

# Learning Shapes Spatiotemporal Brain Patterns for Flexible Categorical Decisions

Sheng Li<sup>1,2,3</sup>, Stephen D. Mayhew<sup>3</sup> and Zoe Kourtzi<sup>3</sup>

<sup>1</sup>Department of Psychology and <sup>2</sup>Key Laboratory of Machine Perception (Ministry of Education), Peking University, Beijing 100871, China and <sup>3</sup>School of Psychology, University of Birmingham, Birmingham B15 2TT, UK

Sheng Li and Stephen D. Mayhew have contributed equally to this work

Address correspondence to Zoe Kourtzi, School of Psychology, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK.  
Email: z.kourtzi@bham.ac.uk.

**Learning is thought to facilitate our ability to perform complex perceptual tasks and optimize brain circuits involved in decision making. However, little is known about the experience-dependent mechanisms in the human brain that support our ability to make fine categorical judgments. Previous work has focused on identifying spatial brain patterns (i.e., areas) that change with learning. Here, we take advantage of the complementary high spatial and temporal resolution of simultaneous electroencephalography–functional magnetic resonance imaging (EEG–fMRI) to identify the spatiotemporal dynamics between cortical networks involved in flexible category learning. Observers were trained to use different decision criteria (i.e., category boundaries) when making fine categorical judgments on morphed stimuli (i.e., radial vs. concentric patterns). Our findings demonstrate that learning acts on a feedback-based circuit that supports fine categorical judgments. Experience-dependent changes in the behavioral decision criterion were associated with changes in later perceptual processes engaging higher occipitotemporal and frontoparietal circuits. In contrast, category learning did not modulate early processes in a medial frontotemporal network that are thought to support the coarse interpretation of visual scenes. These findings provide evidence that learning flexible criteria for fine categorical judgments acts on distinct spatiotemporal brain circuits and shapes the readout of sensory signals that provide evidence for categorical decisions.**

**Keywords:** brain imaging, brain plasticity, EEG, fMRI, visual categorization

## Introduction

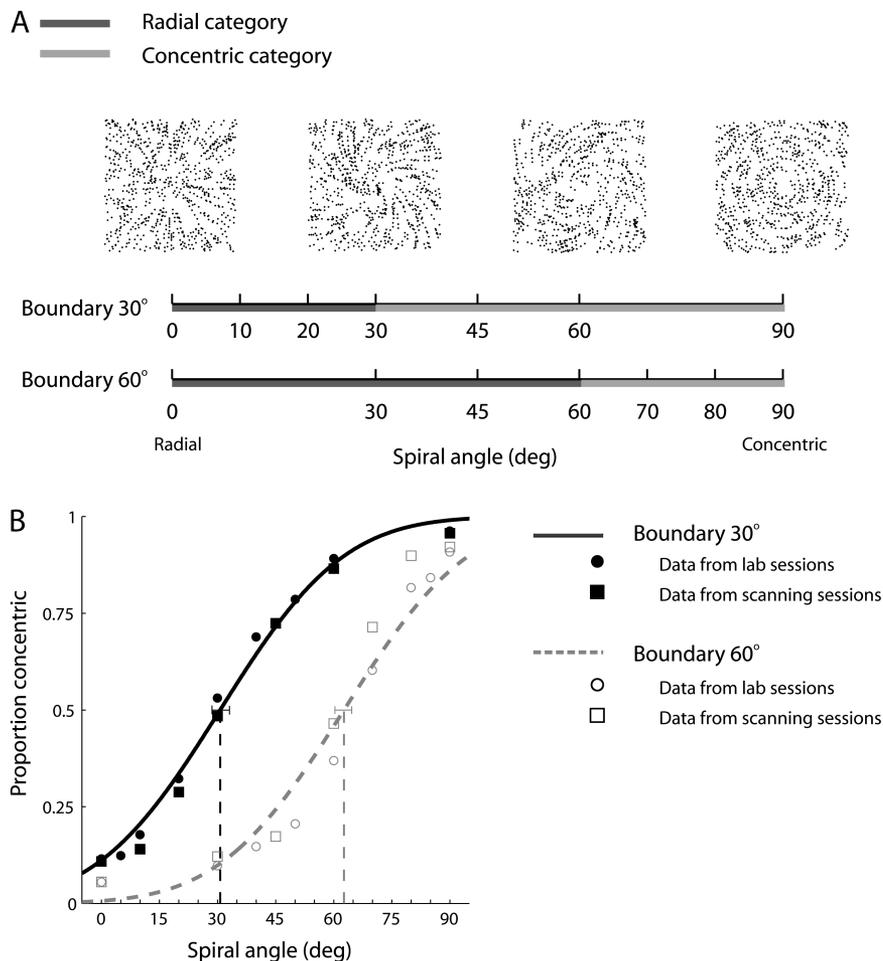
Successful interactions in our complex environments rely on the ability to translate sensory input to meaningful object categories. Despite the apparent ease with which we perform this task, our ability for visual categorization is challenged by sensory uncertainty due to feature similarity between objects that may belong in different categories. The primate brain is thought to overcome this challenge by taking into account previous experience that shapes the behaviorally relevant stimulus dimensions and facilitates our ability for fine categorical judgments (Goldstone et al. 2001; Seger and Miller 2010).

Previous theoretical and experimental investigations (Bullier 2001; Bar 2007; Hegde 2008; Peyrin et al. 2010) have proposed a fast route to visual categorization based on feedback interactions. In particular, a medial frontotemporal network has been suggested to provide a first coarse interpretation of the visual scene that is then propagated to occipitotemporal areas for refined processing of behaviorally relevant visual features. However, little is known about the spatiotemporal interactions between these cortical circuits that mediate our ability for category learning (Scott et al. 2006, 2008; Rossion et al. 2007).

In particular, we have previously shown that category learning supports the ability of observers to adopt different criteria (i.e., decision rules) for the flexible categorization of highly similar visual patterns (Li et al. 2007, 2009). Here, we seek to identify the spatiotemporal brain mechanisms (i.e., cortical areas engaged in specific temporal processes) that support our ability to learn flexible criteria for fine categorical decisions.

Using functional magnetic resonance imaging (fMRI) alone would make it difficult to identify cortical circuits related to different temporal processes involved in category learning due to the low temporal resolution of the technique. We exploit the complementary high temporal and spatial resolution of simultaneous electroencephalography–functional magnetic resonance imaging (EEG–fMRI) recordings to determine experience-dependent changes in categorical decision processes. Such changes in category boundary have been established as the hallmark of category learning by previous behavioral, physiological and brain imaging studies (e.g., Nosofsky 1986; Maddox and Ashby 1993; Goldstone and Steyvers 2001; Ashby et al. 2002; Freedman and Assad 2006; Jiang et al. 2007; Li et al. 2009). Following these previous studies, we used training with feedback to change the category boundary resulting in experience-dependent changes in the decision criterion. In particular, we tested observers' ability to make categorical judgments when presented with linearly morphed stimuli that varied in their similarity between radial and concentric patterns (Fig. 1A). Observers were asked to decide whether the viewed stimulus was radial or concentric. Uncertainty in this task increased as stimuli approached the boundary between the stimulus categories. Observers were trained to use different decision criteria (i.e., category boundaries) when performing categorical decisions in each of 2 scanning sessions, thereby dissociating the physical stimuli from their categorical interpretation. We tested for experience-dependent changes in EEG–fMRI activation patterns that related to behavioral shifts in criteria for fine categorical judgments.

Our findings provide evidence for a feedback-based circuit that supports category learning under uncertainty. In particular, fine categorical judgments were mediated by 1) early processing in middle frontal and medial temporal regions that are known to support contextual or semantic associations (Bar 2007) and 2) later processing in occipitotemporal and frontoparietal regions that are known to be involved in the accumulation of sensory information and performance monitoring for categorical judgments (Newsome et al. 1989; Kim and Shadlen 1999; Shadlen and Newsome 2001; Heekeren et al. 2004; Grinband et al. 2006). Importantly, experience-dependent changes in the decision criterion (category boundary) for categorizing similar visual forms were associated with changes only in later perceptual processes that engaged higher occipitotemporal and frontoparietal circuits. In contrast, learning fine categorical discriminations



**Figure 1.** Stimuli and behavioral data. (A) Stimuli: example Glass pattern stimuli (100% signal) at spiral angles of 0°, 30°, 60°, and 90°. The 2 boundaries and spiral angles tested are shown (black bar: stimuli that resemble radial; gray bar: stimuli that resemble concentric) indicating the categorical membership of the stimuli for each boundary. (B) Behavioral data for the categorization task collected in the lab (circles) and the scanner (squares) are shown for each boundary. Error bars indicate the 95% confidence interval at 50% concentric threshold.

did not modulate early processes that may support the coarser interpretation of visual stimuli. Thus, our findings suggest that learning flexible criteria for fine categorical judgments acts on later decision processes and shapes the readout of fine sensory signals that provide evidence for categorical decisions.

## Materials and Methods

### Observers

Eight observers (5 male, 3 female, mean age: 25.6) participated in the experiment. All observers were from the University of Birmingham, had normal or corrected to normal vision, and gave written informed consent. The study was approved by the local ethics committee.

### Stimuli

We used Glass pattern stimuli defined by white dot pairs (dipoles) displayed within a square aperture ( $7.7^\circ \times 7.7^\circ$ ) on a black background (100% contrast). For all stimulus patterns, the dot density was 3%, and the Glass shift (i.e., the distance between 2 dots in a dipole) was 16.2 arcmin. The size of each dot was  $2.3 \times 2.3$  arcmin<sup>2</sup>. These parameters were chosen based on pilot psychophysical studies and in accordance with previous studies (Li et al. 2009; Mayhew et al. 2010) showing that coherent form patterns are reliably perceived for these parameters. We generated stimuli intermediate between concentric and radial patterns by parametrically varying the dipole angles from 0° (radial pattern) to 90° (concentric pattern) (Fig. 1A). Half of the observers

were presented with clockwise spiral patterns (0–90° spiral angle) and half with anticlockwise patterns (0–90° spiral angle).

### Design

Observers were trained to perform a categorization task (concentric vs. radial) and tested in 2 EEG-fMRI sessions. Each session was preceded by psychophysical training outside the scanner, and the observers' behavioral performance was matched before the 2 EEG-fMRI sessions.

### Psychophysical Training

Before the experiment, observers were familiarized with the task and stimuli in a short practice session during which observers were shown 100% signal Glass patterns and categorized the presented stimuli as either radial (0° spiral angle) or concentric (90° spiral angle) patterns. During the training and test sessions, observers were presented with stimuli embedded in background noise ensuring that the task was demanding across all stimulus conditions (spiral angles). During the pretraining test (270 trials), observers were presented with 40% signal Glass patterns (spiral angles = 0°, 20°, 30°, 40°, 45°, 50°, 60°, 70°, and 90°) and performed a categorization task (concentric vs. radial patterns). This pretraining test allowed us to identify the categorical boundary between radial and concentric for each observer before training. Following the pretraining test, observers were presented with 60% signal stimuli and were trained (self-paced procedure with audio error feedback) to shift this boundary to either 30° or 60° of spiral angle. In the first training session, half of the observers (group 1) were trained to categorize the stimuli based on a boundary at 30°

spiral angle, whereas the other half (group 2) were trained to categorize the stimuli based on 60° boundary. In the second training session, observers from group 1 were trained on the 60° boundary, while observers from group 2 were trained on the 30° boundary. For the 30° boundary session, observers were trained at steps: 0°, 5°, 20°, 25°, 35°, 40°, 55°, and 90° of spiral angle, while for the 60° boundary at steps: 0°, 35°, 50°, 55°, 65°, 70°, 85°, and 90° of spiral angle. Each training session comprised multiple runs with 96 trials per run. Stimuli were presented for 300 ms. A white fixation square was presented at the center of each stimulus. Observers indicated the category to which each stimulus belonged by pressing 1 of 2 keys. Observers were trained until their performance reached a stable level (85% correct twice on the training and 80% correct on the posttraining test). This training procedure ensured that the performance of the observers was similar for both boundaries before scanning. After each training session, observers were tested in a posttraining test (405 trials) during which 40% signal stimuli were presented for 200 ms. For the 30° boundary session, observers were tested at steps: 0°, 5°, 10°, 20°, 30°, 40°, 50°, 60°, and 90° of spiral angle, while for the 60° boundary at steps: 0°, 30°, 40°, 50°, 60°, 70°, 80°, 85°, and 90° of spiral angle. No feedback was given during this posttraining test.

#### *fMRI Measurements*

All observers participated in 2 scanning sessions during which they performed the categorization task on the Glass pattern stimuli. For each observer, we collected data from 7 or 8 event-related runs in each session. The order of trials was matched for history (one trial back) such that each trial was equally likely to be preceded by any of the conditions. The order of the trials differed across runs and observers. Eight conditions (7 stimulus conditions and 1 fixation condition during which only the fixation point was displayed at the center of the screen) with 16 trials per condition were presented in each run. Each run comprised 129 trials (128 trials across conditions and 1 initial trial for balancing the history of the second trial) and two 9 s fixation periods (one in the beginning and one at the end of the run).

Observers were scanned after training in each of the 2 boundaries (30°, 60° spiral angle). For the 30° boundary, the stimulus conditions comprised Glass patterns of 0°, 10°, 20°, 30°, 45°, 60°, and 90° spiral angles. For the 60° boundary, the stimulus conditions comprised Glass patterns of 0°, 30°, 45°, 60°, 70°, 80°, and 90° spiral angles. Five stimulus conditions (0°, 30°, 45°, 60°, and 90° spiral angle) were common between sessions. As in our previous study (Li et al. 2009), the choice of stimulus conditions was constrained by 2 factors. First, we equated the number of conditions and stimuli across categories while avoiding stimulus repetition to ensure that observers were not biased in their responses due to uneven number of conditions (stimuli) in 1 of the 2 categories. Second, we aimed to sample representative points on the psychometric function while selecting a limited but adequate number of conditions to ensure that enough trials were recorded per condition and high quality signals were measured within the time constraints of fMRI scanning.

For fixation trials, the fixation square was displayed for 3 s. For experimental trials, each trial lasted for 3 s and started with 200 ms stimulus presentation followed by 1300 ms delay during which a white fixation square was displayed at the center of the screen. After this fixed delay, the fixation dot changed color to either green or red. This change in fixation color served as a cue for the motor response using 1 of 2 buttons. If the color cue was green, observers indicated concentric versus radial by pressing the left versus right finger key, while if the color was red, the opposite keys were used (e.g., concentric = right key). The fixation color was changed back to white 300 ms before the next trial onset. This procedure dissociated the motor response (key press) from the stimulus categories. Observers were familiarized with this procedure before scanning.

#### *Data Acquisition*

##### *fMRI Scanning*

The experiments were conducted at the Birmingham University Imaging Center (3-T Achieva scanner; Philips, Eindhoven, the Netherlands). Echo planar imaging (EPI) and  $T_1$ -weighted anatomical ( $1 \times 1 \times 1 \text{ mm}^3$ ) data

were collected with an 8-channel sensitivity encoding head coil. EPI data (Gradient echo pulse sequences) were acquired from 24 slices (whole-brain coverage, repetition time [TR]: 1500 ms, echo time [TE]: 35 ms, flip angle: 73°,  $2.5 \times 2.5 \times 4 \text{ mm}^3$  resolution).

##### *EEG Recordings*

We recorded EEG and fMRI signals simultaneously during scanning. EEG data were acquired from 64 electrodes using an MRI compatible cap and amplifiers (Brain Products, Munich, Germany) with current limiting safety resistors of 5 k $\Omega$  at the amplifier input and in each electrode. The EEG cap comprises 62 scalp electrodes distributed in 10-20 system. To identify trials contaminated by eyeblinks, and electroculographic artifacts we recorded signals using an electrode placed over the mid-lower eyelid. To correct for ballistocardiographic (BCG) artifacts, we recorded the electrocardiogram (ECG) from an electrode attached to the subject's chest, below the left collarbone. Data were sampled at 5000 Hz, with a low-pass hardware filter at 250 Hz. Electrode impedances were generally kept below 20 k $\Omega$ . The EEG system clock was synchronized with the MRI scanner clock using SyncBox (Brain Products). A custom-made photosensor was used to measure the precise timing of stimulus onset on the screen inside the scanner. The detected stimulus onsets and the MRI volume triggers were saved as markers together with the recorded EEG signals.

#### *Data Analysis*

##### *Behavioral Data Analysis*

We fitted psychometric (proportion concentric) data collected in the lab with a cumulative Gaussian function using a procedure that implements a maximum-likelihood method (Wichmann and Hill 2001). Confidence intervals were calculated on the fits from 2000 bootstrap iterations of the data. Using this procedure for each individual observer's behavioral data, we identified the perceptual boundary (spiral angle at 50% concentric threshold) for each observer and categorization task. Comparing the mean threshold and slope values between the 2 sessions for the same category boundary did not show any session order effects. For the 30° boundary task, the mean threshold values for 2 sessions were  $31.8 \pm 12.2$  and  $34.7 \pm 5.6$  and the slope values  $0.015 \pm 0.002$  and  $0.019 \pm 0.006$ , respectively. For the 60° boundary task, the mean threshold values for 2 sessions were  $55.6 \pm 10.4$  and  $59.2 \pm 6.8$  and the slope values  $0.017 \pm 0.004$  and  $0.019 \pm 0.002$ , respectively. That is, the threshold and slope values were similar when observers were tested first with stimulus categories defined by the 30° and then the 60° spiral angle boundary and vice versa. Therefore, we analyzed the data based on the trained boundary (boundary 30° and boundary 60°) independent of the order on which the observers were trained on the 2 boundaries.

##### *fMRI Data Processing*

fMRI data were processed using BrainVoyager QX (Brain Innovations, Maastricht, the Netherlands). Anatomical data were used for 3D cortex reconstruction, inflation, and flattening. Preprocessing of functional data included slice scan time correction, head movement correction, temporal high-pass filtering (3 cycles), and removal of linear trends. Trials with head motion larger than 1 mm of translation or 1° of rotation were excluded from the analysis. Spatial smoothing (Gaussian filter; full-width at half maximum, 6 mm) was performed only for group random effect analysis but not for data used for the multivoxel pattern classification analysis. The functional images were aligned to anatomical data, and the complete data were transformed into Talairach space. For each observer, the functional imaging data between the 2 sessions were coaligned registering all volumes of each observer to the first functional volume of the first run and session. This procedure ensured a cautious registration across sessions. To avoid confounds from any remaining registration errors, we compared fMRI signals between stimulus conditions within each session rather than across sessions. A gray matter mask was generated for each observer in Talairach space from the anatomical data for selecting only gray matter voxels for further analyses.

### EEG Data Processing

We focused our EEG analysis on robust event-related signals in the 1–40 Hz frequency range that have previously been shown to reflect visual form processing (Ohla et al. 2005; Pei et al. 2005). The MRI volume triggers were used to identify the onset of each gradient artifact in order to create an artifact template. MRI gradient artifacts were then removed using average artifact subtraction (Allen et al. 2000) in BrainVision Analyzer (Brain Products). EEG data were downsampled to 500 Hz, and BCG artifacts removed using the optimal basis set method (Niazy et al. 2005) available as a plug-in to EEGLAB (Delorme and Makeig 2004). For each imaging session, EEG data from all experimental runs were concatenated, and EEG signals were bandpass filtered between 0.1 and 40 Hz. The filtered data were then analyzed with a FastICA (Hyvarinen and Oja 1997) algorithm to generate 62 independent components (ICs). In each session, ICs containing the transient eyeblink artifact were removed from the data (Jung et al. 2000). These ICs were identified from 1) plots of trial amplitude or event-related potential (ERP) images (Jung et al. 2000) that showed a distinctive pattern of transient deviations with large amplitude occurring at unpredictable latencies relative to the stimulus and 2) scalp maps of electric field distribution with an obvious frontal weighting. These measurements are quite distinct from ERP images and scalp maps from stimulus-related signals. Furthermore, components whose time course significantly correlated with the recorded ECG signal were rejected as residual BCG artifacts (Srivastava et al. 2005; Debener et al. 2007). The remaining ICs were used to reconstruct the EEG signal for further analysis. Single-trial EEG epochs were extracted using a window of 0.7 s (from 200 ms prestimulus to 500 ms poststimulus) based on the stimulus onset markers provided by the photosensor. For each epoch, a baseline correction was performed by subtracting the average of the prestimulus (200 ms) data. Single trials with maximum amplitude difference greater than 100  $\mu\text{V}$  were excluded from further analysis.

### Mutual Information Estimation for EEG Components

We used information theory (Shannon 1948; Cover and Thomas 1991) to estimate mutual information (MI) between categories (radial, concentric) and EEG responses (Montemurro et al. 2008). This measure is driven by the distribution of the EEG signal amplitudes and therefore is more sensitive than the mean ERP signals in identifying informative EEG components related to stimulus conditions. That is, MI between EEG amplitude and a given stimulus condition is a measure of the statistical dependence between these 2 variables. High MI values suggest that the distributions of the 2 variables share common information. We estimated the mutual information  $I(S, R)$  between stimulus conditions ( $N = 7$ ) and EEG responses for each observer, session, and EEG channel as follows:

$$I(S, R) = \sum_{s,r} P(s)P(r|s) \log_2 \frac{P(r|s)}{P(r)} \quad (1)$$

where  $s$  is the stimulus condition,  $r$  is the amplitude of the EEG response, and  $P$  is the probability.

For each channel and trial, the EEG time series was smoothed by averaging 10 ms signal around each time point. We then estimated the distribution of the signal amplitudes using 30 response bins (that is  $N_s/\bar{R} \geq 4$ , where  $N_s$  is number of trials per condition and  $\bar{R}$  is number of bins). For each session, we set the two tailed 95% confidence intervals to the upper and lower bounds of the amplitude distribution. Amplitude values outside this range were set to the upper bound and lower bound, respectively. Note that this amplitude correction was applied only for this MI analysis; all subsequent analyses used the preprocessed EEG signal without this correction. This amplitude correction was performed to preserve the sensitivity of the information measurement given the number of bins necessary for estimating the MI (Panzeri et al. 2007).

We estimated the MI for each EEG channel based on equation (1). We also shuffled the condition labels 500 times and estimated the shuffled MI to create a baseline measurement. Due to the limited number of trials, we corrected the estimated MI following a Bayesian procedure (Panzeri and Treves 1996) and subtracting the shuffled MI. Following this correction, we tested which time points across observers and sessions had MI values that differed significantly from

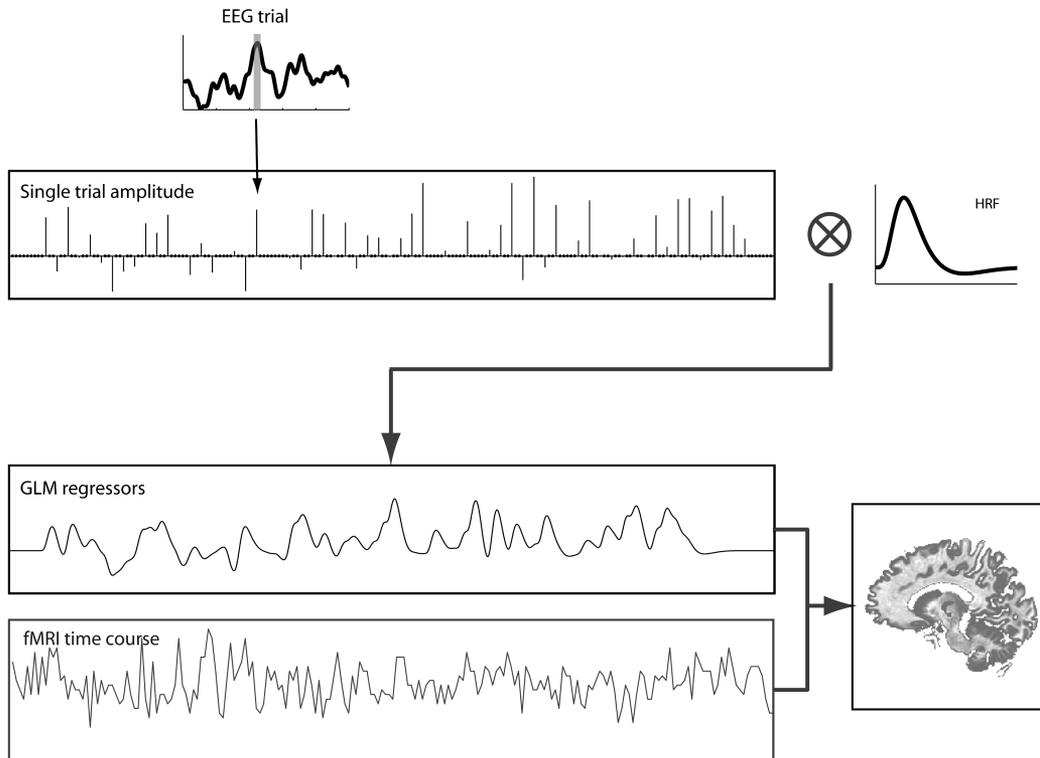
chance ( $P < 0.05$ , paired  $t$ -test of MI values against 0 across all observers and scanning sessions). Using this procedure, we computed the MI separately for each session and observer. No significant differences were observed between the 2 sessions in either the component latency (paired  $t$ -test, Component 1:  $t_7 = 0.12$ ,  $P = 0.91$ ; Component 2:  $t_7 = 1.43$ ,  $P = 0.20$ ) or amplitude (paired  $t$ -test, Component 1:  $t_7 = 1.83$ ,  $P = 0.11$ ; Component 2:  $t_7 = 0.94$ ,  $P = 0.38$ ). Thus, to ensure sufficient signal power for the robust and independent estimation of informative EEG components, we calculated the MI per temporal bin (30 ms) across all channels, stimulus conditions, and EEG single trials in both sessions and observers. In particular, we used the maximum amplitude peaks of the MI time course (averaged across all EEG channels and observers) to identify the temporal components that contained discriminative information across stimulus conditions.

### EEG Channel Selection

To select EEG channels that contained information for discriminating between radial and concentric patterns, we used a receiver operating characteristic (ROC) analysis on the response amplitude of each channel across single trials. We performed this analysis on the data within a 10-ms window around the peak of each of the 2 components and measured the area under the curve that indicated the discriminability of EEG signals related to radial ( $0^\circ$ ) and concentric ( $90^\circ$ ) trials. For each observer, we selected the same number of channels per cortical lobe (occipital, parietal, temporal, central, and frontal) with the highest normalized ROC values (the absolute difference from the chance level of 0.5). As there were only 3 available channels in the occipital lobe, we limited the number of channels to 3 per lobe. Averaging signals across 3 electrodes ensured measurements that are robust against noise occurring at a single electrode. This procedure ensured equal and unbiased use of EEG information from different scalp locations. We then averaged the time course of the selected channels per lobe to generate a mean EEG time course for each component, lobe, and observer.

### EEG-Informed fMRI Mapping

To identify brain regions associated with different processing stages, we used an EEG-informed general linear model (GLM) analysis (Debener et al. 2005; Eichele et al. 2005; Philiastides and Sajda 2007). To test which voxels correlated significantly with the time course of each of the 2 EEG components, we generated separate regressors for each of the 2 EEG temporal components and tested for fMRI responses that correlated with the amplitude of each EEG component across trials. For each individual observer and lobe, a separate regressor for each EEG temporal component was generated based on the single-trial variability in the EEG amplitude at the respective component latency. The regressor amplitude at each trial was calculated by averaging the amplitude of the selected channel time course within a 10-ms window, centered at the component peak latency. The 2 EEG regressors were decorrelated (using Gram-Schmidt orthogonalization for removing the common variance between the 2 from the first or second regressor; Eichele et al. 2005). In particular, correlations between regressors (mean across subjects and sessions:  $R = 0.18$ ; standard deviation: 0.07) were eliminated ( $R = 0.00$ ) after removing any common variance from the second or first component regressor. This procedure ensured that fMRI activations were specific to each component rather than reflecting general features of the visual evoked response. Both regressors were then convolved with a canonical double-gamma hemodynamic response function. These regressors were used to form a GLM along with 6 other regressors derived from the motion correction parameters (Fig. 2). For each lobe-based channel grouping, we performed group random effects analysis and searched for regions across the whole brain that showed significant ( $P < 0.01$ , cluster threshold corrected) correlations between each of the 2 EEG components and the blood oxygen level-dependent (BOLD) signal. We overlaid activation maps from all lobe-based channel groupings to illustrate voxels that correlated significantly with each of the 2 EEG components independent of channel/scalp location. We identified the same regions in individual observers using fixed effects analysis.



**Figure 2.** Schematic of single-trial EEG-informed fMRI mapping To identify brain regions associated with different decision processes, we used an EEG-informed GLM analysis with separate regressors for the 2 EEG temporal components. The regressor amplitude at the time of each stimulus trial was calculated by averaging the amplitude of the EEG channel time course within a 10-ms window, centered at the component peak latency. These time series of EEG component amplitudes were used to generate 2 separate regressors and convolved with a canonical double-gamma hemodynamic response function.

### Multivoxel Pattern fMRI Analysis

To test which brain regions showed experience-dependent changes, we conducted multivoxel pattern analysis (MVPA) on the activation patterns of the regions identified based on the EEG-informed fMRI analysis. This approach has been shown to be more sensitive in revealing experience-dependent differences in the discrimination of visual forms than conventional univariate statistical analyses of fMRI signals (Li et al. 2009). For each observer, we selected voxels in each region of interest (ROI) that were significantly activated in the corresponding EEG-fMRI statistical map ( $P < 0.05$ , uncorrected). We ordered these voxels based on their  $t$  value (in descending order) when comparing all stimulus conditions with fixation. Following this procedure, we selected up to 200 voxels (minimum of 50 voxels) for each ROI and observer for the analysis, as prediction accuracy had saturated at this pattern size across areas, resulting in a dimensionality compatible with previous studies (Haynes and Rees 2005; Kamitani and Tong 2005; Li et al. 2007). The only exception to this procedure was the medial temporal lobe area (MTL, LH), for which we used  $P < 0.10$  to select a sufficient number of voxels (minimum 50) for individual observers. This was possible in 6 of the 8 observers. Each voxel time course was  $z$ -score normalized for each experimental run separately. The data pattern for each trial was generated by shifting the fMRI time series by 3 volumes (4.5 s) to account for the hemodynamic delay.

Finally, we used a linear support vector machine (SVM) and a leave-one-run-out cross-validation procedure for the pattern classification. To investigate the link between fMRI activity and behavioral performance, we tested the classifier's performance in predicting the observer's choice rather than the stimulus condition. We reasoned that shifts in these fMR-metric functions similar to those in the psychometric functions would indicate brain areas that contain information about the participant's behavioral choice. In particular, we trained the classifier to associate fMRI signals with a label (radial vs. concentric) that related to the observer's choices. We averaged the 2 volumes from each trial (trial duration = 3 s, TR = 1.5 s) to generate one training pattern per trial. We then tested whether the classifier predicted the observers' behavioral

choice (radial vs. concentric) using an independent data set. For each observer, we calculated the mean performance of the classifier (proportion of trials classified correctly) in predicting whether each stimulus was radial or concentric across cross-validations. fMR-metric functions were calculated per ROI by averaging the classifier performance across observers and fitting the data using a cumulative Gaussian (Li et al. 2009). It is important to note that the classification comparisons were independent from the voxel selection procedure. The voxel selection was conducted only on the extreme stimulus conditions, and the classification was conducted on the observers' response per trial.

### Multivariate Pattern Analysis for EEG Data

To test which temporal processes associated with each of the 2 EEG components showed experience-dependent changes, we performed a similar pattern classification analysis on the EEG data (15 channels, 112 trials across conditions per run) for each of the 2 components. For each session, we trained a linear SVM to classify single-trial EEG signals associated with the observer's choice (concentric vs. radial) and tested the classifier's accuracy using an independent data set. For each EEG trial, we averaged the signal from a 30-ms window centered at the peak of each of the 2 EEG components. For each cross-validation, 10% of the data was left out as an independent test data set, and the rest 90% of the data was used as the training set. We calculated the classifier performance for each condition across 100 cross-validations and observers and fitted the data using a cumulative Gaussian.

## Results

### Behavioral Results

Following previous studies on category learning (e.g., Nosofsky 1986; Maddox and Ashby 1993; Goldstone and Steyvers 2001; Ashby et al. 2002; Freedman and Assad 2006; Jiang et al. 2007)

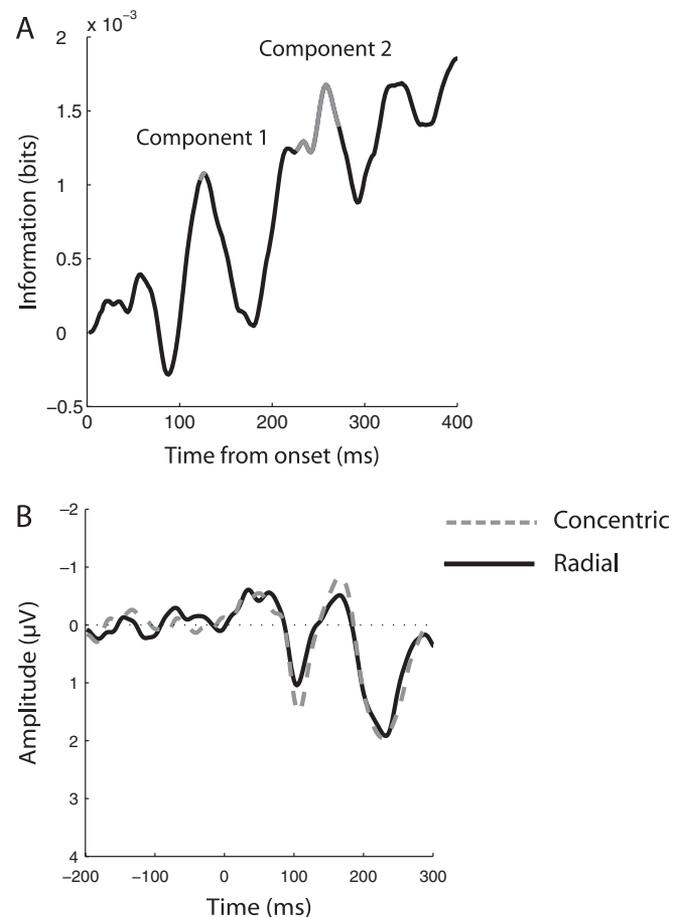
including our previous work (Li et al. 2009), we used training with feedback to change the observers' criterion in a categorization task. This paradigm entails training on one arbitrary category boundary and then shifting to a different boundary through training.

In particular, we tested the observers' ability to categorize global form patterns as radial or concentric (Fig. 1A) and plotted their performance (proportion concentric) as a function of spiral angle. Observers were trained on 1 of 2 new boundaries (30°, 60° spiral angle) and then retrained on the other. To control for possible effects of session order, half of our observers were trained first on boundary 30° and then 60°, while the rest of the observers were trained first on boundary 60° and then 30°. As the new boundary (30° or 60° spiral angle) was not explicitly revealed to the observers, categorization training with feedback was necessary for discriminating fine differences between the stimuli and shifting the category boundary between sessions. The training ended when observers reached a stable level of mean performance across stimulus conditions (85% correct twice on the training and 80% correct on the posttraining test).

Comparing the psychometric functions before and after training showed that categorization training altered the observers' decision criterion (category boundary; Fig. 1B). Our results did not show any significant differences between training sessions (number of training runs for session 1:  $7.9 \pm 2.23$ , session 2:  $5.5 \pm 3.02$ ; paired *t*-test:  $t_7 = 1.57$ ,  $P = 0.16$ ). This was further supported by additional analyses calculating psychometric functions separately for 2 groups of observers: one group that was trained first on the 30° boundary and a second group that was trained first on the 60° boundary. Despite the smaller number of observers (4 in each group), we observed similar shifts in the psychometric functions for both groups and no significant differences between groups ( $t_6 = 0.69$ ,  $P = 0.51$ ). As a result, we pooled the data across sessions for each boundary and measured category learning as the shift in the boundary (50% point at the psychometric function). Before training (pretraining test), the mean categorization boundary was  $43.7^\circ (\pm 5.37^\circ)$  spiral angle, matching closely the mean of the physical stimulus space ( $45^\circ$  spiral angle). After training, the observers' criteria shifted to  $30.7^\circ (\pm 2.49^\circ)$  for the 30° boundary and  $62.6^\circ (\pm 2.21^\circ)$  for the 60° boundary. Fitting the behavioral data with a cumulative Gaussian showed a significant shift in the mean boundary when observers were trained with different categorization boundaries ( $t_7 = 15.93$ ,  $P < 0.001$ ).

### EEG-Informed fMRI Mapping of ROIs

We exploited the high temporal resolution of EEG to identify temporal components that correspond to distinct processes in categorical decisions. We used information theory (Shannon 1948; Cover and Thomas 1991; Montemurro et al. 2008) to identify informative components (i.e., temporal components that contain stimulus or task-related information) in the EEG signal related to the stimulus category. This method provides a sensitive tool for identifying task-relevant temporal components of the EEG signal (Fig. 3A) that may be difficult to discriminate from comparison of standard ERP waveforms between stimulus conditions (Fig. 3B). We identified 2 time intervals that showed significant MI values: 1) 122–128 ms and 2) 226–272 ms after stimulus onset. We identified peak time points with the highest MI value within these significant time intervals that corre-



**Figure 3.** EEG temporal components. (A) MI time course. Information was smoothed using a 30-ms window and averaged across all trials, channels, observers, and scanning sessions. Information values (bits) are shown after subtraction of the MI estimated using shuffled condition labels (500 iterations) and normalization to prestimulus baseline. The latencies of the 2 temporal peaks that showed the highest MI values are indicated in the plots. Shaded parts of the time course indicate significant MI values ( $P < 0.05$ ). (B) Group average visual-evoked potential in response to concentric and radial Glass patterns. The waveforms are averaged across trials, the 15 channels selected for the EEG-informed fMRI analysis and all observers. This analysis showed similar peaks but lacked sufficient sensitivity to discriminate between signals related to concentric and radial patterns, in contrast to the MI-based analysis that allowed us to identify components that discriminated significantly between these signals.

sponded to early and later EEG components. The peak points for these components were at 126 ms ( $\pm 14.8$  ms) and 258 ms ( $\pm 33.2$  ms), respectively. These components were identified reliably in individual participant data, while other nearby peaks (e.g., the early phase of component 2 at mean 216 ms) were not significant. We concentrated on these components for further analyses, as previous studies suggest that they reflect distinct processes (Johnson and Olshausen 2003; Ohla et al. 2005; Pei et al. 2005; Tanskanen et al. 2008; Das et al. 2010). In particular, previous studies showing differential responses to global forms at later rather than early latencies suggest that latencies around the first component relate to visual form integration, while latencies around the second component relate to finer perceptual judgments. The higher MI values for component 2 than 1 are consistent with perceptual differences in the fine discrimination task we employed in contrast to low-level stimulus differences (e.g., local orientation signals) for which the stimuli were on average matched across conditions.

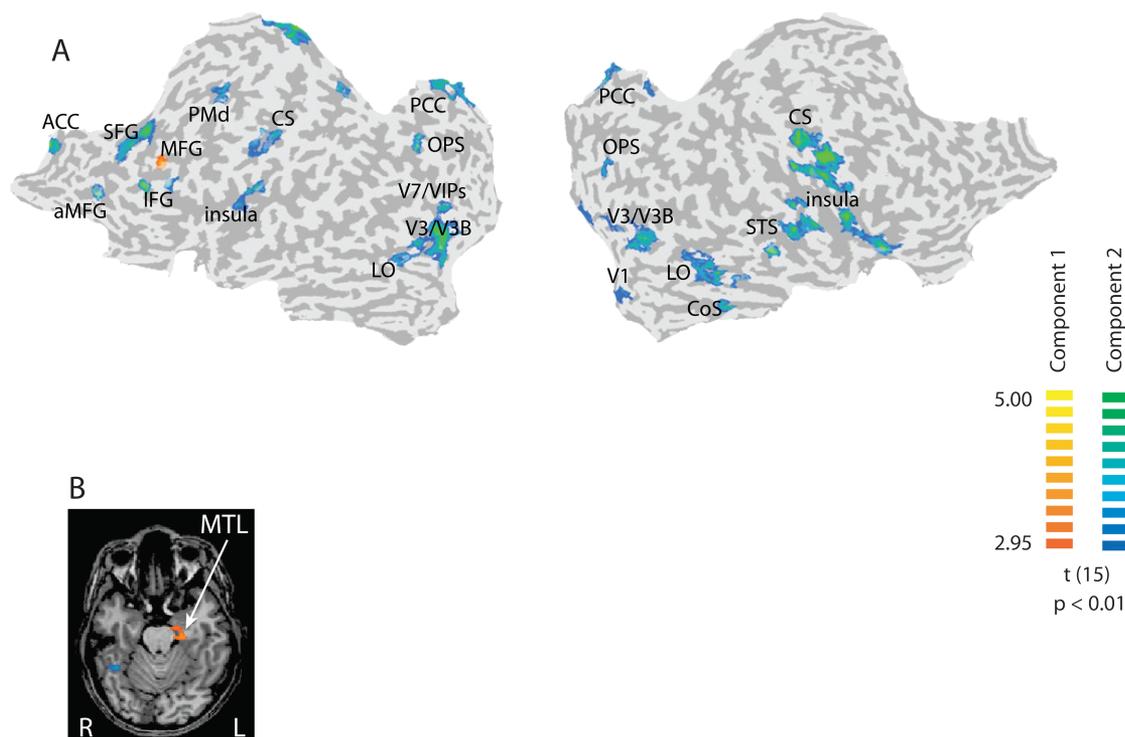
Finally, previous studies (Philiastides et al. 2006) have discriminated between components related to task difficulty (around 220 ms) and decision-related events (later than 300 ms). Following these studies, we explored a third component with average peak latency of 358 ms. However, analysis of peak latencies around the second and third component did not show any significant differences across stimulus difficulty levels ( $F_{1,15} = 0.44$ ,  $P = 0.52$ ), suggesting that the second component could not be discriminated from the third one on the basis of task difficulty. This is consistent with recent work showing that ERP signals at latencies around 220 ms reflect sensory processing of stimuli embedded in noise rather than task difficulty (Banko et al. 2011). Thus, we focus on the first 2 temporal components for the rest of the analyses.

To identify brain regions that are involved in the different temporal processes related to the above EEG components, we conducted an EEG-informed fMRI analysis (Fig. 2), as described in previous studies (Debener et al. 2005; Eichele et al. 2005; Philiastides and Sajda 2007). This analysis (Fig. 4) showed significant ( $P < 0.01$ , cluster threshold corrected) correlations with the amplitude of the first EEG component at middle frontal gyrus (MFG, LH) and MTL (LH). Significant correlations with the amplitude of the second EEG component were observed in V1 (RH), V3/V3B, V7/VIPs (ventral intraparietal sulcus, LH), LO (lateral occipital region anterior to retinotopic area V4), CoS (collateral sulcus, RH, only one observer), OPS (occipitoparietal sulcus), STS (superior temporal sulcus, RH), PCC (posterior cingulate cortex), CS (central sulcus), PMd (premotor dorsal, LH), ACC (anterior cingulate cortex, LH), insula, IFG (inferior frontal gyrus, LH), aMFG (anterior middle

frontal gyrus, LH), and SFG (superior frontal gyrus, LH). Prefrontal activations were more prominent in the left hemisphere. This is consistent with previous work showing left lateralization for verbal working memory (Smith and Jonides 1997) and various categorization tasks (Grossman, Koenig, et al. 2002; Grossman, Smith, et al. 2002; Koenig et al. 2005; Li et al. 2007) when the category labels can be verbally expressed.

These results demonstrate 2 distinct cortical networks engaged in fine categorical decisions. First, activations in MFG and MTL were significantly correlated with the first component, consistent with the role of these areas in extracting contextual shape information that may contribute to fast categorization in cluttered scenes (Bar 2007). Second, processes related to perceptual judgments (i.e., associated with the second component) were significantly correlated with prefrontal areas implicated in categorization and adaptive cognitive processes (Miller 2000; Duncan 2001), parietal and motor regions known to be involved in perceptual categorization (Freedman and Assad 2006), and stimulus-response association processes (Toni et al. 2001) as well as occipito-temporal regions engaged in the integration of visual forms (Ostwald et al. 2008).

In interpreting these results, it is important to take into account possible limitations of the EEG-fMRI methodology. First, the EEG-informed GLM analysis relies on differences in the amplitude rather than the latency of the regressors, as latencies related to the different EEG components overlap in the fMRI time course. Despite this limitation, this approach has been successful in linking fMRI activations to specific temporal components that differ in their response amplitude



**Figure 4.** Single-trial EEG-informed fMRI analysis. Random effect GLM analysis (data grouped across all observers and scanning sessions) using EEG defined regressors corresponding to the 2 temporal components. Regressors were orthogonalized by removing the common variance from the second component regressor. Activations significantly ( $P < 0.01$ , cluster threshold corrected) correlated with the first and second EEG components. (A)  $t$ -statistic maps are superimposed on flattened cortical surfaces of both hemispheres (Supplementary Table S1 for Talairach coordinates). Sulci are shown in dark gray and gyri in light gray. (B) Additional activations correlated with the first EEG component were observed in MTL.

across trials (Debener et al. 2005; Eichele et al. 2005; Philiastides and Sajda 2007). Second, activations associated with channels selected per lobe may reflect processing across brain regions and lobes due to the low spatial resolution of the EEG. Our reasoning for selecting the most informative channels across lobes was to ensure equal and unbiased use of EEG information from different scalp locations in order to identify regions across the whole brain associated with distinct temporal processes. Third, EEG-fMRI allows us to identify cortical areas that are more strongly rather than causally related to one of the processes associated with different temporal components (e.g., stimulus integration vs. perceptual classification). Our results suggest that categorical decisions are mediated by early medial frontotemporal processing for fast categorization and may influence via feedback connections later perceptual judgments supported by a network of frontoparietal and occipitotemporal regions (Bar 2007). However, it is possible that additional interactions across areas are engaged at a finer resolution than can not be measured by EEG-fMRI. That is, recurrent interactions between frontoparietal and occipitotemporal areas may support fine classification judgments at later processing stages.

Finally, despite recent advances in data acquisition and artifact correction techniques that have greatly improved the signal to noise ratio of EEG signals recorded during fMRI (Laufs et al. 2008), small residual artifacts may remain in the EEG and compromise activation maps resulting from EEG-informed fMRI analyses. It is important to consider this limitation when interpreting the extent of activations observed in EEG-informed fMRI maps. For example, our results showed a more extensive network of areas associated with later processes related to fine discrimination (i.e., second component) than early stimulus-related processes (i.e., first component). Although artifact correction techniques would have similar results across the EEG signal time course, small variation in the amplitude of component 1 across stimulus conditions would have resulted in the lower MI values and the lower correlations with the BOLD signal observed in our data. However, these small amplitude variations in component 1 were expected as stimulus statistics (e.g., local orientation signals) were on average matched across conditions. In contrast, larger amplitude variations in component 2 reflect perceptual differences among stimulus conditions related to the fine discrimination task. Furthermore, the activations we observed using EEG-based GLMs correspond to activation patterns in our previous fMRI studies on categorical decisions (Li et al. 2009). This was confirmed by an additional analysis using searchlight multivoxel pattern classification analysis to compare fMRI activations between stimulus categories (Supplementary Fig. S1). Despite the different methodological approaches used for generating EEG-based and searchlight maps, the activations revealed by the EEG-informed fMRI analysis largely correspond to the regions engaged in categorical decisions and are consistent with activations reported by previous studies on visual categorization (for reviews, Keri 2003; Ashby and Maddox 2005; Poldrack and Foerd 2008; Seger and Miller 2010). The advantage of EEG-informed fMRI is that it allows us to discern cortical areas associated with the distinct temporal processes (i.e., early form integration vs. later categorical judgments), in contrast to fMRI only based analyses that test for activity related to the task (i.e., visual categorization) but cannot discern between the processes involved.

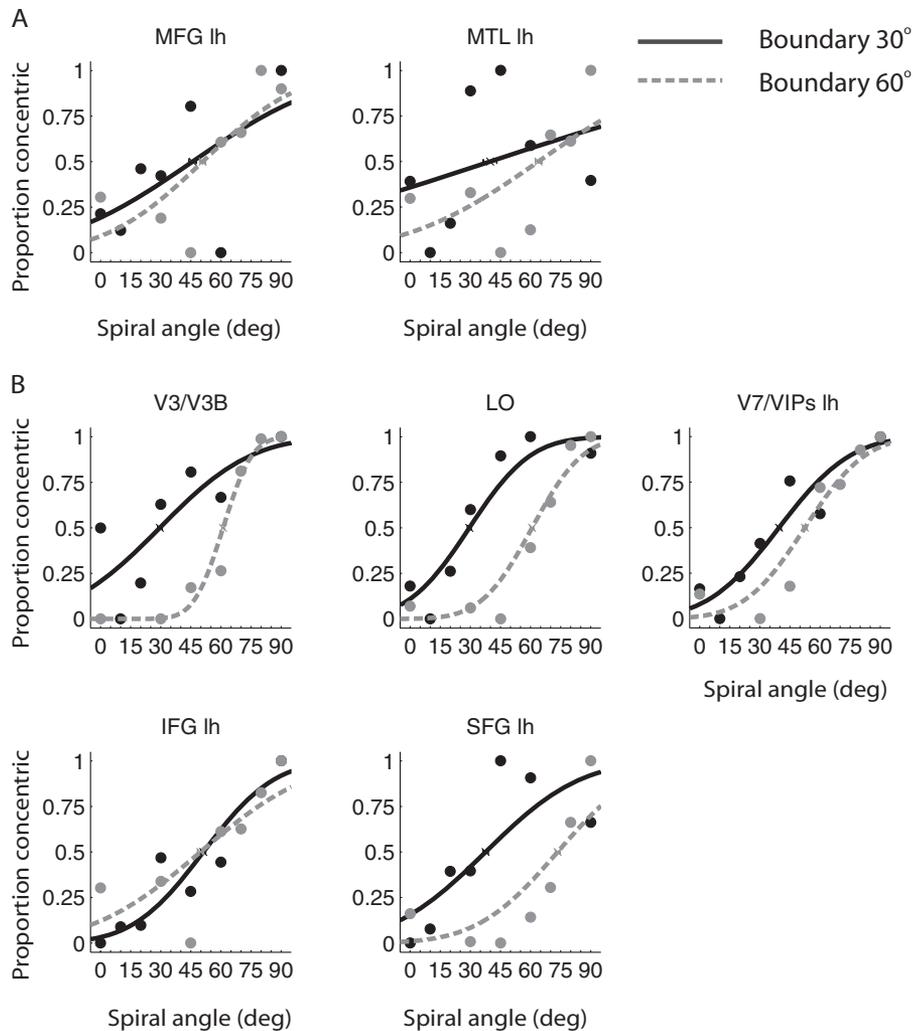
### **Experience-Dependent Changes: fMR-Metric Functions**

We tested which brain regions identified by the EEG-informed fMRI analysis showed experience-dependent changes in their activation patterns. In particular, we tested whether activation patterns in these regions corresponded to the changes in criterion that we observed in behavioral performance with training. We compared psychometric functions obtained from the behavioral data with fMR-metric functions obtained for each of the ROIs (Li et al. 2009).

We observed significant criterion shifts (two-way repeated measures analysis of variance [ANOVA] [Greenhouse-Geisser corrected] for ROI and Boundary) in the fMR-metric functions for areas associated with the second ( $F_{1,5} = 35.87, P = 0.002$ ) but not the first (fMR-metric functions in MFG and MTL were not significantly fitted) EEG component (Fig. 5; Supplementary Fig. S2 for fMR-metric functions based on nonscaled data). To quantify this experience-dependent criterion shift, we compared the mean (i.e., 50% point) of the fMR-metric functions obtained when observers performed the categorization task on the 2 different boundaries (30° vs. 60°). We observed shifts in the mean of the functions that corresponded to the behavioral criterion shifts in higher occipitotemporal (V3/V3B, LO), parietal (V7/VIPs), and frontal (SFG) regions, consistent with the role of these regions in categorization tasks (Grinband et al. 2006; Philiastides and Sajda 2007; Li et al. 2009). These findings suggest that category learning shapes the criterion for fine categorical decisions by altering later processing associated with perceptual judgments in frontoparietal and higher occipitotemporal areas rather than early processing in medial frontotemporal regions that have been implicated in fast but coarse categorization. Finally, the lack of significant experience-dependent effects in early visual or motor areas is possibly due to the fact that representations in these areas are related to the physical stimulus space or decision execution, respectively, and may not change with training in shape discrimination. The rest of the areas that did not show significant experience-dependent changes were primarily frontal regions that are associated with contextual or semantic processing (Grossman, Smith, et al. 2002; Koenig et al. 2005; Badre and D'Esposito 2009) in categorization tasks (i.e., MFG and MTL associated with the first component; dorsal IFG/insula, PCC, ACC, and aMFG associated with the second component). This may be due to our categorization training paradigm that resulted in changes to the perceptual boundary rather than the semantic interpretation of the stimuli.

### **Experience-Dependent Changes: EEG-Metric Functions**

Similar to the analysis for fMR-metric functions, we generated EEG-metric functions (Philiastides and Sajda 2006; Das et al. 2010) for the 2 EEG components (Fig. 6A). We tested whether decoding the observer's choice from single-trial EEG data differed between sessions based on the relevant category boundary. Our results showed that experience-dependent changes in the decision criterion were associated with changes only in later EEG processes related to categorical judgments rather than early processes related to form integration. In particular, comparing EEG-metric functions before and after categorization training showed significant experience-dependent changes only in the second EEG component. Paired sample *t*-test analysis for each component showed a significant shift for the second ( $t_6 = 2.46, P < 0.05$ ) but not the first component ( $t_6 = 1.24, P = 0.26$ ).



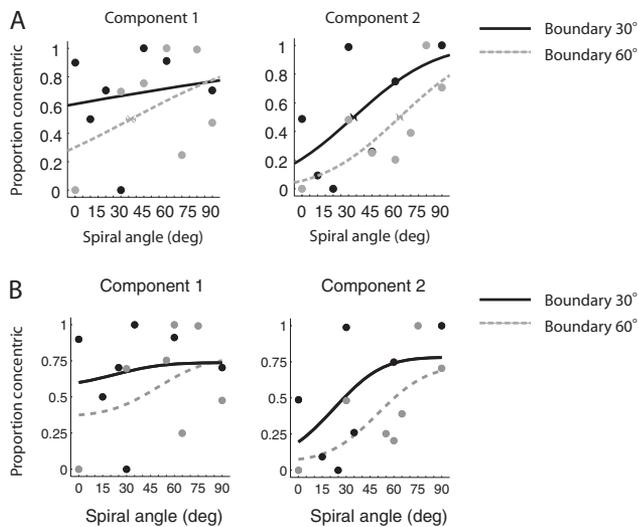
**Figure 5.** fMR-metric functions. fMR-metric functions based on the observers' behavioral choice for each session (30° and 60° boundary) for regions significantly correlated with (A) the early and (B) later EEG component. Mean classifier performance (proportion of trials classified as concentric for each stimulus condition) across observers is plotted after linear scaling (i.e., for each fMR-metric function, we scaled each prediction by subtracting the minimum value across conditions and dividing by the difference between maximum and minimum prediction values) to allow comparison between cortical areas. Average classifier prediction data are scaled from 0 to 1 and fitted with cumulative Gaussian functions for each session: 30° (black solid lines) and 60° (gray dotted lines) boundary. Nonscaled classification values (Supplementary Fig. S2) were consistent with our previous findings (Li et al. 2009). Error bars indicate the 95% confidence interval at 50% concentric threshold. Only ROIs with significantly fitted fMR-metric functions for both sessions are shown (maximum correlation:  $R = 0.98$ ,  $P < 0.001$ ; minimum correlation:  $R = 0.78$ ,  $P = 0.04$ ). The goodness of fit ( $R$ ,  $P$ ) for ROIs with nonsignificantly fitted fMR-metric functions in at least 1 of the 2 scanning sessions is reported in Supplementary Table S2.

### Control Analyses

We performed the following additional analyses to control for possible confounding factors. In particular, to control for the possibility that our results were due to random correlations in the data, we computed the fMR-metric and EEG-metric functions from randomly permuted signal patterns (i.e., we randomized the correspondence between the data and training labels and estimated the classifier prediction for each stimulus condition). The lack of significant correlations in these control analyses supports our interpretation for a link between task-relevant behavioral performance and neural preferences. Supporting evidence for this link comes from an additional analysis. In particular, fitting the fMRI (Fig. 7) and EEG (Fig. 6B) data using a scaled version of the psychometric function showed similar experience-dependent changes. Paired sample  $t$ -test analysis showed significant shift in EEG-metric functions for the second ( $t_6 = 2.59$ ,  $P < 0.05$ ) but not the first component ( $t_6 = 1.25$ ,  $P = 0.26$ ). A two-way repeated measures ANOVA

(ROI  $\times$  Boundary, Greenhouse-Geisser corrected) showed significant shift in fMR-metric functions for areas associated with the second ( $F_{1,5} = 54.14$ ,  $P = 0.001$ ) but not the first component ( $F_{1,5} = 0.95$ ,  $P = 0.374$ ).

The design of our study allowed us to rule out a number of less likely interpretations of our results. First, our design ensured that the observers were not biased in their responses by equating the number of conditions and stimuli across categories. As a result, the stimulus set tested in the 2 sessions could not remain identical when the category boundary changed across sessions. However, as in our previous study (Li et al. 2009), it is unlikely that our fMRI results were due to differences in stimulus conditions between sessions (i.e., boundary 30° vs. 60°). In particular, our design allowed us to directly compare between critical stimulus conditions (0°, 30°, 45°, 60°, and 90°) that were common in the 2 sessions. fMRI responses in these conditions across sessions (i.e., when the same stimuli were interpreted as the category boundary or not)



**Figure 6.** EEG-metric functions. (A) EEG-metric functions are shown for the first and second components for each session (30° and 60° boundary). The classifier performance at each condition was averaged across observers and fitted with cumulative Gaussian functions for each session. (B) Correlating psychometric and EEG-metric functions. As with the fMR-metric functions, we scaled the cumulative Gaussian model obtained from the psychophysical data to fit the classifiers predictions based on single-trial EEG data.

were similar suggesting that differences in the MVPA performance reflect the observers' behavioral choice rather than differences in the stimulus statistics across conditions. Furthermore, 2 additional MVPA analyses provide evidence that the shift observed in the fMR-metric functions reflects neural representations related to the observers' decision criterion rather than differences in the stimulus conditions. In a first analysis, we trained the classifier using the stimulus conditions rather than the behavioral response as labels. As in our previous work (Li et al. 2009), we trained the classifier on the extreme stimulus conditions and tested its accuracy on all the conditions. This analysis showed a shift in the fMR-metric functions, suggesting experience-dependent changes in neural representations related to the perceived stimulus categories (Supplementary Fig. S3). In a second analysis, we labeled the stimuli either based on the 30° or the 60° boundary. In particular, we trained and tested classifiers on all stimulus conditions (excluding the stimulus condition at the boundary) using cross-validation on independent data sets. Our results showed a shift in the fMR-metric functions only when the observers' and classifier's criterion matched (Supplementary Fig. S4), suggesting neural representations that reflect the observers' choice. When the observers' and classifier's criterion did not match (e.g., the observers were trained to use the 30° boundary, but the classifier used the 60° boundary for labeling the data), there was no significant shift in the fMR-metric functions (Supplementary Fig. S5).

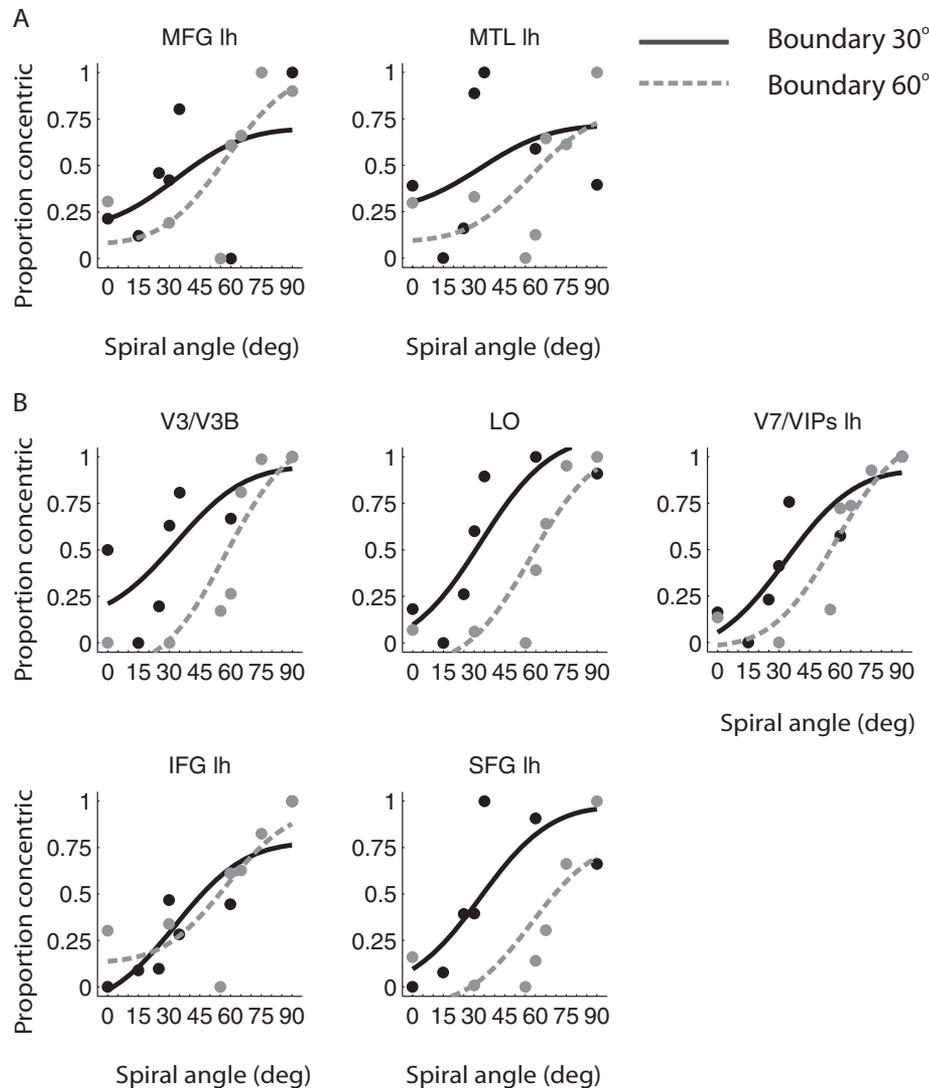
Second, it is unlikely that the experience-induced changes we observed resulted from learning specific category exemplars or local stimulus statistics, as the stimuli tested during scanning were randomized for presentation order and differed in their visual properties (i.e., signal level or spiral angle) from the stimuli presented during training. These manipulations ensured that observers learned the categorization boundary and generalized to new stimuli rather than learning specific stimulus examples for each condition. Further, we controlled

for the possibility that our fMRI results could be due to memorized stimulus-response associations by randomizing the motor responses based on the cue in the main experiment. Third, the experience-dependent changes we observed could not be due to differences in task difficulty across conditions, as the classification analysis compared trials associated with different stimuli (radial vs. concentric) rather than condition. Fourth, the cued-delay paradigm we used controlled for differences in the observers' response time. That is, observers made their decision during the delay after stimulus offset and waited for the cue before they could select the correct motor response, resulting in similar response times across stimulus conditions. As the stimulus-response association was randomized across trials, the motor response could not be anticipated on a given trial. Furthermore, a searchlight-based classification analysis on the button press used by the observers to indicate that their behavioral choice showed significant accuracies in motor regions but not in occipitotemporal, parietal, or prefrontal regions, suggesting that results in these areas cannot be simply explained on the basis of motor responses.

Furthermore, analysis of the fMRI responses (percentage signal change) across areas did not show any significant differences between the 2 fMRI sessions ( $F_{1,5} = 0.46$ ,  $P = 0.53$ ), suggesting that it is unlikely that the experience-dependent fMRI changes we observed were due to differences in attentional allocation between the 2 sessions (i.e., training may result in enhanced target salience and increased fMRI responses, or familiarity with the task may decrease fMRI responses). Finally, eye movement recordings during scanning showed that there were no significant differences in the eye position, number, or amplitude of saccades across stimulus conditions and session (Supplementary Fig. S6). A repeated measures ANOVA (Greenhouse-Geisser corrected) indicated that there was no significant difference between stimulus conditions on mean horizontal eye position (boundary 30°:  $F_{1,6,3,1} = 0.67$ ,  $P = 0.54$ ; boundary 60°:  $F_{1,2,2,5} = 0.73$ ,  $P = 0.50$ ), mean vertical eye position (boundary 30°:  $F_{1,9,3,8} = 3.19$ ,  $P = 0.15$ ; boundary 60°:  $F_{1,3,2,6} = 3.26$ ,  $P = 0.19$ ), mean saccade amplitude (boundary 30°:  $F_{1,8,3,6} = 1.60$ ,  $P = 0.31$ ; boundary 60°:  $F_{1,4,2,7} = 0.60$ ,  $P = 0.55$ ), or the number of saccades per trial per condition (boundary 30°:  $F_{1,5,3,1} = 0.80$ ,  $P = 0.49$ ; boundary 60°:  $F_{1,6,3,2} = 0.54$ ,  $P = 0.59$ ). Furthermore, no significant interactions were observed across boundaries and stimulus conditions on horizontal eye position ( $F_{1,8,3,7} = 1.19$ ,  $P = 0.39$ ), vertical eye position ( $F_{1,1,2,2} = 5.20$ ,  $P = 0.14$ ), saccade amplitude ( $F_{1,4,2,9} = 2.02$ ,  $P = 0.27$ ), and number of saccades ( $F_{1,6,3,2} = 0.09$ ,  $P = 0.88$ ) for the stimulus presentation period. These analyses suggest that it is unlikely that our results were significantly confounded by eye movements.

## Discussion

Combining behavioral measurements and simultaneous EEG-fMRI recordings, we provide evidence for spatiotemporal brain mechanisms that support our ability for flexible categorical decisions. Specifically, we demonstrate that cortical circuits engaged in later decision processes support our ability to learn new decision criteria for fine categorical shape judgments. Our work advances our understanding of the processes that mediate adaptive decision making beyond previous studies in 4 main respects.



**Figure 7.** Correlating psychometric and fMRI-metric functions. Similar to previous studies (Zenger-Landolt and Heeger 2003; Chandrasekaran et al. 2007), we scaled the cumulative Gaussian model obtained from the psychophysical data to fit the fMRI data (i.e., the predictions of the pattern classifier):  $y = B + \frac{S}{1 + \exp(\beta - \alpha x)}$ , where  $B$  is the baseline,  $S$  is the scale of the fitting,  $\beta$  is the intercept, and  $\alpha$  is the slope of the cumulative Gaussian model. Data are shown for regions significantly correlated with (A) the first EEG component and (B) the second EEG component for each session: 30° (black solid lines) and 60° (gray dotted lines) boundary. Only ROIs with significantly fitted fMRI-metric functions for both sessions are shown. The goodness of fit ( $R$ ,  $P$ ) for ROIs with nonsignificantly fitted fMRI-metric functions in at least 1 of the 2 scanning sessions is reported in Supplementary Table S3.

First, previous neuroimaging work has focused on identifying the spatial brain patterns that change with category learning. Previous studies have implicated a network of cortical and subcortical regions in category learning (for reviews, Keri 2003; Ashby and Maddox 2005; Poldrack and Foerd 2008; Seger and Miller 2010). In particular, frontal regions have been implicated in category and rule learning with ventrolateral prefrontal cortex suggested to monitor uncertainty in stimulus detection, discriminability, and probability of reward (Grinband et al. 2006; Philiastides and Sajda 2007; Daniel et al. 2010). Furthermore, we have recently shown that learning supports flexible categorical decisions by shaping the representation of visual categories not only in frontal but also posterior parietal and occipitotemporal areas in accordance with the perceived category boundary (Li et al. 2007, 2009). In contrast, here, we take advantage of the complementary high spatial and temporal resolution of simultaneous EEG-fMRI to identify the spatiotemporal

interactions between the circuits that mediate our ability for flexible category learning. Our findings demonstrate that learning flexible criteria for fine categorical judgments is implemented at later decision stages by tuning the readout of fine sensory signals from higher occipitotemporal areas and the evidence for categorical decisions accumulated by frontoparietal circuits (Newsome et al. 1989; Kim and Shadlen 1999; Shadlen and Newsome 2001; Heekeren et al. 2004; Grinband et al. 2006). In particular, behavioral criterion shifts in fine categorical judgments were related to changes in activation patterns associated with later decision processes in 1) prefrontal regions (SFG) involved in evaluating uncertainty and decision errors in fine judgments (Rushworth and Behrens 2008), 2) parietal regions suggested to contribute to the computation of decision variables (e.g., categorization criterion), and 3) higher occipitotemporal areas (V3B, LO) suggested to maintain information in short-term memory for comparative stimulus judgments

(Philiastides and Sajda 2007) relative to the category boundary (Grinband et al. 2006).

Second, previous work has debated whether visual categorization is mediated by fast feedforward processes (Thorpe et al. 1996; VanRullen and Thorpe 2001; Kirchner and Thorpe 2006) or feedback interactions (Bullier 2001; Bar 2007; Hegde 2008; Peyrin et al. 2010). Taken together, these previous studies and our findings suggest that different spatiotemporal circuits may relate to different categorization levels, that is, a feedforward circuit for coarse categorical judgments (i.e., basic level categorization) and a feedback circuit for fine categorical decisions (i.e., subordinate level categorization). Our findings support a feedback circuit for fine categorical judgments, consistent with previous studies showing that subordinate level categorization is mediated by later rather than early ERP components (Scott et al. 2006, 2008). In particular, we show that early processes recruit middle frontal and medial temporal regions implicated in fast contextual (Bar 2007) and semantic processing (Giesbrecht et al. 2004; Lee and Dapretto 2006; Ebisch et al. 2007), while later categorical judgments engage higher occipitotemporal and frontoparietal circuits known to be involved in fine categorical judgments. Furthermore, we demonstrate that categorization training modulates cortical circuits that support fine categorical processing at later stages rather than the early analysis of visual forms.

Third, our findings showing that category learning modulates shape processing in higher occipitotemporal (LO) but not early visual areas shed light on the contested role of visual cortex in category representations (Op de Beeck et al. 2001, 2006; Freedman et al. 2003; Jiang et al. 2007). We show that category learning modulates later processing in higher occipitotemporal areas related to fine perceptual judgments. This is consistent with previous studies suggesting that learning may rely on top-down influences (Ahissar and Hochstein 2004; Roelfsema and van Ooyen 2005; Seitz and Watanabe 2005; Zhang et al. 2008) that reweight the contribution of sensory areas and optimize the readout of behaviorally relevant signals (Doshier and Lu 1999; Li et al. 2004; Law and Gold 2008; Jacobs 2009). In contrast to previous physiology (Schoups et al. 2001; Li et al. 2004) and imaging studies (Schwartz et al. 2002; Furmanski et al. 2004; Kourtzi et al. 2005; Sigman et al. 2005; Mukai et al. 2007; Yotsumoto et al. 2008; Bao et al. 2010), learning did not modulate processing in V1 significantly. This finding could be due to the task employed (i.e., global form rather than local feature discrimination) and may relate to previous electrophysiological results that demonstrate enhanced learning effects in higher compared with primary visual areas (Yang and Maunsell 2004; Raiguel et al. 2006).

Fourth, our work provides novel methodological advances by combining simultaneous EEG-fMRI with pattern classification and applying this methodology to the study of category learning. Combining the high temporal and spatial resolution of EEG and fMRI allow us to discern interactions between cortical circuits involved in categorical decision making. Although previous studies (Scott et al. 2006, 2008; Philiastides and Sajda 2007; Peyrin et al. 2010) have recorded EEG and fMRI data at different sessions, simultaneous recordings avoid differences across sessions (e.g., alertness, adaptation, and familiarity) that confound learning effects. Furthermore, the EEG-informed fMRI analysis bypasses the source localization limitations of EEG related to the infinite possible source configurations that

may give rise to a given scalp distribution. Finally, comparing the choices of linear classifiers (EEG/fMRI-metric functions) with the observer's choices (psychometric functions) provides us with a sensitive tool for directly comparing brain activity and behavior and determining the link between adaptive human choices and experience-dependent brain plasticity (Pessoa and Padmala 2007; Li et al. 2009).

Using this methodology, we provide evidence for learning mechanisms that modify processing at distinct stages of categorical decision making. It is important to note that EEG-fMRI signals reflect processing at the level of large neural populations and do not allow us to discern whether category learning reflects changes in the selectivity of single neurons or correlations across neural populations. Furthermore, correlations between EEG and fMRI signals do not necessarily imply that the signals have the same underlying physiological source. However, recent work shows that trial-by-trial EEG analysis has the potential to identify correlated fMRI activity providing information about the cortical network that is engaged in specific temporal processes (Debener et al. 2005; Eichele et al. 2005; Mayhew et al. 2010). Despite these potential limitations, our findings make interesting predictions that can be further tested by electrophysiology. In particular, we suggest that making fine categorical judgments (i.e., based on behaviorally relevant features and task context) may benefit from the selective readout of highly sensitive signals in sensory areas that is optimized through top-down frontal influences. In sum, our findings propose that category learning acts on distinct spatiotemporal brain circuits to support our ability for flexible categorical judgment in the face of sensory uncertainty.

### Supplementary Material

Supplementary material can be found at: <http://www.cercor.oxfordjournals.org/>

### Funding

Biotechnology and Biological Sciences Research Council to Z.K. (H012508, D52199X, and E027436).

### Notes

We would like to thank A. P. Bagshaw for help with equipment set up. *Conflict of Interest*: None declared.

### References

- Ahissar M, Hochstein S. 2004. The reverse hierarchy theory of visual perceptual learning. *Trends Cogn Sci*. 8:457-464.
- Allen PJ, Josephs O, Turner R. 2000. A method for removing imaging artifact from continuous EEG recorded during functional MRI. *Neuroimage*. 12:230-239.
- Ashby FG, Maddox WT. 2005. Human category learning. *Annu Rev Psychol*. 56:149-178.
- Ashby FG, Maddox WT, Bohil CJ. 2002. Observational versus feedback training in rule-based and information-integration category learning. *Mem Cognit*. 30:666-677.
- Badre D, D'Esposito M. 2009. Is the rostro-caudal axis of the frontal lobe hierarchical? *Nat Rev Neurosci*. 10:659-669.
- Banko EM, Gal V, Kortvelyes J, Kovacs G, Vidnyanszky Z. 2011. Dissociating the effect of noise on sensory processing and overall decision difficulty. *J Neurosci*. 31:2663-2674.
- Bao M, Yang L, Rios C, He B, Engel SA. 2010. Perceptual learning increases the strength of the earliest signals in visual cortex. *J Neurosci*. 30:15080-15084.

- Bar M. 2007. The proactive brain: using analogies and associations to generate predictions. *Trends Cogn Sci.* 11:280-289.
- Bullier J. 2001. Integrated model of visual processing. *Brain Res Brain Res Rev.* 36:96-107.
- Chandrasekaran C, Canon V, Dahmen JC, Kourtzi Z, Welchman AE. 2007. Neural correlates of disparity-defined shape discrimination in the human brain. *J Neurophysiol.* 97:1553-1565.
- Cover TM, Thomas JA. 1991. *Elements of information theory.* New York: Wiley Interscience.
- Daniel R, Wagner G, Koch K, Reichenbach JR, Sauer H, Schlosser RG. 2011. Assessing the neural basis of uncertainty in perceptual category learning through varying levels of distortion. *J Cogn Neurosci.* 23:1781-1793.
- Das K, Giesbrecht B, Eckstein MP. 2010. Predicting variations of perceptual performance across individuals from neural activity using pattern classifiers. *Neuroimage.* 51:1425-1437.
- Debener S, Strobel A, Sorger B, Peters J, Kranczioch C, Engel AK, Goebel R. 2007. Improved quality of auditory event-related potentials recorded simultaneously with 3-T fMRI: removal of the ballistocardiogram artefact. *Neuroimage.* 34:587-597.
- Debener S, Ullsperger M, Siegel M, Fiehler K, von Cramon DY, Engel AK. 2005. Trial-by-trial coupling of concurrent electroencephalogram and functional magnetic resonance imaging identifies the dynamics of performance monitoring. *J Neurosci.* 25:11730-11737.
- Delorme A, Makeig S. 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods.* 134:9-21.
- Doshier BA, Lu ZL. 1999. Mechanisms of perceptual learning. *Vision Res.* 39:3197-3221.
- Duncan J. 2001. An adaptive coding model of neural function in prefrontal cortex. *Nat Rev Neurosci.* 2:820-829.
- Ebisch SJ, Babiloni C, Del Gratta C, Ferretti A, Perrucci MG, Caulo M, Sitskoorn MM, Romani GL. 2007. Human neural systems for conceptual knowledge of proper object use: a functional magnetic resonance imaging study. *Cereb Cortex.* 17:2744-2751.
- Eichele T, Specht K, Moosmann M, Jongsma ML, Quiroga RQ, Nordby H, Hugdahl K. 2005. Assessing the spatiotemporal evolution of neuronal activation with single-trial event-related potentials and functional MRI. *Proc Natl Acad Sci U S A.* 102:17798-17803.
- 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.
- Furmanski CS, Schluppeck D, Engel SA. 2004. Learning strengthens the response of primary visual cortex to simple patterns. *Curr Biol.* 14:573-578.
- Giesbrecht B, Camblin CC, Swaab TY. 2004. Separable effects of semantic priming and imageability on word processing in human cortex. *Cereb Cortex.* 14:521-529.
- Goldstone RL, Lippa Y, Shiffrin RM. 2001. Altering object representations through category learning. *Cognition.* 78:27-43.
- Goldstone RL, Steyvers M. 2001. The sensitization and differentiation of dimensions during category learning. *J Exp Psychol Gen.* 130:116-139.
- Grinband J, Hirsch J, Ferrera VP. 2006. A neural representation of categorization uncertainty in the human brain. *Neuron.* 49:757-763.
- 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.
- Grossman M, Smith EE, Koenig P, Glosser G, DeVita C, Moore P, McMillan C. 2002. The neural basis for categorization in semantic memory. *Neuroimage.* 17:1549-1561.
- Haynes JD, Rees G. 2005. Predicting the orientation of invisible stimuli from activity in human primary visual cortex. *Nat Neurosci.* 8:686-691.
- Heekeren HR, Marrett S, Bandettini PA, Ungerleider LG. 2004. A general mechanism for perceptual decision-making in the human brain. *Nature.* 431:859-862.
- Hegde J. 2008. Time course of visual perception: coarse-to-fine processing and beyond. *Prog Neurobiol.* 84:405-439.
- Hyvarinen A, Oja E. 1997. A fast fixed-point algorithm for independent component analysis. *Neural Comput.* 9:1483-1492.
- Jacobs RA. 2009. Adaptive precision pooling of model neuron activities predicts the efficiency of human visual learning. *J Vis.* 9(22):1-15.
- 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.
- Johnson JS, Olshausen BA. 2003. Timecourse of neural signatures of object recognition. *J Vis.* 3:499-512.
- Jung TP, Makeig S, Westerfield M, Townsend J, Courchesne E, Sejnowski TJ. 2000. Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects. *Clin Neurophysiol.* 111:1745-1758.
- Kamitani Y, Tong F. 2005. Decoding the visual and subjective contents of the human brain. *Nat Neurosci.* 8:679-685.
- Keri S. 2003. The cognitive neuroscience of category learning. *Brain Res Brain Res Rev.* 43:85-109.
- Kim JN, Shadlen MN. 1999. Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. *Nat Neurosci.* 2:176-185.
- Kirchner H, Thorpe SJ. 2006. Ultra-rapid object detection with saccadic eye movements: visual processing speed revisited. *Vision Res.* 46:1762-1776.
- Koenig P, Smith EE, Glosser G, DeVita C, Moore P, McMillan C, Gee J, Grossman M. 2005. The neural basis for novel semantic categorization. *Neuroimage.* 24:369-383.
- Kourtzi Z, Betts LR, Sarkheil P, Welchman AE. 2005. Distributed neural plasticity for shape learning in the human visual cortex. *PLoS Biol.* 3:e204.
- Laufs H, Daunizeau J, Carmichael DW, Kleinschmidt A. 2008. Recent advances in recording electrophysiological data simultaneously with magnetic resonance imaging. *Neuroimage.* 40:515-528.
- Law CT, Gold JI. 2008. Neural correlates of perceptual learning in a sensory-motor, but not a sensory, cortical area. *Nat Neurosci.* 11:505-513.
- Lee SS, Dapretto M. 2006. Metaphorical vs. literal word meanings: fMRI evidence against a selective role of the right hemisphere. *Neuroimage.* 29:536-544.
- Li S, Mayhew SD, Kourtzi Z. 2009. Learning shapes the representation of behavioral choice in the human brain. *Neuron.* 62:441-452.
- Li S, Ostwald D, Giese M, Kourtzi Z. 2007. Flexible coding for categorical decisions in the human brain. *J Neurosci.* 27:12321-12330.
- Li W, Piech V, Gilbert CD. 2004. Perceptual learning and top-down influences in primary visual cortex. *Nat Neurosci.* 7:651-657.
- Maddox WT, Ashby FG. 1993. Comparing decision bound and exemplar models of categorization. *Percept Psychophys.* 53:49-70.
- Mayhew SD, Li S, Storrar JK, Tsvetanov KA, Kourtzi Z. 2010. Learning shapes the representation of visual categories in the aging human brain. *J Cogn Neurosci.* 22:2899-2912.
- Miller EK. 2000. The prefrontal cortex and cognitive control. *Nat Rev Neurosci.* 1:59-65.
- Montemurro MA, Rasch MJ, Murayama Y, Logothetis NK, Panzeri S. 2008. Phase-of-firing coding of natural visual stimuli in primary visual cortex. *Curr Biol.* 18:375-380.
- Mukai I, Kim D, Fukunaga M, Japee S, Marrett S, Ungerleider LG. 2007. Activations in visual and attention-related areas predict and correlate with the degree of perceptual learning. *J Neurosci.* 27:11401-11411.
- Newsome WT, Britten KH, Movshon JA. 1989. Neuronal correlates of a perceptual decision. *Nature.* 341:52-54.
- Niazy RK, Beckmann CF, Iannetti GD, Brady JM, Smith SM. 2005. Removal of fMRI environment artifacts from EEG data using optimal basis sets. *Neuroimage.* 28:720-737.
- Nosofsky RM. 1986. Attention, similarity, and the identification-categorization relationship. *J Exp Psychol.* 115:39-61.
- Ohla K, Busch NA, Dahlem MA, Herrmann CS. 2005. Circles are different: the perception of Glass patterns modulates early event-related potentials. *Vision Res.* 45:2668-2676.
- Op de Beeck H, Wagemans J, Vogels R. 2001. Inferotemporal neurons represent low-dimensional configurations of parameterized shapes. *Nat Neurosci.* 4:1244.

- 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.
- Ostwald D, Lam JM, Li S, Kourtzi Z. 2008. Neural coding of global form in the human visual cortex. *J Neurophysiol*. 99:2456-2469.
- Panzeri S, Senatore R, Montemurro MA, Petersen RS. 2007. Correcting for the sampling bias problem in spike train information measures. *J Neurophysiol*. 98:1064-1072.
- Panzeri S, Treves A. 1996. Analytical estimates of limited sampling biases in different information measures. *Network-Comput Neural Syst*. 7:87-107.
- Pei F, Pettet MW, Vildavski VY, Norcia AM. 2005. Event-related potentials show configural specificity of global form processing. *Neuroreport*. 16:1427-1430.
- Pessoa L, Padmala S. 2007. Decoding near-threshold perception of fear from distributed single-trial brain activation. *Cereb Cortex*. 17:691-701.
- Peyrin C, Michel CM, Schwartz S, Thut G, Seghier M, Landis T, Marendaz C, Vuilleumier P. 2010. The neural substrates and timing of top-down processes during coarse-to-fine categorization of visual scenes: a combined fMRI and ERP study. *J Cogn Neurosci*. 22:2768-2780.
- Philiastides MG, Ratcliff R, Sajda P. 2006. Neural representation of task difficulty and decision making during perceptual categorization: a timing diagram. *J Neurosci*. 26:8965-8975.
- Philiastides MG, Sajda P. 2006. Temporal characterization of the neural correlates of perceptual decision making in the human brain. *Cereb Cortex*. 16:509-518.
- Philiastides MG, Sajda P. 2007. EEG-informed fMRI reveals spatiotemporal characteristics of perceptual decision making. *J Neurosci*. 27:13082-13091.
- Poldrack RA, Foerd K. 2008. Category learning and the memory systems debate. *Neurosci Biobehav Rev*. 32:197-205.
- Raiguel S, Vogels R, Mysore SG, Orban GA. 2006. Learning to see the difference specifically alters the most informative V4 neurons. *J Neurosci*. 26:6589-6602.
- Roelfsema PR, van Ooyen A. 2005. Attention-gated reinforcement learning of internal representations for classification. *Neural Comput*. 17:2176-2214.
- Rossion B, Collins D, Goffaux V, Curran T. 2007. Long-term expertise with artificial objects increases visual competition with early face categorization processes. *J Cogn Neurosci*. 19:543-555.
- Rushworth MF, Behrens TE. 2008. Choice, uncertainty and value in prefrontal and cingulate cortex. *Nat Neurosci*. 11:389-397.
- Schoups A, Vogels R, Qian N, Orban G. 2001. Practising orientation identification improves orientation coding in V1 neurons. *Nature*. 412:549-553.
- Schwartz S, Maquet P, Frith C. 2002. Neural correlates of perceptual learning: a functional MRI study of visual texture discrimination. *Proc Natl Acad Sci U S A*. 99:17137-17142.
- Scott LS, Tanaka JW, Sheinberg DL, Curran T. 2006. A reevaluation of the electrophysiological correlates of expert object processing. *J Cogn Neurosci*. 18:1453-1465.
- Scott LS, Tanaka JW, Sheinberg DL, Curran T. 2008. The role of category learning in the acquisition and retention of perceptual expertise: a behavioral and neurophysiological study. *Brain Res*. 1210:204-215.
- Seger CA, Miller EK. 2010. Category learning in the brain. *Annu Rev Neurosci*. 33:203-219.
- Seitz AR, Watanabe T. 2005. A unified model for perceptual learning. *Curr Opin Neurobiol*. 9:329-334.
- Shadlen MN, Newsome WT. 2001. Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *J Neurophysiol*. 86:1916-1936.
- Shannon CE. 1948. A mathematical theory of communication. *Bell Syst Tech J*. 27:623-656.
- Sigman M, Pan H, Yang Y, Stern E, Silbersweig D, Gilbert CD. 2005. Top-down reorganization of activity in the visual pathway after learning a shape identification task. *Neuron*. 46:823-835.
- Smith EE, Jonides J. 1997. Working memory: a view from neuroimaging. *Cogn Psychol*. 33:5-42.
- Srivastava G, Crottaz-Herbette S, Lau KM, Glover GH, Menon V. 2005. ICA-based procedures for removing ballistocardiogram artifacts from EEG data acquired in the MRI scanner. *Neuroimage*. 24:50-60.
- Tanskanen T, Saarinen J, Parkkonen L, Hari R. 2008. From local to global: cortical dynamics of contour integration. *J Vis*. 8(7):15.1-15.12.
- Thorpe S, Fize D, Marlot C. 1996. Speed of processing in the human visual system. *Nature*. 381:520-522.
- Toni I, Rushworth MF, Passingham RE. 2001. Neural correlates of visuomotor associations. Spatial rules compared with arbitrary rules. *Exp Brain Res*. 141:359-369.
- VanRullen R, Thorpe SJ. 2001. The time course of visual processing: from early perception to decision-making. *J Cogn Neurosci*. 13:454-461.
- Wichmann FA, Hill NJ. 2001. The psychometric function: II. Bootstrap-based confidence intervals and sampling. *Percept Psychophys*. 63:1314-1329.
- Yang T, Maunsell JH. 2004. The effect of perceptual learning on neuronal responses in monkey visual area V4. *J Neurosci*. 24:1617-1626.
- Yotsumoto Y, Watanabe T, Sasaki Y. 2008. Different dynamics of performance and brain activation in the time course of perceptual learning. *Neuron*. 57:827-833.
- Zenger-Landolt B, Heeger DJ. 2003. Response suppression in v1 agrees with psychophysics of surround masking. *J Neurosci*. 23:6884-6893.
- Zhang J, Kuai S, Xiao L, Klein S, Levi D, Yu C. 2008. Stimulus coding rules for perceptual learning. *PLoS Biol*. 6:e197.