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

Supplementary Web Material

Hesheng Liu[1,3], Yigal Agam[1,3], Joseph R. Madsen[2], Gabriel Kreiman[1,4,5]

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

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

[3]Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Harvard Medical School

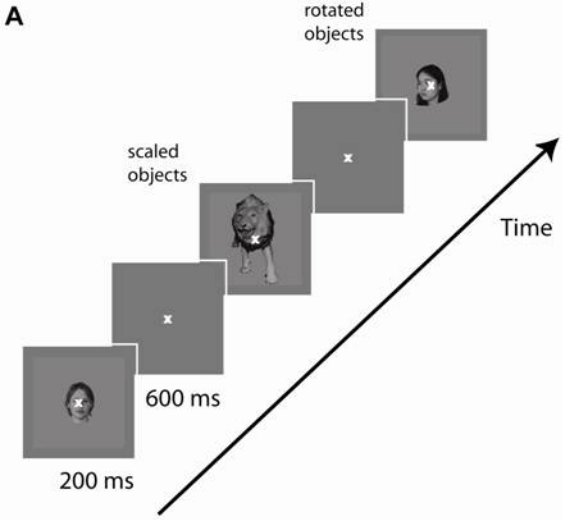
[4]Center for Brain Science, Harvard University

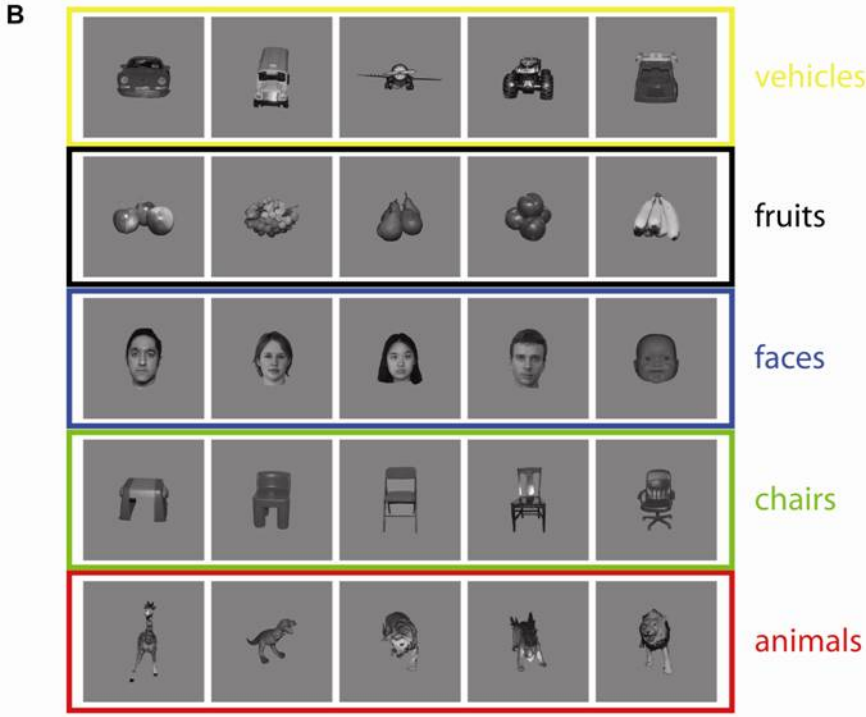
[5]Swartz Center for Theoretical Neuroscience, Harvard University

---

Figure S1: Stimulus presentation scheme

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. **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. 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 very consistent across sessions; throughout the manuscript we merged the responses across all sessions for each subject.


A
rotated objects
scaled objects
Time
600 ms
200 ms


B
vehicles
fruits
faces
chairs
animals


C

Figure S2: Reproducibility across repetitions

A. For each electrode and each object, we computed the correlation coefficient between the IFP responses across repetitions (Pearson correlation of the point-by-point IFP responses between 0 and 800 ms post-stimulus after resampling with a 20 ms rolling average window). In this plot, each column shows one electrode, the dashed vertical lines separate electrodes in different subjects. For each electrode, the gray bars show the average correlation coefficient across all objects and the black curve shows the correlation coefficient for the best object (i.e. the object yielding the highest correlation coefficient). **B.** To assess the statistical significance of the correlation coefficients in part **A**, we computed the null expected value by performing 500 random shuffles of the object labels. The y-axis shows the *z*-score for the average correlation coefficient compared to this null hypothesis. The y-axis was cut off at $z=6$ for display purposes but there were several electrodes with $z>6$. The arrows point to the example electrodes shown in Figures 1**A,** 1**F** and 4 in the main text.

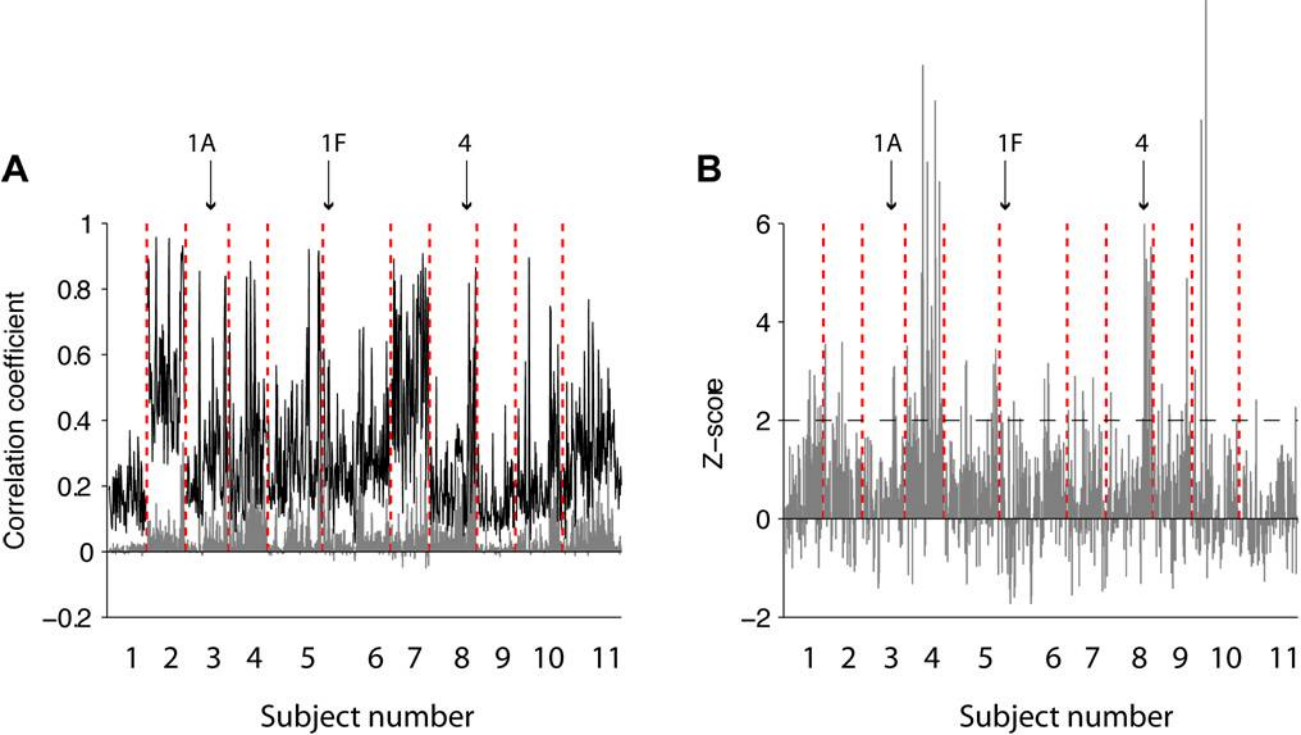

A
1A
1F
4
Correlation coefficient
1
0.8
0.6
0.4
0.2
0
−0.2
1 2 3 4 5 6 7 8 9 10 11
Subject number
B
1A
1F
4
Z−score
6
4
2
0
−2
1 2 3 4 5 6 7 8 9 10 11
Subject number

Figure S3: Multiple examples showing selectivity in IFP recordings
Examples from six electrodes that showed visual selectivity. The format and conventions are the same as in Figure 1**A**. The gray rectangle indicates the image presentation time and the electrode positions are indicated by the arrows in the small insets. Electrode locations: (**A**) left medial temporal (Talairach coordinates: -12.1 -58.3 -7); (**B, C**) right posterior subtemporal (Talairach coordinates: 43.4 -51.3 -13.1 and 48.3 -44.5 -19.3); (**D, E**) right posterior temporal (Talairach coordinates: 27.0 -66.0 3.0 and 38.0 -62.0 15.0); (**F**) left lateral temporal (Talairach coordinates: -39.7 -42.8 -20.0).

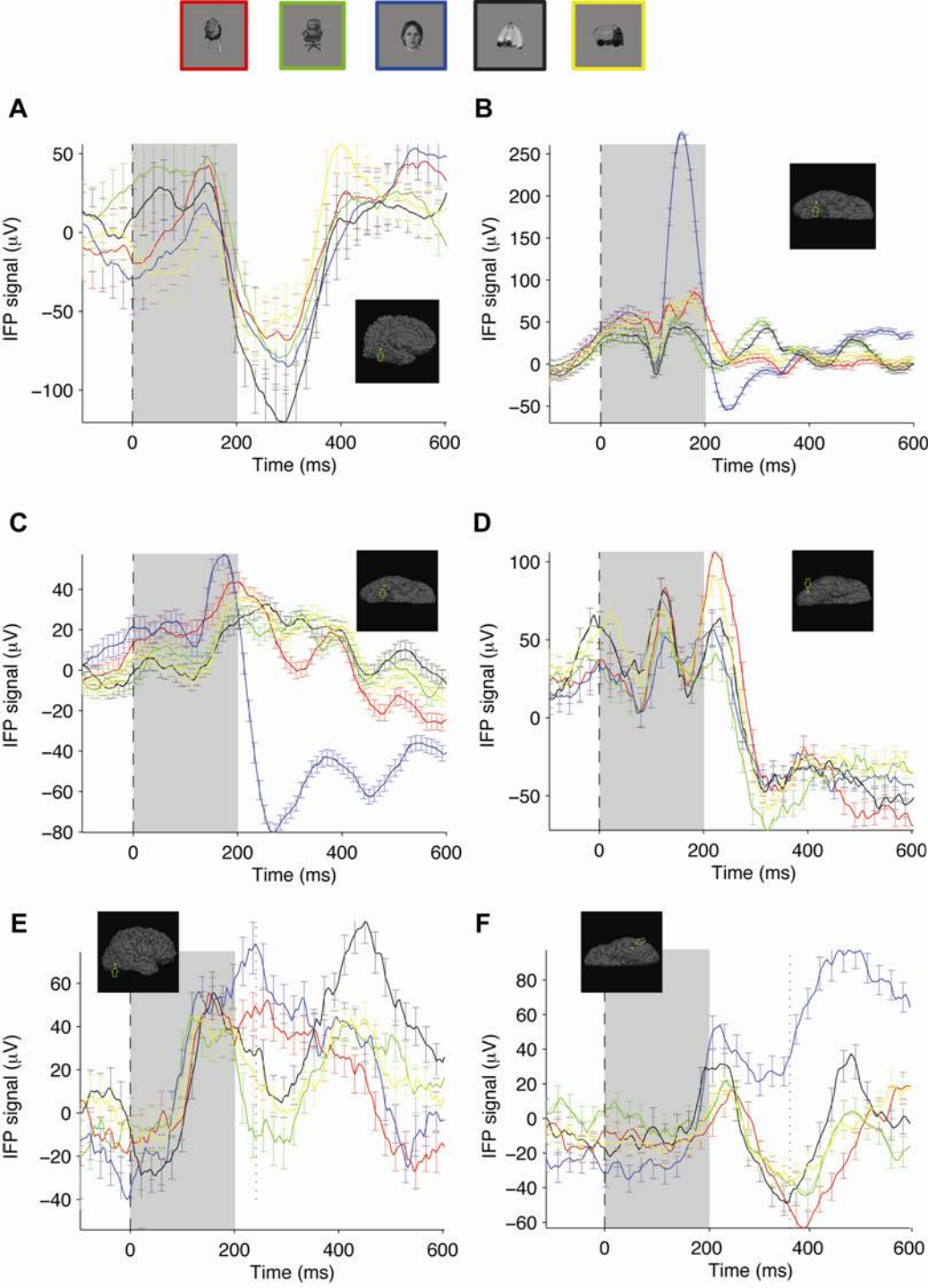

A
B
C
D
E
F
IFP signal (μV)
Time (ms)

Figure S4: Categorization of neural data using statistical classifiers
The figure shows the responses of two electrodes (electrode "1" in the x-axis and electrode "2" on the y-axis) to objects from two categories (denoted by red (animals) and blue (faces)). The aim of the classifier is to separate the data from the two object categories. Here we show the separation boundary computed by an SVM classifier (Vapnik, 1995) with a linear kernel (the classifier used throughout the text; see also Figure S7). The black dashed line indicates the classification boundary: examples above the dashed line are assigned to the blue category and examples below the dashed line are assigned to the red category. The classifier makes mistakes by incorrectly assigning some blue examples to red (indicated here by filled blue circles) and some red examples to blue (indicated here by filled red circles). The measure of classification performance used throughout the text is the overall fraction of examples in the test data which were correctly classified (*single trial responses*). It is important to note that in all cases, separate repetitions were used to train and test the classifier, therefore avoiding the common problem of overfitting in statistical learning. In all cases, the location of the dashed line is determined using training data that is separate from the test data used to evaluate the performance of the classifier. For display purposes, we use two electrodes in this Figure but it is easy to extend the math and procedure to higher dimensions (see Figure S13; for more information on statistical learning, see (Hung et al., 2005a; Vapnik, 1995)).

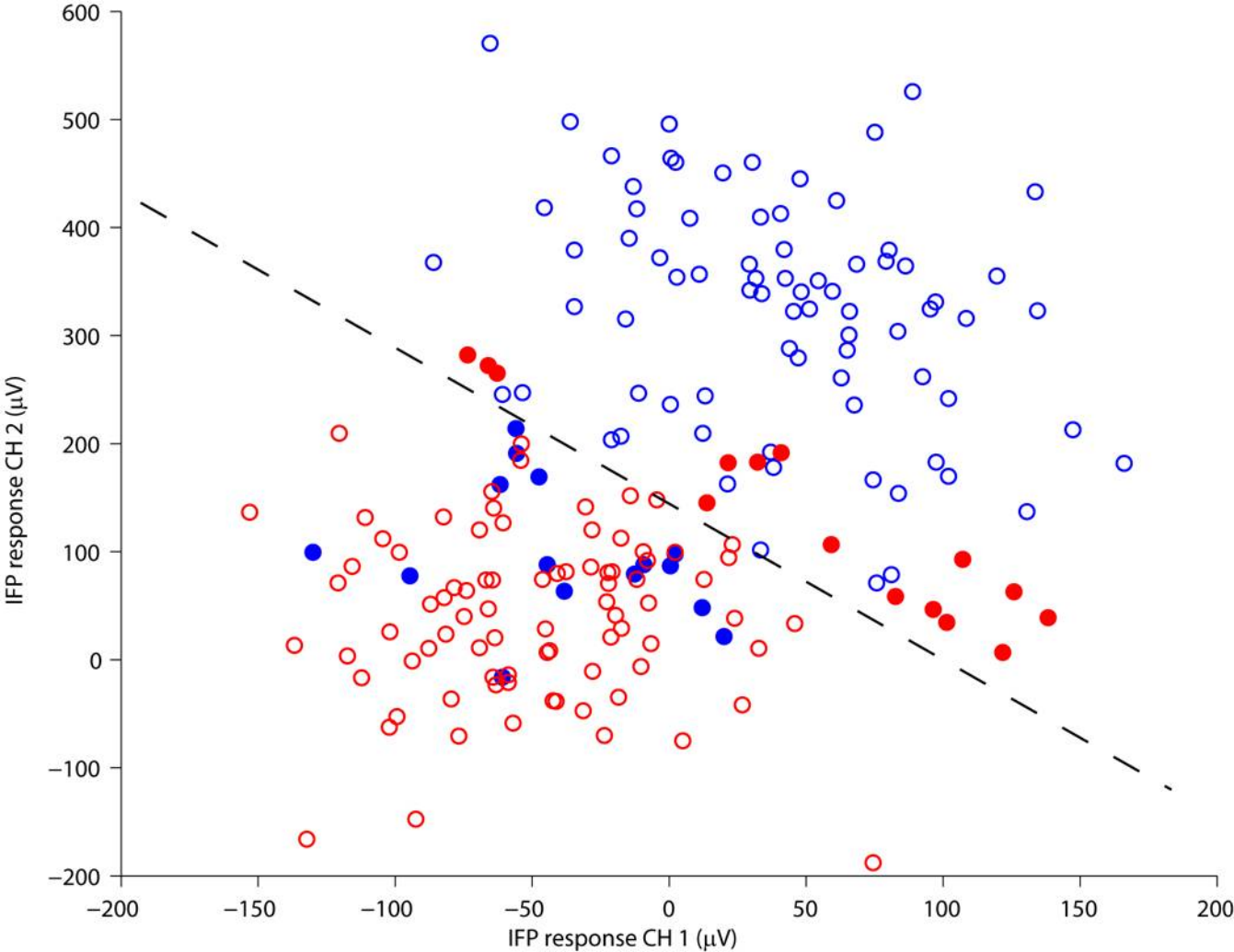

600
500
400
300
200
100
0
−100
−200
IFP response CH 2 (μV)
−200
−150
−100
−50
0
50
100
150
200
IFP response CH 1 (μV)

Figure S5: Classification performance values under the null hypothesis
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) 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 0.57 as the threshold for statistical significance. This value corresponds to a difference of 3 standard deviations from the null hypothesis.

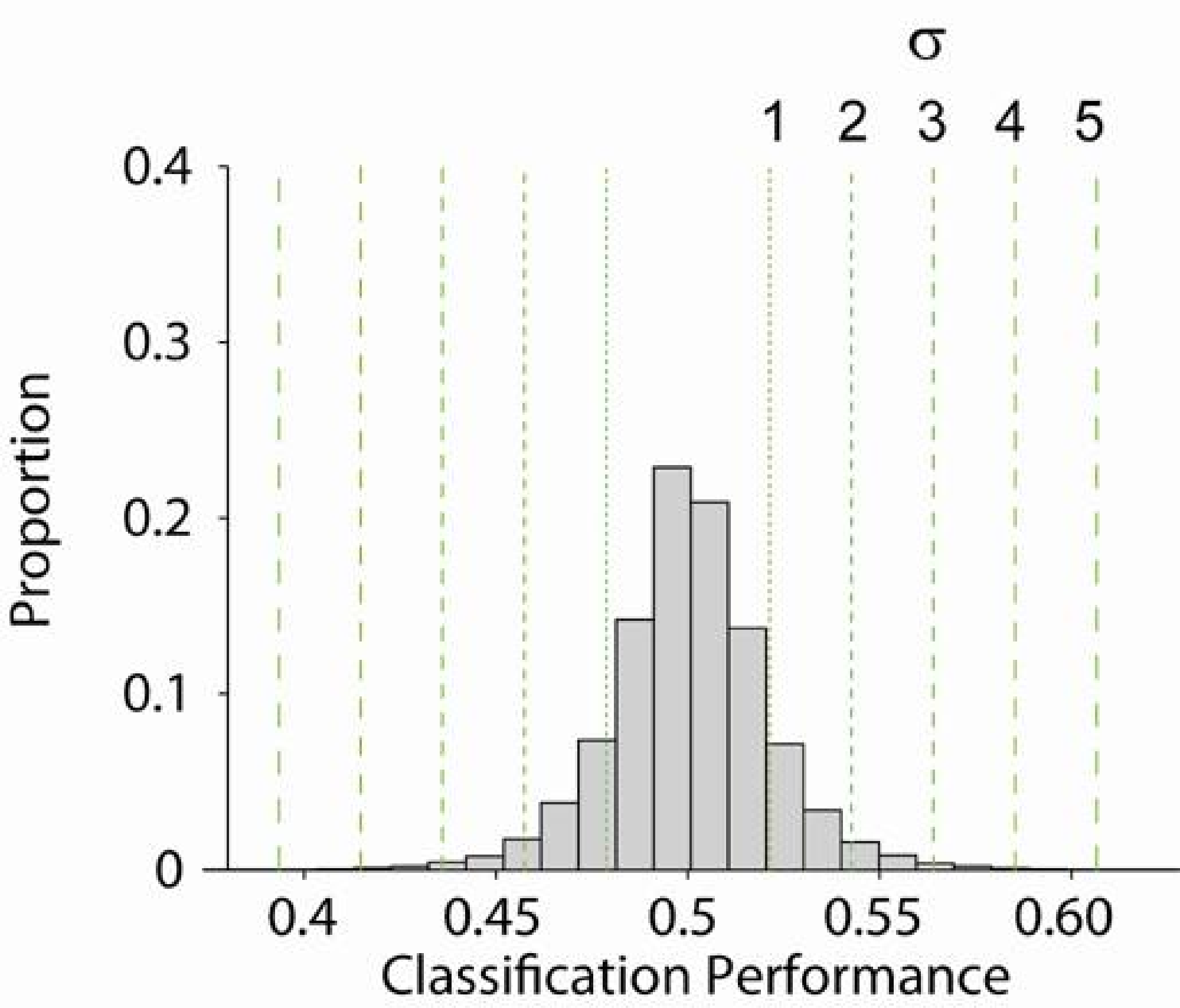

σ
1
2
3
4
5
0.4
0.3
0.2
0.1
0
Proportion
0.4
0.45
0.5
0.55
0.60
Classification Performance

Figure S6: Summary of selective responses

A. Proportion of selective electrodes that responded to 1, 2, 3, 4 or 5 categories. **B.** Proportion of selective electrodes that responded to each category. **C.** Overall distribution of classification performance values for the selective electrodes. The vertical dashed lines indicate chance level (0.5) and the statistical significance threshold (0.57; see Figure S5). The vertical arrow shows the mean of the distribution and the two other arrows indicate the examples shown in Figure 1**A** and 1**F**.

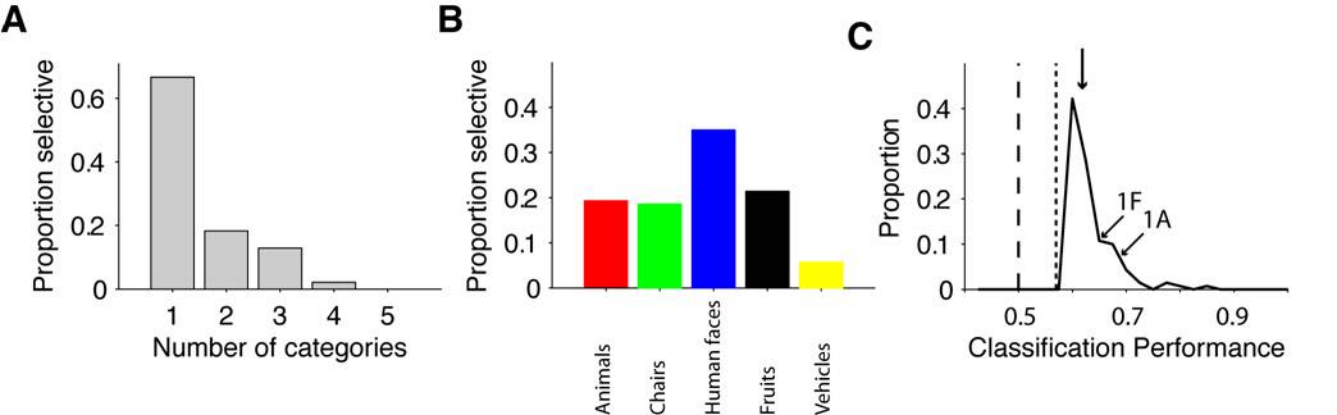

A
Proportion selective
0.6
0.4
0.2
0
1
2
3
4
5
Number of categories
B
Proportion selective
0.4
0.3
0.2
0.1
0
Animals
Chairs
Human faces
Fruits
Vehicles
C
Proportion
0.4
0.3
0.2
0.1
0
1F
1A
0.5
0.7
0.9
Classification Performance

Figure S7: Comparison among different statistical classifiers and selectivity criteria
A, B. Comparison of classification performance levels obtained using different statistical classifiers. Throughout the text, we report the performance of a Support Vector Machine (SVM) classifier with a linear kernel. We compared the performance of this classifier against other possible machine learning classifiers (Bishop, 1995; Hung et al., 2005a; Vapnik, 1995). Here we show a direct comparison between the linear SVM (x-axis) and an SVM classifier with a Gaussian kernel (red) or a nearest neighbor classifier (blue). The diagonal dashed line would correspond to identical performance across classifiers. We compare the values only for those electrodes and conditions that yielded a performance above 3 standard deviations of the null hypothesis, which was our threshold throughout the text for statistical significance (see Figure S5). Overall, the Gaussian SVM showed a slightly better performance and the nearest neighbor classifier showed a slightly worse performance than the linear SVM; yet, none of the conclusions in the current manuscript would be significantly modified if we used these other classifiers. In **A**. we compare the results for multiple IFP response definitions (see Figure S8) whereas in **B**. we restrict the comparison to the IFP range in the 50 to 300 ms window. **C**, **D**. Comparison between selectivity defined by using a linear SVM classifier (y-axis) and using a one-way ANOVA on the IFP responses to define selectivity (Thorpe et al., 1996). The x-axis shows the *p* value of obtained in the ANOVA analysis in log scale for all the electrodes (**C**) or only those electrodes that showed $p<0.01$ (**D**). The green line shows a linear fit to the data.

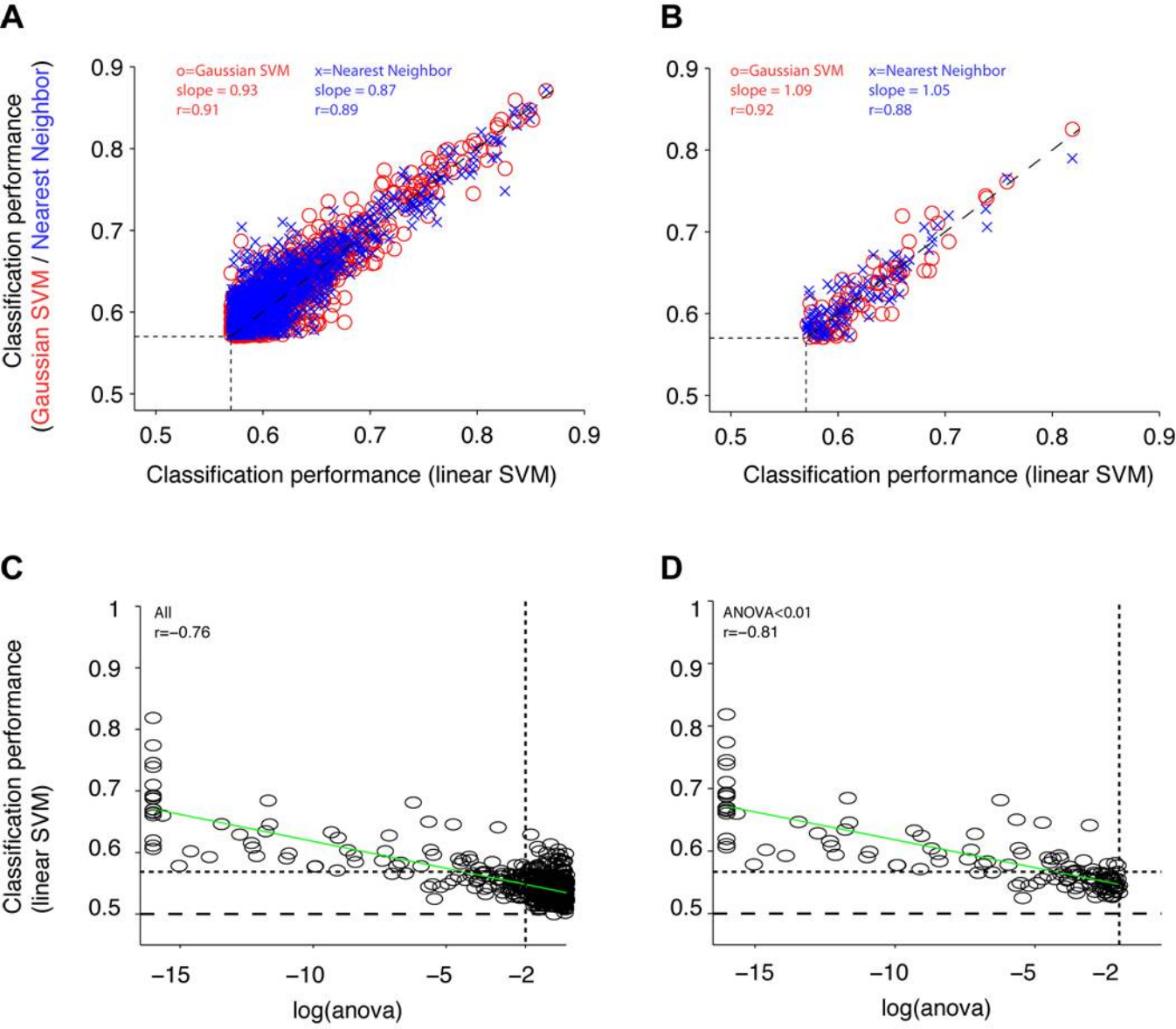
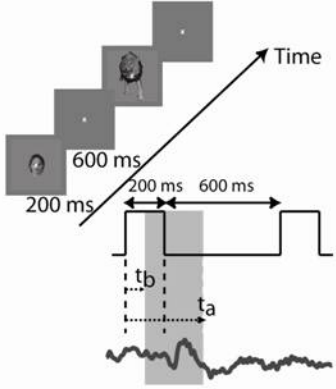
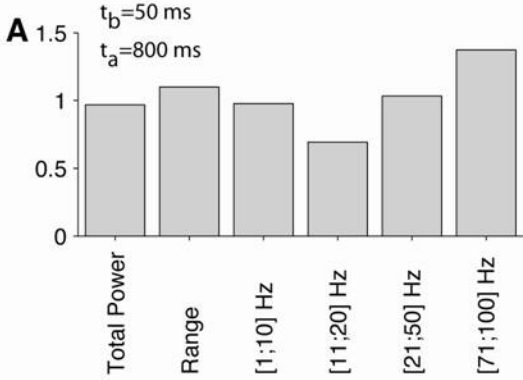

A
B
C
D
o=Gaussian SVM
slope = 0.93
r=0.91
x=Nearest Neighbor
slope = 0.87
r=0.89
o=Gaussian SVM
slope = 1.09
r=0.92
x=Nearest Neighbor
slope = 1.05
r=0.88
Classification performance
(Gaussian SVM / Nearest Neighbor)
Classification performance (linear SVM)
All
r=−0.76
ANOVA<0.01
r=−0.81
Classification performance
(linear SVM)
log(anova)
log(anova)

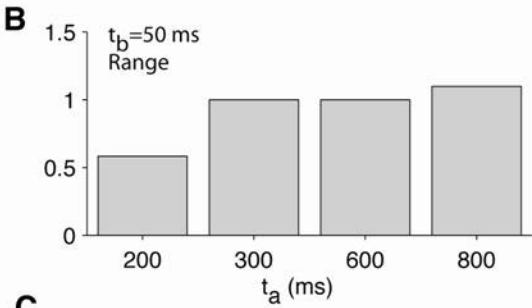
Figure S8: Neural codes and classifier input

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).





Fraction selective
(with respect to range in [50;300] ms)



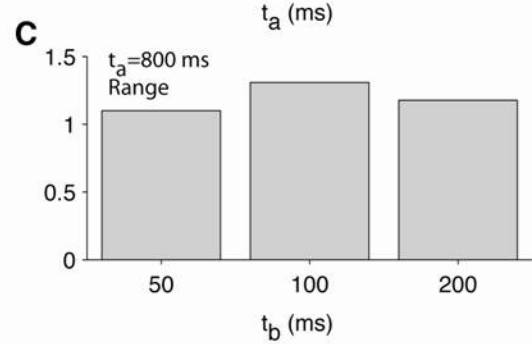


Figure S9: Correlation with basic image properties

To quantify the extent to which the IFP responses could be explained by low-level characteristics of the images, we considered a list of 15 basic image properties: mean pixel grayscale value, standard deviation of the pixel grayscale values, median of the pixel grayscale values, minimum pixel grayscale value, maximum pixel grayscale value, number of pixels different from the background gray, number of pixels below the background gray, number of pixels above the background gray, number of very dark pixels (grayscale value < 64), number of very bright pixels (grayscale > 192), number of gray pixels, number of boxes above background gray (box size = 20 pixels), number of boxes below background gray, number of very dark boxes (mean intensity within the 20x20 pixel box < 64), number of very bright boxes (mean intensity within the 20x20 pixel box > 192). For each electrode (n=912), and each image property (n=15), we computed the Pearson correlation coefficient between the IFP responses to each image and the corresponding property for each image. We considered the property that yielded the maximum correlation (that is, the image property that accounts for most of the variance in the IFP signals for that electrode). Here we show the distribution of the Pearson correlation coefficients for all the electrodes. Bin size = 0.02, the thick bar shows the average.

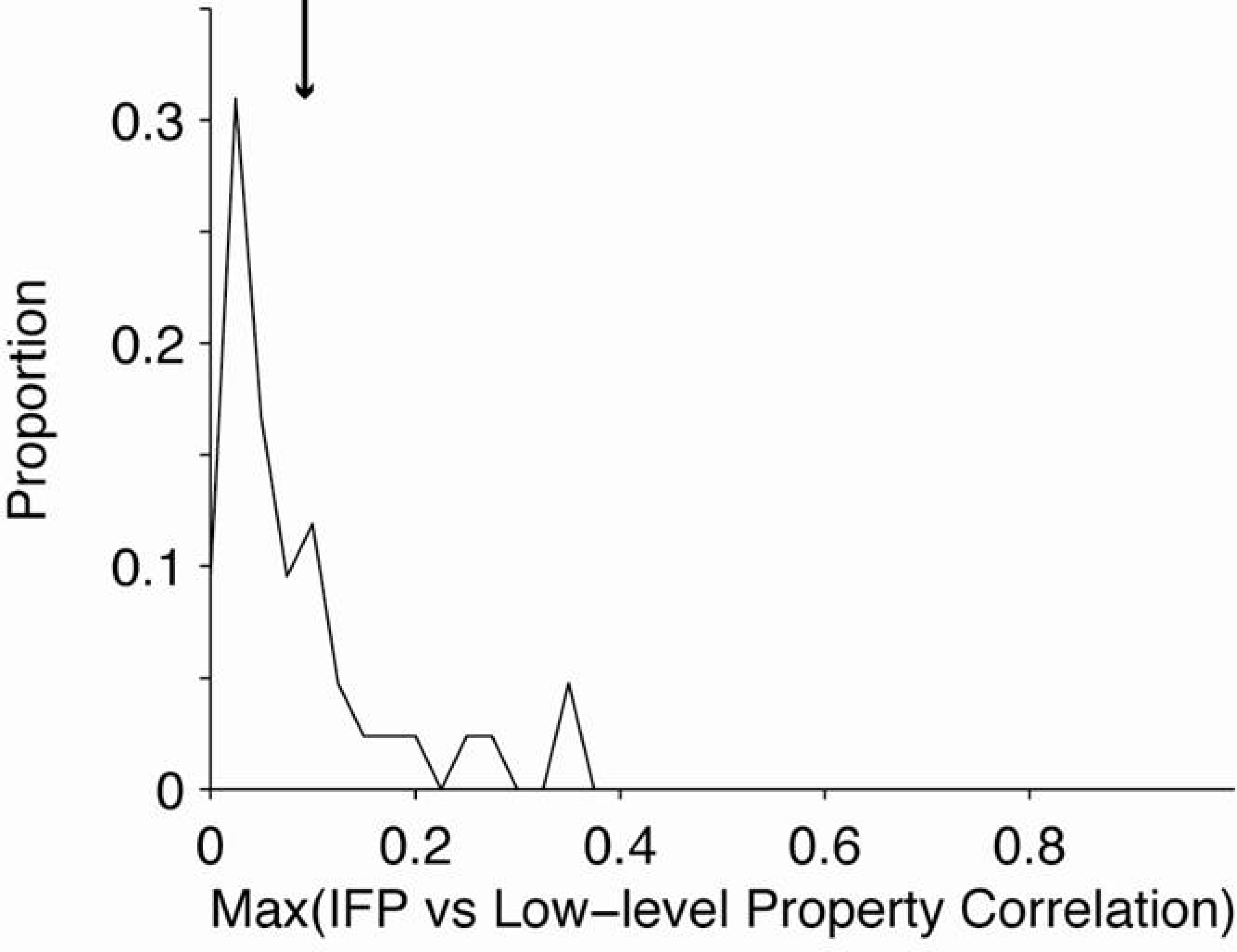

0.3
0.2
0.1
0
Proportion
0
0.2
0.4
0.6
0.8
Max(IFP vs Low−level Property Correlation)

Figure S10: Identification versus Categorization
A. Proportion of electrodes in each subject that showed significant *identification* performance (see Supplementary Material). **B.** Proportion of electrodes in each subject that showed significant categorization performance when categories were defined by randomly selecting 5 exemplars. **C.** Proportion of electrodes in each subject showing category selectivity when the categories were defined as in Figure S1**B** (i.e., the default definition throughout the text).

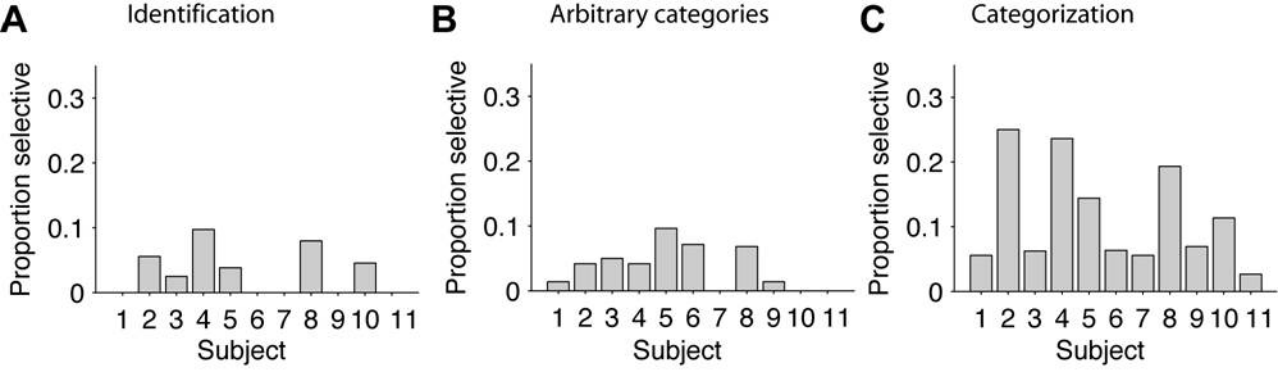

A
Identification
B
Arbitrary categories
C
Categorization
Proportion selective
0
0.1
0.2
0.3
1 2 3 4 5 6 7 8 9 10 11
Subject

Figure S11: Parcellation maps to localize the electrodes
To localize the electrodes, we co-registered pre-operative MR images with post-operative CT images (see Supplementary Material). For each electrode, we obtained the Talairach coordinates and assigned its location to one of 80 possible regions defined in (Dale et al., 1999; Desikan et al., 2006; Fischl et al., 2004). The ventral, medial and lateral views illustrated here show the approximate boundaries for these brain areas (see Table S2 for the full list and average Talairach coordinates).

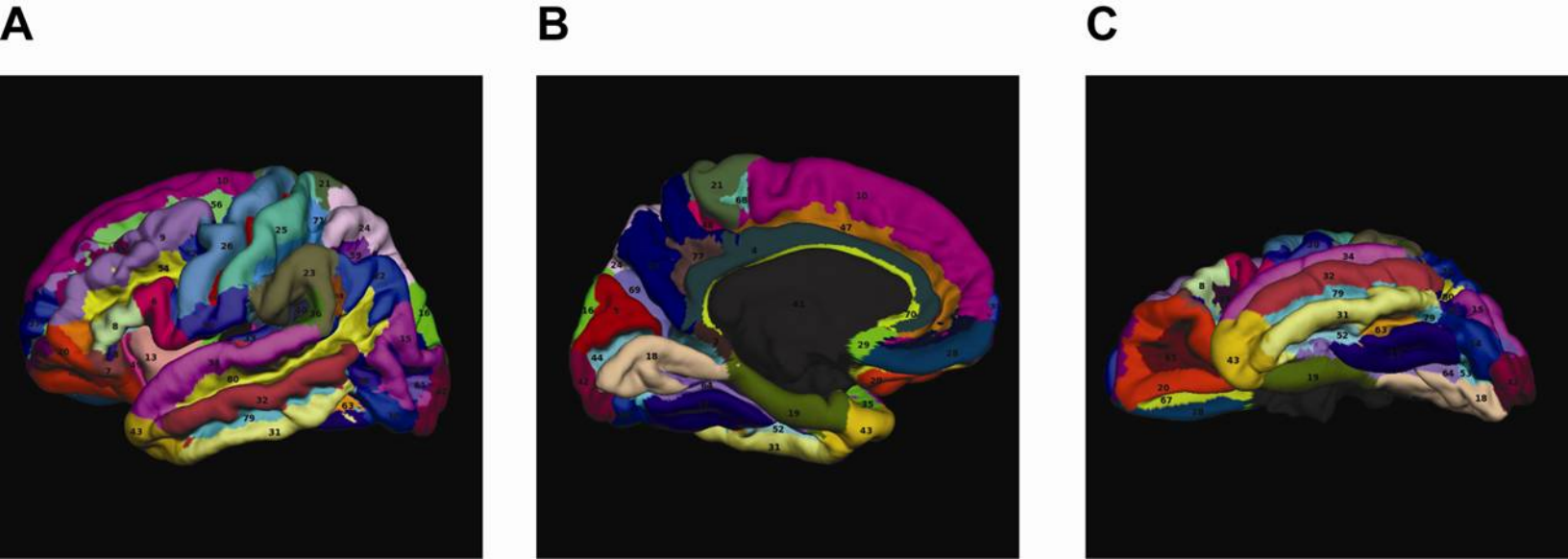

A
B
C

Figure S12: Location of individual electrodes showing selectivity

A. Fraction of electrodes that showed selectivity in each area where we recorded from at least 10 electrodes. The location of each electrode is based on co-registering the MR and CT scans for each subject (see Figure S11, Table S2 and Supplementary Material). The number before each location abbreviation indicates the "area code" within the parcellation maps in (Desikan et al., 2006) (Table S2). Locations in the occipital lobe have are depicted with gray bars and locations in the temporal lobe are depicted with a black bar. Here and throughout the manuscript, locations in the right and left hemisphere were pooled together. **B.** Mean classification performance for each location. Error bars indicate SEM. The dashed horizontal line shows chance performance (0.5) and the dotted line shows the selectivity threshold (see Figure S5).

A
Fraction of selective electrodes
1
0.8
0.6
0.4
0.2
0
6 F–OperInf 34
8 F–TriInf 33
9 F–Mid 49
10 F–Sup 23
11 F–Marg 16
20 F–Orb 54
26 P–PreCen 34
30 F–SubCen 24
14 O–Inf 21
15 O–Mid 18
69 O–Parietal 14
17 T–LTFus 24
18 T–MTLing 29
19 T–MTPhip 27
31 T–IT 88
32 T–MT 117
34 T–STLat 80
43 T–Pole 51
22 P–AngInf 28
23 P–SupraMargInf 34
25 P–PostCen 22
27 P–PreCun 14
B
Classification performance (sel)
0.8
0.7
0.6
0.5
14 O–Inf 18
17 T–LTFus 9
18 T–MTLing 6
19 T–MTPhip 7
31 T–IT 16
32 T–MT
43 T–Pole 7

Figure S13: 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., 2005a; 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 0.5 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 0.2 (5 object categories).

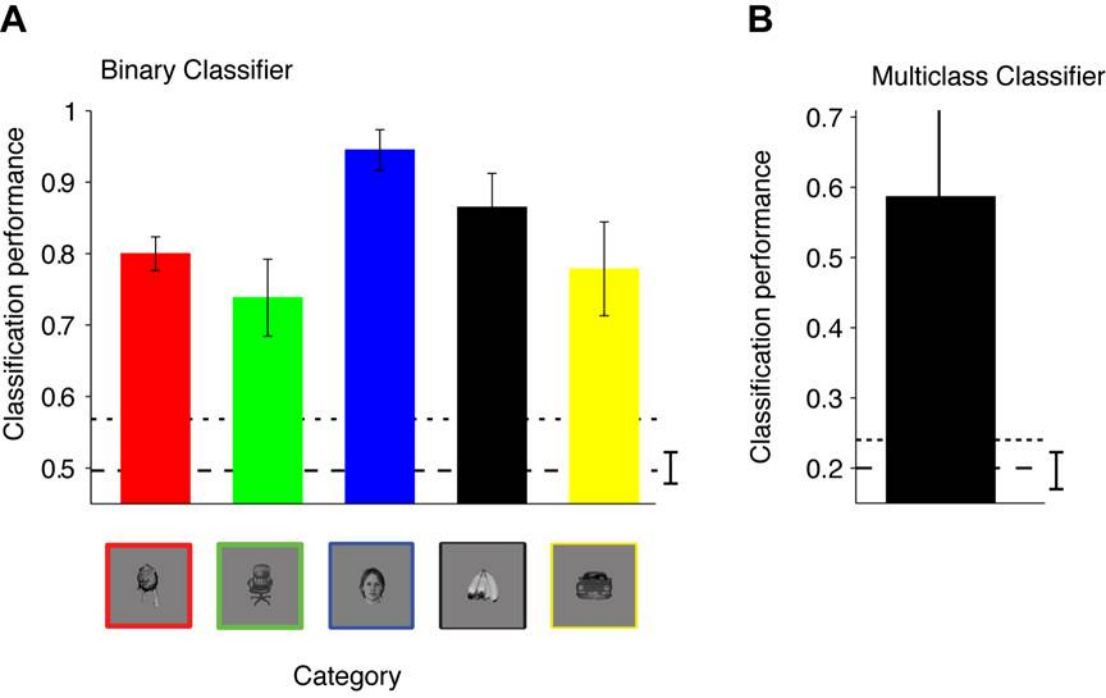

A
Binary Classifier
Classification performance
1
0.9
0.8
0.7
0.6
0.5
Category
B
Multiclass Classifier
Classification performance
0.7
0.6
0.5
0.4
0.3
0.2

Figure S14: Location analysis
In contrast to Figure S13, here the members of the neural ensemble were chosen based on the location of the electrode, randomly sampling to obtain a total of 10 electrodes in each location. The locations are separated by lobes: each row of subplots corresponds to a different brain lobe. The color and other conventions follow Figure S13. 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 in the main text.

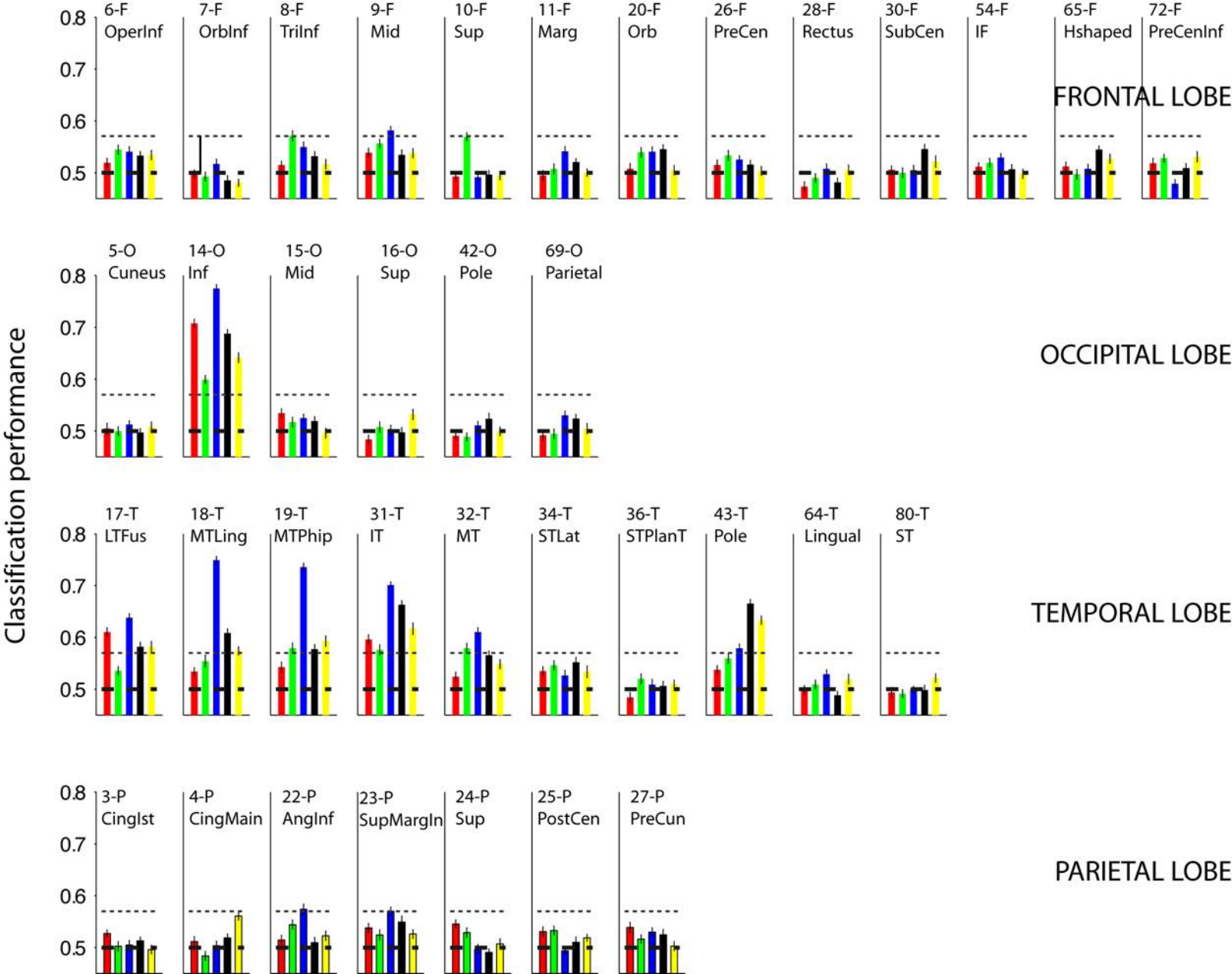

Classification performance
0.8
0.7
0.6
0.5
6-F OperInf
7-F OrbInf
8-F TriInf
9-F Mid
10-F Sup
11-F Marg
20-F Orb
26-F PreCen
28-F Rectus
30-F SubCen
54-F IF
65-F Hshaped
72-F PreCenInf
FRONTAL LOBE
5-O Cuneus
14-O Inf
15-O Mid
16-O Sup
42-O Pole
69-O Parietal
OCCIPITAL LOBE
17-T LTFus
18-T MTLing
19-T MTPhip
31-T IT
32-T MT
34-T STLat
36-T STPlanT
43-T Pole
64-T Lingual
80-T ST
TEMPORAL LOBE
3-P CingIst
4-P CingMain
22-P AngInf
23-P SupMargIn
24-P Sup
25-P PostCen
27-P PreCun
PARIETAL LOBE

Figure S15: Definition of latency and parameter dependency

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 (same electrode shown in Figure 1**F**), 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 $_{sel}n_{cons}$ = 25 consecutive time points (bin size = 2 ms) with $p<0.01$. The response latency was defined as the first time point where $_{lat}n_{cons}$ 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 3**A** in the main text. **B, C**. Dependence of the mean latency on $_{lat}n_{cons}$, $_{lat}n_{cons}$ and *p* value threshold. The horizontal dashed line shows the mean value reported in the text (the mean of the distribution in Figure 3**A** which corresponds to the parameters indicated by the arrow in the fifth subplot). 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.

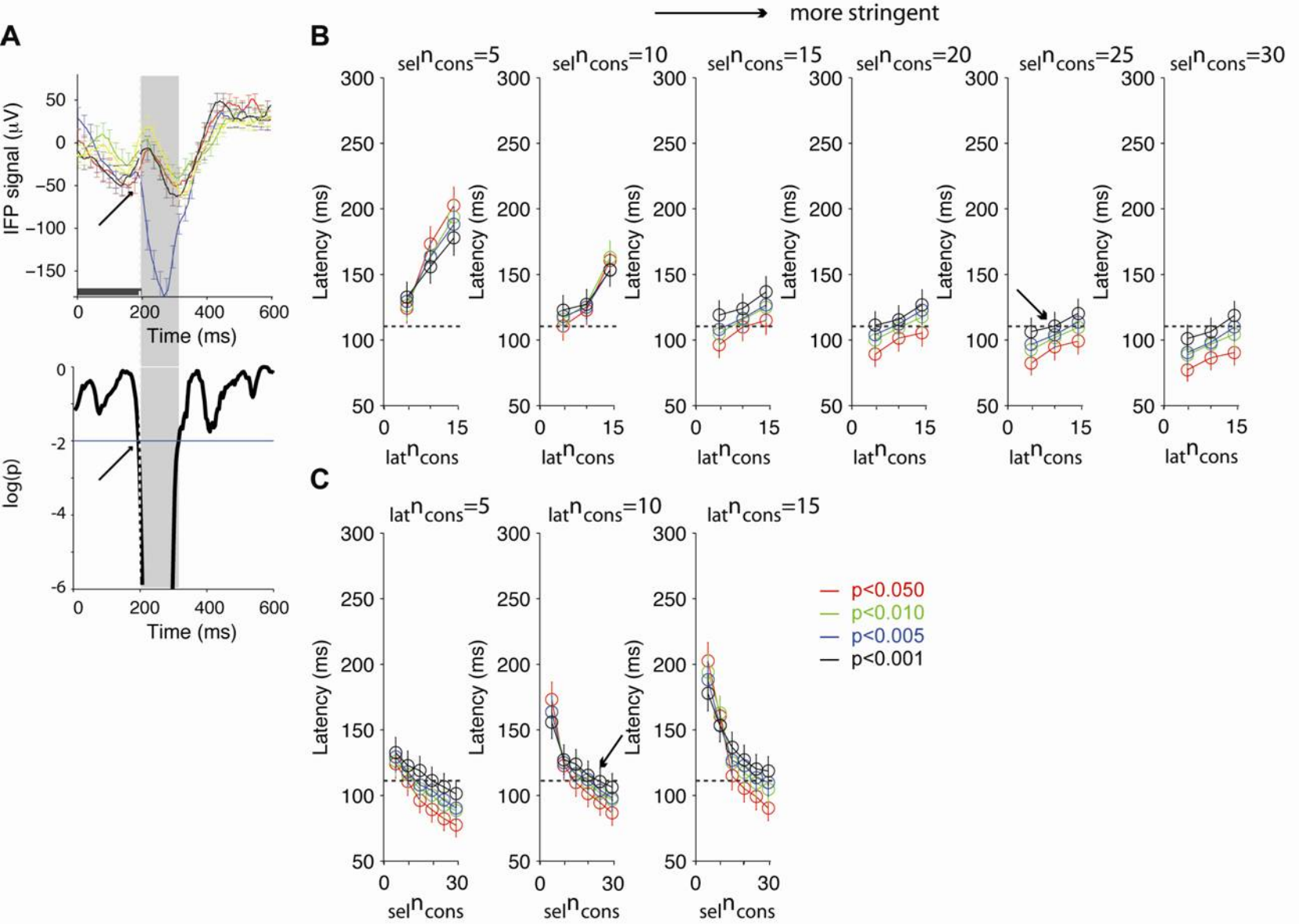

A
IFP signal (μV)
Time (ms)
log(p)
Time (ms)
B
more stringent
selncons=5
selncons=10
selncons=15
selncons=20
selncons=25
selncons=30
Latency (ms)
latncons
C
latncons=5
latncons=10
latncons=15
selncons
p<0.050
p<0.010
p<0.005
p<0.001

Figure S16: Latencies for locations with selective electrodes
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 are SEM.

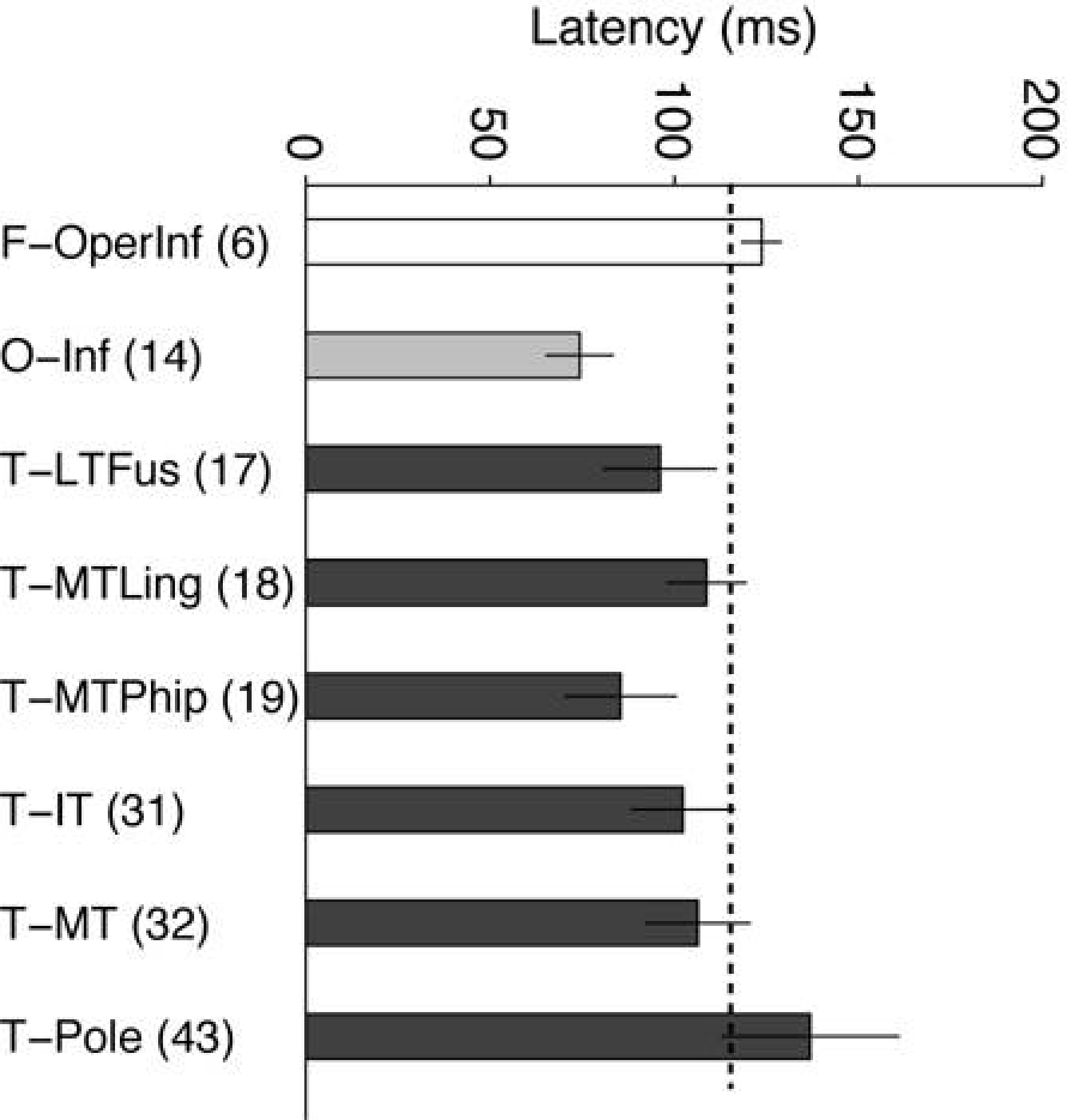

Latency (ms)
0
50
100
150
200
F–OperInf (6)
O–Inf (14)
T–LTFus (17)
T–MTLing (18)
T–MTPhip (19)
T–IT (31)
T–MT (32)
T–Pole (43)

Figure S17: 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 ô (**A, D**. ô=25 ms, **B, E**. ô=50 ms, c, f. ô=100 ms). In **A-C**, we used the power of the IFP signal in each bin whereas in **D-F** we used the range of the IFP signal in each bin. 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 S5** 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 **B** is reported in Figure 3**B** in the main text.

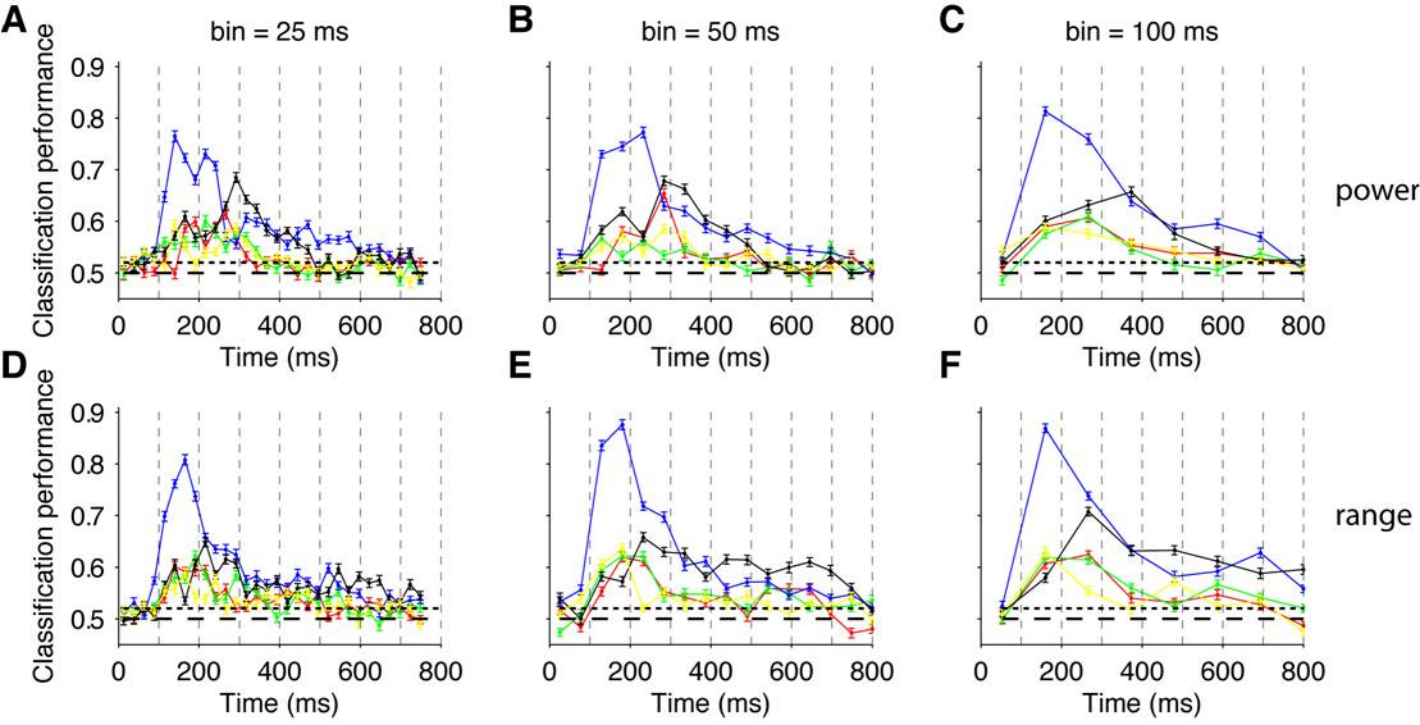
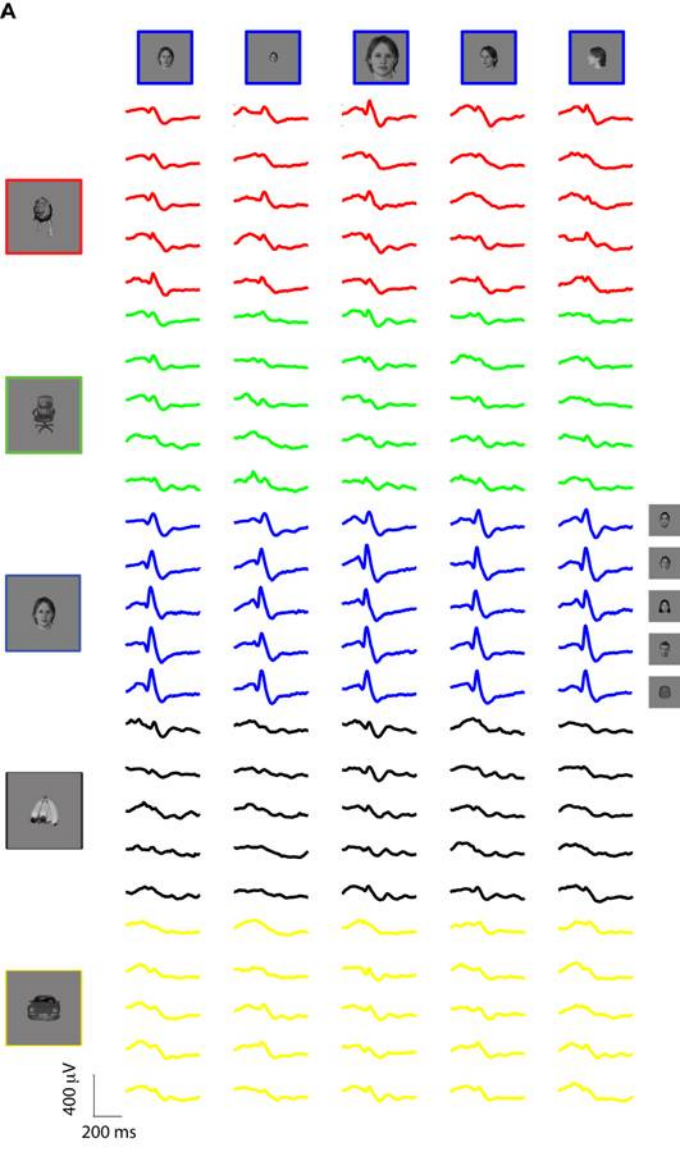

A
bin = 25 ms
B
bin = 50 ms
C
bin = 100 ms
power
D
E
F
range
Classification performance
0.9
0.8
0.7
0.6
0.5
0
200
400
600
800
Time (ms)

Figure S18: 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).


A
400 µV
200 ms

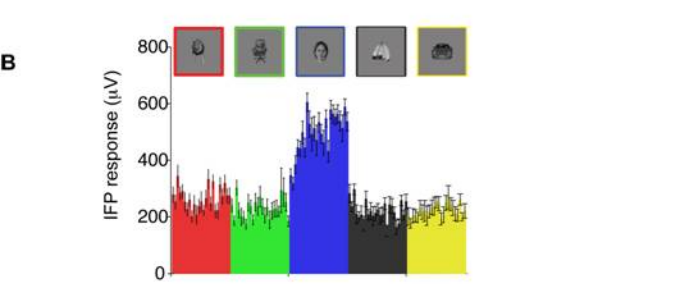

B
IFP response (µV)
800
600
400
200
0

Figure S19: Distribution of scale- and rotation-invariant responses
A. Distribution of classification performance values for the rotation-invariant electrodes (same format as in Figure S6**C**). 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. **B.** Distribution of classification performance values for the scale invariant electrodes. We emphasize that in both **A** and **B**, 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 (**A**) or scales (**B**). 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.

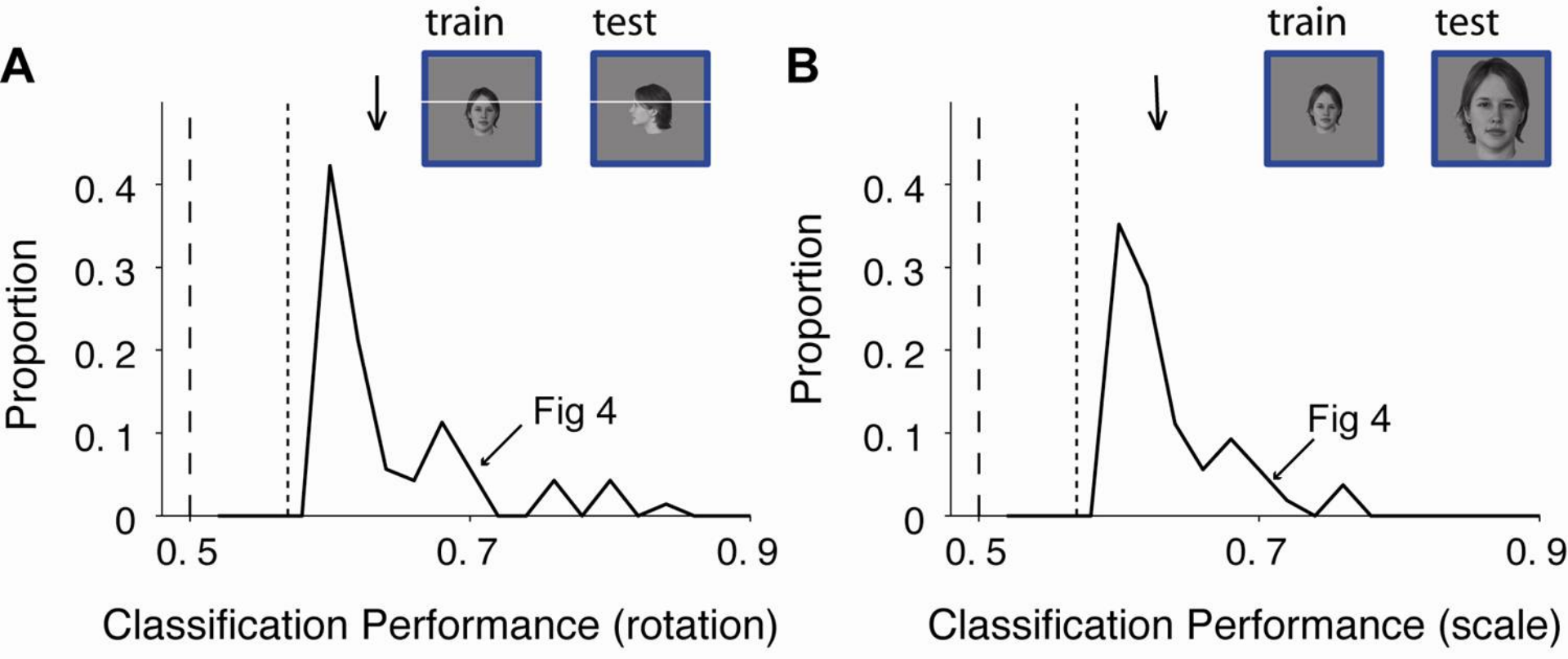

A
train
test
0. 4
0. 3
0. 2
0. 1
0
Proportion
Fig 4
0. 5
0. 7
0. 9
Classification Performance (rotation)
B
train
test
0. 4
0. 3
0. 2
0. 1
0
Proportion
Fig 4
0. 5
0. 7
0. 9
Classification Performance (scale)

Figure S20: Location of individual electrodes showing invariance

A, B. Fraction of electrodes showing invariance to scale changes (**A**) or rotation changes (**B**). **C**, **D**. Mean classification performance for each location with at least 5 electrodes showing invariance to scale (**C**) or rotation (**D**). The number next to the location descriptor indicates the "area code" in the parcellation in (Desikan et al., 2006) (**Table S2**). The color indicates the lobe (white=frontal, gray=occipital, black=temporal). Error bars indicate SEM. The dashed horizontal line shows chance classification performance (0.5) and the dotted line shows the classification performance threshold (see Figure S5).

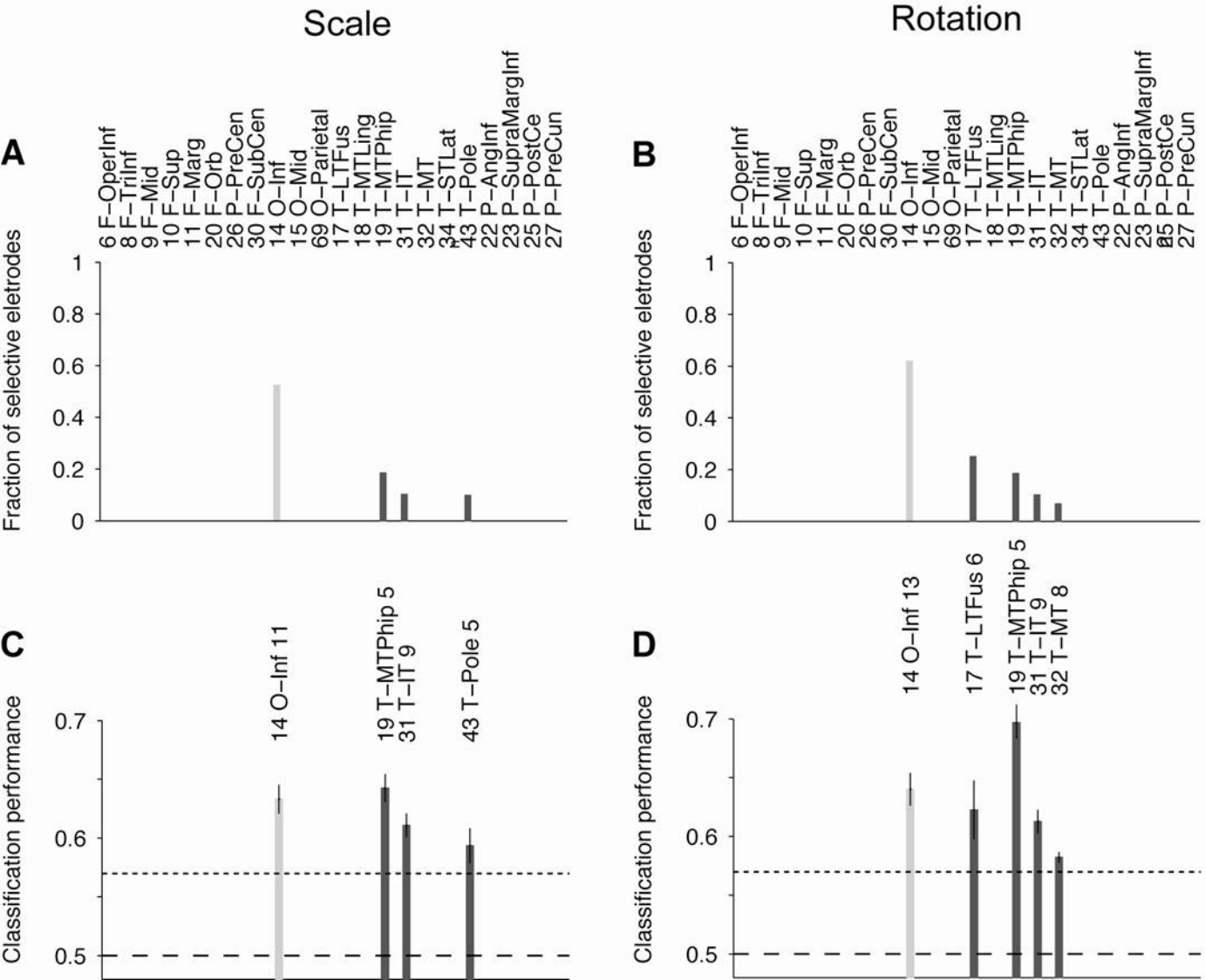

Scale
Rotation
A
B
C
D
Fraction of selective eletrodes
Classification performance
6 F–OperInf
8 F–TriInf
9 F–Mid
10 F–Sup
11 F–Marg
20 F–Orb
26 P–PreCen
30 F–SubCen
14 O–Inf
15 O–Mid
69 O–Parietal
17 T–LTFus
18 T–MTLing
19 T–MTPhip
31 T–IT
32 T–MT
34 T–STLat
43 T–Pole
22 P–AngInf
23 P–SupraMargInf
25 P–PostCe
27 P–PreCun
14 O–Inf 11
19 T–MTPhip 5
31 T–IT 9
43 T–Pole 5
14 O–Inf 13
17 T–LTFus 6
19 T–MTPhip 5
31 T–IT 9
32 T–MT 8

Figure S21: The neural ensemble extrapolates across rotations/scales
The format and conventions are the same as in Figure S13. **A.** Scale invariance. The classifier was trained with the neural responses obtained upon presenting objects at the default scale (3 degrees) and testing its 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 scales. **B.** Rotation invariance. The classifier was trained with the neural responses obtained upon presenting objects at the default rotation (Figure S1) and testing its 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.

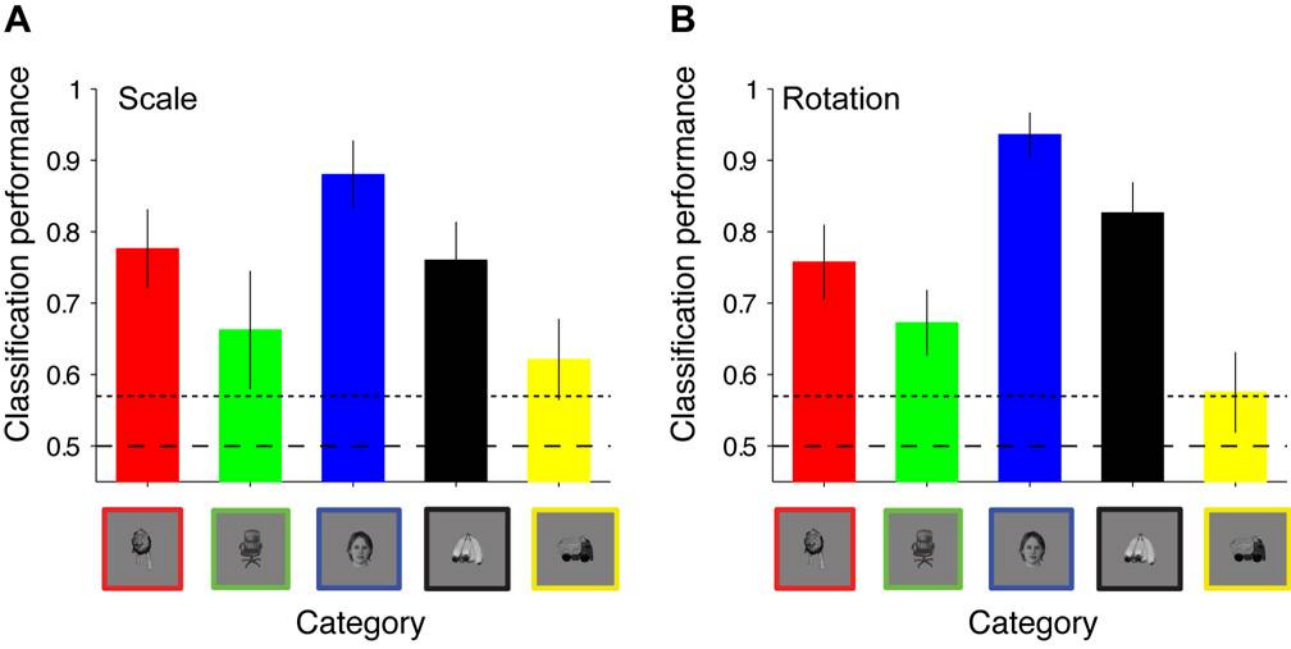

A
Scale
Classification performance
1
0.9
0.8
0.7
0.6
0.5
Category
B
Rotation
Classification performance
1
0.9
0.8
0.7
0.6
0.5
Category

Figure S22: Location analysis (invariance)
In contrast to Figure S21, here the members of the neural ensemble were chosen based on the location of the electrode, randomly sampling to obtain a total of 10 electrodes in each location. The format and conventions are the same as the ones in Figure S14. Here there are two bars for each location and category, one for rotation invariance (left) and one for scale invariance (right). The areas that showed the highest classification performance values from this figure are shown in Figure 5 in the main text.

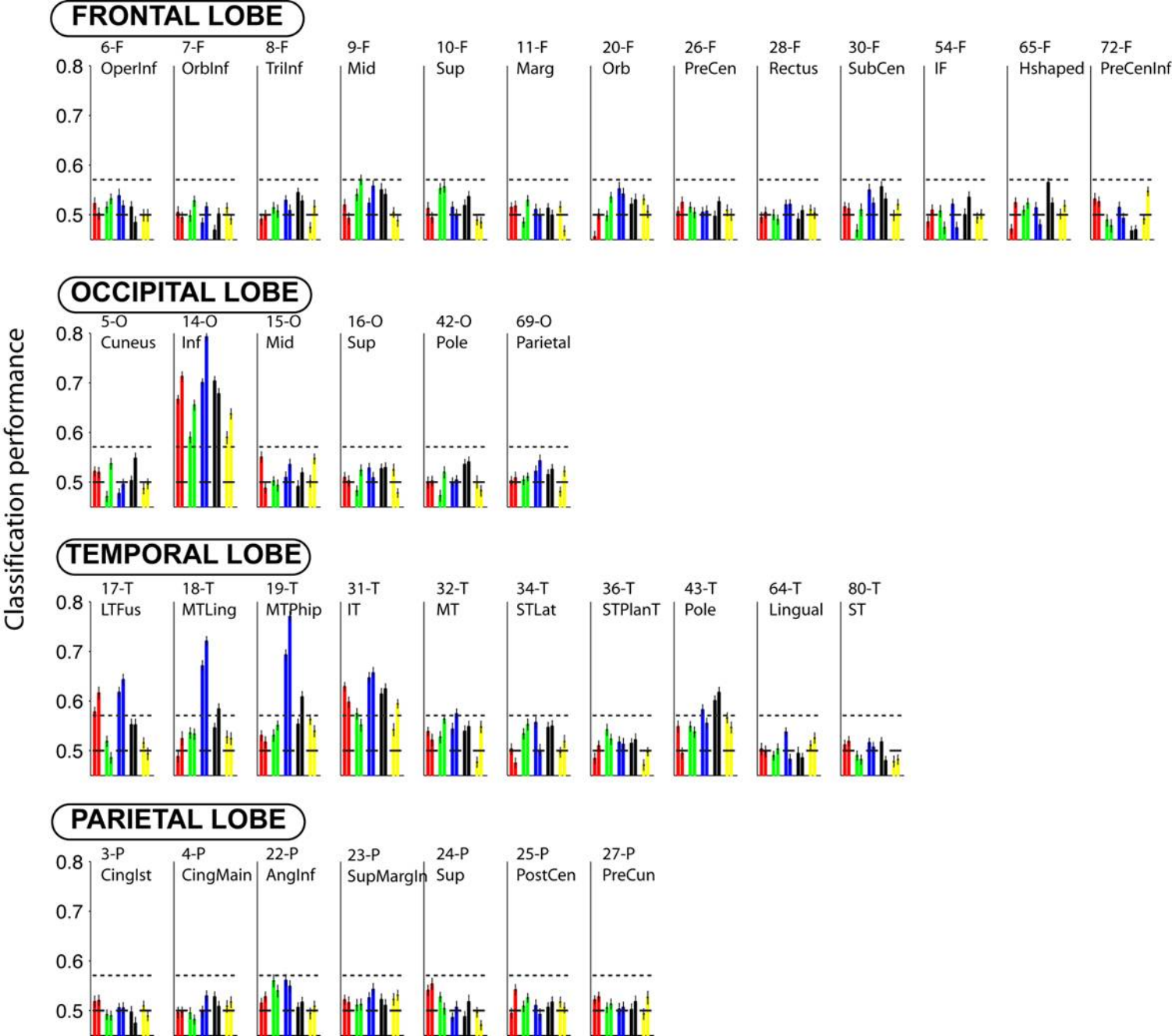

FRONTAL LOBE
6-F OperInf
7-F OrbInf
8-F TriInf
9-F Mid
10-F Sup
11-F Marg
20-F Orb
26-F PreCen
28-F Rectus
30-F SubCen
54-F IF
65-F Hshaped
72-F PreCenInf
OCCIPITAL LOBE
5-O Cuneus
14-O Inf
15-O Mid
16-O Sup
42-O Pole
69-O Parietal
TEMPORAL LOBE
17-T LTFus
18-T MTLing
19-T MTPhip
31-T IT
32-T MT
34-T STLat
36-T STPlanT
43-T Pole
64-T Lingual
80-T ST
PARIETAL LOBE
3-P CingIst
4-P CingMain
22-P AngInf
23-P SupMargIn
24-P Sup
25-P PostCen
27-P PreCun
0.8
0.7
0.6
0.5
Classification performance

Figure S23: Distribution of latencies for the scaled and rotated images
Distribution of IFP latency for the "default" images (black), rotated images (dark gray) and scaled images (light gray). The latency was defined as illustrated in Figure S15. The vertical lines show the mean of each distribution.

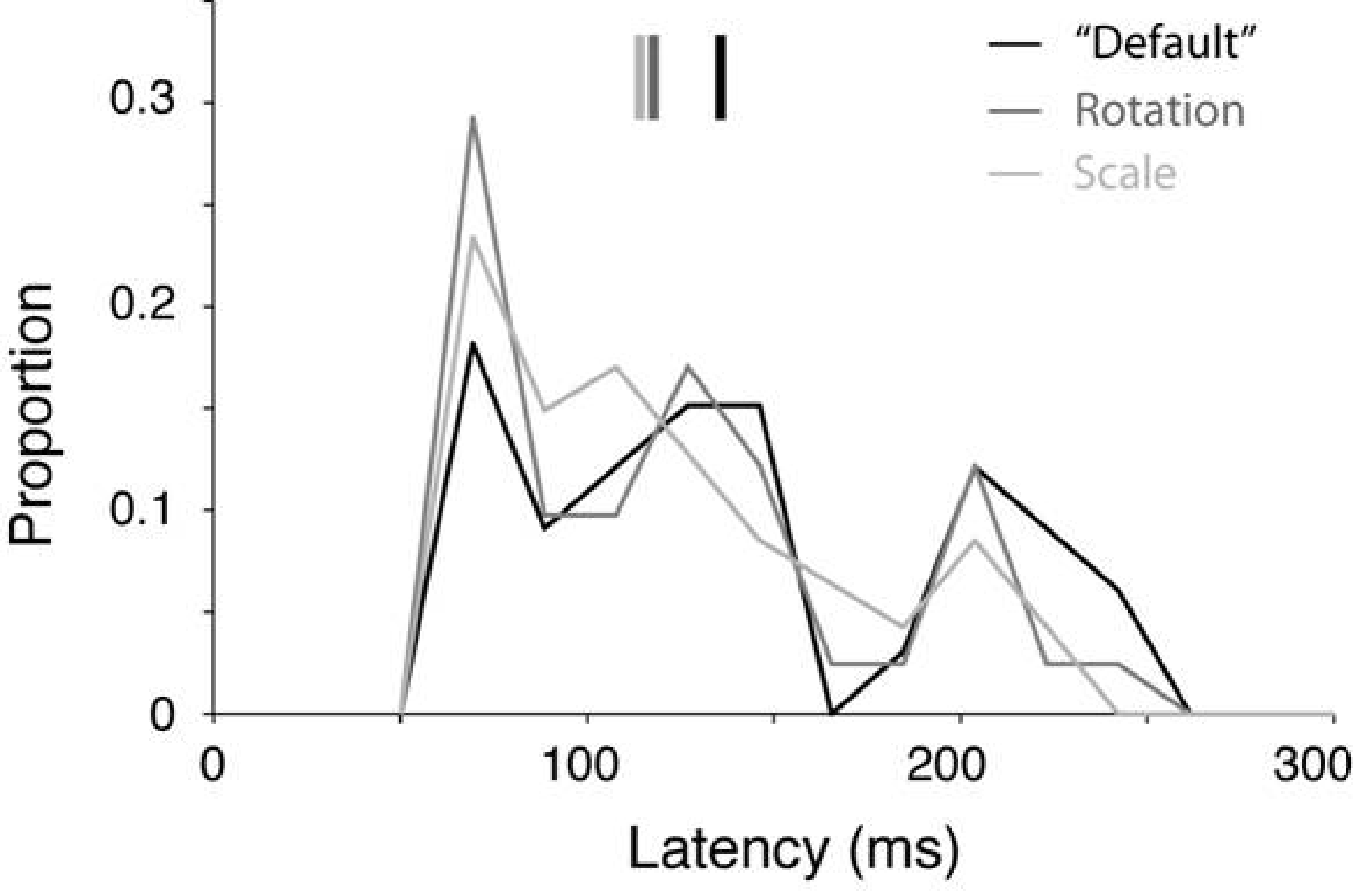
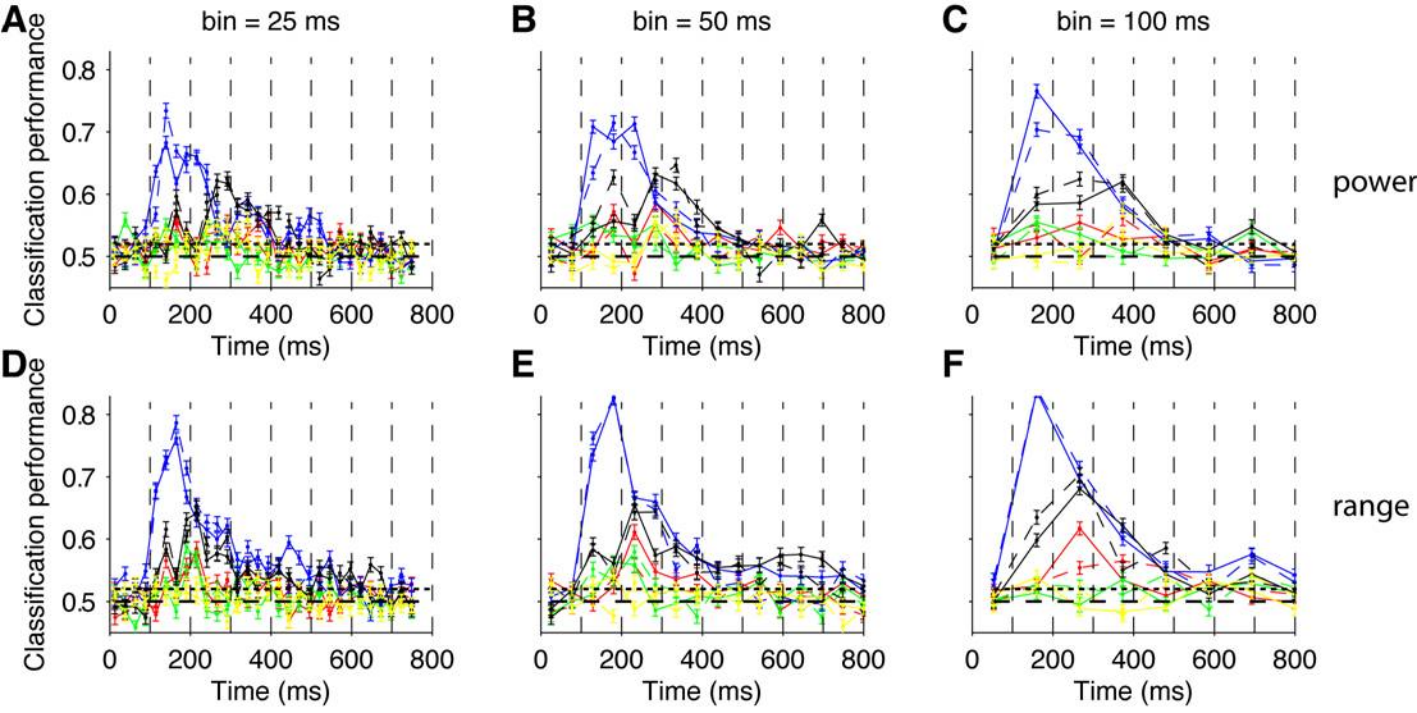

"Default"
Rotation
Scale
Proportion
0.3
0.2
0.1
0
0
100
200
300
Latency (ms)

Figure S24: Dynamics of invariant responses
Dynamics of decoding performance when extrapolating across scales and rotations. The format and conventions are the same as in **Figure S17**. Here there are two curves for each category, one corresponds to rotation invariance (solid line) and the other one to scale invariance (dashed line).


A
bin = 25 ms
B
bin = 50 ms
C
bin = 100 ms
power
D
E
F
range
Classification performance
0.8
0.7
0.6
0.5
0
200
400
600
800
Time (ms)

References

Ashburner, J., and Friston, K. (1997). Multimodal image coregistration and partitioning--a unified framework. Neuroimage *6*, 209-217.

Ashburner, J., and Friston, K.J. (2000). Voxel-based morphometry--the methods. Neuroimage *11*, 805-821.

Bishop, C.M. (1995). Neural Networks for Pattern Recognition (Oxford: Clarendon Press).

Carmena, J.M., Lebedev, M.A., Crist, R.E., O'Doherty, J.E., Santucci, D.M., Dimitrov, D.F., Patil, P.G., Henriquez, C.S., and Nicolelis, M.A. (2003). Learning to control a brain-machine interface for reaching and grasping by primates. PLoS Biol *1*, E42.

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).

Dale, A.M., Fischl, B., and Sereno, M.I. (1999). Cortical surface-based analysis. I. Segmentation and surface reconstruction. Neuroimage *9*, 179-194.

Desikan, R.S., Segonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T., *et al*. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage *31*, 968-980.

Fell, J., Klaver, P., Lehnertz, K., Grunwald, T., Schaller, C., Elger, C.E., and Fernandez, G. (2001). Human memory formation is accompanied by rhinal-hippocampal coupling and decoupling. Nat Neurosci *4*, 1259-1264.

Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Segonne, F., Salat, D.H., Busa, E., Seidman, L.J., Goldstein, J., Kennedy, D., *et al*. (2004). Automatically parcellating the human cerebral cortex. Cereb Cortex *14*, 11-22.

Green, D., and Swets, J. (1966). Signal detection theory and psychophysics (New York: Wiley).

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

Hung, C., Kreiman, G., Poggio, T., and Dicarlo, J. (2005b). Ultra-fast object recognition from few spikes. (Boston, MIT).

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.

McCarthy, G., Puce, A., Belger, A., and Allison, T. (1999). Electrophysiological studies of human face perception. II: Response properties of face-specific potentials generated in occipitotemporal cortex. Cerebral Cortex *9*, 431-444.

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).

Privman, E., Nir, Y., Kramer, U., Kipervasser, S., Andelman, F., Neufeld, M., Mukamel, R., Yeshurun, Y., Fried, I., and Malach, R. (2007). Enhanced category tuning revealed by iEEG in high order human visual areas. Journal of Neuroscience *6*, 6234-6242.

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.
Santhanam, G., Ryu, S.I., Yu, B.M., Afshar, A., and Shenoy, K.V. (2006). A high performance brain-computer interface. Nature *442*, 195-198.
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).