

Timing, timing, timing: Fast decoding of object information from intracranial field potentials in human visual cortex

Hesheng Liu^{1,3}, Yigal Agam¹, Joseph R. Madsen², Gabriel Kreiman^{1,4,5}

¹Department of Neuroscience and Ophthalmology, Children's Hospital Boston, Harvard Medical School

²Department of Neurosurgery, Children's Hospital Boston, Harvard Medical School

³Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital

⁴Center for Brain Science, Harvard University

⁵Swartz Center for Theoretical Neuroscience, Harvard University

Correspondence: gabriel.kreiman@tch.harvard.edu

SUPPLEMENTARY MATERIAL

1. Supplementary Methods
2. Supplementary Tables
3. Supplementary Figures
4. See also Supplementary Web Figures at:

http://klab.tch.harvard.edu/resources/timing_timing_timing/index.htm

1. Supplementary Methods

Stimulus presentation

Subjects were presented with contrast-normalized grayscale photographs of objects from 5 different categories: “animals”, “chairs”, “human faces”, “fruits” and “vehicles” (Figure S1). Throughout the text, the physiological responses to the different object categories are denoted by different colors. In 4 subjects, there were also 2 additional object categories: “artificial Lego objects” and “shoes”. For consistency across subjects, here we describe only the responses to the 5 categories that were common to all subjects (none of the conclusions about selectivity, invariance and latencies would be altered upon addition of those extra categories but there would be more selective electrodes in Table S1. Images were presented on a laptop computer (using the MATLAB Psychophysics toolbox (Pelli, 1997)). Images were shown in pseudo-random order. Within each category, there were 5 different exemplars. A key aspect of visual recognition involves combining selectivity with invariance to object transformations. To assess the robustness to object changes in the physiological responses, each exemplar was shown in one of 5 possible transformations: (1) 3 degrees size, “default” viewpoint, (2) 1.5 degrees size, “default” viewpoint, (3) 6 degrees size, “default” viewpoint, (4) 3 degrees size, ~45 degree depth rotation with respect to the “default” viewpoint, (5) 3 degrees size, ~90 degree depth rotation with respect to the “default” viewpoint. In the case of faces, the “default” viewpoint was a front view; for the other categories, the “default” viewpoint is defined by the images in Figure S1B. An example of these 5 different transformations for one exemplar object is shown in Figure S1C. Each image was presented for 200 ms and was followed by a 600 ms gray screen. In order to ensure that the subjects were paying attention to the object presentation, they performed a one-back task, indicating whether an exemplar was repeated or not (regardless of the object transformations) by pressing one of two pre-assigned mouse buttons. There were 60 image presentations per block and the first 2 presentations in each block were discarded from analyses to avoid potential non-stationary transient effects. Images were repeated

multiple times; the exact number of times depended on the subject's fatigue. We only analyzed data where we had at least 5 repetitions per image and where each recording session was at least 5 minutes long (session duration statistics: min=7 mins, max=40 mins, average=19.6±7.9 mins; number of sessions statistics: min=1, max=7, average=2.6±1.6). Data from all sessions for a given subject were pooled together for analyses.

Data Analyses

Data preprocessing

We followed the procedures in (Kreiman et al., 2006) to reduce the impact of potential electrical and movement artifacts. The raw data was digitally filtered between 1 Hz and 100 Hz (4-pole Butterworth filter). We computed the power spectrum of the intracranial field potential (IFP) signal for each electrode (Press et al., 1996). Electrodes that showed evidence of electrical noise (such as sharp peaks at 60 Hz or harmonics) were excluded from further analyses. In order to remove potential movement artifacts, we computed the overall distribution of the total IFP power in all trials for each electrode (regardless of object category or transformation) and excluded from further analyses those trials where the power was more than 4 standard deviations from the mean in at least 3 electrodes.

Electrode localization

To localize the electrodes on the brain surface, we integrated the anatomical information of the brain provided by Magnetic Resonance Imaging (MRI) and the spatial information of the electrode positions provided by Computer Tomography (CT). MR images were usually acquired several weeks before the intracranial electrode implantation and CT images were acquired ~1-2 days after the implantation. The 3-D brain surface was reconstructed for each subject. An automatic parcellation was performed using Freesurfer (Dale et al., 1999; Fischl et al., 2004). This brain parcellation contains 80 regions in each hemisphere (Table S2 and (Desikan et al., 2006)). A preliminary co-registration of CT and MRI was implemented using SPM (Ashburner and Friston, 1997;

Ashburner and Friston, 2000). The co-registration was then fine-tuned using Freesurfer. In the final stage, we recorded the Talairach coordinates of the electrodes and determined the brain region for each electrode according to the parcellation. Based on these coordinates, the electrodes were superimposed on the reconstructed brain surface for visualization purposes in the Figures. We report Talairach coordinates for the example electrodes and Talairach coordinates statistics for each brain region in Table S2. Coordinates for each one of the 912 electrodes are available upon request. In all the figures and tables that discuss electrode positions throughout the text, we pooled together locations in the right and left hemispheres (except when reporting Talairach coordinates in Table S2).

Classifier analysis

Classifier input. Previous studies of IFP data focused on computing average responses over large number of presentations (e.g. (Fell et al., 2001; McCarthy et al., 1999; Privman et al., 2007)). This averaging approach occludes the challenges due to the intrinsic variability in the data. Furthermore, the brain must be able to extract information about the visual world *in real time*. Therefore, a single-trial analysis provides a more realistic description of the encoding and decoding of visual information. We used a statistical learning approach (Hung et al., 2005; Vapnik, 1995) to read out from the IFP signals, on a single-trial basis, information about the visual stimulus presented to the subjects. Let ${}^r x(t)$ represent the IFP signal recorded from electrode u ($u=1, \dots, N_{elec}$) on trial r ($r=1, \dots, N_{rep}$) where t denotes the time from stimulus onset, N_{elec} is the number of electrodes and N_{rep} is the number of repetitions. We define the IFP response over a time interval T as ${}^r s_T = {}^r f_T(x)$. The function f allows us to quantitatively compare different possible neural codes (see “Neural codes” below and Figure S3). Unless stated otherwise, throughout the text we refer to the “IFP response”, $f_T(x)$, defined as the range of the signal ($\max(x) - \min(x)$) in the interval T . The default T was [50;300] ms with respect to stimulus onset. Figure S3 compares the results for different definitions of the interval T and the function $f_T(x)$. The classifier approach allows us to consider each electrode independently

(Figures 1, 4 and Figures S1-S3) or to consider the encoding of information by an ensemble of multiple electrodes (Figures 2, 3, 5, 6 and Figures S4, S5 and S7). For a population of N_{elec} electrodes, we assumed independence and constructed a population vector by concatenating the responses of each electrode: ${}^r\mathbf{S}_T = [{}^r_1s_T, \dots, {}^r_{N_{elec}}s_T]$. We considered ensembles consisting of data from multiple electrodes recorded simultaneously in one subject or ensembles consisting of data taken from a pool of electrodes across subjects (Carmena et al., 2003; Hung et al., 2005a; Santhanam et al., 2006) (see “Feature Selection” below).

For a given repetition, in addition to the physiological response ${}^r\mathbf{S}_T$, we define an indicator variable ${}^r\lambda$ that denotes the label of the object presented to the subject. Unless stated otherwise, the results shown throughout the manuscript correspond to binary classification between a given object category and the other object categories. In this case, ${}^r\lambda = +1$ for the given category and ${}^r\lambda = -1$ for the other object categories. In the case of multiclass problems (when ${}^r\lambda$ can take more than 2 possible values), we use a one-versus-all approach (Hung et al., 2005; Rifkin and Klautau, 2004). In Figure 1D, 1I and 4B), we show the distribution of responses for those trials with ${}^r\lambda = +1$ (e.g. “fruits” in the example in Figure 1D) and those trials with ${}^r\lambda = -1$ (“rest”). We also build an ROC curve (e.g. Figure 1E, Figure 1J, Figure 4B insets), showing the proportion of correct detections (P_{CD} , y-axis) as a function of the probability of false alarms (P_{FA} , x-axis) for different thresholds on the distributions shown in Figures 1D, 1I and 4B (Green and Swets, 1966; Kreiman et al., 2000).

The input to the classifier consisted of the example pairs $\{{}^r\mathbf{S}_T; {}^r\lambda\}$. Importantly, the data were divided into two non-overlapping sets, a *training* set and a *test* set. The division into a training set and test set depended on the specific question being asked (see “Separation of train/test data” below for details). The classifier was trained to learn the map between the physiological signals and the labels. The performance of the classifier was evaluated by comparing the classifier predictions on the test data with the actual labels for the test data. Throughout the text, we report the proportion of *test* repetitions

correctly labeled as “Classification Performance”. In the case of binary classification, the chance level is 50% and is indicated in the plots by a dashed line. In the case of multiclass problems (e.g. Figure S4B), the chance level is 20% (5 categories). The maximum classification performance is 100% both for binary classification and multiclass classification.

Classifiers. Unless stated otherwise, we used a Support Vector Machine (SVM) classifier with a linear kernel (Cristianini and Shawe-Taylor, 2000; Hung et al., 2005b; Vapnik, 1995). Results for other machine-learning classifiers are presented in the Supplementary Web Material. For a linear kernel, the classifier boundary can be expressed as $c(\mathbf{S}) = \text{sign}[\mathbf{w} \cdot \mathbf{S}]$ where the vector \mathbf{S} denotes the ensemble response as defined above, the weights \mathbf{w} are learnt during training and the dot indicates a dot product. An example of the discrimination of data from two different categories using a linear classifier is shown in the Supplementary Web Material. The classifier finds a line (a hyperplane in higher dimensions) that will separate the two classes of data. The SVM algorithm finds this line by minimizing the “margin” between the classification boundary and the data (Cristianini and Shawe-Taylor, 2000; Vapnik, 1995). We chose to use a linear SVM because it is known to show good performance while the results remain relatively easy to interpret. We also explored several other classifiers. In the Supplementary Web Material, we compare the classification performance values for a the linear SVM against the classification performance values of an SVM classifier with a Gaussian Kernel or a Nearest Neighbor classifier (Bishop, 1995). Overall, the performance for the different classifiers that we examined was quite similar.

Separation of train/test data. In *all* cases, the training and test data were separated to avoid overfitting. The way of separating the data into a training set and a test set depends on the specific question asked in each analysis. In Figures 1 through 3 in the main text, we randomly chose 70% of the responses for each object category to train the classifier and used the remaining 30% of the responses to test the classification performance. This

analysis evaluates the selectivity in the data across categories by assessing whether the responses *in single trials* during some of the repetitions (70%) are sufficiently reliable to support discrimination of the object categories in the remaining 30% of the data. In Figures 4 through 6, the training data consisted of the responses to one of the 5 possible object transformations and the test data corresponded to the responses to the other object transformations. There are $\binom{5}{2} = 10$ possible combinations. For simplicity, throughout the text, we show only the classification performance when training on the IFP responses to the default viewpoint and size and testing on the responses to the other 4 transformations. The classification performance results for the other possible train/test combinations were similar. This analysis evaluates the degree of invariance to changes in the objects' scale and rotation.

Neural codes. As noted above, the function f allows us to quantitatively compare different decoding schemes through ${}^r_u s_T = f_T({}^r_u x)$. We considered several different possible definitions for f (always restricted to the interval T ; Figure S3A): (i) total power of the IFP signal, (ii) range of the IFP signal (default throughout the text), (iii) power in the (0,10] Hz frequency band, (iv) power in the (10,20] Hz frequency band, (v) power in the (20,50] Hz frequency band, (vi) power in the (70;100] Hz frequency band. We observed that there was more selectivity in the high gamma frequency band (Figure S3A; 71 to 100 Hz). We also considered several different possibilities for the interval $T=[t_b, t_a]$ where the parameter t_b could be 50, 100 or 200 ms after stimulus onset and the parameter t_a could be 200, 400, 600 or 800 ms after stimulus onset (Figure S3B-C). Unless stated otherwise, throughout the text we report the classification performance for code (ii) above (signal range) in the window $T=[50;300]$ ms with respect to stimulus onset. The time window from 50 to 300 ms was chosen so as to account for the approximate minimum latency of the neural responses (Richmond et al., 1983; Schmolesky et al., 1998) while remaining within the average saccade time (Rayner, 1998).

Given the high temporal resolution of our recordings, we were particularly interested in estimating how fast we could decode category information (Figure 3B, 6B and Figure S7). For this purpose, we considered windows of size τ ($\tau = 25, 50$ or 100 ms). In each window, we computed the IFP signal range and we assessed the performance of the classifier as described above by moving the individual windows from 0 to 800 ms after stimulus onset (we also show the performance using the signal power in each window in the Supplementary Web Material).

Null hypothesis. In order to assess the probability that a given classification performance level could be achieved by chance, we defined the null hypothesis by randomizing the object category labels and repeating the same analysis. We performed 100 iterations for each electrode. The distribution of classification performance values under the null hypothesis for all electrodes is shown in (Figure S2A). Throughout the text, we use a threshold criterion of 3 standard deviations from the null hypothesis for statistical significance.

Feature selection. The classifier analysis approach outlined above can be applied to any number of electrodes. We defined neural ensembles in several different ways. In some cases, we analyzed neural ensembles by drawing the electrodes according to their location (e.g. Figure 2, Figure 5B, Figure S5). In other cases (Figure 3B, Figure 6B, Figure S4, Figure S7), we selected electrodes by performing a simple feature selection step: for each electrode, we computed the ratio of the variance across categories to the variance within categories, r_v , which is also the measure used in ANOVA tests (Keeping, 1995). For each subject, electrodes were sorted based on their r_v values and the ensemble was formed by selecting the electrodes with the higher r_v values in each subject. The feature selection step was performed using only the training data.

Identification versus categorization. The analyses in the text correspond to *categorization* performance. When asking about the category presented in each trial, the

indicator variable λ only depends on the category of the object present in the image and not on the particular exemplar. We compared the classification performance for categorization against the performance for *identification* where λ denoted the identity of the specific exemplars within the preferred category (Supplementary Web Material). Except for the indicator variable, the other steps of the classifier analysis were the same as the ones described above for categorization. The training data consisted of 70% of the repetitions of each exemplar and the test data consisted of the remaining 30%. We also considered arbitrary categories defined by randomly drawing exemplars from all categories (Supplementary Web Material).

Other statistical analyses

Reproducibility across repetitions. We assessed the degree of reproducibility in the IFP responses across multiple repetitions of the same images. For each electrode and each exemplar object, we considered the time course of the IFP signal between 0 and 800 ms in each repetition after resampling with a 20 ms rolling average window. We computed the Pearson correlation coefficient between the resampled IFP signals for every pair of repetitions for each exemplar. To assess the statistical significance of these Pearson correlation coefficients, we computed their expected distribution according to a null hypothesis defined by performing 500 random shuffles of the object labels. The results of these analyses are shown in the Supplementary Web Material.

Selectivity based on ANOVA. The key analyses in the current manuscript involve assessing the performance of a statistical classifier in single trials (see main text and “Classifier analysis” above). We also performed a one-way analysis of variance (ANOVA) on the IFP responses. For example, in the case of assessing visual selectivity (Figure 1), we asked whether the variance in the responses *across* object categories was significantly larger than the variance in the responses *within* object categories. The IFP response was defined as the signal range (maximum – minimum) in the [50;300] ms interval with respect to stimulus onset. The ANOVA results were consistent with the

single electrode statistical classifier results (Supplementary Web Material). Throughout the text, we report the results from the statistical classifier because they provide a more rigorous quantitative approach and a stricter criterion for selectivity, they allow considering ensembles of electrodes in a straightforward way and they provide a direct measure of performance in single trials.

Latency based on ANOVA. As described above, we used the classifier approach to determine the earliest time point when we could discriminate among the object categories in single trials (Figure 3B and Figure 6B). In addition, we used a point-by-point ANOVA test across object categories to define the latency of the physiological responses (Thorpe et al., 1996). We defined an electrode to be “selective” if there were $seln_{cons}=25$ consecutive time points where the ANOVA test yielded $p<0.01$. The latency was defined, for the selective electrodes, as the first time point where $latn_{cons}=10$ consecutive time points yielded an ANOVA $p<0.01$ (Figure 6A). The latencies are reported in Figure 3A and Figure 6A. The dependency of the latency on $seln_{cons}$, $latn_{cons}$ and p is shown in Figure S6B-C. To assess the statistical significance of the latencies reported in the text, we randomly shuffled the object categories and repeated the point-by-point ANOVA tests. With the parameters above ($seln_{cons}=25$ and $p<0.01$), there was no selective electrode in 1000 iterations.

Identification versus categorization, another look. In order to assess whether category selectivity could be explained by selectivity to particular exemplars within a category, we rank ordered the responses to all the individual exemplars (e.g. Figure 1C). As a null hypothesis, we assumed that the distribution of rank orders across categories was uniform and we assessed deviations from this null hypothesis by randomly assigning exemplars to categories in 10,000 iterations. The p values quoted in the text for this analysis indicate the odds that m exemplars from any category would rank among the top n best responses.

2. Supplementary Tables

Table S1: Location of selective and invariant electrodes

We report here the number of electrodes from each area where we recorded from at least 10 electrodes. The areas and the “area code” numbers correspond to the parcellation maps in (Dale et al., 1999; Fischl et al., 2004) (see “Electrode localization” above and Table S2). Because there were other areas with < 10 electrodes which are not included in this table, the sum of the number of electrodes reported for each lobe (column labeled “Sum”) is somewhat less than the total number of electrodes in that lobe (column labeled “Total”). The column labeled “Fract” is the ratio of “Sum”/ “Total”. An electrode was defined as selective if the classification performance was more than 3 standard deviations from the null hypothesis; see Experimental Procedures and Supplementary Methods). Examples of the responses from selective electrodes are shown in Figure 1 and Figure S1D-G. “% sel” refers to the percentage of selective electrodes with respect to the total number of electrodes in each area. “inv scale” and “inv rotation” indicate the number of electrodes that showed invariance to scale and rotation respectively based on a classification performance beyond 3 standard deviations of the null hypothesis mean when extrapolating across object scales or rotations (see Experimental Procedures and Supplementary Methods and an example electrode in Figure 4). The “% inv scale” and “% inv rotation” refer to the percentage of electrodes that showed invariance to these two transformations with respect to the number of selective electrodes in each area. We only report these percentages in the table when there were at least 5 selective electrodes (those columns are bolded). Data from the right and left hemispheres were pooled in this Table.

Table S2: List of areas and codes

To localize the electrodes on the brain surface, we integrated the anatomical information of the brain provided by Magnetic Resonance Imaging (MRI) and the spatial information of the electrodes provided by Computer Tomography (CT). For each subject, the 3-D brain surface was reconstructed and then an automatic parcellation was performed using

Freesurfer (Dale et al., 1999; Fischl et al., 2004). This brain parcellation contains 80 regions in each hemisphere (Desikan et al., 2006). This table lists the code, name and abbreviation for each of these 80 regions in each hemisphere. These area code numbers and abbreviations are used throughout the text and figures. We also computed the Talairach coordinates for each one of the 912 electrodes. Here we show some basic statistics about the Talairach coordinates for the electrodes in each brain region. This should be taken only as a rough approximation because it is not clear that computing means, standard deviations, and similar operations are valid in Talairach space. The number of electrodes in this Table is 890 instead of 912; there were 22 electrodes for which we could not obtain reliable Talairach coordinates.

3. Supplementary Figures

Figure S1: Stimulus presentation scheme and further examples of selectivity in intracranial field potentials

A. Grayscale images were presented for 200 ms, with a 600 ms gray screen in between images. Subjects were instructed to fixate and performed a one-back task indicating whether the same object was repeated (regardless of changes in scale or viewpoint). Image order was pseudo-randomized. B. Objects belonged to one of 5 possible categories. Here and throughout the manuscript, each category is indicated by a separate color (**red=animals**, **green=chairs**, **blue=faces**, **black=fruits**, **yellow=vehicles**). There were 5 exemplar objects per category. These images show the “default” viewpoint for each exemplar. The duration of each session depended on clinical constraints and subject fatigue (min duration = 7 min., max duration = 40 min, mean=19.6±7.9 mins). In many cases we were able to record several sessions per subject (min=1, max=7, mean=2.6±1.6). Results were consistent across sessions; throughout the manuscript we merged the responses across all sessions for each subject. C. Each object was presented in one of 5 possible transformations, illustrated here for only one object. There were 3 depth rotations (0, ~45 and ~90 degrees) at the same scale and 3 scales (visual angles of 1.5 degrees, 3 degrees and 6 degrees) with the same rotation for each object. D-G Examples of visual responses from four electrodes (in different subjects). The format and conventions are the same as the ones in Figure 1A. The gray rectangle indicates the image presentation time and the electrode positions are indicated by the arrows in the small insets. Electrode locations: (D) left medial temporal (Talairach coordinates: -12.1 -58.3 -7); (E) right posterior subtemporal (Talairach coordinates: 48.3 -44.5 -19.3); (F) right posterior temporal (Talairach coordinates: 38.0 -62.0 15.0); (G) right posterior subtemporal (Talairach coordinates: 43.4 -51.3 -13.1).

Figure S2: Classification performance values under the null hypothesis and summary of selective responses

A. In order to assess whether a given classification performance value was statistically significant or not, we compared the values against those obtained from the null hypothesis. The null hypothesis states that there is no statistically significant difference across the different object categories. The performance level under the null hypothesis was obtained by randomly shuffling the object categories ($n=100$ iterations for each electrode) and repeating the same training/testing procedure used with the real data. Here we show the distribution of classification performance values obtained under the null hypothesis (bin size=0.01). The multiple green vertical dashed lines indicate 1 through 5 standard deviations. Throughout the text, we used 3 standard deviations from the null hypothesis as a threshold for statistical significance. **B.** Proportion of selective electrodes that responded to 1, 2, 3, 4 or 5 categories. **C.** Proportion of selective electrodes that responded to each one of the 5 categories. **D.** Overall distribution of classification performance values for the selective electrodes. The vertical dashed lines indicate chance level (50%) and the statistical significance threshold (see **A**). The vertical arrow shows the mean of the distribution and the two other arrows indicate the examples shown in Figure 1A and 1F. **E.** Distribution of classification performance values for the rotation-invariant electrodes (same format as in **D**). The object transformation is illustrated for one exemplar image on the top. The vertical dashed lines indicate the chance classification performance value and the significance threshold used throughout the text. The arrows indicate the mean of the distribution and the position of the example shown in Figure 4 in the main text. Bin size = 0.02. **F.** Distribution of classification performance values for the scale-invariant electrodes. We emphasize that in both **E** and **F**, the classifier was trained using the IFP responses to images at one scale and rotation and its performance was tested using IFP responses to images at different rotations (**E**) or scales (**F**). The classification performance values on the x-axis therefore reflect the degree of invariance in the neural responses to scale and rotation changes in the images.

Figure S3: Neural codes and classifier input

Classification performance for different IFP response definitions. The inset on the top shows the definition of the analysis windows $[t_b, t_a]$ where the times are defined with respect to the image onset. **A.** Comparison among different IFP response definitions: total power, range and power in different frequency bands. The y-axis shows the relative number of selective electrodes with respect to the values reported in the text (which correspond to the range in the [50;300) ms window. The high gamma frequency band (71 to 100 Hz) yielded a higher number of selective electrodes. **B.** Relative fraction of selective electrodes as a function of t_a , the end time for the analysis window (for a fixed $t_b=50$ ms). **C.** Relative fraction of selective electrodes as a function of t_b , the start time for the analysis window (for a fixed $t_a=800$ ms; see Supplementary Methods for details).

Figure S4: Decoding the activity of neural ensembles

Classification performance using an ensemble of 11 electrodes. For each subject, one electrode was chosen based on the rank of r_v values. r_v is the ratio of the variance *across* categories divided by the variance *within* categories (see Supplementary Methods and (Hung et al., 2005; Kreiman et al., 2006)). r_v was computed using only the training data. **A.** Binary classification performance. The colors correspond to different object categories. The horizontal dashed lines denote the chance performance value of 50% and the significance threshold value. Next to the chance level line we show the range of classification performance values obtained after randomly shuffling the object category labels (100 iterations). **B.** Multiclass classification performance. Here the chance level is 20% (5 object categories). **C.** Scale invariance. The classifier was trained with the neural responses obtained upon presenting objects at the default scale (3 degrees) and rotation; we tested the classifier's performance with the neural responses obtained upon presenting objects at half size or twice the size. The classification performance values reported here show the average between the two scale transformations. **D.** Rotation invariance. The classifier was trained with the neural responses obtained upon presenting objects at the default rotation (Figure S1) and scale; we tested the classifier's performance with the neural responses obtained upon presenting objects at 45 and 90 degrees rotation. The

classification performance values reported here show the average between the two rotations. Electrodes were selected based on the r_v values (ratio of the variance *across* categories to the variance *within* categories) using only the training data (see Supplementary Methods).

Figure S5: Location analysis

In contrast to Figure S4, here the electrodes forming the neural ensemble were chosen based on the electrode location, randomly sampling to obtain a total of 10 electrodes in each location. Only locations with ≥ 10 electrodes were used in this analysis. The locations are separated by lobes: each row of subplots corresponds to a different brain lobe. The color and other conventions follow Figure S4. The location names and codes are shown in Table S2. The areas that showed the highest classification performance values from this figure are shown in Figure 2 and 5B in the main text. There are three bars for each location and category: (i) training and testing on the IFP responses to different repetitions of the images at the default scale and viewpoint; (ii) training the classifier on the IFP responses to the default scale and viewpoint and testing its performance with the responses to rotated images; (iii) training the classifier on the IFP responses to the default scale and viewpoint and testing its performance with the responses to scaled images.

Figure S6: Definition of latency and parameter/location dependencies

A. We followed the definition of latency used in (Thorpe et al., 1996). At each time point, we computed a one-way ANOVA on the IFP signals across object categories. For the example electrode shown in the top part (from Figure 1F), we show the point-by-point p values in the bottom figure (in log scale). An electrode was defined as “selective” if there was a period of $selN_{cons} = 25$ consecutive time points (bin size = 2 ms at BWH and 3.9 ms at CHB) with $p < 0.01$. The response latency was defined as the first time point where $latN_{cons} = 10$ consecutive points yielded $p < 0.01$. The latency for this electrode is indicated by an arrow and the shaded area shows the interval of consecutive points where

$p < 0.01$. The y-axis was cut at $\log(p) = -6$ for graphic purposes; the $\log(p)$ values between ~ 200 and ~ 300 ms were below -6 . The distribution of latency values for all the selective electrodes is shown in Figure 3A in the main text. **B, C.** Dependence of the mean latency across all selective electrodes on $selN_{cons}$, $latN_{cons}$ and p -value threshold. The horizontal dashed line shows the mean value reported in the text (the mean of the distribution in Figure 3A which corresponds to the parameters indicated by the arrow in the fifth subplot in **B** and the second subplot **C**). The color corresponds to the p value thresholds (red: 0.05, green: 0.01, blue: 0.005, black: 0.001). The error bars correspond to one SEM of the corresponding distribution across electrodes. **D.** The distribution of latencies shown in Figure 3A in the main text includes all the selective electrodes. Here we show the mean latencies for those electrodes in each location where there were at least 5 selective electrodes. The horizontal dashed line indicates the mean value reported in the main text. The numbers in parenthesis indicate the “area code” for each location. The location names, codes and Talairach coordinates are indicated in Table S2. Frontal lobe locations are shown in white, occipital lobe locations in gray and temporal lobe locations are shown in black (there was no location with >5 selective electrodes in the parietal lobe). Error bars represent SEM.

Figure S7: Fast decoding of object category

Classification performance as a function of time from stimulus onset. Each color corresponds to an object category (**red=animals**, **green=chairs**, **blue=faces**, **black=fruits**, **yellow=vehicles**). Here the classifier was trained using individual bins of size τ (**A, D.** $\tau=25$ ms, **B, E.** $\tau=50$ ms, **C, F.** $\tau=100$ ms). We used the range of the IFP signal in each bin (in the Supplementary Web Material we show results using the power of the IFP signal in each bin). In **A-C**, we use the responses to the default scale and position to train and test the classifier’s performance (using different repetitions for training and testing). In **D-F**, we show the dynamics of decoding performance when extrapolating across scales and rotations. The classifier was trained using the responses to the objects at the default scale and rotation and its performance was evaluated with the

responses to the rotated images (solid line) or the scaled images (dashed line). The horizontal dashed lines denote the chance performance value of 0.5 and the significance threshold value (note that this threshold is different from the one in Figure S2A because we are using a much smaller window here). The vertical dashed lines mark 100 ms intervals to facilitate visualizing the dynamics of the responses. The average across all the categories in part **A** is reported in Figure 3B in the main text and the average across all categories in part **D** is reported in Figure 6B in the main text.

Figure S8: Responses to all objects/transformations for the electrode in Figure 4

A. Expanding on the presentation in Figure 4, here we show the IFPs of the same electrode to all the objects (each row corresponds to a separate object) and transformations (each column corresponds to a separate transformation). The top row shows an example of the transformations for only one of the objects. Columns 2 and 3 correspond to two scaled versions (visual angle of 1.5 degrees and 6 degrees respectively) of the standard image (visual angle of 3 degrees) and columns 4 and 5 correspond to two rotated versions of the original image (~45 and 90 degree rotation). Object category is indicated by the response color. The strong response and invariance across objects and transformations shows that the results are not due to selectivity or robustness for one particular exemplar object only. **B.** Average response to each object and transformation. Same data as in part **A**, showing a single response value for each electrode (signal range in the interval from 50 to 300 ms).

References

- Bishop, C. M. (1995). *Neural Networks for Pattern Recognition* (Oxford, Clarendon Press).
- Cristianini, N., and Shawe-Taylor, J. (2000). *An introduction to support vector machines: and other kernel-based learning methods* (Cambridge, U.K.; New York, Cambridge University Press).
- Green, D., and Swets, J. (1966). *Signal detection theory and psychophysics* (New York, Wiley).

Hung, C., Kreiman, G., Poggio, T., and DiCarlo, J. (2005). Fast Read-out of Object Identity from Macaque Inferior Temporal Cortex. *Science* 310, 863-866.

Keeping, E. S. (1995). *Introduction to Statistical Inference* (New York, Dover).

Kreiman, G., Hung, C., Quian Quiroga, R., Kraskov, A., Poggio, T., and DiCarlo, J. (2006). Object selectivity of local field potentials and spikes in the inferior temporal cortex of macaque monkeys. *Neuron* 49, 433-445.

Kreiman, G., Koch, C., and Fried, I. (2000). Category-specific visual responses of single neurons in the human medial temporal lobe. *Nature Neuroscience* 3, 946-953.

Pelli, D. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision* 10, 437-442.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (1996). *Numerical Recipes in C*, 2nd edn (Cambridge, Cambridge University Press).

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychol Bull* 124, 372-422.

Richmond, B., Wurtz, R., and Sato, T. (1983). Visual responses in inferior temporal neurons in awake Rhesus monkey. *Journal of Neurophysiology* 50, 1415-1432.

Rifkin, R., and Klautau, A. (2004). In defense of one-vs-all classification. *Journal of Machine Learning Research* 5, 101-110.

Schmolesky, M., Wang, Y., Hanes, D., Thompson, K., Leutgeb, S., Schall, J., and Leventhal, A. (1998). Signal timing across the macaque visual system. *Journal of Neurophysiology* 79, 3272-3278.

Thorpe, S., Fize, D., and Marlot, C. (1996). Speed of processing in the human visual system. *Nature* 381, 520-522.

Vapnik, V. (1995). *The Nature of Statistical Learning Theory* (New York, Springer).

Figure S1

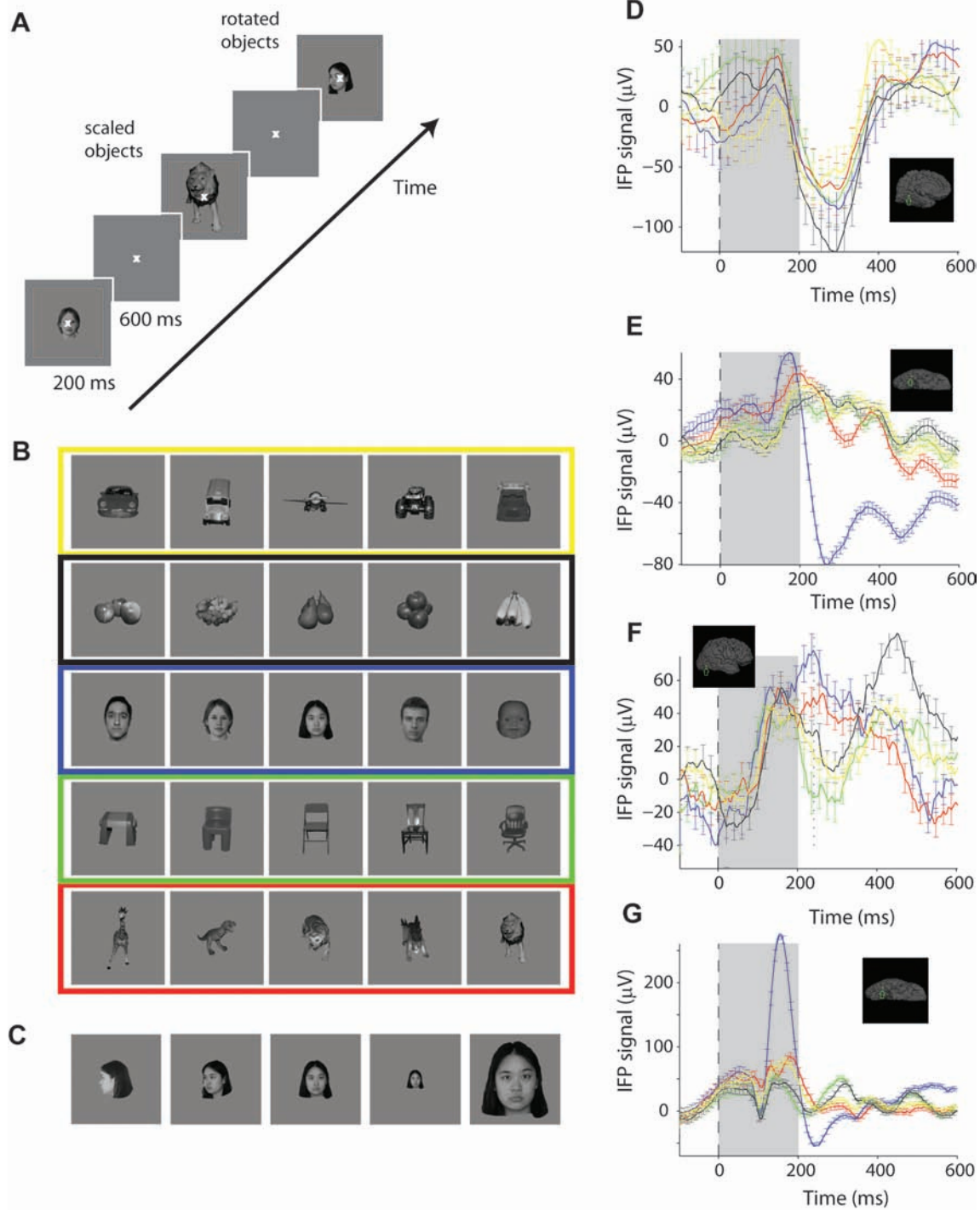


Figure S2

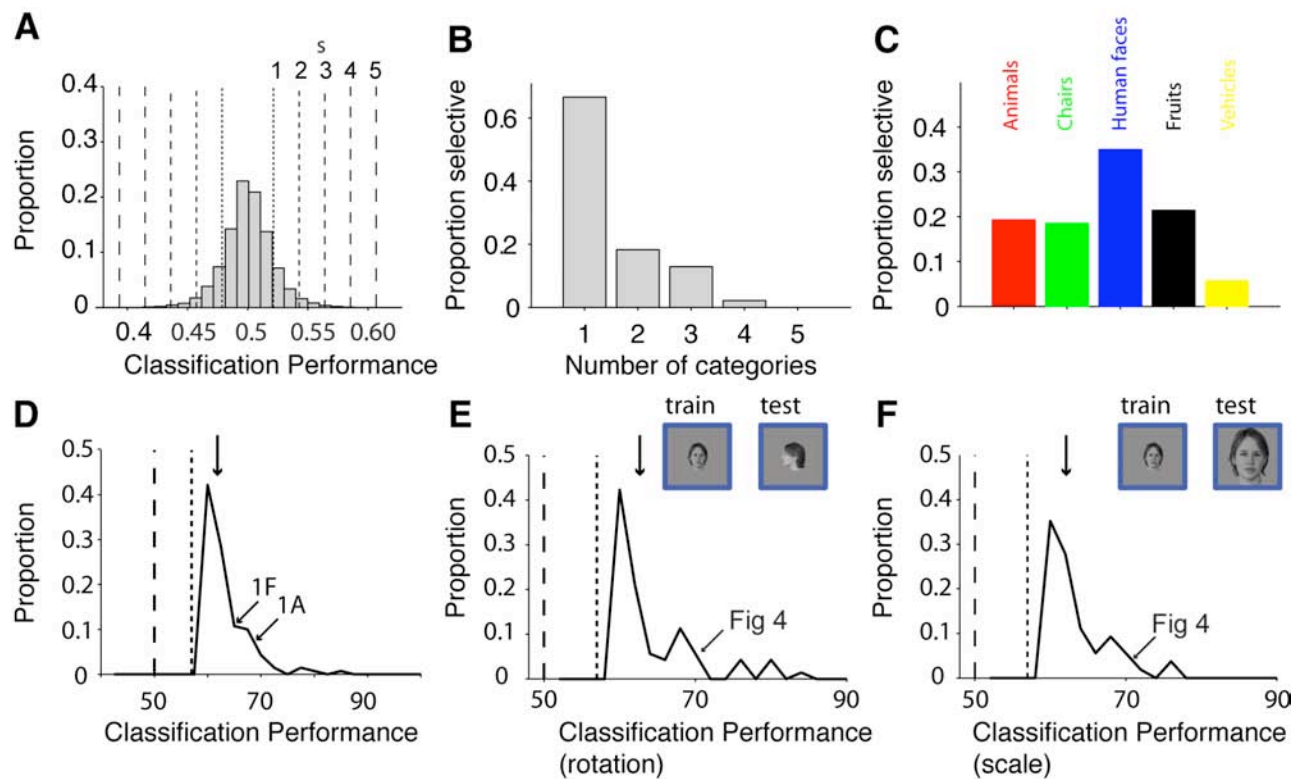


Figure S3

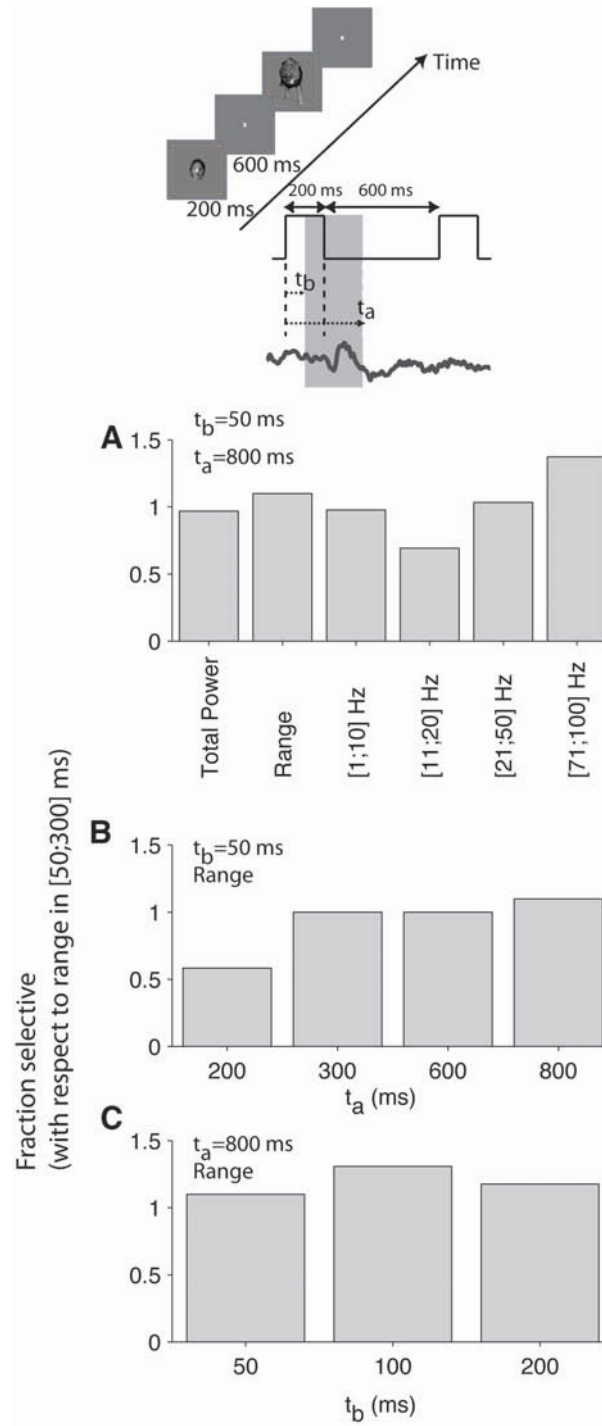


Figure S4

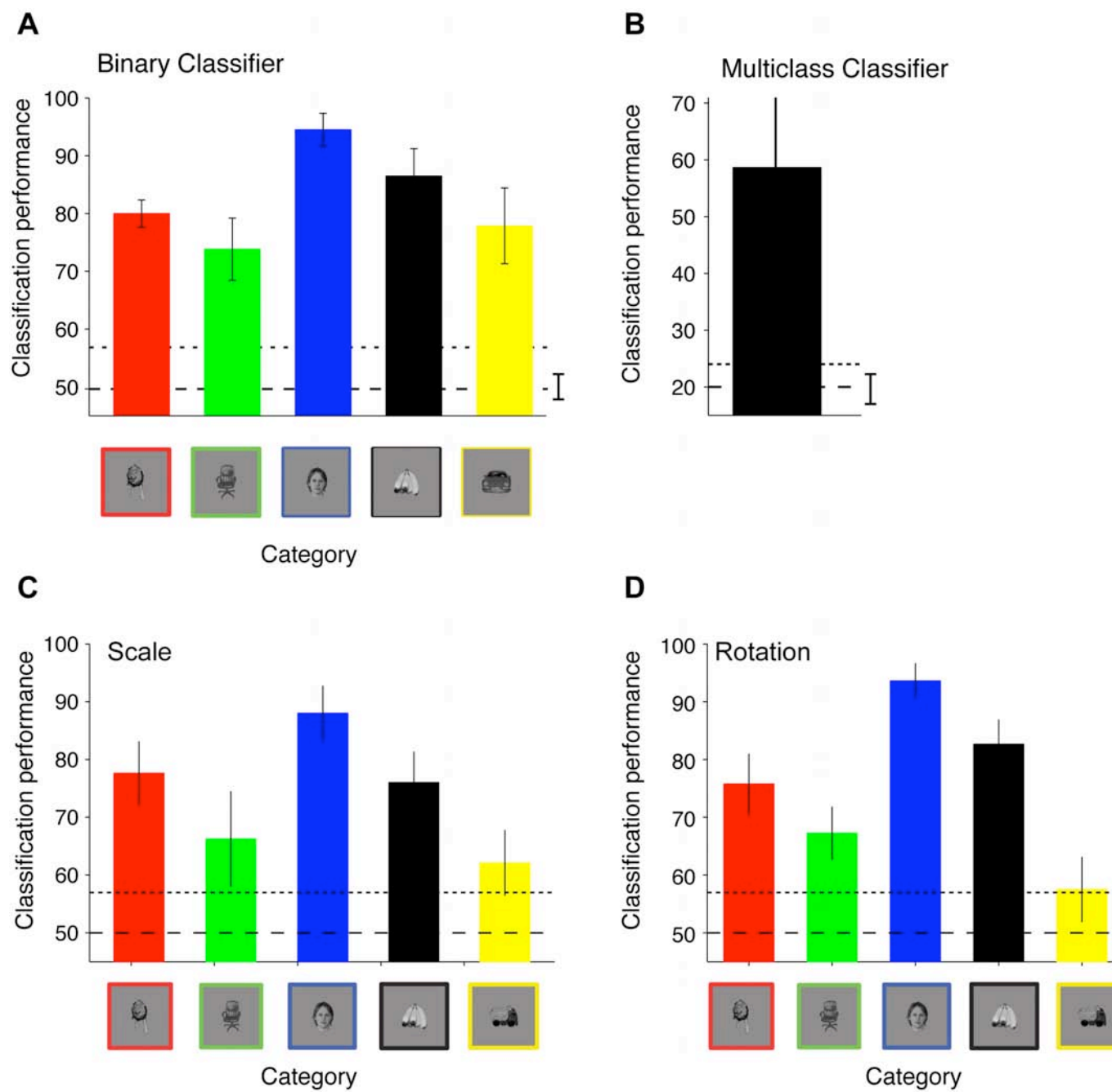


Figure S5

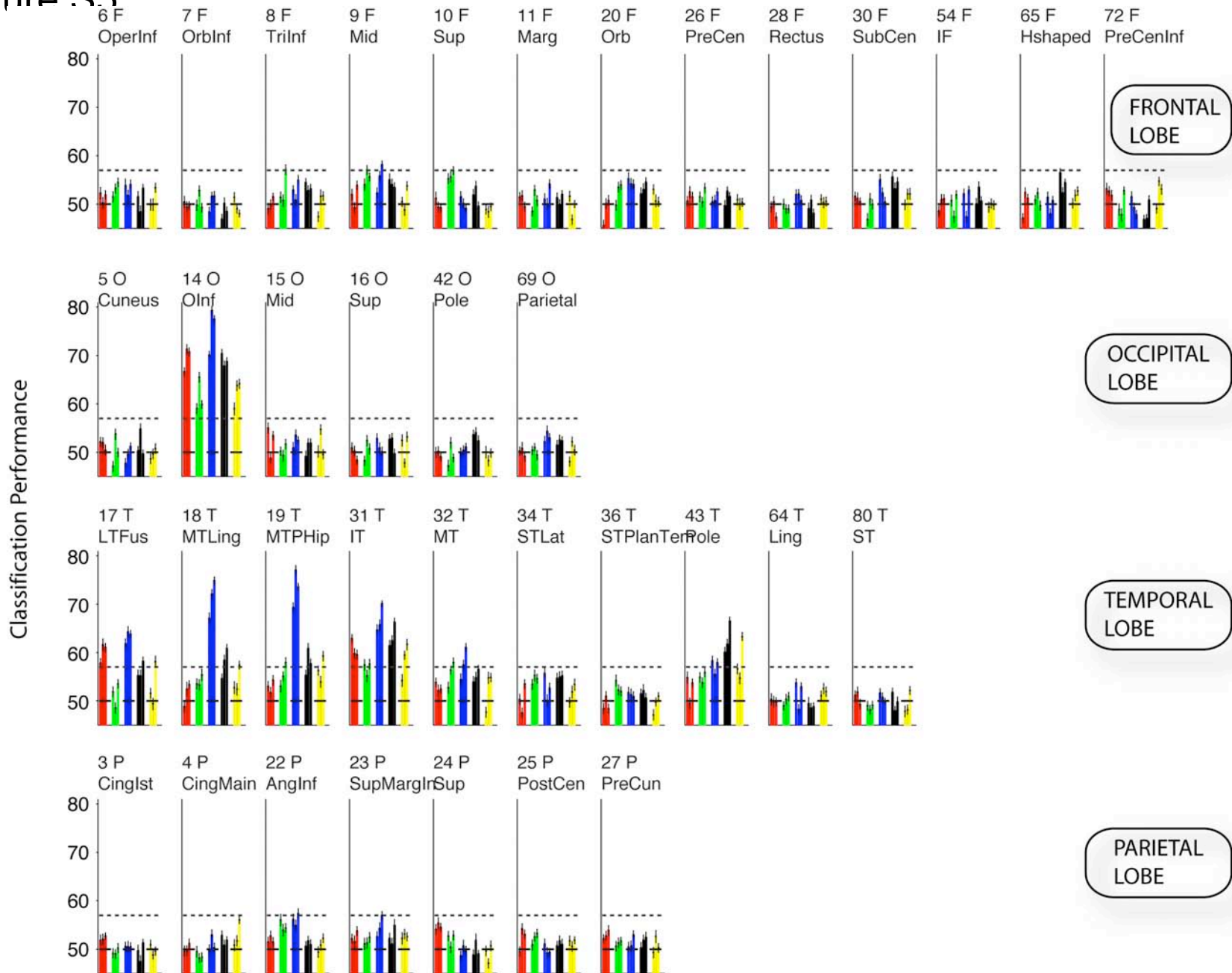


Figure S6

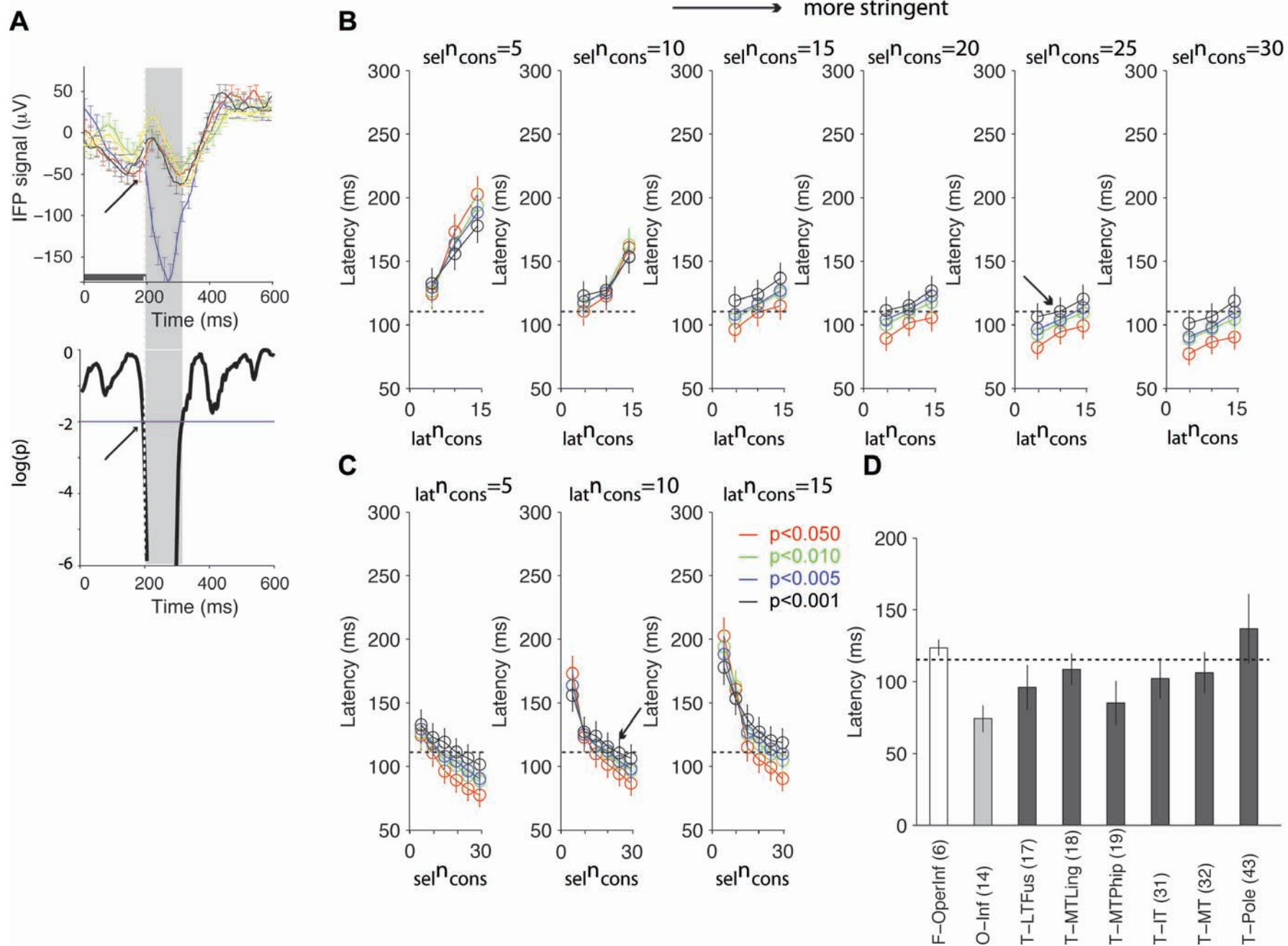


Figure S7

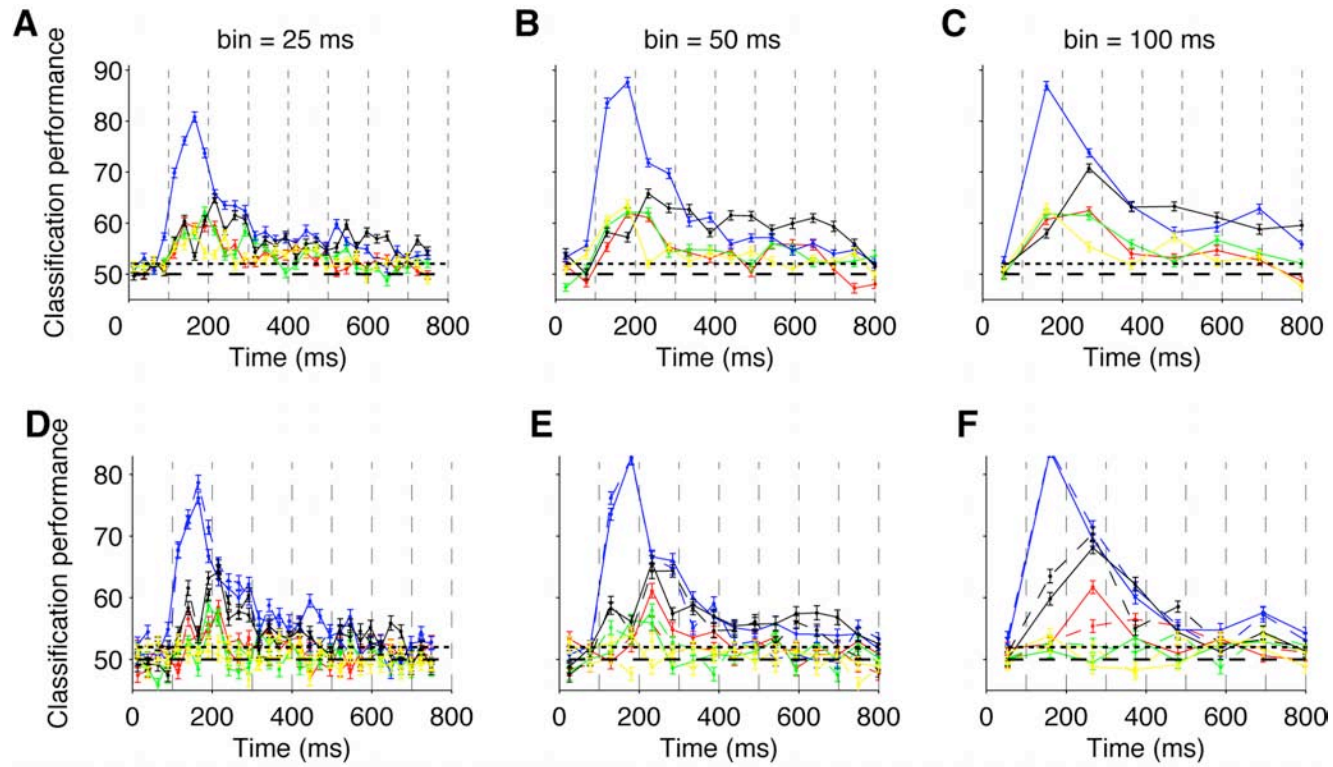


Figure S8

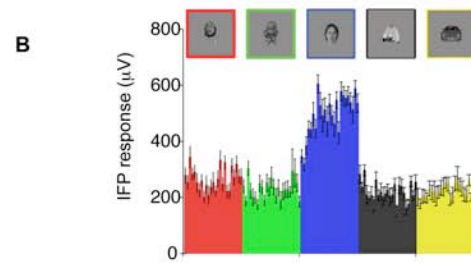
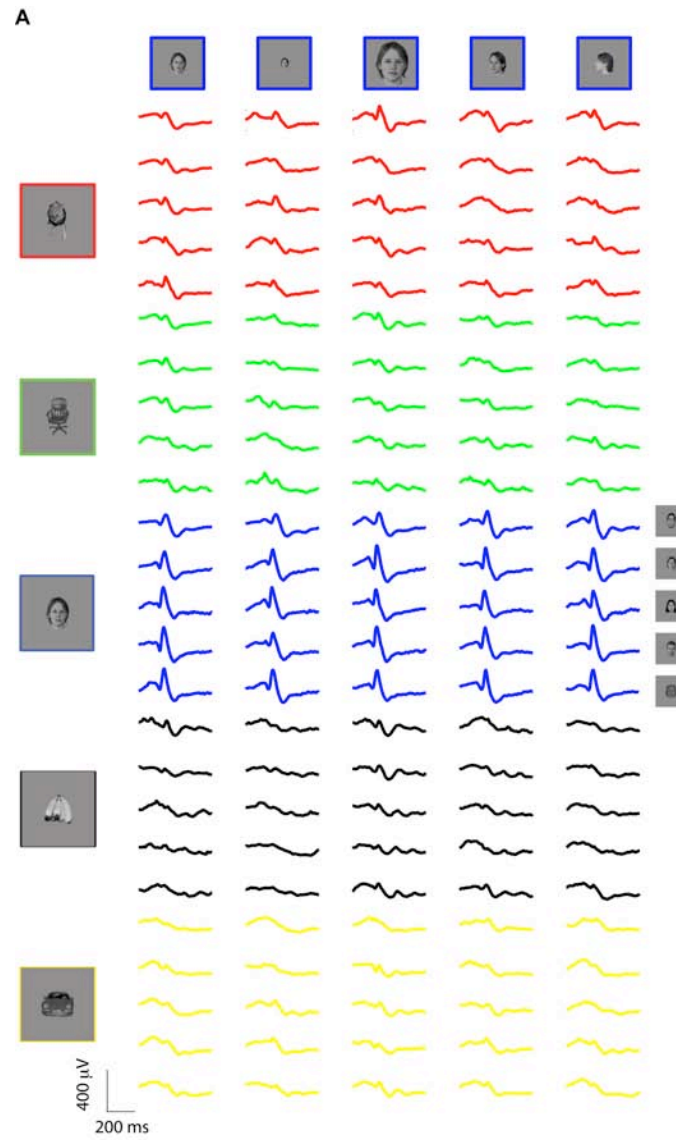


TABLE S1

	F-OperInf	F-TrInf	F-Mid	F-Sup	F-Marg	F-Orb	P-PreCen	F-SubCen	Total*	Sum	Fract			
Frontal Lobe														
Area code	6	8	9	10	11	20	26	30						
Total electrodes	34	33	49	23	16	54	34	24	298	267	90%			
Selective	2	1	5	2	0	1	2	1	22	14	64%			
% sel	6%	3%	10%	9%	0%	2%	6%	4%	7%					
inv scale	2	1	1	1	0	1	1	0	7	7	100%			
% inv scale			20%						32%					
inv rotation	1	1	2	1	0	1	1	0	8	7	88%			
% inv rotation			40%						36%					
Occipital Lobe	O-Inf	O-Mid	O-Par								Total*	Sum	Fract	
Area code	14	15	69											
Total electrodes	21	18	14								74	53	72%	
Selective	18	4	1								26	23	88%	
% sel	86%	22%	7%								35%			
inv scale	11	3	0								14	14	100%	
% inv scale	61%										54%			
inv rotation	13	4	0								17	17	100%	
% inv rotation	72%										65%			
Tempotal Lobe	T-LTFus	T-MTLing	T-MTPhip	T-IT	T-MT	T-STLat	T-Pole					Total*	Sum	Fract
Area code	17	18	19	31	32	34	43							
Total electrodes	24	29	27	88	117	80	51					423	416	98%
Selective	9	6	7	16	9	3	7					58	57	98%
% sel	38%	21%	26%	18%	8%	4%	14%					14%		
inv scale	5	2	5	9	0	1	5					27	27	100%
% inv scale	56%	33%	71%	56%	0%		71%					47%		
inv rotation	6	4	5	9	8	3	3					38	38	100%
% inv rotation	67%	67%	71%	56%	89%		43%					66%		
Parietal Lobe	Gparietal inferiorA	Gparietal inferiorS	Gpostce ntral	Gprecun eus								Total*	Sum	Fract
Area code	22	23	25	27										
Total electrodes	28	36	22	14								117	100	85%
Selective	1	3	1	0								5	5	100%
% sel	4%	8%	5%	0%								4%		
inv scale	1	0	0	0								1	1	
% inv scale												20%		
inv rotation	1	0	0	0								1	1	100%
% inv rotation												20%		

TABLE S2

Area Code	Area abbreviation	Area name	Talairach coordinates (LEFT HEMISPHERE)					Talairach coordinates (LEFT HEMISPHERE)							
			means	std	mins	maxs	medians	n	means	std	mins	maxs	medians	n	
0	U-UNKN	Unknown													
1	U-CCL	Corpus-callosum													0
2	T-IAvg	G-and-S-insula-only-average													0
3	P-CingIst	G-cingulate-Isthmus							0	[-2 -46 6]	[0 0 4]	[-3 -46 3]	[-1 -46 9]	[-2 -46 6]	2
4	P-CingMain	G-cingulate-Main-part							0	[2 -41 27]	[2 16 11]	[-0 -75 15]	[7 -30 48]	[2 -33 25]	7
5	O-Cuneus	G-cuneus	[2 -68 12]	[0 0 0]	[2 -68 12]	[2 -68 12]	[2 -68 12]	1	[1 -70 13]	[0 4 4]	[0 -76 9]	[1 -64 20]	[1 -70 12]	6	
6	F-OperInf	G-frontal-inf-Opercular-part	[-36 16 14]	[40 11 12]	[-61 0 -9]	[59 47 32]	[-52 12 14]	26	[57 14 1]	[6 2 9]	[47 11 -11]	[63 18 10]	[61 14 -0]	5	
7	F-OrbInf	G-frontal-inf-Orbital-part	[7 42 -8]	[50 17 4]	[-40 25 -13]	[51 66 -3]	[8 38 -9]	4	[54 32 -4]	[0 7 3]	[54 27 -8]	[55 41 -2]	[55 30 -2]	3	
8	F-TrilInf	G-frontal-inf-Triangular-part	[-33 31 6]	[39 13 9]	[-58 10 -11]	[61 58 23]	[-48 26 7]	27	[56 30 10]	[2 6 4]	[54 23 6]	[59 35 18]	[56 34 8]	5	
9	F-Mid	G-frontal-middle	[-39 36 30]	[8 20 16]	[-53 3 -16]	[-27 73 52]	[-36 32 33]	30	[40 38 30]	[9 16 17]	[21 10 -5]	[52 66 58]	[43 41 33]	19	
10	F-Sup	G-frontal-superior	[-24 25 51]	[13 21 22]	[-53 7 9]	[-9 72 68]	[-21 20 60]	8	[6 60 27]	[4 9 15]	[0 39 -8]	[14 76 50]	[7 57 30]	15	
11	F-Marg	G-frontomarginal	[-33 57 -7]	[0 0 0]	[-33 57 -7]	[-33 57 -7]	[-33 57 -7]	1	[34 61 18]	[12 8 23]	[14 44 -18]	[52 74 44]	[34 62 29]	15	
12	F-InsLong	G-insular-long						0						0	
13	F-InsShort	G-insular-short	[-52 5 -5]	[6 1 3]	[-57 4 -8]	[-47 7 -3]	[-52 5 -5]	2						0	
14	O-Inf	G-and-S-occipital-inferior	[5 -67 -5]	[50 11 8]	[-50 -89 -15]	[62 -48 10]	[7 -67 -6]	16	[7 -69 7]	[38 11 10]	[-36 -84 1]	[43 -56 25]	[27 -66 3]	5	
15	O-Mid	G-occipital-middle	[-42 -73 22]	[5 11 23]	[-51 -86 -4]	[-35 -48 84]	[-43 -75 16]	12	[13 -68 11]	[36 7 2]	[-36 -77 7]	[45 -58 15]	[32 -70 11]	6	
16	O-Sup	G-occipital-superior						0	[-0 -76 25]	[1 3 10]	[-1 -80 10]	[1 -68 36]	[-0 -78 29]	10	
17	T-LTFus	G-occipit-temp-lat-Or-fusiform	[-20 -34 -17]	[21 13 6]	[-47 -55 -26]	[38 -11 -9]	[-26 -34 -17]	11	[11 -42 -12]	[27 23 5]	[-23 -78 -28]	[37 -12 -5]	[29 -40 -11]	13	
18	T-MTLing	G-occipit-temp-med-Lingual-part	[-8 -55 -1]	[5 7 5]	[-13 -61 -6]	[-1 -44 5]	[-8 -57 -3]	4	[0 -65 -0]	[8 13 6]	[-13 -86 -15]	[25 -28 10]	[-0 -67 0]	25	
19	T-MTPhip	G-occipit-temp-med-Parahippocampal-p	[-22 -0 -18]	[7 13 7]	[-39 -25 -41]	[-12 20 -7]	[-21 -0 -18]	14	[14 -31 -13]	[25 13 6]	[-24 -53 -23]	[51 -9 -3]	[20 -34 -13]	13	
20	F-Orb	G-orbital	[-16 39 -13]	[28 12 13]	[-55 14 -35]	[50 66 15]	[-21 40 -15]	40	[39 31 -14]	[11 14 4]	[15 12 -24]	[51 59 -7]	[44 25 -14]	13	
21	P-ParaCen	G-paracentral						0						0	
22	P-AngInf	G-parietal-inferior-Angular-part	[-49 -55 35]	[12 21 21]	[-67 -75 2]	[-31 -13 61]	[-48 -61 44]	13	[21 -54 44]	[40 16 12]	[-34 -79 21]	[63 -31 59]	[44 -54 44]	11	
23	P-SupraMargInf	G-parietal-inferior-Supramarginal-part	[-27 -40 36]	[52 13 14]	[-71 -63 12]	[69 -7 64]	[-48 -42 32]	32	[63 -18 46]	[0 8 2]	[62 -24 44]	[63 -12 48]	[63 -18 46]	2	
24	P-Sup	G-parietal-superior	[-28 -35 63]	[4 19 6]	[-35 -66 53]	[-23 -16 69]	[-29 -32 65]	5	[-35 -50 60]	[0 7 4]	[-35 -55 57]	[-35 -44 63]	[-35 -50 60]	2	
25	P-PostCen	G-postcentral	[-53 -6 44]	[7 17 12]	[-70 -32 20]	[-46 16 61]	[-52 -2 48]	12	[36 -15 56]	[31 12 10]	[-35 -36 38]	[63 -0 67]	[46 -15 58]	8	
26	P-PreCen	G-precentral	[-33 1 36]	[42 16 13]	[-66 -24 12]	[65 34 59]	[-48 -1 35]	26	[36 -0 53]	[32 13 16]	[-34 -27 27]	[62 16 68]	[47 4 60]	7	
27	P-PreCun	G-precuneus						0	[0 -55 22]	[0 9 10]	[-0 -74 9]	[2 -43 35]	[1 -52 20]	14	
28	F-Rectus	G-rectus	[9 30 -25]	[0 0 0]	[9 30 -25]	[9 30 -25]	[9 30 -25]	1	[3 40 -25]	[3 3 0]	[1 38 -25]	[5 43 -25]	[3 40 -25]	2	
29	U-SCCL	G-subcallosal						0						0	
30	F-SubCen	G-subcentral	[2 -7 17]	[65 7 9]	[-70 -24 -1]	[69 4 36]	[6 -6 18]	22						0	

31	T-IT	G-temporal-inferior	[-18 -31 -22]	[50 20 9]	[-72 -75 -41]	[71 11 3]	[-45 -32 -22]	69	[5 -44 -7]	[59 20 6]	[-65 -72 -18]	[70 -11 7]	[45 -45 -7]	15
32	T-MT	G-temporal-middle	[-22 -23 -6]	[58 23 12]	[-74 -68 -34]	[72 41 20]	[-56 -25 -7]	88	[40 -27 -3]	[50 19 9]	[-67 -61 -18]	[72 13 16]	[62 -29 -3]	28
33	T-Interm	G-temp-sup-G-temp-transv-and-interm-S						0						0
34	T-STLat	G-temp-sup-Lateral-aspect	[-26 -4 2]	[53 18 13]	[-73 -39 -21]	[72 38 36]	[-54 -4 2]	64	[60 6 -5]	[7 16 6]	[44 -21 -14]	[70 33 8]	[60 7 -7]	14
35	T-STPlanPol	G-temp-sup-Planum-polare						0						0
36	T-STPlanTemp	G-temp-sup-Planum-tempolare	[-57 -53 20]	[2 10 1]	[-59 -60 19]	[-55 -46 22]	[-57 -53 20]	2						0
37	F-FrontoPol	G-and-S-transverse-frontopolar						0						0
38	F-RamusH	Lat-Fissure-ant-sgt-ramus-horizontal						0						0
39	F-RamusV	Lat-Fissure-ant-sgt-ramus-vertical						0						0
40	P-LatFiss	Lat-Fissure-post-sgt						0						0
41	U-MWL	Medial-wall						0	[8 -8 3]	[0 6 7]	[8 -13 -2]	[9 -4 8]	[8 -8 3]	2
42	O-Pole	Pole-occipital						0	[-4 -73 0]	[14 28 14]	[-23 -92 -18]	[11 -31 16]	[-2 -85 1]	4
43	T-Pole	Pole-temporal	[-7 4 -31]	[36 14 7]	[-55 -15 -43]	[58 49 -13]	[-23 2 -32]	44	[46 3 -19]	[20 25 6]	[20 -22 -28]	[67 36 -11]	[55 -12 -22]	7
44	O-Calc	S-calcarine						0						0
45	P-Cen	S-central						0						0
46	P-InsCen	S-central-insula						0						0
47	F-CingMain	S-cingulate-Main-part-and-Intracingulate						0						0
48	F-CingMarg	S-cingulate-Marginalis-part						0						0
49	F-InsAnt	S-circular-insula-anterior						0	[43 -51 -13]	[0 0 0]	[43 -51 -13]	[43 -51 -13]	[43 -51 -13]	1
50	F-InsSup	S-circular-insula-inferior						0						0
51	T-InsSup	S-circular-insula-superior						0						0
52	F-CollAnt	S-collateral-transverse-ant	[-27 -1 -21]	[0 0 0]	[-27 -1 -21]	[-27 -1 -21]	[-27 -1 -21]	1						0
53	F-CollPost	S-collateral-transverse-post						0	[44 -16 -35]	[0 0 0]	[44 -16 -35]	[44 -16 -35]	[44 -16 -35]	1
54	F-IF	S-frontal-inferior	[-48 38 12]	[1 5 2]	[-49 34 9]	[-46 45 14]	[-49 36 13]	3	[54 28 28]	[3 2 6]	[51 26 24]	[56 29 33]	[54 28 28]	2
55	F-MF	S-frontal-middle						0	[30 59 14]	[0 0 0]	[30 59 14]	[30 59 14]	[30 59 14]	1
56	F-SF	S-frontal-superior						0						0
57	F-FrontMarg	S-frontomarginal						0						0
58	P-Jensen	S-intermedius-primus-Jensen						0						0
59	P-Intra	S-intraparietal-and-Parietal-transverse						0	[39 -27 62]	[0 0 0]	[39 -27 62]	[39 -27 62]	[39 -27 62]	1
60	O-Ant	S-occipital-anterior						0						0
61	O-Lunatus	S-occipital-middle-and-Lunatus						0						0
62	O-Transversalis	S-occipital-superior-and-transversalis						0						0
63	T-LT	S-occipito-temporal-lateral						0						0
64	T-Lingual	S-occipito-temporal-medial-and-S-Lingual						0	[20 -26 -6]	[0 1 14]	[20 -27 -16]	[21 -26 3]	[20 -26 -6]	2
65	F-Hshaped	S-orbital-H-shapped	[-22 37 -16]	[19 6 9]	[-43 25 -25]	[20 44 7]	[-25 39 -18]	8						0
66	F-OrbLat	S-orbital-lateral						0						0
67	F-Olf	S-orbital-medial-Or-olfactory						0						0
68	P-ParaCen	S-paracentral						0						0

69	O-Parietal	S-parieto-occipital						0	[1 -63 27]	[0 6 5]	[-0 -74 19]	[2 -54 36]	[1 -64 26]	14
70	P-PeriCallosal	S-pericallosal						0						0
71	P-PostCen	S-postcentral						0						0
72	F-PreCenInf	S-precentral-Inferior-part	[-50 5 33]	[5 1 8]	[-53 3 27]	[-46 6 39]	[-50 5 33]	2						0
73	F-PreCenSup	S-precentral-Superior-part						0						0
74	P-SubCenAnt	S-subcentral-ant						0						0
75	P-SubCenPos	S-subcentral-post						0						0
76	P-SubOrb	S-suborbital						0						0
77	P-SubPar	S-subparietal						0						0
78	F-CingSupra	S-supracingulate						0						0
79	T-IT	S-temporal-inferior						0						0
80	T-ST	S-temporal-superior	[-62 -17 14]	[1 7 1]	[-63 -23 13]	[-61 -12 15]	[-62 -17 14]	2						0