Behavioral/Systems/Cognitive

Flexible Coding for Categorical Decisions in the Human Brain

Sheng Li,[1] **Dirk Ostwald,**[1] **Martin Giese,**[2,3] **and Zoe Kourtzi**[1]

[1]School of Psychology, University of Birmingham, Birmingham B15 2TT, United Kingdom, [2]School of Psychology, University of Wales, Bangor LL57 2AS, United Kingdom, and [3]Hertie Institute for Clinical Brain Research, D-72076 Tübingen, Germany

Despite the importance of visual categorization for interpreting sensory experiences, little is known about the neural representations that mediate categorical decisions in the human brain. Here, we used psychophysics and pattern classification for the analysis of functional magnetic resonance imaging data to predict the features critical for categorical decisions from brain activity when observers categorized the same stimuli using different rules. Although a large network of cortical and subcortical areas contain information about visual categories, we show that only a subset of these areas shape their selectivity to reflect the behaviorally relevant features rather than simply physical similarity between stimuli. Specifically, temporal and parietal areas show selectivity for the perceived form and motion similarity, respectively. In contrast, frontal areas and the striatum represent the conjunction of spatiotemporal features critical for complex and adaptive categorization tasks and potentially modulate selectivity in temporal and parietal areas. These findings provide novel evidence for flexible neural coding in the human brain that translates sensory experiences to categorical decisions by shaping neural representations across a network of areas with dissociable functional roles in visual categorization.

Key words: visual categorization; fMRI; multivoxel pattern classification; learning; motion; form

Introduction

Our ability to group diverse sensory events into meaningful categories is a cognitive skill critical for adaptive behavior and survival in a dynamic, complex world (Miller and Cohen, 2001). Extensive behavioral work on visual categorization (Nosofsky, 1986; Schyns et al., 1998; Goldstone et al., 2001) has involved a strong debate between the traditional view suggesting that the brain solves this challenging task by representing physical similarity between low-level features and the diagnostic-feature view suggesting that the representations of visual feature are shaped by their relevance for categorical decisions. Converging evidence from neurophysiology suggests that the prefrontal cortex represents the behaviorally relevant stimulus features for categorization and modulates their processing in sensory areas (Miller, 2000; Duncan, 2001; Miller and D'Esposito, 2005). Isolating this flexible code for sensory information in the human brain is limited at the typical functional magnetic resonance imaging (fMRI) resolution that does not allow us to discern selectivity for features represented by overlapping neural populations. As a result, important questions about the neural representations that mediate

categorical decisions in the human brain remain open. Which human brain areas contain information about diagnostic features that define visual categories and the rules that guide the observers' categorical decisions? Is neural selectivity in human visual areas shaped by frontal circuits suggested to encode the behavioral relevance of visual features for categorical judgments?

We addressed these questions using combined psychophysical and fMRI measurements. We used complex movement sequences (i.e., artificial skeleton models defined by dots and animated with sinusoidal movements) (see Fig. 1A) (supplemental information, available at www.jneurosci.org as supplemental material) that differed parametrically in their spatial configuration (e.g., closer to prototype A vs B) and temporal profile (i.e., slow–fast, fast–slow). We trained participants to categorize these stimuli based on two different rules (see Fig. 1B): (1) a simple (one-dimensional) rule based on the spatial similarity of the stimuli, and (2) a complex (two-dimensional) rule (Ashby and Maddox, 2005) that entailed taking into account both the spatial configuration and the temporal profile of the stimuli.

To determine human brain regions that carry information about the diagnostic stimulus features for the different categorization tasks (Nosofsky, 1986; Schyns et al., 1998; Goldstone et al., 2001), we exploited the sensitivity of multivoxel pattern analysis (MVPA) methods (Cox and Savoy, 2003; Haynes and Rees, 2006; Norman et al., 2006) that take advantage of information across spatial patterns in the brain. We reasoned that signals from brain areas encoding behaviorally relevant information would be decoded more reliably when we classify brain responses for stimulus categories based on the categorization rule used by the observers rather than a rule that does not match the perceived stimulus categories. Consistent with this prediction, we demonstrate that

the temporal and parietal areas encode the perceived form and motion similarity, respectively. In contrast, frontal areas and the striatum represent task-relevant conjunctions of spatiotemporal features critical for complex and adaptive categorization tasks and potentially modulate selectivity in temporal and parietal areas. These findings provide novel evidence that flexible coding is implemented in the human brain by shaping neural representations in a network of areas with dissociable roles in visual categorization.

Materials and Methods
Observers
Six students from the University of Birmingham participated in the experiments. All observers had normal or corrected-to-normal vision and gave written informed consent. The study was approved by the local ethics committee.

Stimuli
Articulated movement stimuli were generated by linear interpolation and time warping between prototypical trajectories (Giese and Lappe, 2002; Giese and Poggio, 2003; Jastorff et al., 2006) (for more details on the generation and properties of the stimuli, see supplemental information, available at www.jneurosci.org as supplemental material). We generated movies (21 frames/1.3 s) of artificial point-light skeleton models (9.4 × 6.1° of visual angle) that were animated by sinusoidal movements. We tested two sets of prototypical stimuli (A and B, C and D) that differed in their kinematics, i.e., their segment lengths and the offsets of their joint angles. Half of the participants were tested on each stimulus set. To create a stimulus space that could be separated into different categories based on different stimulus dimensions, we manipulated the spatial similarity of the stimuli by spatial morphing and the temporal motion profile by time warping. In particular, by morphing between pairs of stimuli, we generated stimuli that differed in their spatial configuration: closer to A versus B or closer to C versus D. By time warping these morphs, we generated stimuli that differed in their temporal profile: slow–fast (i.e., the stimulus movement speed changed from slow to fast) and fast–slow (i.e., the stimulus movement speed changed from fast to slow).

Design and procedure
All observers participated in four fMRI sessions (two mapping and two MVPA sessions). Each of these sessions was preceded by psychophysical training with feedback outside the scanner, and the participants' performance was matched across tasks and scanning sessions.

Psychophysical training. Participants were trained to categorize the stimuli based on two different rules: (1) a simple rule based on the similarity of the stimuli to the kinematics of the prototype movements (spatial similarity rule), and (2) a complex rule based on similarity in two stimulus dimensions: kinematics and temporal profile. Based on the simple categorization rule, category 1 comprised stimuli closer to A, whereas category 2 comprised stimuli closer to B for this pair of prototypes. Based on the complex categorization rule, category 1 comprised stimuli closer to A slow–fast and closer to B fast–slow, whereas category 2 comprised stimuli closer to A fast–slow and closer to B slow–fast (Fig. 1*B*). Similar category groupings were used for prototypes C and D.

Training comprised several sessions consisting of multiple training runs. To control for stimulus-specific training effects and ensure generalization of the learnt categories, we partitioned the stimulus spaces into two stimulus sets: stimuli for training and stimuli for scanning (i.e., one-quarter of the stimuli presented during scanning were not presented during training). At the beginning of the session, participants were instructed which categorization rule to use in each session and were trained on this rule based on a self-paced procedure with audio error feedback. Ninety-six stimuli randomly selected from the training stimulus set were presented in each training run. Stimuli were presented for 1.3 s, and the participants were instructed to indicate in which category the stimulus belonged by pressing one of two keys. Observers were trained until they reached 85% correct responses in two consecutive training runs and after a test run. The training runs were preceded and followed by test runs
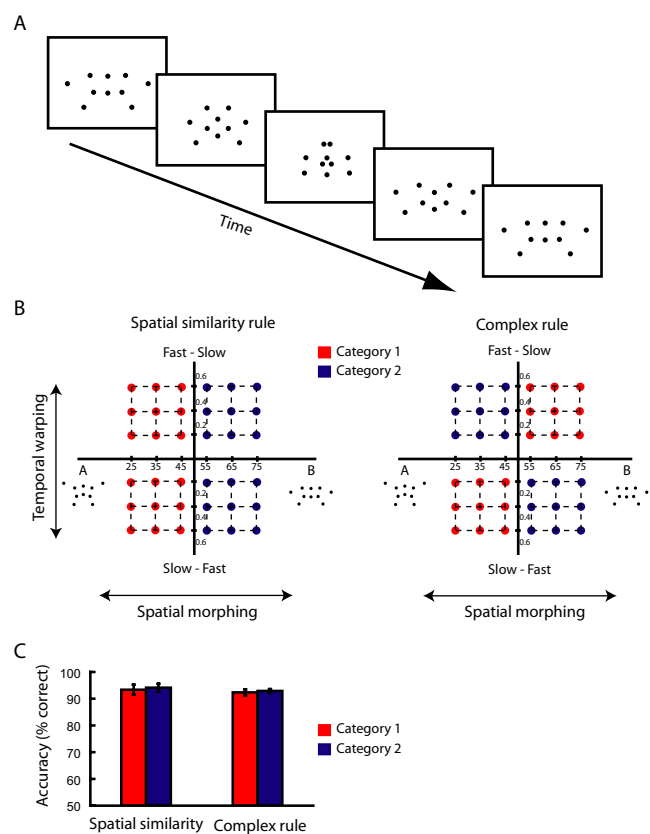

Figure 1. Stimulus space and behavioral responses. *A*, Stimuli: five sample frames of a prototypical stimulus. *B*, Stimulus space and categorization tasks. Stimuli were generated by applying spatial morphing (steps of percentage stimulus B) between prototypical trajectories (e.g., A–B) and temporal warping (steps of time warping constant). Stimuli were assigned to one of four groups: A fast–slow (AFS), A slow–fast (ASF), B fast–slow (BFS), and B slow–fast (BSF). For the simple categorization task (left), the stimuli were categorized according to their spatial similarity: category 1 (red dots) consisted of AFS, ASF; category 2 (blue dots) consisted of BFS, BSF. For the complex task (right), the stimuli were categorized based on their spatial and temporal similarity: category 1 (red dots) consisted of ASF, BFS; category 2 (blue dots) consisted of AFS, BSF. *C*, Behavioral data. Average response accuracy (percentage correct responses) across all blocks included in the multivoxel pattern analysis and participants.

(without feedback) that used the same protocol as the categorization task scans to assess the overall training and generalization effects. This training ensured that the performance of the participants in the scanner was similar across categories and tasks. Response accuracy and time did not differ significantly between categorization rules and stimulus categories during fMRI scanning. No significant (repeated-measures ANOVA) effects of stimulus category ($F_{(1,5)} = 0.14$; $p = 0.73$) or categorization task ($F_{(1,5)} = 4.3$; $p = 0.09$) were observed for response accuracy. Similarly, for response times, no significant effects of categorization rule ($F_{(1,5)} = 0.024$; $p = 0.88$) or stimulus category ($F_{(1,5)} = 0.505$; $p = 0.51$) were observed.

fMRI measurements
All observers participated in four scanning sessions: two for mapping regions of interest and two MVPA sessions (one for the simple and one for the complex categorization task).

Region of interest mapping sessions. We identified areas involved in visual categorization, retinotopic areas, and areas involved in motion [human middle temporal area (hMT+/V5); kinetic occipital area (V3b/KO)] and form [lateral occipital complex (LOC)] processing based on standard mapping procedures (supplemental information, available at www.jneurosci.org as supplemental material). We identified areas involved in the categorization of movement sequences by comparing fMRI responses when participants performed categorization versus spatial discrimination tasks on the same stimuli (rotated and not rotated). In the

categorization task, participants judged whether the stimuli belonged to category 1 or 2 or neither of the two categories based on the spatial similarity or the complex rule (Fig. 1 B). In the spatial discrimination task, participants judged whether the stimuli were rotated to the left or to the right or not rotated. Stimuli were presented in epochs (blocked design) of 32 s. Each block started with written instructions (2 s) about the task to be performed during the block (categorization or spatial discrimination task) and continued with 12 experimental trials of 2.5 s (1.3 s stimulus presentation, 1.2 s fixation, and response interval). Each experimental run comprised four categorization and four spatial discrimination task blocks in counterbalanced order and interleaved fixation blocks (20 s) after every four blocks. We collected data from four to eight experimental runs for each categorization task (simple and complex).

MVPA sessions. Participants were scanned on eight runs in each session. Each run started with a fixation period of 20 s and continued with 16 blocks of 10 trials (1.3 s stimulus presentation, 0.7 s fixation, and response interval). All stimuli presented during one block were drawn randomly without replacement from one quadrant of the stimulus space (Fig. 1 B), except for two stimuli, which were presented randomly during the block and were chosen from a different category than the rest of the stimuli. The four different block types relating to the four stimulus space quadrants (A fast–slow, B fast–slow, A slow–fast, and B slow–fast) were presented sequentially in four sets of four blocks (each set was followed by a fixation period of 20 s), and their order was counterbalanced across runs. This ensured that the participants could not predict the order of the stimuli within or across blocks and had to pay attention to each individual stimulus to achieve high performance at the categorization task. Participants performed the simple or the complex categorization task on the same stimuli but in different scanning sessions.

fMRI data acquisition

The experiments were conducted at the Birmingham University Imaging Centre (3T Achieva scanner; Philips, Eindhoven, The Netherlands). Echo planar imaging (EPI) and T1-weighted anatomical ($1 \times 1 \times 1$ mm) data were collected with an eight-channel SENSE head coil. EPI data (gradient echo-pulse sequences) were acquired from 33 slices (whole brain coverage; repetition time, 2000 ms; echo time, 35 ms; flip angle, 80°; $2.5 \times 2.5 \times 3$ mm resolution).

fMRI data analysis

fMRI data were processed using Brain Voyager QX (Brain Innovations, Maastricht, The Netherlands). Preprocessing of functional data included slice-scan time correction, head movement correction, temporal high-pass filtering (three cycles), and removal of linear trends. Spatial smoothing (Gaussian filter; full-width at half maximum, 6 mm) was performed only for mapping cortical areas involved in categorization (group analysis) but not for data used for the multivoxel pattern classification analysis. The functional images were aligned to anatomical data under careful visual inspection, and the complete data were transformed into Talairach space. In addition, anatomical data were used for three-dimensional cortex reconstruction, inflation, and flattening.

Mapping regions of interest. We identified categorization-related regions that showed significantly [$p < 0.05$, Bonferroni corrected; $q < 0.001$, false discovery rate (FDR)] higher activations for the simple or complex categorization task than the spatial discrimination task across participants.

MVPA. We selected voxels (gray matter only) that were activated significantly stronger during the categorization task than fixation baseline ($p < 0.05$, uncorrected) in both MVPA sessions for each individual subject in regions defined based on independent data for each participant: categorization-related regions, retinotopic areas, LOC, hMT+/V5, and V3b/KO. As in previous MVPA studies (Kamitani and Tong, 2005), we ranked the voxels in these regions (based on t values for stimulus vs fixation). Each voxel time course was z-score normalized for each experimental run separately. We generated the data patterns for MVPA by shifting the fMRI time series by 4 s to account for the hemodynamic delay. We then averaged the volumes for which the participants responded correctly in each block (Kamitani and Tong, 2005). This average signal per block comprised a training pattern resulting in a total of 16

training patterns per run. We used a linear support vector machine (SVM) and an eightfold cross-validation procedure (leave-one-run-out) for classification of these data (supplemental information, available at www.jneurosci.org as supplemental material). In particular, we trained the SVM on seven runs (112 training patterns per subject) and tested on the data from one run (16 patterns). We repeated this procedure eight times and calculated the prediction accuracy of the SVM (number of correct predictions/number of total predictions) for each cross-validation and then averaged across cross-validations to obtain an overall prediction accuracy value for the classification for each region of interest. We trained the SVM on 1–200 voxels in each area (if the number of voxels in the area was <200, we used the total number of voxels in the area). Prediction accuracy across brains areas improved with increasing numbers of voxels (supplemental Fig. 1, available at www.jneurosci.org as supplemental material) and reached an asymptote at a maximum of 100 voxels, allowing us to use the same number of voxels for comparison of pattern classification across areas.

Results

Identifying areas involved in visual categorization

We first identified human brain areas involved in the visual categorization of articulated movement stimuli. After training (i.e., when participants were above 85% correct for both categorization tasks) (Fig. 1C) (supplemental Fig. 2, available at www.jneurosci.org as supplemental material), participants were scanned (two mapping sessions) while performing a categorization task (simple or complex task) and a spatial discrimination task in which they judged the direction of rotation of the stimuli in the plane of the screen (Vogels et al., 2002). Consistent with previous imaging studies that have used a variety of categorization tasks and stimuli (for review, see Keri, 2003; Ashby and Maddox, 2005), we observed significantly stronger activations for the categorization than the spatial discrimination task in a network of temporal, parietal, and frontal cortical areas, as well as subcortical regions (thalamus and striatum) (Fig. 2) (supplemental Table 1, available at www.jneurosci.org as supplemental material). This procedure allowed us to identify independently in each individual participant areas involved in the categorization of articulated movements. We further identified regions in the occipital and temporal cortex known to be involved in the processing of basic visual features (retinotopic areas), global patterns and shapes (LOC), and motion (V3b/KO and hMT+/V5) based on standard procedures (supplemental information, available at www.jneurosci.org as supplemental material).

fMRI pattern classification: decoding simple versus complex categorization rules

We asked whether activation patterns in human brain areas involved in visual categorization reflect the rule that observers use when making categorical decisions (i.e., the behaviorally relevant stimulus dimensions for each categorization task) rather than simply the physical similarity of the stimuli. We used MVPA methods (Cox and Savoy, 2003; Haynes and Rees, 2006; Norman et al., 2006) to discern activation patterns across voxels for different perceptual categories. These advanced methods for the analysis of fMRI data provide a sensitive tool for investigating the neural representations of features in the human brain that are encoded at a higher spatial resolution than the typical resolution of fMRI and thus cannot be distinguished when considering single voxels in isolation. Pattern classification analyses have been previously used successfully for decoding neural selectivity for basic visual features (e.g., orientation) (Haynes and Rees, 2005; Kamitani and Tong, 2005), motion direction (Kamitani and Tong, 2006), and object categories (Haxby et al., 2001; Hanson et
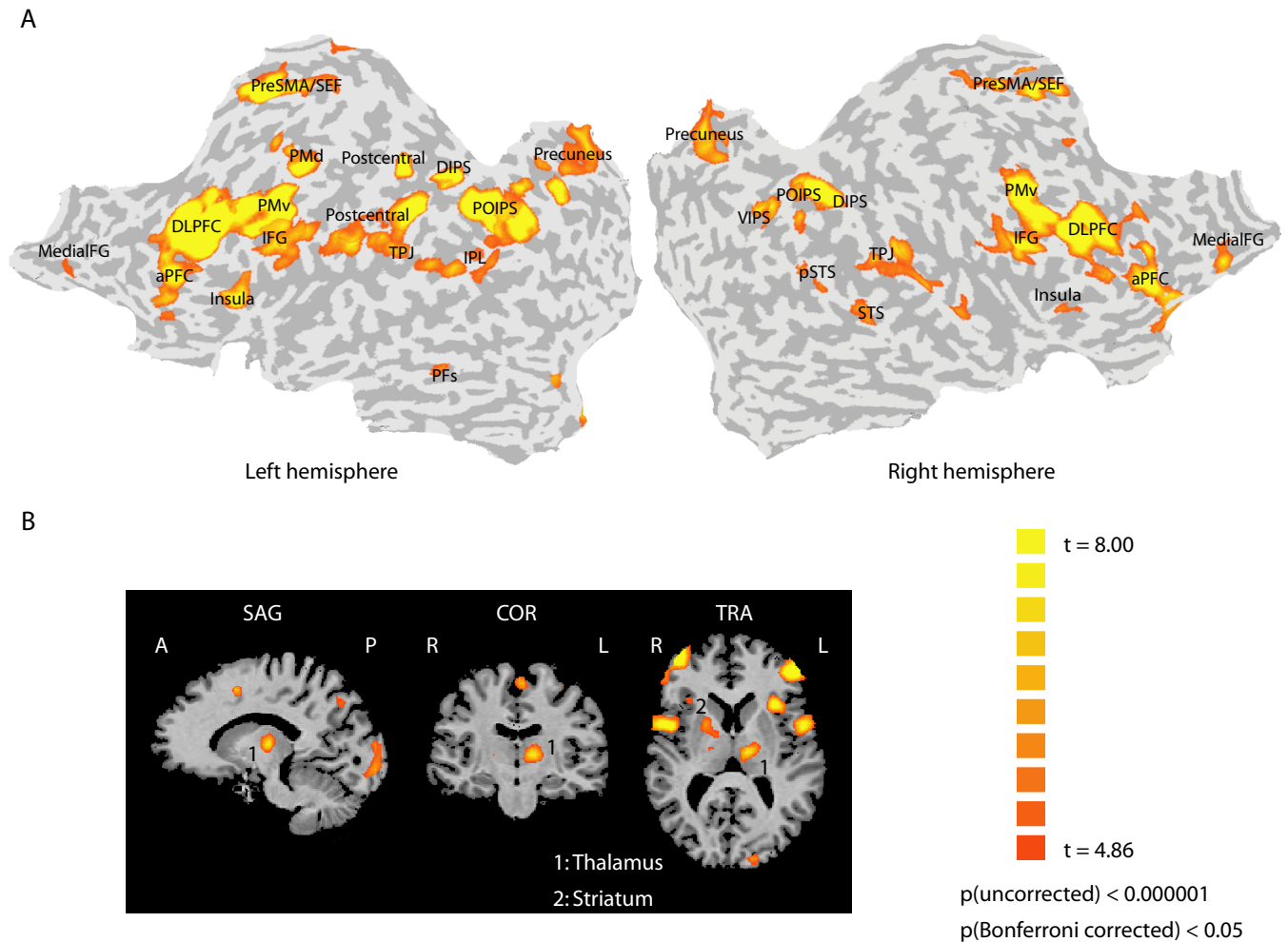
A

Left hemisphere

Right hemisphere

B

SAG   COR   TRA

1: Thalamus
2: Striatum

t = 8.00

t = 4.86

p(uncorrected) < 0.000001

p(Bonferroni corrected) < 0.05

Figure 2. Functional mapping of areas involved in categorization. Functional activation maps (group analysis) of areas that showed significantly ($p < 0.05$, Bonferroni corrected; $q < 0.001$, FDR) stronger activations for the categorization task (based on either the simple or the complex rule) than the spatial discrimination task (for Talairach coordinates, see supplemental Table 1, available at www.jneurosci.org as supplemental material). **A**, Cortical activations superimposed on flattened cortical surfaces of the right and left hemispheres. Sulci are shown in dark gray, and gyri are shown in light gray (for ROI definitions, see supplemental Table 1, availabe at www.jneurosci.org as supplemental material). **B**, Subcortical areas, primarily activated for the complex categorization task, are shown superimposed on the three-dimensional anatomy.

al., 2004; O'Toole et al., 2005; Williams et al., 2007) in the human brain. Here, we use multivoxel pattern classification for the first time to determine whether neural selectivity for visual features in areas involved in categorization changes depending on the behavioral relevance of these features for simple and complex categorization tasks.

Specifically, for each participant, we collected fMRI measurements when participants performed the simple or complex categorization task (two MVPA sessions for each participant). We tested whether we could predict the perceived stimulus categories (Fig. 1B, category 1 vs category 2) for each categorization task from fMRI data in all regions of interest (visual areas and categorization-related regions). Critically, we compared the prediction accuracy of two different classification schemes (spatial similarity MVPA rule and complex MVPA rule) in discerning stimulus categories from fMRI data for each categorization task. We reasoned that brain areas encoding the relevant dimensions for the categorization task would show higher prediction accuracy when the MVPA rule matched the rule used by observers for categorization and lower accuracy when the MVPA rule differed from the relevant categorization rule. That is, we tested for interactions in the prediction accuracy across categorization tasks for

different MVPA rules rather than differences in the absolute accuracy that may vary across brain areas depending on their responsiveness (supplemental Fig. 3, available at www.jneurosci. org as supplemental material).

In particular, we used a linear SVM classifier (supplemental information, available at www.jneurosci.org as supplemental material) and an eightfold cross-validation procedure (leave-one-run-out). For each region of interest, we averaged the volumes for which the participants responded correctly in each block (Kamitani and Tong, 2005) to generate training and test patterns for the pattern classification. We trained the pattern classifier on data from seven scanning runs (112 training patterns, 16 patterns per run) and tested on independent data from one run (16 test patterns) in the same session for eight times. After this procedure, we tested classification between activation patterns evoked by the two stimulus categories (Fig. 1B, category 1 vs category 2) for each categorization task. For the spatial similarity MVPA rule, training patterns for category 1 were evoked by stimuli closer to prototype A (A fast–slow and A slow–fast), whereas for category 2 training patterns were evoked by stimuli closer to prototype B (B fast–slow and B slow–fast). For the complex MVPA rule, training patterns for category 1 were evoked by
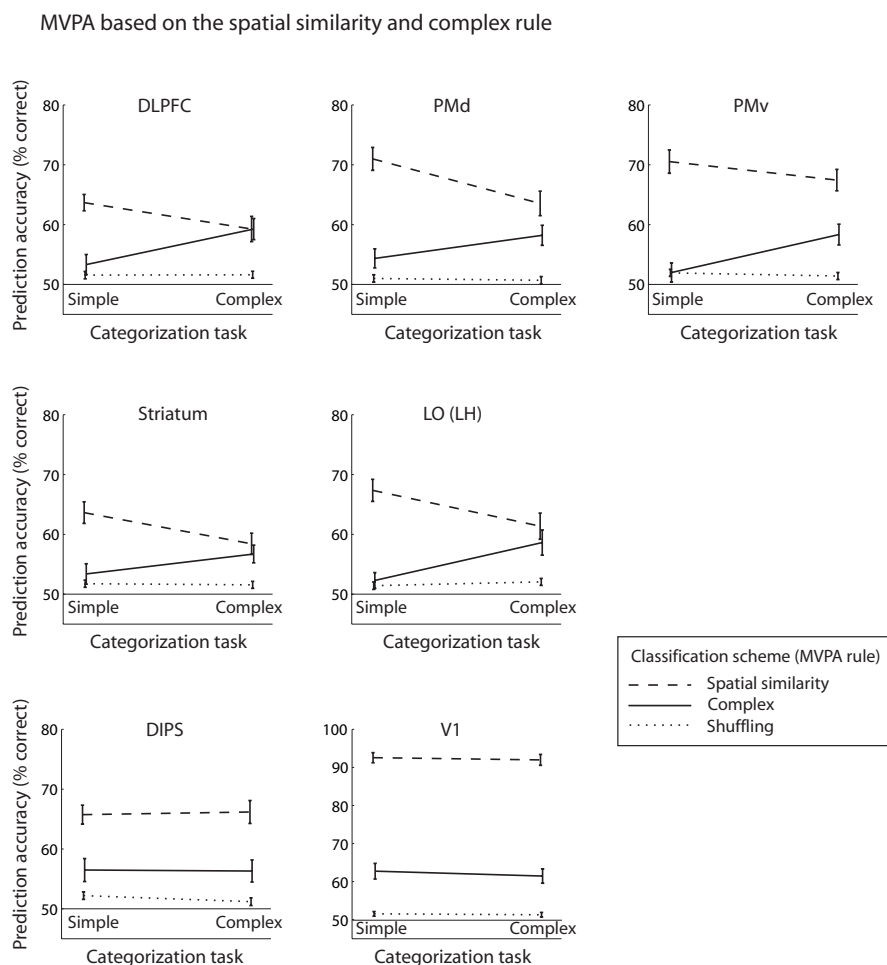
MVPA based on the spatial similarity and complex rule

Figure 3. Prediction accuracy for MVPA on the spatial similarity and complex rule. Prediction accuracy for the spatial similarity and complex classification schemes across categorization tasks (simple and complex tasks). Prediction accuracies for these MVPA rules are compared with accuracy for the shuffling rule (baseline prediction accuracy, dotted line). Interactions are shown for representative areas (for ROI definitions, see supplemental Table 1, available at www.jneurosci.org as supplemental material): DLPFC, $F_{(1,5)} = 7.33, p < 0.05$; PMd, $F_{(1,5)} = 5.83, p = 0.06$; PMv, $F_{(1,5)} = 6.11, p = 0.056$; left LO, $F_{(1,5)} = 8.55, p < 0.05$; and striatum, $F_{(1,5)} = 10.97, p < 0.05$. Error bars indicate the SEM across participants.

stimuli closer to A slow–fast and B fast–slow, whereas for category 2 training patterns were evoked by stimuli closer to A fast–slow and B slow–fast. Importantly, only one of these two schemes matched the rule used by the observers in each of the two categorization tasks. For the simple task, the spatial similarity MVPA rule matched the rule used by the observers, whereas the complex MVPA rule assigned the stimuli in different categories than the observers. In contrast, for the complex task, the complex MVPA rule matched the rule used by the observers for stimulus categorization based on both the spatial and temporal profile, whereas the spatial similarity MVPA rule performed the classification based only on the spatial similarity of the stimuli.

Results of this analysis showed significant interactions (supplemental Table 2, available at www.jneurosci.org as supplemental material) in prefrontal [dorsolateral prefrontal cortex (DLPFC)] and temporal [left posterior LOC, consistent with left lateralization for category knowledge in the temporal cortex (Seger et al., 2000; Grossman et al., 2002)] areas and the striatum, indicating neural representations that are selective for the behaviorally relevant stimulus dimensions. As shown in Figure 3, when participants categorized stimuli based on spatial similarity (simple categorization task), prediction accuracy was high for the

spatial similarity MVPA rule. In contrast, when we used the complex MVPA rule for classification of the same data (simple categorization task), prediction accuracy did not differ significantly (DLPFC, $t_{(5)} < 1$, $p = 0.59$; left LO, $t_{(5)} < 1$, $p = 0.85$; striatum, $t_{(5)} = 1.21$, $p = 0.28$) from baseline (i.e., prediction accuracy when we randomly assigned category labels to the activation patterns). Interestingly, when we used the spatial similarity MVPA rule for classification of the data obtained when participants performed the complex categorization task, prediction accuracy decreased but it did not reach baseline accuracy. This result suggests that spatial similarity could be predicted from activation patterns related to both categorization tasks (simple and complex tasks). This is consistent with the behavioral relevance of spatial similarity for categorical judgments in both tasks (i.e., spatial similarity was relevant for the complex categorization task) and indicates that this dimension was encoded when participants performed the complex categorization task. Analysis of the same data with conventional univariate methods showed that the visual categories perceived by the observers could not be distinguished reliably when considering the average signal across voxels (supplemental Fig. 4, Table 3, available at www.jneurosci.org as supplemental material), supporting the advantage of MVPA methods in decoding feature selectivity from multivariate brain patterns across voxels. Finally, no significant interactions in prediction accuracy across categorization tasks for different MVPA rules were observed in retinotopic and parietal areas (supplemental Table 2, available at www.jneurosci.org as supplemental material). Prediction accuracy in these areas was highest for the simple MVPA rule independent of the categorization task, indicating activation patterns that reflect differences in the physical stimulus properties (average Euclidean distance of stimulus trajectories between different categories, 1.04).

Figure 4 (for individual subject data, see supplemental Fig. 5, available at www.jneurosci.org as supplemental material) summarizes the data across all brain areas of interest by means of an accuracy interaction index. We computed this index by calculating for each categorization task the normalized difference in prediction accuracy for the two MVPA rules implemented for the analysis of the fMRI data. When computing this index, we normalized prediction accuracies for the two MVPA rules to accuracy for the shuffling rule (baseline prediction accuracy) that was computed by randomly assigning category labels to the activation patterns. This procedure ensured that the classification results could not be simply attributed to random statistical regularities in the data and/or the high efficiency of the classification algorithm used. An index higher than zero indicates higher prediction accuracy when the MVPA rule matched the rule used by the observers for categorization than when the MVPA rule did not
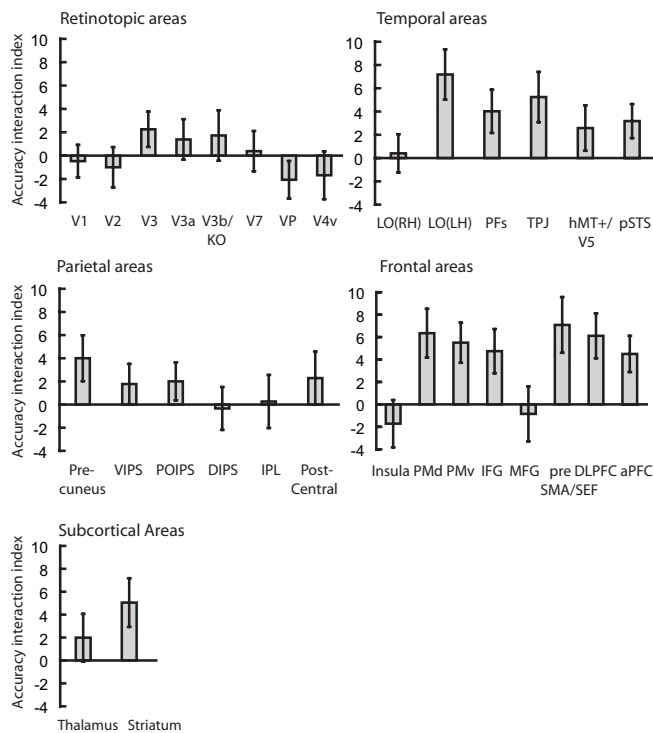
Figure 4. Accuracy interaction index for all regions of interest. The accuracy interaction index (AII) is plotted for all regions of interest (for ROI definitions, see supplemental Table 1, available at www.jneurosci.org as supplemental material), indicating the interaction between classification schemes and categorization tasks. The AII is computed based on the normalized difference in prediction accuracy obtained for two classification schemes: the MVPA rule used by the observers minus an MVPA rule not used by the observers for each categorization task:

$$ \text{AII} = \left(\frac{S_S - C_S}{S_S + C_S + R_S} + \frac{C_C - S_C}{S_C + C_C + R_C} \right) \times 100, $$

where S, C, and R refer to the accuracies obtained for the spatial similarity (S), complex (C), and shuffling (R) MVPA rules, and the subscript indices denote the respective categorization task (S, simple; C, complex). The mean AII is plotted across observers. Error bars indicate SEM across participants.

match the rule used by the observers. These differences in prediction accuracy between tasks were higher in frontal, temporal cortex, and striatum than in retinotopic and parietal areas, suggesting distinct functional roles for these areas in visual categorization.

Next, we tested whether we could decode the temporal profile dimension that was behaviorally relevant for the complex categorization task. We compared prediction accuracies for MVPA classification based on the spatial similarity rule and a temporal similarity rule (i.e., similarity of the stimuli in their temporal profile, slow–fast vs fast–slow). For the spatial similarity MVPA rule, training patterns for category 1 were evoked by stimuli closer to prototype A (A fast–slow and A slow–fast), whereas for category 2 training patterns were evoked by stimuli closer to prototype B (B fast–slow and B slow–fast). For the temporal similarity MVPA rule, training patterns for category 1 were evoked by stimuli closer to fast–slow (A and B), whereas for category 2 training patterns were evoked by stimuli closer to slow–fast.

Figure 5 shows significant interactions for MVPA prediction accuracies across tasks (simple and complex tasks) for frontal (DLPFC), lateral occipitotemporal (left posterior LOC) areas, and the striatum, providing additional evidence that activation patterns in these areas reflect the stimulus dimensions that are

behaviorally relevant for categorical decisions. The overall high prediction accuracy in prefrontal areas and the striatum for the temporal similarity MVPA rule may reflect involvement of working memory processes in matching labels (slow–fast vs fast–slow movement) to category members (Ashby and Maddox, 2005) or attention (Duncan and Owen, 2000) to the unique stimulus dimension (i.e., the temporal profile because spatial similarity was relevant for both categorization tasks). Interestingly, we observed increased prediction accuracy in inferior–parietal areas for the temporal similarity MVPA rule in the complex categorization task and interactions between tasks for the spatial and temporal similarity MVPA rules. In contrast, prediction accuracy in the LOC for the temporal similarity MVPA rule did not increase in the complex categorization task. These findings suggest involvement of parietal rather than temporal areas in the categorical representation of temporal movement properties, consistent with recent neurophysiological findings (Freedman and Assad, 2006). A control experiment (supplemental Fig. 6, available at www.jneurosci.org as supplemental material) in which participants performed the simple categorization task based on the similarity of the stimuli in their temporal profile (i.e., participants judged whether the stimuli moved fast–slow or slow–fast) corroborated these findings. We compared prediction accuracy for two orthogonal MVPA rules (i.e., spatial vs temporal similarity) when participants performed the simple categorization task based on one of two stimulus properties (i.e., participants categorized the stimuli based on similarity in their spatial configuration or temporal profile). Prediction accuracy in frontal, parietal areas and the striatum was higher for the temporal similarity MVPA rule when participants categorized stimuli based on their temporal than spatial similarity, suggesting that neural representations in these areas reflect the behavioral relevance of motion properties for categorical judgments.

Is it possible that MVPA prediction accuracies for the simple and complex categorization tasks relate to the participants' motor responses rather than the relevant dimensions for categorization? Although our experimental design does not allow us to separate fMRI signals for the stimulus from signals for the participants' responses, this hypothesis could not explain prediction accuracy higher than baseline when the classification scheme and the rule used by the participants for the categorization task did not match, for example, when predicting spatial similarity (Fig. 3) or temporal similarity (Fig. 5) from fMRI data for the complex categorization task. These findings suggest classification based on the relevant stimulus dimensions for the categorization task rather than the participants' motor response (supplemental Fig. 6, available at www.jneurosci.org as supplemental material). Furthermore, our findings could not be simply attributed to differences in eye movements, attention, or task difficulty. Our MVPA analysis compared prediction accuracy for the classification of two stimulus categories rather than fMRI signals across categorization tasks. Participants were trained with feedback to a high level of performance (above 85% correct) in both tasks, and their performance during scanning was similar between categories and tasks (Fig. 1C) (supplemental Fig. 2, available at www.jneurosci.org as supplemental material). Eye movement measurements did not differ significantly across sessions (supplemental Fig. 7, available at www.jneurosci.org as supplemental material), suggesting that fMRI findings could not be significantly confounded by differences in eye movements across sessions.
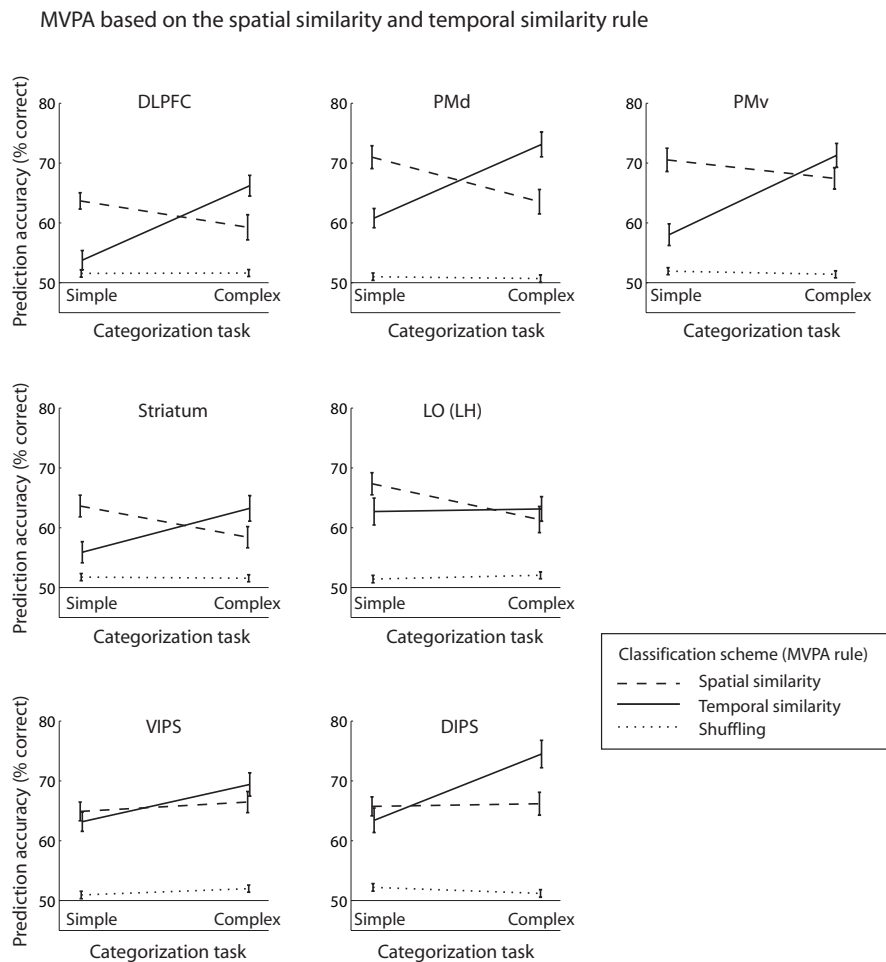
MVPA based on the spatial similarity and temporal similarity rule

Figure 5. Prediction accuracy for MVPA on the spatial and temporal similarity rule. Prediction accuracy for the spatial and temporal classification schemes across categorization tasks (simple and complex tasks) compared with accuracy for the shuffling rule. Interactions are shown for representative areas (for ROI definitions, see supplemental Table 1, available at www.jneurosci.org as supplemental material): DLPFC, $F_{(1,5)} = 18.82, p < 0.01$; PMd, $F_{(1,5)} = 16.93, p < 0.01$; PMv, $F_{(1,5)} = 23.59, p < 0.01$; left LO, $F_{(1,5)} = 10.58, p < 0.05$; striatum, $F_{(1,5)} = 10.96, p < 0.05$; IPS (ventral and dorsal), $F_{(1,5)} = 12.44, p = 0.017$.

fMRI pattern classification: decoding categorical membership

We further investigated whether human brain areas involved in visual categorization contain information about categorical membership by testing for classification performance within categories and across the categorical boundary. We reasoned that prediction accuracy would be lower for the classification of activation patterns related to stimuli that are members of the same category than stimuli that belong to different categories. We performed this pattern classification analysis on fMRI data obtained when observers categorized stimuli using the spatial similarity rule, allowing us to quantify the physical (one-dimensional) similarity and identify the categorical boundary between stimuli. This analysis showed that primarily in frontal areas [DLPFC, $F_{(1,2)} = 283.97, p < 0.01$; dorsal premotor area (PMd), $F_{(1,2)} = 125.12, p < 0.01$; and ventral premotor area (PMv), $F_{(1,2)} = 361.64, p < 0.01$] perceptual differences between stimuli from different categories could be predicted with significantly higher accuracy than differences between stimuli from the same category (Fig. 6). In contrast, no significant differences were observed for small stimulus distances close to the boundary and small distances within category in LO left hemisphere ($F_{(1,2)} = 2.08; p = 0.28$), dorsal intraparietal sulcus (DIPS) ($F_{(1,2)} < 1; p = 0.60$), the striatum ($F_{(1,2)} = 2.81; p = 0.23$), or primary visual cortex (V1)

($F_{(1,2)} = 6.29; p = 0.13$). Furthermore, we compared mean prediction accuracy between small and large distances within category. Significant differences were observed in V1 ($F_{(1,2)} = 142.36; p < 0.01$), consistent with physical differences between the stimuli. However, no significant differences were observed in DLPFC ($F_{(1,2)} = 4.07; p = 0.18$), PMd ($F_{(1,2)} = 1.28; p = 0.37$), PMv ($F_{(1,2)} = 2.26; p = 0.27$), LO left hemisphere ($F_{(1,2)} < 1; p = 0.50$), DIPS ($F_{(1,2)} < 1; p = 0.67$), or the striatum ($F_{(1,2)} = 2.41; p = 0.26$), suggesting generalization of responses in these areas to stimuli from the same perceptual category.

These results are consistent with previous studies showing fMRI-selective adaptation in human frontal cortex for perceptual categories (Jiang et al., 2007) (i.e., higher fMRI responses for stimuli between than within categories) and classification accuracy in discriminating between different objects that decreased as visual similarity between objects increased (O'Toole et al., 2005). Because our study was not designed to directly test for categorical membership, we only tested a small number of morphing steps around the boundary and did not evaluate perceptual discriminability between stimuli from the same versus different categories. Despite this, the analysis provides supporting evidence for the role of human frontal cortex in representing perceptual rather than physical differences between members of different categories while generalizing among members of the same category.

Discussion

Our study provides novel evidence that the human brain uses a flexible code across stages of processing to translate sensory experiences to categorical decisions. Our findings advance our understanding of the neural representations that mediate categorical decisions in several respects. First, previous behavioral work on categorization has involved a long-standing debate between two views on the representation of visual categories: the traditional view that visual stimuli are represented based on physical similarity and the diagnostic-feature view that visual representations are shaped to reflect behavioral relevance for categorical decisions. The role of learning in shaping visual categories within these two frameworks is radically different: either tuning for low-level differences (traditional view) or flexible tuning for diagnostic features that are critical for categorical judgments (diagnostic features view). Dissociating the predictions of these two approaches regarding the neural mechanisms that mediate categorical decisions is not trivial, and, as a result, the neurobiological basis of diagnostic features in the human brain has remained essentially unknown. The conjunction of our experimental design and analysis methods allowed us to dissociate these approaches and provide novel evidence that flexible coding is implemented in the human brain. In particular, unlike previous imaging studies that typically use familiar object categories, we were able to dissociate physical (i.e.,
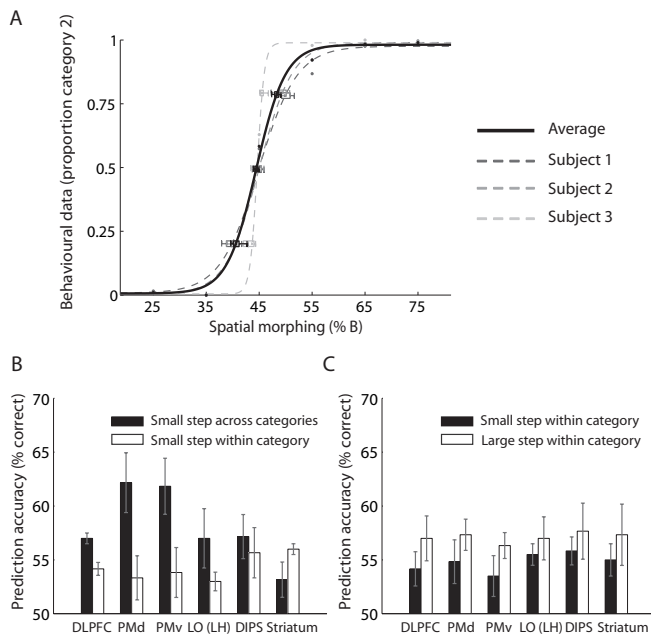
Figure 6. Prediction accuracy within and across perceptual categories. Behavioral performance and prediction accuracy are shown for three observers that performed the simple categorization task based on spatial similarity for the A–B stimulus set, allowing us to evaluate perceptual differences between stimuli that were matched for their distances in the physical space. Note that data from the C–D stimulus set was not used in the analysis as the morphing steps used were not equidistant (20, 30, 40, 60, 70, and 80%); morphing steps in this set were selected to match performance between stimulus sets (A–B, C–D). *A*, Behavioral performance for categorization based on spatial similarity (proportion of trials in which observers assigned the stimuli to category 2) across spatial morphing steps is shown for each observer and averaged across observers. Performance was fitted with a logistic function to estimate the 50% categorization threshold for each observer. The average threshold across observers was 44.82 (thresholds for individual observers were very similar: s1, 44.87; s2, 44.97; s3, 44.62), placing the categorization boundary approximately at the morphing step of 45. *B*, We compared the mean prediction accuracy for small stimulus distances close to the boundary (morphing steps, 35 vs 45 and 45 vs 55) and small distances within category (morphing steps, 25 vs 35 and 55 vs 65). The Euclidean distances were very similar across these morphing steps (25–35 = 0.33; 35–45 = 0.34; 45–55 = 0.32; 55–65 = 0.35). A significant interaction between region of interest and classification ($F_{(29,58)} = 2.21$; $p < 0.01$) was observed (for ROI definitions, see supplemental Table 1, available at www.jneurosci.org as supplemental material). *C*, We compared mean prediction accuracy between small distances within category (morphing steps, 55 vs 65 and 65 vs 75; Euclidean distances, 0.35 and 0.36) and a large distance within category (morphing steps, 55 vs 75; Euclidean distance, 0.51). A significant interaction between region of interest and classification ($F_{(29,58)} = 1.78$; $p < 0.05$) was observed.

the Euclidean distance between motion trajectories) and perceptual (categorical) similarity using novel stimulus patterns defined in a two-dimensional space. Furthermore, unlike simple categorization rules that separate physically different stimuli in different perceptual categories, the complex categorization rule used in our study necessitated that physically disparate stimuli be grouped into the same perceptual category, allowing us to distinguish between physical and perceptual stimulus similarity. Finally, using advanced analysis methods (pattern classification), we provide the first insights into the changes of neural representation for visual categories as a result of learning and, critically, provide strong evidence for the representation of diagnostic stimulus features in the human brain.

Second, previous neuroimaging studies have identified a large network of cortical and subcortical areas involved in visual categorization (Keri, 2003; Ashby and Maddox, 2005) and have revealed a distributed pattern of activations for object categories in the temporal cortex (Haxby et al., 2001; Hanson et al., 2004;

O'Toole et al., 2005; Williams et al., 2007). Our study extends significantly beyond the localization of areas involved in categorization; it characterizes their neural representations and discerns their functional roles. In particular, we show that, although activation patterns in the network of human brain areas involved in categorization contain information about visual categories, only in prefrontal, lateral occipitotemporal, inferior–parietal areas, and the striatum do these distributed representations mediate categorical decisions. That is, neural representations in these areas are shaped by the behavioral relevance of sensory features and previous experience to reflect the perceptual (categorical) rather than the physical similarity between stimuli. These flexible representations reflect selectivity for diagnostic stimulus features that determine the rule that observers use for categorization (Nosofsky, 1986; Schyns et al., 1998; Goldstone et al., 2001; Sigala et al., 2002; Palmeri and Gauthier, 2004; Smith et al., 2004) rather than features fixed by low-level processes (i.e., similarity in the physical input).

Third, our findings reveal distinct neural representations and functional roles for the different human brain areas involved in visual categorization. In particular, activation patterns in LOC and intraparietal sulcus (IPS) reflect selectivity for visual features (form and motion, respectively) that is modulated by the perceptual similarity of the stimuli in their spatial or temporal properties. In contrast activation patterns in retinotopic areas reflect physical similarity in the sensory input. These findings suggest that categorization based on visual form (spatial configuration) shapes processing in the LOC, whereas categorization based on visual motion (temporal profile of movement) shapes activations in the IPS. Furthermore, these findings are highly relevant to the contested role of temporal cortex in categorization. Previous neurophysiological studies have proposed that the temporal cortex represents primarily the visual similarity between stimuli and their identity (Op de Beeck et al., 2001; Thomas et al., 2001; Freedman et al., 2003), although other studies show that it represents diagnostic dimensions for categorization (Sigala and Logothetis, 2002; Mirabella et al., 2007) and is modulated by task demands (Koida and Komatsu, 2007) and experience (Miyashita and Chang, 1988; Logothetis et al., 1995; Gauthier et al., 1997; Booth and Rolls, 1998; Kobatake et al., 1998; Baker et al., 2002; Kourtzi et al., 2005; Op de Beeck et al., 2006). Our results in the human are consistent with findings in support of flexible representations of diagnostic features in the monkey inferotemporal cortex (Sigala and Logothetis, 2002; Mirabella et al., 2007). However, our experimental design allows us to draw a firm distinction between physical and categorical similarity, extending our findings beyond these previous studies. In particular, we show that the neural representations of form-related features in temporal cortex are shaped by their relevance for categorization and reflect the perceptual rather than physical similarity between visual stimuli.

Fourth, these feature-based representations in temporal and parietal areas contrast with responses in the frontal cortex and the striatum, the only brain areas in which activation patterns were shaped by categorization based on complex rules that depend on the conjunction of spatiotemporal features. These findings are consistent with the role of these areas in adaptive behavior and abstract rule learning (Duncan and Owen, 2000; Miller and Cohen, 2001; Moore et al., 2006). In particular, there is accumulating evidence proposing that the prefrontal cortex plays a fundamental role in adaptive neural coding (Duncan, 2001). The prefrontal cortex has been suggested to resolve the competition for processing of the plethora of input features in complex scenes

by guiding visual attention to behaviorally relevant information (Desimone and Duncan, 1995; Desimone, 1998; Reynolds and Chelazzi, 2004; Maunsell and Treue, 2006), representing the task-relevant features and shaping selectivity in sensory areas according to task demands (for review, see Miller, 2000; Duncan, 2001; Miller and D'Esposito, 2005). Our findings are consistent with the role of prefrontal cortex in controlling feature-based attention and selectivity in the visual cortex, providing novel insights in understanding the mechanisms for adaptive neural coding in the human brain. Specifically, previous imaging studies have shown that selective attention to visual features and objects enhances responsiveness in areas involved in their processing (Rees et al., 1997; Kastner et al., 1998; O'Craven et al., 1999; Kanwisher and Wojciulik, 2000; Corbetta and Shulman, 2002; Saenz et al., 2002). Recent fMRI studies using multivariate approaches (Kamitani and Tong, 2005, 2006; Serences and Boynton, 2007) have extended this work showing that attention shapes selectivity for visual features encoded at a finer scale than the typical fMRI resolution within visual areas specialized for their analysis (e.g., orientation in V1 and motion direction in hMT+/V5). However, the challenge of studying adaptive coding in the prefrontal cortex is that information about multiple visual features (e.g., orientation, motion, color, etc.) is potentially represented in a distributed manner across overlapping neural populations (Duncan, 2001). Our study provides the first evidence that, using multivariate methods for the analysis of neuroimaging data, we can decode selectivity in the human prefrontal cortex for multiple visual features (i.e., temporal and spatial movement profile) that is shaped by task context and feature-based attention (i.e., whether the observers attend and categorize the stimuli based on one or both stimulus features). Furthermore, the enhanced selectivity observed for form similarity in temporal areas and for motion similarity in parietal areas is consistent with the proposal that prefrontal cortex shapes neural representations in sensory areas specialized for the analysis of these features in a top-down manner.

Thus, our findings provide insights important for understanding the interactions within the neural circuit that mediates categorical decisions in the human brain. Although fMRI measurements are limited in temporal resolution and do not allow us to discern bottom-up from top-down processes, our findings are consistent with neurophysiological evidence for recurrent processes in visual categorization. It is possible that information about spatial and temporal stimulus properties in temporal and parietal cortex is combined with motor responses to form associations and representations of meaningful categories in the striatum and frontal cortex (Toni et al., 2001; Muhammad et al., 2006). In turn, these category formation and decision processes modulate selectivity for perceptual categories along the behaviorally relevant stimulus dimensions in a top-down manner (Freedman et al., 2003; Smith et al., 2004; Rotshtein et al., 2005; Mirabella et al., 2007), resulting in enhanced selectivity for form similarity in temporal areas but temporal similarity in parietal areas.

Finally, our study provides the first insights into the neural mechanisms that mediate our ability to assign meaning to articulated movements and perceive categories of actions. Recent findings provide evidence for categorical representations during the planning of action sequences in the monkey lateral prefrontal cortex (Shima et al., 2007). Our work demonstrates a flexible neural code for the visual categorization of movement sequences in the human brain that is fundamental for understanding the neural mechanisms that translate our everyday sensory experiences into perceptual decisions and social interactions.

References

Ashby FG, Maddox WT (2005) Human category learning. Annu Rev Psychol 56:149–178.

Baker CI, Behrmann M, Olson CR (2002) Impact of learning on representation of parts and wholes in monkey inferotemporal cortex. Nat Neurosci 5:1210–1216.

Booth MCA, Rolls ET (1998) View-invariant representations of familiar objects by neurons in the inferior temporal visual cortex. Cereb Cortex 8:510–523.

Corbetta M, Shulman GL (2002) Control of goal-directed and stimulus-driven attention in the brain. Nat Rev Neurosci 3:201–215.

Cox DD, Savoy RL (2003) Functional magnetic resonance imaging (fMRI) "brain reading": detecting and classifying distributed patterns of fMRI activity in human visual cortex. NeuroImage 19:261–270.

Desimone R (1998) Visual attention mediated by biased competition in extrastriate visual cortex. Philos Trans R Soc Lond B Biol Sci 353:1245–1255.

Desimone R, Duncan J (1995) Neural mechanisms of selective visual attention. Annu Rev Neurosci 18:193–222.

Duncan J (2001) An adaptive coding model of neural function in prefrontal cortex. Nat Rev Neurosci 2:820–829.

Duncan J, Owen AM (2000) Common regions of the human frontal lobe recruited by diverse cognitive demands. Trends Neurosci 23:475–483.

Freedman DJ, Assad JA (2006) Experience-dependent representation of visual categories in parietal cortex. Nature 443:85–88.

Freedman DJ, Riesenhuber M, Poggio T, Miller EK (2003) A comparison of primate prefrontal and inferior temporal cortices during visual categorization. J Neurosci 23:5235–5246.

Gauthier I, Anderson AW, Tarr MJ, Skudlarski P, Gore JC (1997) Levels of categorization in visual recognition studied using functional magnetic resonance imaging. Curr Biol 7:645–651.

Giese MA, Lappe M (2002) Measurement of generalization fields for the recognition of biological motion. Vision Res 42:1847–1858.

Giese MA, Poggio T (2003) Neural mechanisms for the recognition of biological movements. Nat Rev Neurosci 4:179–192.

Goldstone RL, Lippa Y, Shiffrin RM (2001) Altering object representations through category learning. Cognition 78:27–43.

Grossman M, Koenig P, DeVita C, Glosser G, Alsop D, Detre J, Gee J (2002) The neural basis for category-specific knowledge: an fMRI study. NeuroImage 15:936–948.

Hanson SJ, Matsuka T, Haxby JV (2004) Combinatorial codes in ventral temporal lobe for object recognition: Haxby (2001) revisited: is there a "face" area? NeuroImage 23:156–166.

Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P (2001) Distributed and overlapping representations of faces and objects in ventral temporal cortex. Science 293:2425–2430.

Haynes JD, Rees G (2005) Predicting the orientation of invisible stimuli from activity in human primary visual cortex. Nat Neurosci 8:686–691.

Haynes JD, Rees G (2006) Decoding mental states from brain activity in humans. Nat Rev Neurosci 7:523–534.

Jastorff J, Kourtzi Z, Giese MA (2006) Learning to discriminate complex movements: biological versus artificial trajectories. J Vis 6:791–804.

Jiang X, Bradley E, Rini RA, Zeffiro T, Vanmeter J, Riesenhuber M (2007) Categorization training results in shape- and category-selective human neural plasticity. Neuron 53:891–903.

Kamitani Y, Tong F (2005) Decoding the visual and subjective contents of the human brain. Nat Neurosci 8:679–685.

Kamitani Y, Tong F (2006) Decoding seen and attended motion directions from activity in the human visual cortex. Curr Biol 16:1096–1102.

Kanwisher N, Wojciulik E (2000) Visual attention: insights from brain imaging. Nat Rev Neurosci 1:91–100.

Kastner S, De Weerd P, Desimone R, Ungerleider LG (1998) Mechanisms of directed attention in the human extrastriate cortex as revealed by functional MRI. Science 282:108–111.

Keri S (2003) The cognitive neuroscience of category learning. Brain Res Brain Res Rev 43:85–109.

Kobatake E, Wang G, Tanaka K (1998) Effects of shape-discrimination training on the selectivity of inferotemporal cells in adult monkeys. J Neurophysiol 80:324–330.

Koida K, Komatsu H (2007) Effects of task demands on the responses of

color-selective neurons in the inferior temporal cortex. Nat Neurosci 10:108–116.

Kourtzi Z, Betts LR, Sarkheil P, Welchman AE (2005) Distributed neural plasticity for shape learning in the human visual cortex. PLoS Biol 3:e204.

Logothetis NK, Pauls J, Poggio T (1995) Shape representation in the inferior temporal cortex of monkeys. Curr Biol 5:552–563.

Maunsell JH, Treue S (2006) Feature-based attention in visual cortex. Trends Neurosci 29:317–322.

Miller BT, D'Esposito M (2005) Searching for "the top" in top-down control. Neuron 48:535–538.

Miller EK (2000) The prefrontal cortex and cognitive control. Nat Rev Neurosci 1:59–65.

Miller EK, Cohen JD (2001) An integrative theory of prefrontal cortex function. Annu Rev Neurosci 24:167–202.

Mirabella G, Bertini G, Samengo I, Kilavik BE, Frilli D, Della Libera C, Chelazzi L (2007) Neurons in area V4 of the macaque translate attended visual features into behaviorally relevant categories. Neuron 54:303–318.

Miyashita Y, Chang HS (1988) Neuronal correlate of pictorial short-term memory in the primate temporal cortex. Nature 331:68–70.

Moore CD, Cohen MX, Ranganath C (2006) Neural mechanisms of expert skills in visual working memory. J Neurosci 26:11187–11196.

Muhammad R, Wallis JD, Miller EK (2006) A comparison of abstract rules in the prefrontal cortex, premotor cortex, inferior temporal cortex, and striatum. J Cogn Neurosci 18:974–989.

Norman KA, Polyn SM, Detre GJ, Haxby JV (2006) Beyond mind-reading: multi-voxel pattern analysis of fMRI data. Trends Cogn Sci 10:424–430.

Nosofsky RM (1986) Attention, similarity, and the identification-categorization relationship. J Exp Psychol Gen 115:39–61.

O'Craven KM, Downing PE, Kanwisher N (1999) fMRI evidence for objects as the units of attentional selection. Nature 401:584–587.

O'Toole AJ, Jiang F, Abdi H, Haxby JV (2005) Partially distributed representations of objects and faces in ventral temporal cortex. J Cogn Neurosci 17:580–590.

Op de Beeck H, Wagemans J, Vogels R (2001) Inferotemporal neurons represent low-dimensional configurations of parameterized shapes. Nat Neurosci 4:1244–1252.

Op de Beeck HP, Baker CI, DiCarlo JJ, Kanwisher NG (2006) Discrimination training alters object representations in human extrastriate cortex. J Neurosci 26:13025–13036.

Palmeri TJ, Gauthier I (2004) Visual object understanding. Nat Rev Neurosci 5:291–303.

Rees G, Frackowiak R, Frith C (1997) Two modulatory effects of attention that mediate object categorization in human cortex. Science 275:835–838.

Reynolds JH, Chelazzi L (2004) Attentional modulation of visual processing. Annu Rev Neurosci 27:611–647.

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.

Saenz M, Buracas GT, Boynton GM (2002) Global effects of feature-based attention in human visual cortex. Nat Neurosci 5:631–632.

Schyns PG, Goldstone RL, Thibaut JP (1998) The development of features in object concepts. Behav Brain Sci 21:1–17; discussion 17–54.

Seger CA, Poldrack RA, Prabhakaran V, Zhao M, Glover GH, Gabrieli JD (2000) Hemispheric asymmetries and individual differences in visual concept learning as measured by functional MRI. Neuropsychologia 38:1316–1324.

Serences JT, Boynton GM (2007) Feature-based attentional modulations in the absence of direct visual stimulation. Neuron 55:301–312.

Shima K, Isoda M, Mushiake H, Tanji J (2007) Categorization of behavioural sequences in the prefrontal cortex. Nature 445:315–318.

Sigala N, Logothetis NK (2002) Visual categorization shapes feature selectivity in the primate temporal cortex. Nature 415:318–320.

Sigala N, Gabbiani F, Logothetis NK (2002) Visual categorization and object representation in monkeys and humans. J Cogn Neurosci 14:187–198.

Smith ML, Gosselin F, Schyns PG (2004) Receptive fields for flexible face categorizations. Psychol Sci 15:753–761.

Thomas E, Van Hulle MM, Vogels R (2001) Encoding of categories by noncategory-specific neurons in the inferior temporal cortex. J Cogn Neurosci 13:190–200.

Toni I, Rushworth MF, Passingham RE (2001) Neural correlates of visuomotor associations. Spatial rules compared with arbitrary rules. Exp Brain Res 141:359–369.

Vogels R, Sary G, Dupont P, Orban GA (2002) Human brain regions involved in visual categorization. NeuroImage 16:401–414.

Williams MA, Dang S, Kanwisher NG (2007) Only some spatial patterns of fMRI response are read out in task performance. Nat Neurosci 10:685–686.