel SA, Glover GH, Heeger DJ (1996) Linear systems analysis of functional magnetic resonance imaging in human V1. *J Neurosci* 16:4207–4221.
14. Friston KJ, et al. (1995) Statistical parametric maps in functional imaging: A general linear approach. *Hum Brain Mapp* 2:189–210.
15. Genovese CR, Lazar NA, Nichols T (2002) Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage* 15:870–878.
16. Friston KJ, Jezzard P, Turner R (1994) Analysis of functional MRI time-series. *Hum Brain Mapp* 1:153–171.
17. Hsu CW, Lin CJ (2002) A comparison of methods for multiclass support vector machines. *IEEE Trans Neural Netw* 13:415–425.
18. Pereira F, Botvinick M (2011) Information mapping with pattern classifiers: A comparative study. *Neuroimage* 56:476–496.
19. Misaki M, Kim Y, Bandettini PA, Kriegeskorte N (2010) Comparison of multivariate classifiers and response normalizations for pattern-information fMRI. *Neuroimage* 53: 103–118.
20. Salton G, McGill M (1983) *Introduction to Modern Information Retrieval* (McGraw-Hill, New York).
21. Kriegeskorte N, Cusack R, Bandettini P (2010) How does an fMRI voxel sample the neuronal activity pattern: Compact-kernel or complex spatiotemporal filter? *Neuroimage* 49:1965–1976.
22. Kriegeskorte N, Goebel R, Bandettini P (2006) Information-based functional brain mapping. *Proc Natl Acad Sci USA* 103:3863–3868.
23. Chang CC, Lin CJ (2011) LIBSVM: A library for support vector machines. *ACM Trans Intell Syst Technol* 2:27.
24. Mourão-Miranda J, Friston KJ, Brammer M (2007) Dynamic discrimination analysis: A spatial-temporal SVM. *Neuroimage* 36:88–99.
25. Formisano E, Goebel R (2003) Tracking cognitive processes with functional MRI mental chronometry. *Curr Opin Neurobiol* 13:174–181.
26. Lindquist MA, Meng Loh J, Atlas LY, Wager TD (2009) Modeling the hemodynamic response function in fMRI: Efficiency, bias and mis-modeling. *Neuroimage* 45(1, Suppl): S187–S198.
27. Mitchell TM, et al. (2004) Learning to decode cognitive states from brain images. *Mach Learn* 57:145–175.
28. Johnson JD, McDuff SG, Rugg MD, Norman KA (2009) Recollection, familiarity, and cortical reinstatement: A multivoxel pattern analysis. *Neuron* 63:697–708.
29. Handwerker DA, Ollinger JM, D'Esposito M (2004) Variation of BOLD hemodynamic responses across subjects and brain regions and their effects on statistical analyses. *Neuroimage* 21:1639–1651.
30. Guyon I, Weston J, Barnhill S, Vapnik V (2002) Gene selection for cancer classification using support vector machines. *Mach Learn* 46:389–422.
31. De Martino F, et al. (2008) Combining multivariate voxel selection and support vector machines for mapping and classification of fMRI spatial patterns. *Neuroimage* 43:44–58.
32. Hanson SJ, Halchenko YO (2008) Brain reading using full brain support vector machines for object recognition: There is no "face" identification area. *Neural Comput* 20:486–503.
33. Andrews TJ, Davies-Thompson J, Kingstone A, Young AW (2010) Internal and external features of the face are represented holistically in face-selective regions of visual cortex. *J Neurosci* 30:3544–3552.
34. Bindemann M, Burton AM (2009) The role of color in human face detection. *Cogn Sci* 33:1144–1156.
35. Nestor A, Tarr MJ (2008) Gender recognition of human faces using color. *Psychol Sci* 19:1242–1246.
36. Yip AW, Sinha P (2002) Contribution of color to face recognition. *Perception* 31: 995–1003.
37. Nestor A, Tarr MJ (2008) The segmental structure of faces and its use in gender recognition. *J Vis* 8:7–.1–12.
38. Cover TM, Thomas JA (1991) *Elements of Information Theory* (Wiley, New York).

Fig. S1. Experimental orthographic form (OF) stimuli (four pseudowords × four types of font). All stimuli were five-letter pronounceable nonwords with the same syllabic structure but different orthographic properties (they contain different letters in a given position).

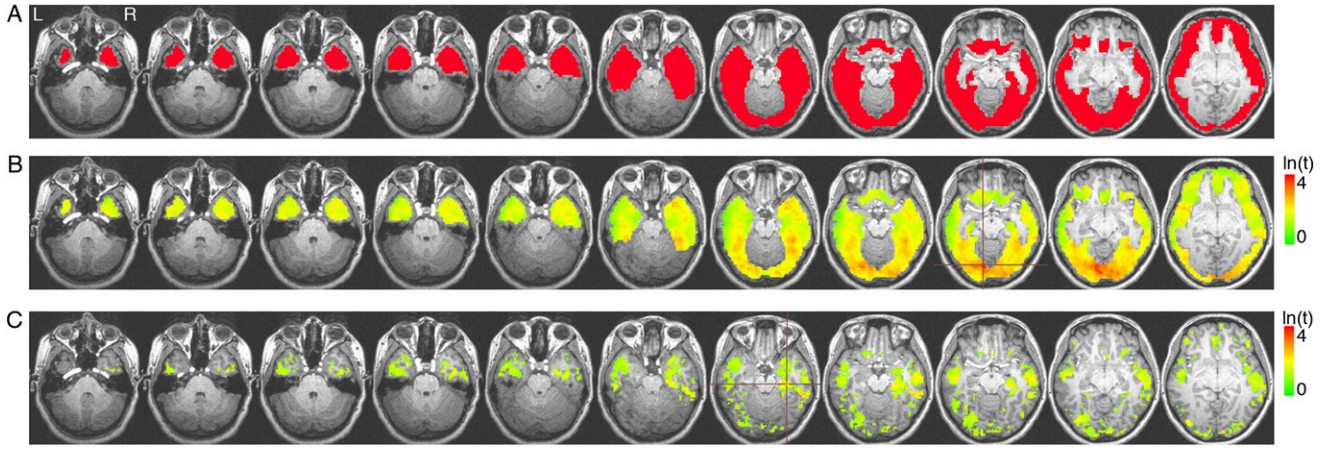

Fig. S2. (*A*) Group cortical mask and group information-based maps of category-level discrimination (faces vs. OFs; $q < 0.05$) derived through (*B*) multivariate searchlight and (*C*) its univariate analog. Effect size is scaled logarithmically. Crosshairs mark the discrimination peaks in each map (Talairach coordinates −11, −76, −11 and 31, −24, −16 for *B* and *C*, respectively). The differences between the two types of map indicate that univariate analysis underestimates the amount and expanse of category information in ventral cortex compared with its multivariate counterpart.

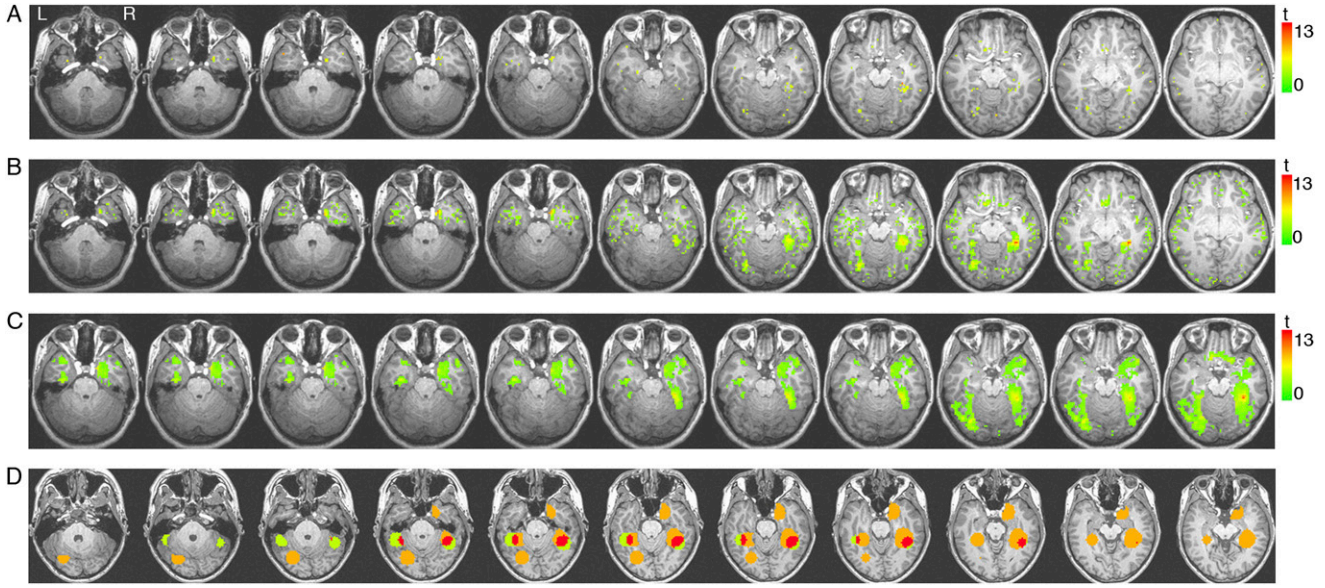

Fig. S3. (*A–C*) Group information-based maps of face individuation obtained with SL masks of different spatial radii (4, 5, and 6 voxels). Maps were thresholded at a liberal level ($q < 0.10$) for better comparison with each other. Larger masks yielded a larger number of significant voxels as expected on the basis of the loss of spatial specificity. Despite such variation, the three maps agree on the presence of individuation effects in the bilateral FG and the right aMTG. (*D*) Examples of SL ROIs (orange), bilateral FFA (green), and their overlap (red) within a single subject.

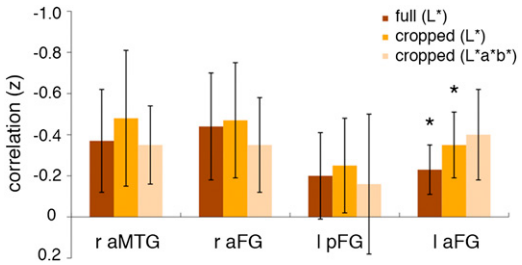

Fig. S4. Correlation of discrimination performance with low-level image distances across different SL ROIs (Fisher's *z* scores). Low-level distances were computed in three different ways: across the entire image using only luminance (full L*), across internal features using only luminance (cropped L*), and across internal features using all color channels (cropped L*a*b*). The laFG produced correlations of discrimination performance and image distance for luminance-based measurements ($P < 0.05$ uncorrected for multiple comparisons)—the effect was less reliable across all color channels ($P < 0.06$ uncorrected).

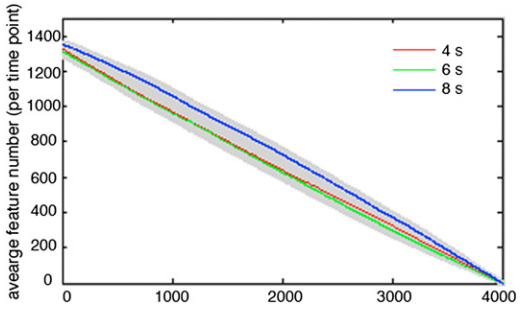

Fig. S5. The time course of feature elimination by time point for 4,000 features—top 1,000 features for each of four ROIs. Features tend to be eliminated at similar rates across different time points although we note a small advantage for features extracted at 8 s poststimulus onset (shaded areas correspond to ±1 SE across subjects).

Fig. S6. Examples of image cropping used for stimulus selection and for measurements of low-level image similarity. Images were cropped to show only internal features. Face images courtesy of the Face-Place Face Database Project (http://www.face-place.org/) Copyright 2008, Michael J. Tarr. Funding provided by NSF Award 0339122.

Table S1. Estimates of low-level image distances across facial identity pairs (*z* scores)

Identity pair	Full image (L*)	Cropped image (L*)	Cropped image (L*a*b*)
1–2	−0.31	−0.54	−1.07
1–3	−0.19	0.22	0.31
1–4	0.58	0.08	0.04
2–3	−0.15	0.13	−0.08
2–4	−0.91	−0.91	−0.38
3–4	0.98	1.02	1.19