Supplemental Material

Golomb, J.D, and Kanwisher, N. (2012). Higher-level visual cortex represents retinotopic, not spatiotopic, object location. *Cerebral Cortex.*

Contents:

- Supplemental Figures S1-S3
- Supplemental Methods
- Supplemental Results/Discusion including results from individual topographically-defined ROIs, an alternative MVPA technique (Support Vector Machine analysis), a variation of Experiment 3 confounding spatiotopic location and motor code, and an exploratory eye position information analysis.

Supplemental Figures

Figure S1. Example Confusion Matrices. SVM confusion matrices (illustrating the percentage of time the model classified each of the 12 conditions with each of the 12 labels) are shown for Experiment 1 for the same few sample regions as Figure 4. Data are shown from the "test" datasets; the model was trained on separate "training" datasets. LOC exhibits a combination of category and retinotopic location information, FFA exhibits primarily category information, and early visual cortex primarily retinotopic location information. N=8.

Figure S2. SVM Results, Experiment 1. A-B, Same and different category and location confusion values for each region. **A,** Comparisons for conditions sharing the same eye position; location is the combined retinotopic and spatiotopic location. All regions show main effects of category and location. **B,** Comparisons for conditions differing in eye position, where retinotopic and spatiotopic locations can be dissociated. In addition to the category effect, all regions exhibit greater confusion when retinotopic location is preserved. **C,** The amount of information about each type of information was calculated by subtracting "same" minus "different" values for that type of information. Compare to Figure 5.

Figure S3. Searchlight Analysis, Manual Response confound. Results from searchlight analysis for the Experiment 3 variation (direct report of spatiotopic position) projected onto inflated brains. Group-averaged maps of significant category information (yellow), retinotopic information (red), and spatiotopic information (blue) are overlaid on top-views of both hemispheres. Large "spatiotopic" clusters can be seen around left motor cortex, but these are confounded with right hand motor output. N=7.

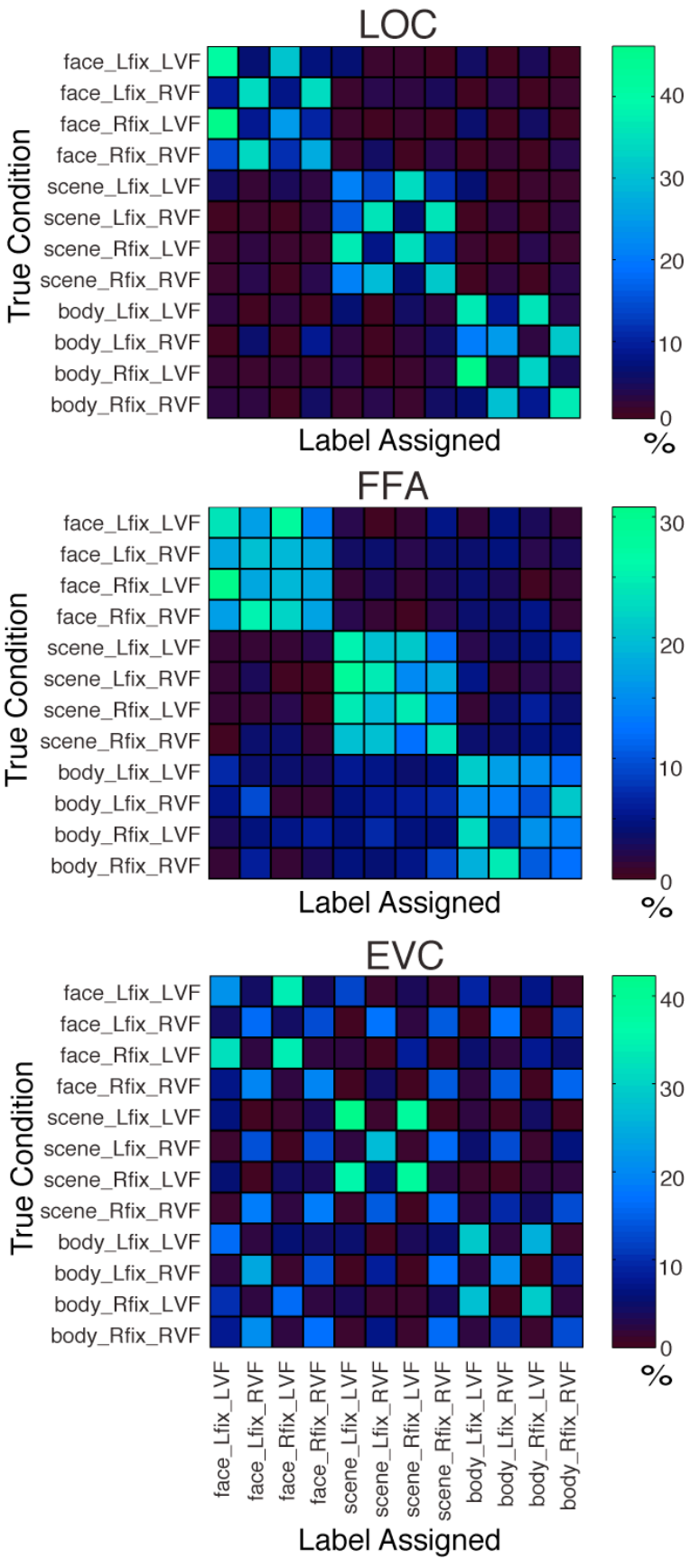
LOC
FFA
EVC
True Condition
Label Assigned
face_Lfix_LVF
face_Lfix_RVF
face_Rfix_LVF
face_Rfix_RVF
scene_Lfix_LVF
scene_Lfix_RVF
scene_Rfix_LVF
scene_Rfix_RVF
body_Lfix_LVF
body_Lfix_RVF
body_Rfix_LVF
body_Rfix_RVF
%

Figure S1

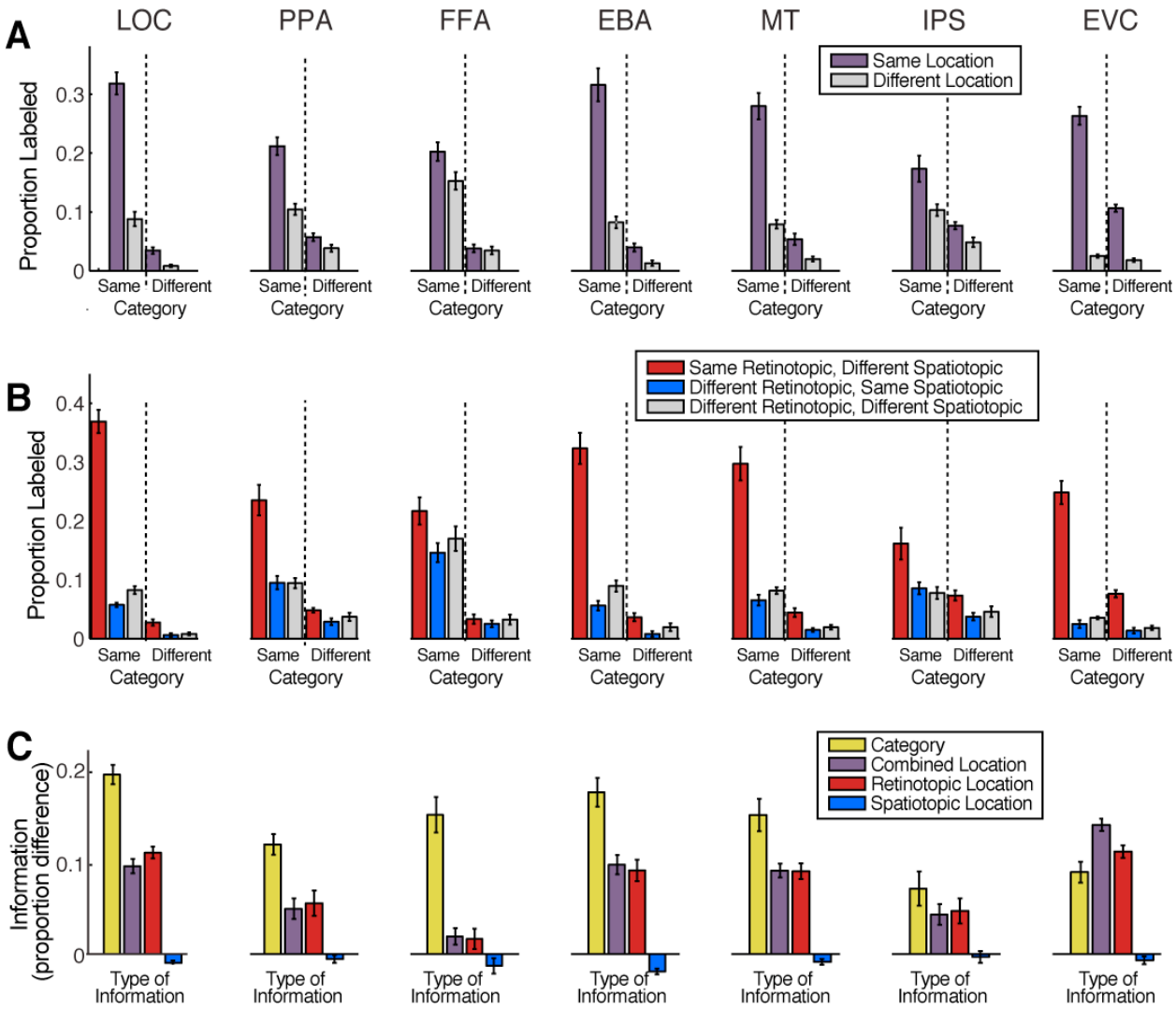
A
LOC
PPA
FFA
EBA
MT
IPS
EVC
Same Location
Different Location
Proportion Labeled
0
0.1
0.2
0.3
Same
Different
Category
B
Same Retinotopic, Different Spatiotopic
Different Retinotopic, Same Spatiotopic
Different Retinotopic, Different Spatiotopic
Proportion Labeled
0
0.1
0.2
0.3
0.4
Same
Different
Category
C
Category
Combined Location
Retinotopic Location
Spatiotopic Location
Information (proportion difference)
0
0.1
0.2
Type of Information

Figure S2

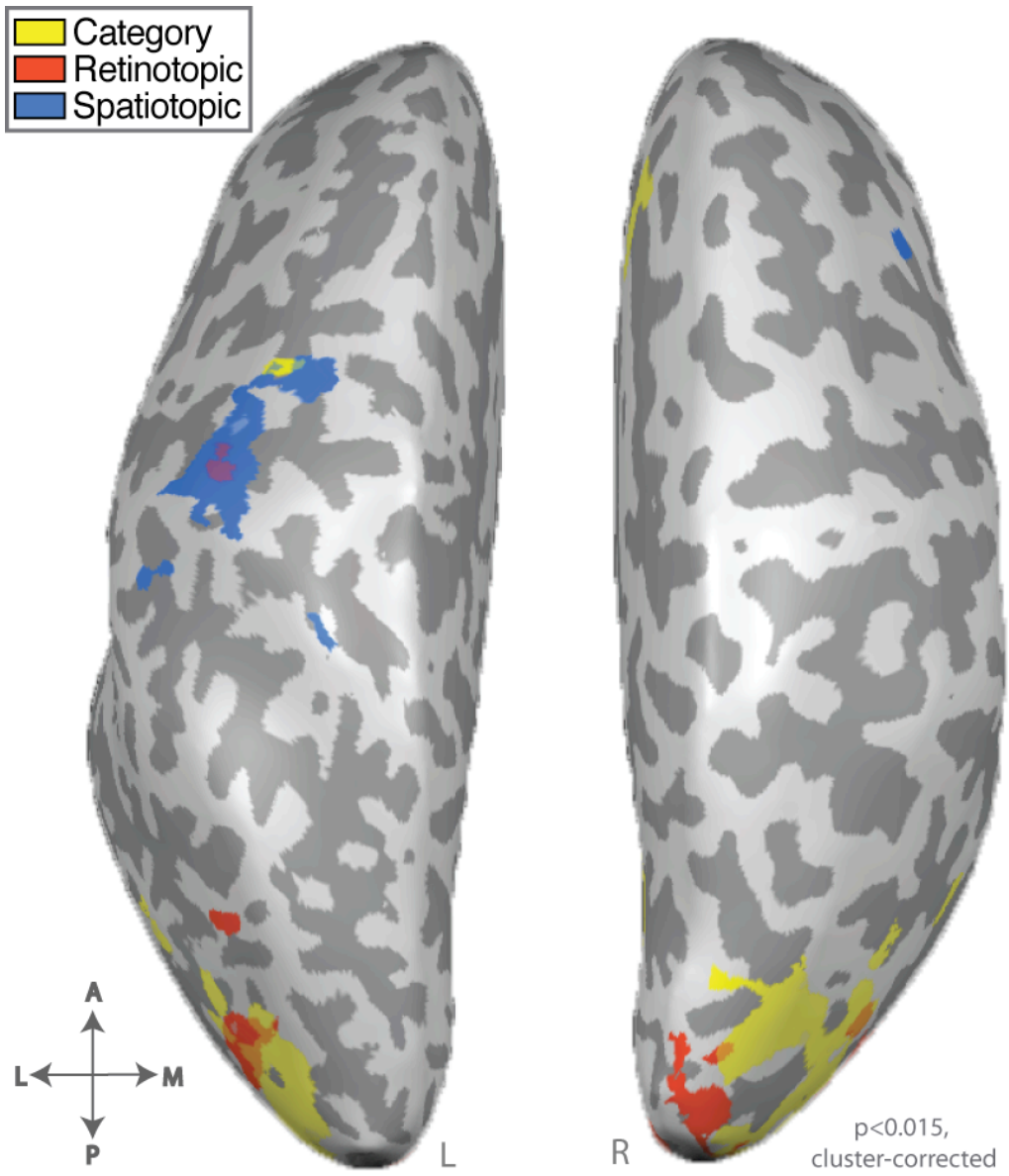
Category
Retinotopic
Spatiotopic
A
L
M
P
L
R
p<0.015,
cluster-corrected

Figure S3

Supplemental Methods

Support Vector Machine (SVM) analysis. As an alternate multivariate pattern analysis technique, an SVM model was trained to discriminate between the 12 conditions using a training dataset, and then tested on the remaining test dataset. The SVM used a leave-one-out technique, where on each iteration of the model, one run (containing a block of each condition) was reserved for the test dataset, and the remaining runs were used for the training dataset (Kamitani and Tong 2005). The SVM was implemented using the OSU-SVM and LIBSVM Matlab toolbox (Chang and Lin 2001). On each iteration, the trained model was presented with 12 unlabeled trials (1 per condition), and asked to predict which condition each belonged to. The classification was based on a series of binary pairwise comparisons (*one-versus-one*), with the winning class determined using a "max-wins" voting strategy. Averaging across iterations, the percentage of time the model classified each of the 12 conditions with each of the 12 possible labels was calculated as a confusion matrix. The 12x12 confusion matrix is analogous to the 12x12 correlation matrix, and all "same" vs "different" comparisons were conducted analogously on the SVM results as described for the correlation analysis in the main text.

Supplemental Results/Discussion

Individual topographically-defined regions

MVPA data from Experiment 1 for each of the individual occipital and parietal regions defined by retintotopic mapping is shown in Figure 4D in the main text. In each of these regions, there were significant main effects of same versus different category (V1: $F(1,7)=54.0$, $p<0.001$; V2: $F(1,7)=63.5$, $p<0.001$; V3: $F(1,7)=240.8$, $p<0.001$; V4: $F(1,7)=97.3$, $p<0.001$; V3A: $F(1,7)=85.0$, $p<0.001$; V7: $F(1,7)=115.6$, $p<0.001$; IPS1: $F(1,7)=46.7$, $p<0.001$; IPS2: $F(1,7)=54.7$, $p<0.001$; IPS3: $F(1,7)=21.5$, $p=0.002$; IPS4: $F(1,7)=15.1$, $p=0.006$) and same versus different combined location (V1: $F(1,7)=88.0$, $p<0.001$; V2: $F(1,7)=49.2$, $p<0.001$; V3: $F(1,7)=53.8$, $p<0.001$; V4: $F(1,7)= 45.2$, $p<0.001$; V3A: $F(1,7)=37.9$, $p=0.001$; V7: $F(1,7)=46.5$, $p<0.001$; IPS1: $F(1,7)=15.1$, $p=0.006$; IPS2: $F(1,7)=13.5$, $p=0.008$; IPS3: $F(1,7)=7.08$, $p=0.032$; IPS4: $F(1,7)=7.32$, $p=0.030$).

Retinotopic location information was significant in each of these regions (V1: $t(7)=8.21$, $p<0.001$; V2: $t(7)=6.17$, $p<0.001$; v3: $t(7)=6.26$, $p<0.001$; V4: $t(7)=5.58$, $p=0.001$; V3A: $t(7)=5.74$, $p=0.001$; V7: $t(7)=5.22$, $p=0.001$; IPS1: $t(7)=2.73$, $p=0.029$; IPS2: $t(7)=3.05$, $p=0.019$; IPS3: $t(7)=2.45$, $p=0.044$; IPS4: $t(7)=3.23$, $p=0.014$). In none of the regions was spatiotopic information present; if anything, correlations were weaker when spatiotopic position was preserved than when both spatiotopic and retinotopic position differed (V1: $t(7)=-1.80$, $p=0.116$; V2: $t(7)=-2.62$, $p=0.034$; v3: $t(7)=-1.76$, $p=0.123$; V4: $t(7)=-1.74$, $p=0.126$; V3A: $t(7)=-2.11$, $p=0.073$; V7: $t(7)=-1.84$, $p=0.108$; IPS1: $t(7)=-1.63$, $p=0.147$; IPS2: $t(7)=-0.95$, $p=0.374$; IPS3: $t(7)=-1.50$, $p=0.178$; IPS4: $t(7)=-0.93$, $p=0.382$). A direct comparison of retinotopic to spatiotopic information also revealed significantly more retinotopic than spatiotopic information (V1: $t(7)=9.94$, $p<0.001$; V2: $t(7)=7.04$, $p<0.001$; v3: $t(7)=7.06$, $p<0.001$; V4: $t(7)=6.53$,

$p<0.001$; V3A: $t(7)=6.60$, $p<0.001$; V7: $t(7)=6.53$, $p<0.001$; IPS1: $t(7)=2.88$, $p=0.024$; IPS2: $t(7)=2.84$, $p=0.025$; IPS3: $t(7)=3.22$, $p=0.015$; IPS4: $t(7)=2.81$, $p=0.026$).

Alternative MVPA Technique: Support Vector Machine (SVM) Analysis

The data in the main text were derived from a type of multivariate pattern analysis using a cross-correlation technique: splitting the data into two halves ("odd" and "even") and correlating the voxel-wise patterns for each pair of odd conditions with those of each pair of even conditions. To confirm these results with an alternative MVPA technique, we also conducted an SVM analysis (see Supplemental Methods), where a support vector machine model was trained on a "training" dataset and then tested on a "test" dataset. For the correlation analysis, we can plot the correlations between each pair of conditions as a 12x12 *correlation matrix* (Figure 4). For the SVM analysis, an analogous 12x12 *confusion matrix* can also be plotted (Figure S1), summarizing the percentage of time the model classified each of the 12 conditions with each of the 12 possible labels.

The two techniques produce quite similar results. We quantified the SVM data for Experiment 1 in the same way we did the correlation data (Figure S2, compare to Figure 5). Although the scales are different – correlations can be negative because they are normalized by the mean across all conditions (Haxby et al. 2001), while the minimum confusion value is zero – the overall patterns are remarkably similar. As with the correlation technique, the highest confusions in the ventral and dorsal regions were for the same category of image presented in the same combined location, followed by same category presented in a different location, different category in the same location, and finally conditions differing in both category and location. Early visual cortex was more likely to confuse images in the same location than of the same

category. In each of these regions, there were significant main effects of same versus different category (LOC: $F(1,7)=225.9$, $p<0.001$; PPA: $F(1,7)=95.6$, $p<0.001$; FFA: $F(1,7)=62.0$, $p<0.001$; EBA: $F(1,7)=92.3$, $p<0.001$; MT: $F(1,7)=66.8$, $p<0.001$; IPS: $F(1,7)=17.2$, $p=0.004$; $F(1,7)=58.9$, $p<0.001$) and same versus different combined location (LOC: $F(1,7)=116.9$, $p<0.001$; PPA: $F(1,7)=23.8$, $p=0.002$; FFA: $F(1,7)=6.3$, $p=0.04$; EBA: $F(1,7)=71.6$, $p<0.001$; MT: $F(1,7)=110.5$, $p<0.001$; IPS: $F(1,7)=12.4$, $p=0.010$; $F(1,7)=362.9$, $p<0.001$).

There was significant information about retinotopic location in every region tested except FFA (LOC: $t(7)=17.1$, $p<0.001$; PPA: $t(7)=4.04$, $p=0.005$; FFA: $t(7)=1.47$, $p=0.184$; EBA: $t(7)=7.77$, $p<0.001$; MT: $t(7)=10.4$, $p<0.001$; IPS: $t(7)=3.45$, $p=0.011$; EVC: $t(7)=16.4$, $p<0.001$), and in none of the regions was positive spatiotopic information present (LOC: $t(7)=-3.68$, $p=0.008$; PPA: $t(7)=-1.30$, $p=0.235$; FFA: $t(7)=-1.54$, $p=0.168$; EBA: $t(7)=-6.15$, $p<0.001$; MT: $t(7)=-2.64$, $p=0.033$; IPS: $t(7)=-0.487$, $p=0.641$; EVC: $t(7)=-1.57$, $p=0.161$). A direct comparison of retinotopic to spatiotopic information also revealed significantly more retinotopic than spatiotopic information in every region tested (LOC: $t(7)=17.3$, $p<0.001$; PPA: $t(7)=4.28$, $p=0.004$; FFA: $t(7)=2.83$, $p=0.026$; EBA: $t(7)=9.46$, $p<0.001$; MT: $t(7)=10.26$, $p<0.001$; IPS: $t(7)=4.00$, $p=0.005$; EVC: $t(7)=13.7$, $p<0.001$).

The two techniques thus reach converging conclusions; in every region, for all of the above statistical tests, the two techniques produced the same results. Interestingly, the SVM technique seemed to be more sensitive than the correlation technique at detecting category information in early visual cortex. Although there was significant category information detected by both techniques, it clearly influenced the SVM confusion pattern more than the correlation pattern. Other regions exhibited relatively consistent proportions of category and location information across the two techniques, so there does not seem to be an overall difference in

sensitivity, but potential differences in early visual category information could be an interesting question for future research.

Experiment 3 Variation: Spatiotopic Location / Motor Confound

To test whether the amount of retinotopic and spatiotopic location information would be influenced by task, Experiment 3 used two separate tasks: one that required subjects to report the category of the image, and the other the spatiotopic location of the image. As described in the main text, these tasks used an indirect report technique. However, we also tested a variation of Experiment 3 using a direct report of spatiotopic position. In this variation, subjects held a button box in their right hand and pressed one of three buttons corresponding with spatiotopic position: the left button with their index finger for "left", the middle button with middle finger for "middle", and the right button with ring finger for "right". Seven subjects participated in this experiment.

Unlike the other experiments that revealed negligible spatiotopic information, when a searchlight analysis was run on this data (Figure S3), robust spatiotopic information was present even at strict statistical thresholds. It is tempting to conclude that the spatiotopic task-relevance increased the prevalence of spatiotopic information. However, given the locus of this spatiotopic information near left motor cortex, it seems more likely that the searchlight was reflecting a decoding of motor response, not spatiotopic visual code. In other words, the voxel-wise response pattern in left motor cortex was distinguishing which of the right-hand fingers was used to make the response on a given trial. While this is an interesting proof of concept, it says nothing about whether spatiotopic location was also represented. Because of the direct, one-to-one mapping between spatiotopic location and response finger, these two aspects of the task cannot be differentiated. Thus the importance of using the indirect report technique described in the main

text Experiment 3, where location and response were unconfounded and we could search for true spatiotopic visual information.

Eye position information

As suggested in the main text, it is possible that information about spatiotopic position is in fact present in our visual object processing regions, just not in an explicit form. In other words, in lieu of an explicit spatiotopic map of the world, spatiotopic position could be computed on the fly by combining retinotopic position with information about the current eye position. If this is the case, we might also expect to find information about eye position in these same regions. Although our experiments were not originally designed to test this idea, we could nonetheless look within our ROIs for information about fixation position. In Experiment 1, eye position was represented in many of the regions (LOC: $t(7)=2.47$, $p=0.043$; PPA: $t(7)=4.36$, $p=0.003$; FFA: $t(7)=1.65$, $p=0.142$; EBA: $t(7)=1.99$, $p=0.087$; MT: $t(7)=2.31$, $p=0.054$; IPS: $t(7)=1.89$, $p=0.101$; EVC: $t(7)=2.70$, $p=0.030$). In Experiment 2 eye position information was present in early visual cortex (4 out of 4 subjects), but was not apparent in the other regions, perhaps due to the vertical stimulus arrangement and/or the fact that the fixation positions were located closer together (5.2 deg vs 7 deg). Interestingly, in Experiment 3 eye position information was quite robust, equaling or even exceeding the magnitude of retinotopic information in most regions. For the category task, in every region eye position information was present in 4 out of 4 subjects, with the exception of one subject in IPS; in the spatiotopic location task, eye position information was present in 4 out of 4 subjects for every region except the FFA and EBA (3 out of 4). One intriguing explanation could be that because stimulus location was no longer blocked and eye position was less predictable in Experiment 3, this information became more relevant for the task. However, this eye position information did not simply reflect an

artifact of saccade related activity: in the event-related experiment we excluded the first two trials after each eye position change to allow sufficient time for saccade related activity to dissipate, and found no reduction in eye position information.

As noted above, the experiments were not designed to directly look for eye position information, and thus these results may be subject to possible confounds and should be viewed as exploratory. In particular, the constant presence of the three placeholder boxes and an inability to ensure a completely dark room could have created unbalanced visual stimulation at the two fixation positions that could have been mistaken for true eye position information. Although we eccentricity-restricted our early visual ROIs to cover only the immediately neighboring stimulus positions, it is still possible the responses could have been influenced by surrounding stimulation, and future work under more controlled conditions is needed to truly assess the nature of eye position information in these regions. However, the presence of eye position representations would be consistent with previous findings from human fMRI, primate neurophysiology, and computational models (see main text), and would suggest that this information may be organized on a large enough scale to be manifested robustly throughout visual cortex, perhaps compensating for the lack of similarly large-scale explicit spatiotopic representations.

Supplemental References

Chang C-C, Lin C-J. 2001. LIBSVM : A library for support vector machines. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm.

Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P. 2001. Distributed and overlapping representations of faces and objects in ventral temporal cortex. Science. 293: 2425-430.

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