Unraveling the distributed neural code of facial identity through spatiotemporal pattern analysis

Adrian Nestor[1], David C. Plaut, and Marlene Behrmann

Department of Psychology, Carnegie Mellon University, and Center for the Neural Basis of Cognition, Pittsburgh, PA 15213

Edited by Charles G. Gross, Princeton University, Princeton, NJ, and approved May 12, 2011 (received for review February 13, 2011)

Face individuation is one of the most impressive achievements of our visual system, and yet uncovering the neural mechanisms subserving this feat appears to elude traditional approaches to functional brain data analysis. The present study investigates the neural code of facial identity perception with the aim of ascertaining its distributed nature and informational basis. To this end, we use a sequence of multivariate pattern analyses applied to functional magnetic resonance imaging (fMRI) data. First, we combine information-based brain mapping and dynamic discrimination analysis to locate spatiotemporal patterns that support face classification at the individual level. This analysis reveals a network of fusiform and anterior temporal areas that carry information about facial identity and provides evidence that the fusiform face area responds with distinct patterns of activation to different face identities. Second, we assess the information structure of the network using recursive feature elimination. We find that diagnostic information is distributed evenly among anterior regions of the mapped network and that a right anterior region of the fusiform gyrus plays a central role within the information network mediating face individuation. These findings serve to map out and characterize a cortical system responsible for individuation. More generally, in the context of functionally defined networks, they provide an account of distributed processing grounded in information-based architectures.

The neural basis of face perception is the focus of extensive research as it provides key insights both into the computational architecture of visual recognition (1, 2) and into the functional organization of the brain (3). A central theme of this research emphasizes the distribution of face processing across a network of spatially segregated areas (4–10). However, there remains considerable disagreement about how information is represented and processed within this network to support tasks such as individuation, expression analysis, or high-level semantic processing.

One influential view proposes an architecture that maps different tasks to distinct, unique cortical regions (6) and, as such, draws attention to the specificity of this mapping (11–20). As a case in point, face individuation (e.g., differentiating Steve Jobs from Bill Gates across changes in expression) is commonly mapped onto the fusiform face area (FFA) (6, 21). Although recent studies have questioned this role of the FFA (14, 15), overall they agree with this task-based architecture as they single out other areas supporting individuation.

However, various distributed accounts have also been considered. One such account ascribes facial identity processing to multiple, independent regions. Along these lines, the FFA's sensitivity to individuation has been variedly extended to areas of the inferior occipital gyrus (5), the superior temporal sulcus (12), and the temporal pole (22). An alternative scenario is that identity is encoded by a network of regions rather than by any of its separate components—such a system was recently described for subordinate-level face discrimination (23). Still another distributed account attributes individuation to an extensive ventral cortical area rather than to a network of smaller separate regions (24). Clearly, the degree of distribution of the information supporting face individuation remains to be determined.

Furthermore, insofar as face individuation is mediated by a network, it is important to determine how information is distributed across the system. Some interesting clues come from the fact that right fusiform areas are sensitive to both low-level properties of faces (16, 25) and high-level factors (26, 27), suggesting that these areas may mediate between image-based and conceptual representations. If true, such an organization should be reflected in the pattern of information sharing among different regions.

The current work investigates the nature and the extent of identity-specific neural patterns in the human ventral cortex. We examined functional MRI (fMRI) data acquired during face individuation and assessed the discriminability of activation patterns evoked by different facial identities across variation in expression. To uncover the neural correlate of identity recognition, we performed dynamic multivariate mapping by combining information-based mapping (28) and dynamic discrimination analysis (29). The results revealed a network of fusiform and anterior temporal regions that respond with distinct spatiotemporal patterns to different identities. To elucidate the distribution of information, we examined the distribution of diagnostic information across these regions using recursive feature elimination (RFE) (30) and related the information content of different regions to each other. We found that information is evenly distributed among anterior regions and that a right fusiform region plays a central role within this network.

Results

Participants performed an individuation task with faces (Fig. 1) and orthographic forms (OFs) (Fig. S1). Specifically, they recognized stimuli at the individual level across image changes introduced by expression (for faces) or font (for OFs). Response accuracy was at ceiling (>95%) as expected given the familiarization with the stimuli before scanning and the slow rate of stimulus presentation. Thus, behavior exhibits the expected invariance to image changes, and the current investigation focuses on the neural codes subserving this invariance.

Dynamic Multivariate Mapping. The analysis used a searchlight (SL) with a 5-voxel radius and a 3-TR (Time to Repeat) temporal envelope to constrain spatiotemporal patterns locally. These patterns were submitted to multivariate classification on the basis of facial identity (*Methods* and *SI Text*). The outcome of the analysis is a group information-based map (28) revealing the strength of discrimination (Fig. 2). Each voxel in this map represents an entire region of neighboring voxels defined by the SL mask.

Author contributions: A.N., D.C.P., and M.B. designed research; A.N. performed research; A.N., D.C.P., and M.B. analyzed data; and A.N., D.C.P., and M.B. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The fMRI dataset has been deposited with the XNAT Central database under the project name "STSL."

[1]To whom correspondence should be addressed. E-mail: anestor@andrew.cmu.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1102433108/-/DCSupplemental.

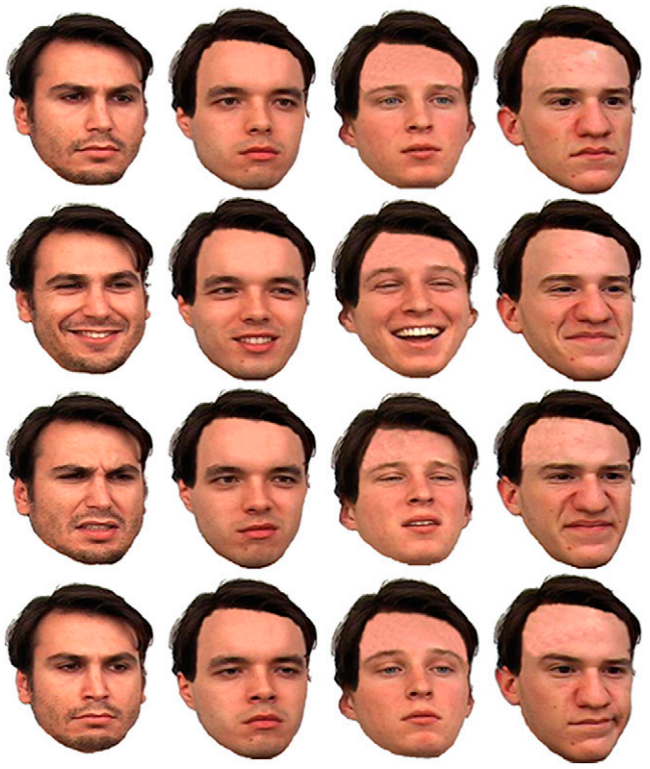

Fig. 1. Experimental face stimuli (4 identities × 4 expressions). Stimuli were matched with respect to low-level properties (e.g., mean luminance), external features (hair), and high-level characteristics (e.g., sex). Face images courtesy of the Face-Place Face Database Project (http://www.face-place.org/) Copyright 2008, Michael J. Tarr. Funding provided by NSF Award 0339122.

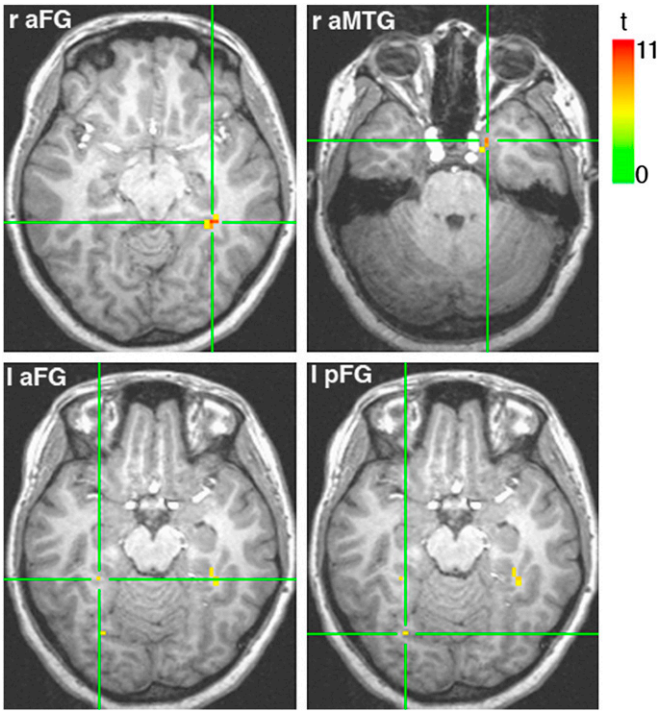

Fig. 2. Group information-based map of face individuation. The map is computed using a searchlight (SL) approach and estimates the discriminability of facial identities across expression ($q < 0.05$). Each voxel in the map represents the center of an SL-defined region supporting identity discrimination. The four slices show the sensitivity peaks of the four clusters revealed by this analysis.

Our mapping revealed four areas sensitive to individuation: two located bilaterally in the anterior fusiform gyrus (aFG), one in the right anterior medial temporal gyrus (aMTG), and one in the left posterior fusiform gyrus (pFG). The largest of these areas corresponded to the right (r)aFG whereas the smallest corresponded to its left homolog (Table 1).

To test the robustness of our mapping, the same cortical volume (Fig. S2*A*) was examined with SL masks of different sizes. These alternative explorations produced qualitatively similar results (Fig. S3 *A–C*). In contrast, a univariate version of the mapping (*SI Text*) failed to uncover any significant regions, even at a liberal threshold ($q < 0.10$), attesting to the strength of multivariate; mapping.

To further evaluate these results, we projected the four areas from the group map back into the native space of each subject and expanded each voxel to the entire SL region centered on it. The resulting SL clusters (Fig. S3*D*) mark entire regions able to support above-chance classification. Examination of subject-specific maps revealed that the bilateral aFG clusters were consistently located anterior to the FFA [rFFA peak coordinates, 39, −46, and −16; left (l)FFA, −36, −47, and −18]. However, we also found they consistently overlapped with the FFA (mean ± SD: 24 ± 14% of raFG volume and 35 ± 20% of laFG).

Finally, multivariate mapping was applied to other types of discrimination: expression classification (across identities) and category-level classification (faces versus OFs). Whereas the former analysis did not produce any significant results, the latter found reliable effects extensively throughout the cortical volume analyzed (Fig. S2*B*). In contrast, a univariate version of the latter analysis revealed considerably less sensitivity than its multivariate counterpart (Fig. S2*C*).

The results above suggest that identity coding relies on a distributed cortical system. Clarifying the specificity of this system to face individuation is addressed by our region-of-interest (ROI) analyses.

ROI Analyses. First, we examined whether the FFA supports reliable face individuation as tested with pattern classification. Bilateral FFAs were identified in each subject using a standard face localizer, and discrimination was computed across all features in a region—given the use of spatiotemporal patterns, our features are voxel *X* time-point pairings rather than voxels alone. The analysis revealed above-chance performance for rFFA (Fig. 3).

To reduce overfitting, we repeated the analysis above using subsets of diagnostic features identified by multivariate feature selection, specifically RFE. The method works by systematically removing features, one at a time, on the basis of their impact on classification (*SI Text*). Following this procedure, we found above-chance performance bilaterally in the FFA (Fig. 3). In contrast, early visual cortex (EVC) did not exhibit significant sensitivity either before or after feature selection.

Second, we reversed our approach by using multivariate mapping to localize clusters and univariate analysis to assess face se-

Table 1. Areas sensitive to face individuation

	Coordinates (peak)				
Region	*x*	*y*	*z*	SL centers (voxels)	Peak *t* value
raFG	33	−39	−9	16	11.31
raMTG	19	6	−26	8	9.90
lpFG	−26	−69	−14	2	7.11
laFG	−29	−39	−14	1	7.24

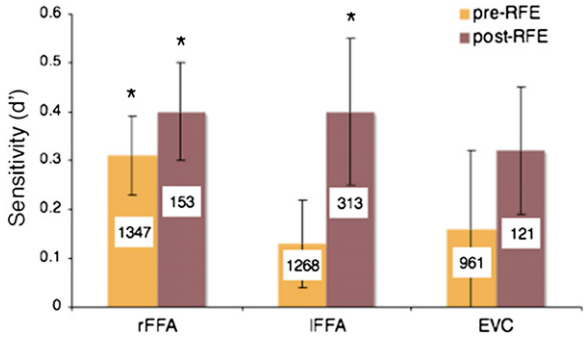

Fig. 3. Sensitivity estimates in three ROIs. Facial identity discrimination was computed using both the entire set of features in an ROI and a subset of diagnostic features identified by multivariate feature selection (i.e., RFE), the two types of classification are labeled as pre- and post-RFE. The average number of features involved in classification is superimposed on each bar. The results indicate that the bilateral FFA, in contrast to an early visual area, contains sufficient information to discriminate identities above chance ($P < 0.05$).

lectivity. Concretely, we examined the face selectivity of the SL clusters using the data from our functional localizers. No reliable face selectivity was detected for any cluster.

Third, SL clusters along with the FFA were tested for their ability to discriminate expressions across changes in identity. Above-chance discrimination was found in rFFA and raMTG ($P < 0.05$).

Finally, we tested our clusters for OF individuation across variation in font. The analysis found sensitivity in two regions: rFFA and lpFG.

These findings are important in several respects. They suggest conventionally defined face selectivity, although informative, may not be enough to localize areas involved in fine-level face representation. Also, they show that the identified network is not exclusively dedicated to individuation or even to face processing per se. One hypothesis, examined below, may explain this involvement in multiple types of perceptual discrimination simply by appeal to low-level image properties.

Impact of Low-Level Image Similarity on Individuation. To determine the engagement of the network in low-level perceptual processing, image similarity was computed across images of different individuals using an L_2 metric (Table S1). For each pair of face identities, the average distance was correlated with the corresponding discrimination score produced by every ROI (including the FFA). The only ROI susceptible to low-level image sensitivity was laFG ($P < 0.05$ uncorrected) (Fig. S4).

These results, along with the inability of the EVC to support individuation, suggest that low-level similarity is unlikely to be the main source of the individuation effects observed here.

Feature Ranking and Mapping. Having located a network of regions sensitive to face identity, we set out to determine the spatial and temporal distribution of the features diagnostic of face individuation. Specifically, we performed RFE analysis jointly across all SL clusters and recorded the ranking of the features within each subject. To eliminate any spatial bias within the initial feature set, we started with an equally large number of features for each cluster: the 1,000 top-ranked ones based on single-cluster RFE computation.

Fig. 4*A* shows the average ranking of the 400 most informative features across subjects and Fig. 4*B* summarizes their distribution across clusters and time points. We found a significant effect of cluster (two-way analysis of variance, $P < 0.05$) but no effect of time point and no interaction. Further comparisons revealed that lpFG contains fewer features than other clusters ($P < 0.05$), which did not differ among each other. The time course of feature elimination revealed that lpFG features were consistently eliminated at a higher rate than features from other clusters (Fig. 4*C*). A

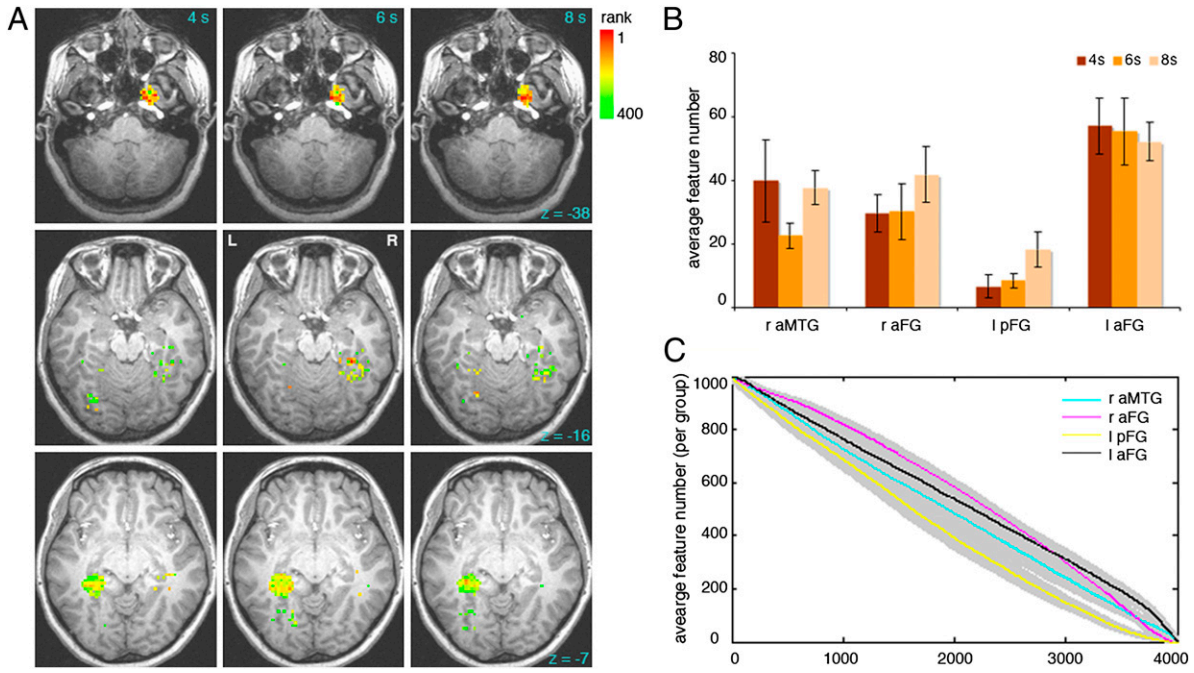

Fig. 4. Spatiotemporal distribution of information diagnostic for face individuation. (*A*) Group map of average feature ranking for the top 400 features—rows show different slices and columns different time points. Color codes the ranking of the features across space (the four regions identified by our SL analysis) and time (4–8 s poststimulus onset). The map shows a lower concentration of features in the lpFG relative to other regions but a comparable number of features across time. (*B*) Average feature distribution across subjects by cluster and time point (the bar graph quantifies the results illustrated in *A*). (*C*) Time course of feature elimination by ROI for 4,000 features (top 1,000 features for each ROI). This analysis confirms that lpFG features are eliminated at a higher rate, indicative of their reduced diagnosticity (shaded areas show ±1 SE across subjects).

similar analysis across time points showed no substantial differences across time (Fig. S5).

Feature mapping provides a bird's eye view of information distribution across regions. In our case, it reveals a relatively even division of diagnostic information among anterior regions. Further pairwise comparisons of SL clusters were deployed to examine how areas share information with each other.

Information-Based Pairwise Cluster Analysis. Whereas activation patterns in different areas are not directly comparable with each other (e.g., because they have different dimensionalities, there is no obvious mapping between them, etc.), the classification results they produce serve as convenient proxies (*Methods*). Here, we compared classification patterns for pairs of regions while controlling for the pattern of correct labels. Thus, the analysis focuses on common biases in misclassification.

First, we computed the partial correlation between classification patterns corresponding to different clusters. Group results are displayed as a graph in Fig. 5. Similarity scores for all pairs of regions tested above chance (one-sample t test, $P < 0.01$). Similarity scores within the network were not homogeneous (one-way analysis of variance, $P < 0.05$) mainly because raFG evinced higher similarity estimates than the rest ($P < 0.01$). In addition, we computed similarity scores between the four clusters and the EVC. Average similarity scores within the network were significantly higher than scores with the EVC (paired-sample t test, $P < 0.01$). Second, to verify our findings using an information-theoretic measure, we computed the conditional mutual information between classification patterns produced by different clusters (Fig. 5). Examination of information estimates revealed a relational structure qualitatively similar to that obtained using correlation.

We interpret the results above as evidence for the central role of the right FG in face individuation. More generally, they support the idea of redundant information encoding within the network.

Discussion

The present study investigates the encoding of facial identity in the human ventral cortex. Our investigation follows a multivariate approach that exploits multivoxel information at multiple stages from mapping to feature selection and network analysis. We favor this approach because multivariate methods are more successful at category discrimination than univariate tests (31) and possibly more sensitive to "subvoxel" information than adaptation techniques (32). In addition, we extend our investigation to take advantage of spatiotemporal information (29) and, thus, optimize the discovery of small fine-grained pattern differences underlying the perception of same-category exemplars.

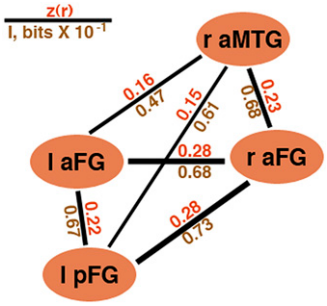

Fig. 5. Pairwise ROI relations. The pattern of identity (mis)classifications is separately compared for each pair of regions using correlation-based scores (red) and mutual information (brown). Specifically, we relate classification results across regions while controlling for the pattern of true labels. These measures are used as a proxy for assessing similarity in the encoding of facial identity across regions. Of the four ROIs, the raFG produced the highest scores in its relationship with the other regions (connector width is proportional to z values).

Multiple Cortical Areas Support Face Individuation. Multivariate mapping located four clusters in the bilateral FG and the right aMTG encoding facial identity information. These results indicate that individuation relies on a network of ventral regions that exhibit sensitivity to individuation independently of each other. This account should be distinguished both from local ones (6, 21) and from other versions of distributed processing (23, 24).

With regard to the specific clusters identified, previous work uncovered face-selective areas in the vicinity of the FFA, both posterior (9) and anterior (33) to it. Our clusters did not exhibit face selectivity when assessed with a univariate test. However, this lack of selectivity may reflect the variability of these areas (9, 33) and/or the limitations of univariate analysis (34). Alternatively, it is possible that face processing does not necessarily entail face selectivity (35). More relevantly here, face individuation, rather than face selectivity, was previously mapped to an area in the anterior vicinity of the FFA (14). Overall, our present results confirm the involvement of these areas in face processing and establish their role in individuation.

On a related note, the proximity of these fusiform clusters to the FFA may raise questions as to whether they are independent clusters or rather extensions of the FFA (7, 9). On the basis of differences in peak location and the lack of face selectivity, we treat them here as distinct from the FFA although further investigation is needed to fully understand their relationship with this area.

Unlike the FG areas discussed above, an anterior region of the right middle temporal cortex (36, 37) or temporal pole (22) was consistently associated with identity coding. Due to its sensitivity to higher-level factors, such as familiarity (22), and its involvement in conceptual processing (38, 39), the anterior temporal cortex is thought to encode biographical information (6). Whereas our ability to localize this region validates our mapping methodology, the fact that our stimuli were not explicitly associated with any biographical information suggests that the computations hosted by this area also involve a perceptual component. Consistent with this, another mapping attempt, based on perceptual discrimination (15), traced face individuation to a right anterior temporal area. Also, primate research revealed encoding of face–space dimensions in the macaque anterior temporal cortex (40) as well as different sensitivity to perceptual and semantic processing of facial identity (41). In light of these findings, we argue that this area is part of the network for *perceptual* face individuation although the representations it hosts are likely to also comprise a conceptual component.

In sum, our mapping results argue for a distributed account of face individuation that accommodates a multitude of experimental findings. Previous imaging research may have failed to identify this network due to limits in the sensitivity of the methods used in relation with the size of the effect. Inability to find sensitivity in more than one region can easily lead to a local interpretation. At the same time, combining information from multiple regions may counter the limitations of one method but overestimate the extent of distributed processing. Our use of dynamic multivariate mapping builds upon these previous findings and is a direct attempt to increase the sensitivity of these mapping methods at the cost of computational complexity.

Importantly, the regions uncovered by our analysis may represent only a subset of the full network of regions involved in identity processing. We allow for this possibility given our limited coverage (intended to boost imaging resolution in the ventral cortex) as well as the lower sensitivity associated with the imaging of the inferior temporal cortex. In particular, regions of the superior temporal sulcus and prefrontal cortex (4, 10) are plausible additions to the network uncovered here.

Individuation Effects Are Not Reducible to Low-Level Image Processing. To distinguish identity representations from low-level image differences, we appealed to a common source of intraindividual image variations: emotional expressions. Also, identity was not predictable in our case by prominent external features, such as hair, or by image properties associated with higher-level characteristics (e.g., sex or age). Furthermore, we assessed the contribution of interindividual low-level similarity to discrimination performance. Of all regions examined, only laFG showed potential reliance on low-level properties. Finally, an examination of an early visual area did not produce any evidence for identity encoding. Taken together, these results render unlikely an explanation of individuation effects based primarily on low-level image properties.

The Face Individuation System Does Not Support General-Purpose Individuation. To address the domain specificity of the system identified, we examined whether "abstract" OFs (i.e., independent of font) can be individuated within the regions mapped for faces. Although highly dissimilar from faces, OFs share an important attribute with them, by requiring fine-grained perceptual discrimination at the individual level. In addition, they appear to compete with face representations (42) and to rely on similar visual processing mechanisms (43). Thus, they may represent a more suitable contrast category for faces than other familiar categories, such as houses. Our attempt to classify OF identities revealed that two regions, the lpFG and the rFFA, exhibited sensitivity to this kind of OF information.

Furthermore, to examine the task specificity of the network, we evaluated the ability of its regions to support expression discrimination and found that the rFFA, along with the aMTG, was able to perform this type of discrimination.

Thus, it appears that the mapped network does not support general object individuation, although it may share resources with the processing of other tasks as well as of other visual categories.

FFA Responds to Different Face Identities with Different Patterns. The FFA (44, 45) is one of the most intensely studied functional areas of the ventral stream. However, surprisingly, its role in face processing is far from clear. At one extreme, its involvement in face individuation (6, 18, 21) has been called into question (14–16); at the other, it has been extended beyond the face domain to visual expertise (17) and even to general object individuation (13).

Previous studies of the FFA using multivariate analysis have not been successful in discovering identity information (15, 24) or even subordinate-level information (23) about faces. The study of patient populations is also not definitive. FG lesions associated with acquired prosopagnosia (46) adversely impact face individuation, confirming the critical role of the FFA. However, individuals with congenital prosopagnosia appear to exhibit normal FFA activation profiles (47) in the presence of compromised fiber tracts connecting the FG to anterior areas (48). Thus, the question is more pertinent than ever: Does the FFA encode identity information?

Our results provide evidence that the FFA responds consistently across different images of the same individual, but distinctly to different individuals. In addition, we show that the right FFA can individuate OFs and decode emotional expressions, consistent with its role in expression recognition (12).

Thus, we confirm the FFA's sensitivity to face identity using pattern analysis. Moreover, we show that it extends beyond both a specific task, i.e., individuation, and a specific domain, i.e., faces. Further research is needed to determine how far its individuation capabilities extend and how they relate with each other.

Informative Features Are Evenly Distributed Across Anterior Regions. How uniformly is information distributed across multiple regions? At one extreme, the system may favor robustness as a strategy and assign information evenly across regions. At the other, its structure may be shaped by the feed-forward flow of information and display a clear hierarchy of regions from the least to the most diagnostic. The latter alternative is consistent, for instance, with a posterior-to-anterior accumulation of information culminating in the recruitment of the aMTG as the endpoint of identity processing.

Our results fall in between these two alternatives. Anterior regions appear to be at an advantage compared with the left pFG. At the same time, there was no clear differentiation among anterior regions in terms of the amount of information represented, suggesting that information is evenly distributed across them.

The Right aFG May Be a Hub in the Facial Identity Network. As different regions are not directly comparable as activation patterns, we used their classification results as a proxy for their comparison. Using this approach, we found that different regions do share information with each other, consistent with redundancy in identity encoding. Furthermore, we found that pairwise similarities are more prominent in relation with the right aFG than among other network regions, suggesting that the raFG plays a central role within the face individuation network. Thus, the raFG mirrors the role played by the right FFA among face-selective regions as revealed by functional connectivity (4). One explanation of this role is that a right middle/anterior FG area serves as an interface between low- and high-level information.

Two lines of evidence support this hypothesis. Recent results show, surprisingly, that the FFA exhibits sensitivity to low-level face properties (16, 25). Additionally, the right FG is subject to notable top–down effects (26, 27). Maintaining a robust interface between low-level image properties and high-level factors is likely a key requirement for fast, reliable face processing. Critical for our argument, this requirement would lead to the formation of an FG activation/information hub. Future combinations of information and activation-based connectivity analyses might be able to assess such hypotheses and provide full-fledged accounts of the flow of information in cortical networks.

Summary. A broad body of research suggests that face perception relies on an extensive network of cortical areas. Our results show that a single face-processing task, individuation, is supported by a network of cortical regions that share resources with the processing of other visual categories (OFs) as well as other face-related (expression discrimination) tasks. Detailed investigation of this network revealed an information structure dominated by anterior cortical regions, and the right FG in particular, confirming its central role in face processing. Finally, we suggest that a full understanding of the operation of this system requires a combination of conventional connectivity analyses and information-based explorations of network structure.

Methods

An extended version of this section is available in *SI Text*.

Design. Eight subjects were scanned across multiple sessions using a slow event-related design (10-s trials). Subjects were presented with a single face or OF stimulus for 400 ms and were asked to identify the stimulus at the individual level using a pair of response gloves. We imaged 27 oblique slices covering the ventral cortex at 3T (2.5-mm isotropic voxels, 2-s TR).

Dynamic Information-Based Brain Mapping. The SL was walked voxel-by-voxel across a subject-specific cortical mask. The mask covered the ventral cortex (Fig. S2) and was temporally centered on 6-s poststimulus onset. At each location within the mask, spatiotemporal patterns (29) were extracted for each stimulus presentation. To boost their signal, these patterns were averaged within runs on the basis of stimulus identity. Pattern classification was performed using linear support vector machines (SVM) with a trainable c term followed by leave-one-run-out cross-validation. Classification was separately applied to each pair of identities (six pairs based on four identities). Discrimination performance for each pair was encoded using d' and an

average information map (28) was computed across all pairs. For the purpose of group analysis, these maps were normalized into Talairach space and examined for above-chance sensitivity ($d' > 0$), using voxelwise t tests across subjects [false discovery rate (FDR) corrected]. Expression discrimination followed a similar procedure.

Analyses were carried out in Matlab with the Parallel Processing Toolbox running on a ROCKS+ multiserver environment.

ROI Localization. Three types of ROIs were localized as follows: (*i*) We identified locations of the group information map displaying above-chance discriminability and projected their coordinates in the native space of each subject. ROIs were constructed by placing spherical masks at each of these locations—a set of overlapping masks gave rise to a single ROI. (*ii*) Bilateral FFAs were identified for each subject by standard face–object contrasts. ROIs were constructed by applying a spherical mask centered on the FG face-selective peak. (*iii*) Anatomical masks were manually drawn around the calcarine sulcus of each subject and a mask was placed at the center of these areas. The results serve as a rough approximation of EVC. All masks used for ROI localization had a 5-voxel radius.

RFE-Based Analysis. SVM-based RFE (30) was used for feature selection and ranking—the order of feature elimination provides an estimate of feature diagnosticity for a given type of discrimination. To obtain unbiased estimates of performance, we executed two types of cross-validation. We performed cross-validation, first, at each RFE iteration step to measure performance and, second, across iteration steps to find the best number of features. RFE analysis was applied to all types of ROIs described above.

Pairwise ROI Analysis. We computed the similarity of classification patterns produced by each pair of ROIs, namely the patterns of classification labels obtained with test instances during cross-validation. However, patterns are likely correlated across regions by virtue of the ability of SVM models to approximate true labels. Therefore, we measured pattern similarity with partial correlation while controlling for the pattern of true labels. Correlations were computed for each pair of facial identities, transformed using Fisher's z, and averaged within subjects. Critically, to eliminate common biases based on spatial proximity between regions (because of spatially correlated noise) all activation patterns were z-scored before classification. To obtain estimates of the information shared between ROIs we also computed the conditional mutual information between ROI-specific classification patterns given the pattern of true labels.

ACKNOWLEDGMENTS. This work was funded by National Science Foundation Grant SBE-0542013 to the Temporal Dynamics of Learning Center and by National Science Foundation Grant BCS0923763. M.B. was partially supported by a Weston Visiting Professorship at the Weizmann Institute of Science.

1. DiCarlo JJ, Cox DD (2007) Untangling invariant object recognition. *Trends Cogn Sci* 11: 333–341.
2. Jiang X, et al. (2006) Evaluation of a shape-based model of human face discrimination using FMRI and behavioral techniques. *Neuron* 50:159–172.
3. Kanwisher N (2010) Functional specificity in the human brain: A window into the functional architecture of the mind. *Proc Natl Acad Sci USA* 107:11163–11170.
4. Fairhall SL, Ishai A (2007) Effective connectivity within the distributed cortical network for face perception. *Cereb Cortex* 17:2400–2406.
5. Gauthier I, et al. (2000) The fusiform "face area" is part of a network that processes faces at the individual level. *J Cogn Neurosci* 12:495–504.
6. Haxby JV, Hoffman EA, Gobbini MI (2000) The distributed human neural system for face perception. *Trends Cogn Sci* 4:223–233.
7. Pinsk MA, et al. (2009) Neural representations of faces and body parts in macaque and human cortex: A comparative FMRI study. *J Neurophysiol* 101:2581–2600.
8. Rossion B, et al. (2003) A network of occipito-temporal face-sensitive areas besides the right middle fusiform gyrus is necessary for normal face processing. *Brain* 126: 2381–2395.
9. Weiner KS, Grill-Spector K (2010) Sparsely-distributed organization of face and limb activations in human ventral temporal cortex. *Neuroimage* 52:1559–1573.
10. Tsao DY, Schweers N, Moeller S, Freiwald WA (2008) Patches of face-selective cortex in the macaque frontal lobe. *Nat Neurosci* 11:877–879.
11. Calder AJ, Young AW (2005) Understanding the recognition of facial identity and facial expression. *Nat Rev Neurosci* 6:641–651.
12. Fox CJ, Moon SY, Iaria G, Barton JJ (2009) The correlates of subjective perception of identity and expression in the face network: An fMRI adaptation study. *Neuroimage* 44:569–580.
13. Haist F, Lee K, Stiles J (2010) Individuating faces and common objects produces equal responses in putative face-processing areas in the ventral occipitotemporal cortex. *Front Hum Neurosci* 4:181.
14. Nestor A, Vettel JM, Tarr MJ (2008) Task-specific codes for face recognition: How they shape the neural representation of features for detection and individuation. *PLoS ONE* 3:e3978.
15. Kriegeskorte N, Formisano E, Sorger B, Goebel R (2007) Individual faces elicit distinct response patterns in human anterior temporal cortex. *Proc Natl Acad Sci USA* 104: 20600–20605.
16. Xu X, Yue X, Lescroart MD, Biederman I, Kim JG (2009) Adaptation in the fusiform face area (FFA): Image or person? *Vision Res* 49:2800–2807.
17. Gauthier I, Skudlarski P, Gore JC, Anderson AW (2000) Expertise for cars and birds recruits brain areas involved in face recognition. *Nat Neurosci* 3:191–197.
18. Freiwald WA, Tsao DY, Livingstone MS (2009) A face feature space in the macaque temporal lobe. *Nat Neurosci* 12:1187–1196.
19. Andrews TJ, Ewbank MP (2004) Distinct representations for facial identity and changeable aspects of faces in the human temporal lobe. *Neuroimage* 23:905–913.
20. Pourtois G, Schwartz S, Seghier ML, Lazeyras F, Vuilleumier P (2005) View-independent coding of face identity in frontal and temporal cortices is modulated by familiarity: An event-related fMRI study. *Neuroimage* 24:1214–1224.
21. Kanwisher N, Yovel G (2006) The fusiform face area: A cortical region specialized for the perception of faces. *Philos Trans R Soc Lond B Biol Sci* 361:2109–2128.
22. Rotshtein P, Henson RN, Treves A, Driver J, Dolan RJ (2005) Morphing Marilyn into Maggie dissociates physical and identity face representations in the brain. *Nat Neurosci* 8:107–113.
23. Op de Beeck HP, Brants M, Baeck A, Wagemans J (2010) Distributed subordinate specificity for bodies, faces, and buildings in human ventral visual cortex. *Neuroimage* 49:3414–3425.
24. Natu VS, et al. (2010) Dissociable neural patterns of facial identity across changes in viewpoint. *J Cogn Neurosci* 22:1570–1582.
25. Yue X, Cassidy BS, Devaney KJ, Holt DJ, Tootell RB (2011) Lower-level stimulus features strongly influence responses in the fusiform face area. *Cereb Cortex* 21:35–47.
26. Bar M, et al. (2006) Top-down facilitation of visual recognition. *Proc Natl Acad Sci USA* 103:449–454.
27. Cox D, Meyers E, Sinha P (2004) Contextually evoked object-specific responses in human visual cortex. *Science* 304:115–117.
28. Kriegeskorte N, Goebel R, Bandettini P (2006) Information-based functional brain mapping. *Proc Natl Acad Sci USA* 103:3863–3868.
29. Mourão-Miranda J, Friston KJ, Brammer M (2007) Dynamic discrimination analysis: A spatial-temporal SVM. *Neuroimage* 36:88–99.
30. 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.
31. Haxby JV, et al. (2001) Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293:2425–2430.
32. Sapountzis P, Schluppeck D, Bowtell R, Peirce JW (2010) A comparison of fMRI adaptation and multivariate pattern classification analysis in visual cortex. *Neuroimage* 49:1632–1640.
33. Rajimehr R, Young JC, Tootell RB (2009) An anterior temporal face patch in human cortex, predicted by macaque maps. *Proc Natl Acad Sci USA* 106:1995–2000.
34. Kriegeskorte N, Bandettini P (2007) Analyzing for information, not activation, to exploit high-resolution fMRI. *Neuroimage* 38:649–662.
35. Hadj-Bouziane F, Bell AH, Knusten TA, Ungerleider LG, Tootell RBH (2008) Perception of emotional expressions is independent of face selectivity in monkey inferior temporal cortex. *Proc Natl Acad Sci USA* 105:5591–5596.
36. Leveroni CL, et al. (2000) Neural systems underlying the recognition of familiar and newly learned faces. *J Neurosci* 20:878–886.
37. Sugiura M, et al. (2001) Activation reduction in anterior temporal cortices during repeated recognition of faces of personal acquaintances. *Neuroimage* 13:877–890.
38. Damasio H, Tranel D, Grabowski T, Adolphs R, Damasio A (2004) Neural systems behind word and concept retrieval. *Cognition* 92:179–229.
39. Simmons WK, Reddish M, Bellgowan PS, Martin A (2010) The selectivity and functional connectivity of the anterior temporal lobes. *Cereb Cortex* 20:813–825.
40. Leopold DA, Bondar IV, Giese MA (2006) Norm-based face encoding by single neurons in the monkey inferotemporal cortex. *Nature* 442:572–575.
41. Eifuku S, Nakata R, Sugimori M, Ono T, Tamura R (2010) Neural correlates of associative face memory in the anterior inferior temporal cortex of monkeys. *J Neurosci* 30:15085–15096.
42. Dehaene S, et al. (2010) How learning to read changes the cortical networks for vision and language. *Science* 330:1359–1364.
43. Hasson U, Levy I, Behrmann M, Hendler T, Malach R (2002) Eccentricity bias as an organizing principle for human high-order object areas. *Neuron* 34:479–490.
44. Puce A, Allison T, Asgari M, Gore JC, McCarthy G (1996) Differential sensitivity of human visual cortex to faces, letterstrings, and textures: A functional magnetic resonance imaging study. *J Neurosci* 16:5205–5215.
45. Kanwisher N, McDermott J, Chun MM (1997) The fusiform face area: A module in human extrastriate cortex specialized for face perception. *J Neurosci* 17:4302–4311.
46. Barton JJ, Press DZ, Keenan JP, O'Connor M (2002) Lesions of the fusiform face area impair perception of facial configuration in prosopagnosia. *Neurology* 58:71–78.
47. Avidan G, Behrmann M (2009) Functional MRI reveals compromised neural integrity of the face processing network in congenital prosopagnosia. *Curr Biol* 19:1146–1150.
48. Thomas C, et al. (2009) Reduced structural connectivity in ventral visual cortex in congenital prosopagnosia. *Nat Neurosci* 12:29–31.

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).

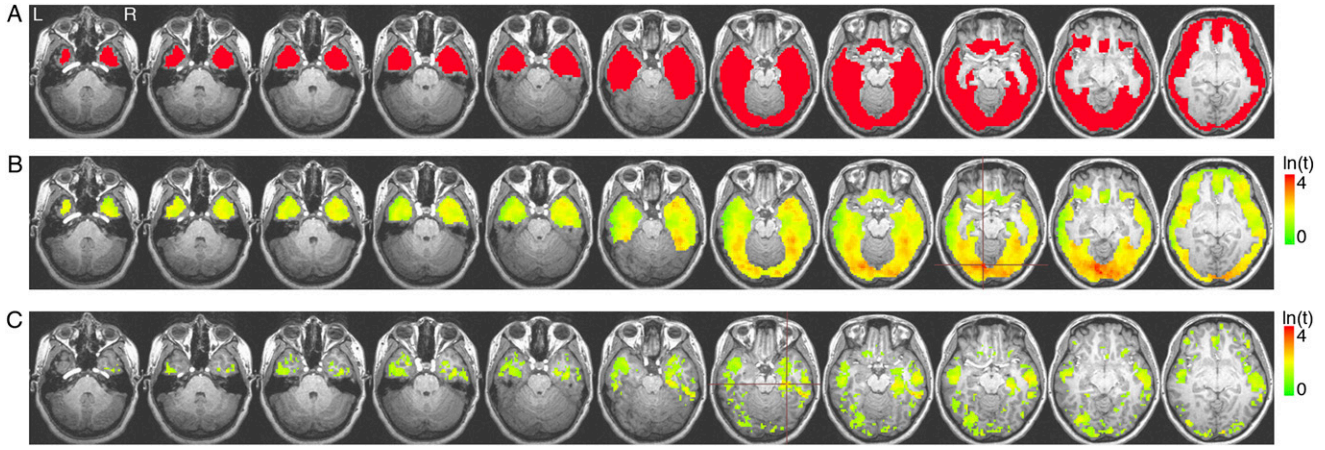

Fig. S2. (*A*) Group cortical mask and group information-based maps of category-level discrimination (faces vs. OFs; $q < 0.05$) derived through (*B*) multivariate searchlight and (*C*) its univariate analog. Effect size is scaled logarithmically. Crosshairs mark the discrimination peaks in each map (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.

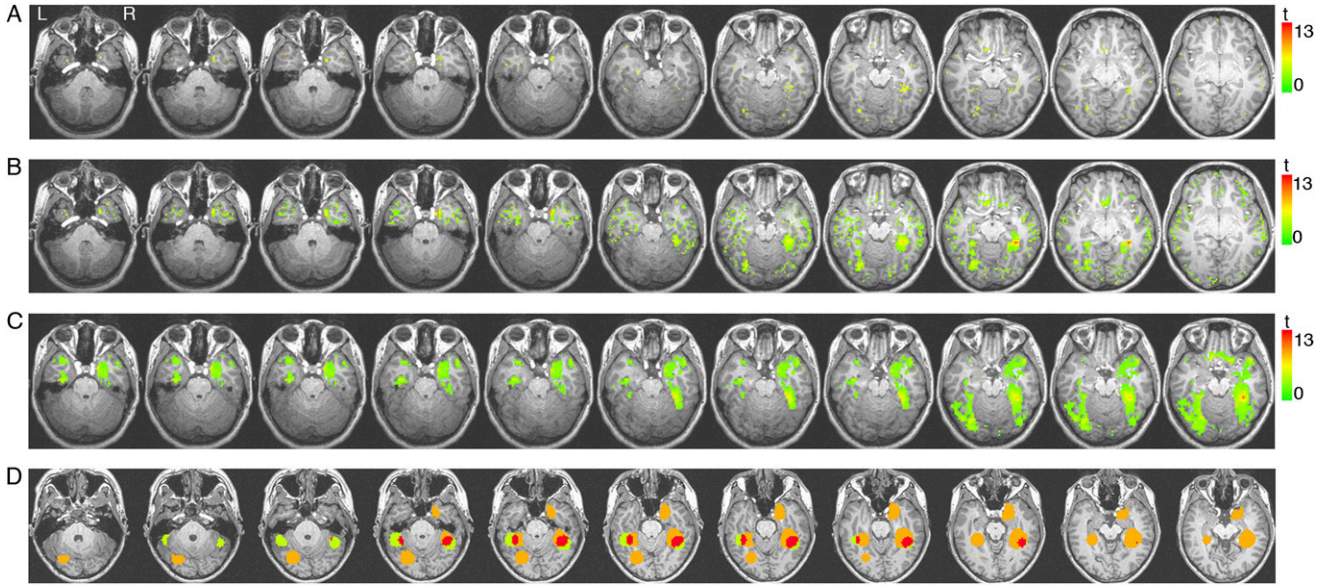

Fig. S3. (*A–C*) Group information-based maps of face individuation obtained with SL masks of different spatial radii (4, 5, and 6 voxels). Maps were thresholded at a liberal level ($q < 0.10$) for better comparison with each other. Larger masks yielded a larger number of significant voxels as expected 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.

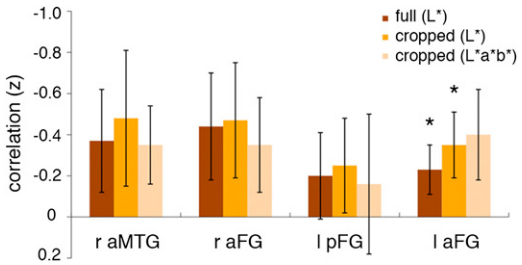

Fig. S4. Correlation of discrimination performance with low-level image distances across different SL ROIs (Fisher's *z* scores). Low-level distances were computed in three different ways: across the entire image using only luminance (full L*), across internal features using only luminance (cropped L*), 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).

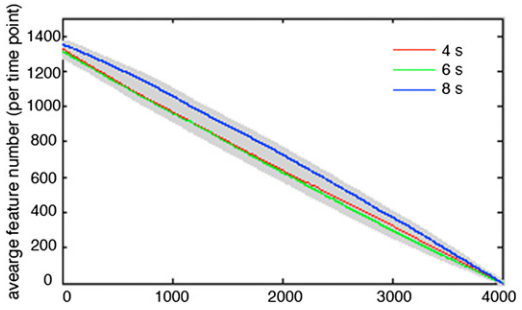

Fig. S5. The time course of feature elimination by time point for 4,000 features—top 1,000 features for each of four ROIs. Features tend to be eliminated at similar rates across different time points although we note a small advantage for features extracted at 8 s poststimulus onset (shaded areas correspond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