Supplemental Data

Decoding the Representation of Multiple Simultaneous Objects in Human Occipitotemporal Cortex

Sean P. MacEvoy and Russell A. Epstein

Department of Psychology and Center for Cognitive Neuroscience, University of Pennsylvania, Philadelphia PA 19104 USA

Supplemental Material

Supplemental Experimental Procedures

Subjects

Twelve subjects (5 female, aged 19 to 30 years) with normal or corrected-to-normal vision were recruited from the University of Pennsylvania community and gave written informed consent in compliance with procedures approved by the University of Pennsylvania Institutional Review Board. Subjects were paid for their participation. An additional subject was also scanned, but was excluded from the study prior to data analysis because of excessive head motion.

MRI acquisition

Scans were performed at the Center for Functional Neuroimaging at the University of Pennsylvania on a 3T Siemens Trio scanner equipped with a Siemens body coil and an eight-channel multiple-array Nova Medical head coil. Structural T1* weighted images for anatomical localization were acquired using a 3D MPRAGE pulse sequences (TR = 1620 ms, TE = 3 ms, TI = 950 ms, voxel size = 0.9766 x 0.9766 x 1mm, matrix size = 192 x 256 x 160). T2* weighted scans sensitive to blood oxygenation level-dependent (BOLD) contrasts were acquired using a gradient-echo echo-planar pulse sequence (TR = 3000ms, TE = 30ms, voxel size = 3x3x3mm, matrix size = 64 x 64 x 45). Visual stimuli were rear projected onto a mylar screen at the head end of the scanner bore with an Epson 8100 3-LCD projector equipped with a Buhl long-throw lens and viewed through a mirror affixed to the head coil. The entire projected field subtended 22.9 x 17.4° and was viewed at 1024 x 768 pixel resolution.

The scanning session for each subject consisted of six experimental scans and two functional localizer scans. Experimental scans were 6 minutes 30 seconds in length,

and were divided into 20 18-second stimulus blocks, with additional 15-second fixation periods at the beginning of each scan and after every fourth stimulus block. Localizer scans were 6 minutes 15 seconds long and were divided into blocks during which subjects viewed color photographs of scenes, faces, common objects, and scrambled objects presented at a rate of 1.33 pictures per second as described previously [S1].

MRI Analysis

Functional images were corrected for differences in slice timing by resampling slices in time to match the first slice of each volume, realigned with respect to the first image of the scan, spatially normalized to the Montreal Neurological Institute (MNI) template. Data for localizer scans were spatially smoothed with an 8 mm FWHM Gaussian filter; all other data were left unsmoothed. Data were analyzed using the general linear model as implemented in VoxBo (www.voxbo.org) including an empirically-derived 1/f noise model, filters that removed high and low temporal frequencies, regressors to account for global signal variations, and nuisance regressors to account for between-scan differences.

Regions of Interest

The lateral occipital complex (LOC) was defined by stronger responses ($t > 3.5$) to objects than to scrambled objects during the functional localizer scans. LOC was trimmed to exclude any voxels that had 1) significantly greater responses to scenes than objects or 2) had significantly greater responses to faces than to objects. This was the same procedure used to define object selective cortex by Reddy and Kanwisher [S2], to whose results we wished to compare our own.

The parahippocampal place area (PPA) was defined as voxels in the posterior parahippocampal/collateral sulcus region that responded more strongly ($t > 3.5$) to scenes than to common objects. To focus on voxels characterized by selectivity for scenes, the PPA was further refined to exclude any voxels with significantly higher responses to intact objects than to scrambled objects. (Voxels that had significantly higher responses to scenes than to objects *and* significantly higher responses to objects than to scrambled objects were not included in any ROI.) The fusiform face area (FFA) was defined by voxels responding more strongly to faces than to objects. We did not exclude voxels from the FFA that also met the criteria for inclusion in LOC since this often produced exceedingly small FFA voxel counts.

To provide a baseline from which to judge multi-voxel pattern classification accuracy, we also defined a non-brain ROI for each subject by selecting 100 contiguous voxels within a supraorbital portion of the skull.

Supplemental Results

Additional classification analyses

We performed several additional analyses to validate our classification results. First, to ensure that differences in classification accuracy between ROIs did not result

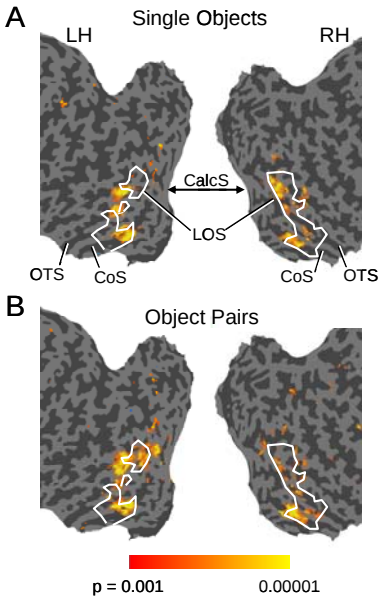

Figure S1. Group analysis of searchlight classification accuracy for single objects (A) and object pairs (B). Colored voxels are those whose associated searchlight clusters had classification accuracies above chance ($p < 0.001$, uncorrected), determined from random-effects analysis across searchlight volumes from all subjects. White contour indicates boundaries of LOC defined from a random-effects group analysis of functional localizer data, thresholded at $p < 0.01$, uncorrected. Maps for single objects and pairs are highly overlapping, indicating that the same clusters encoded information about both single objects and object pairs. CoS: collateral sulcus; OTS: occipitotemporal sulcus; LOS: lateral occipital sulcus CalcS: apex of calcarine sulcus.

from differences in ROI size, we repeated pattern classification for randomly-drawn voxel subsets matched in size to the smallest ROI for each subject, usually the FFA. The median voxel count across subjects for the smallest ROI was 53. Average classification accuracy was computed over 200 draws (light-shaded bars in Figure 1A). Accuracy remained significantly above chance in LOC ($t(11) = 5.54$, $p = 0.00006$) and the PPA ($t(11) = 2.80$, $p = 0.016$), indicating that the superior classification accuracy for single objects in LOC and the PPA relative to the FFA was not simply a result of greater ROI size. Similarly, pair classification accuracy in LOC ($t(11) = 5.25$, $p = 0.0002$) for voxel subsets was significantly above chance.

Second, we adopted a "searchlight" analysis approach to assess local classification accuracy for both single objects and object pairs throughout the whole brain [S3]. This analysis was particularly important given that LOC was defined based on responses to single objects in the localizer, leaving open the possibility that our ROI-based analysis may have overlooked voxels outside of LOC that preferentially carried information about object pairs. Supplemental Figure 1 shows the results of group random-effects analysis of local classification accuracy for both single objects and pairs. The maps show considerable overlap between regions exhibiting above-chance accuracy for pair and single object classification. Importantly, we observe no regions of high classification performance for pairs that are outside the boundaries of LOC *and* are absent from the map for singles. These results, together with our ROI analyses, indicate that LOC is the primary region of the brain involved in representing object pairs.

Position-specificity of classification

Our main interest in conducting these experiments was to quantify the relationship between responses evoked by object pairs and by their constituent objects. A potential

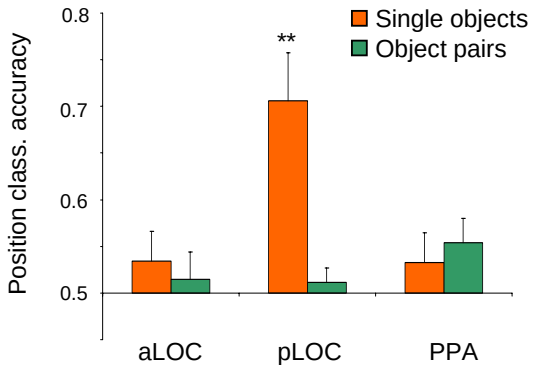

Figure S2. Spatial configuration classification accuracy. Patterns evoked by each single object and object pair were split by stimulus position (singles) or spatial configuration (pairs). Data are accuracy in classifying position/configuration in one half of the data, based on patterns for the same object or pair in the opposite half of the data. ROIs are defined in Supplemental Results. Patterns in aLOC did not discriminate between different positions for single objects or different configurations for pairs. Although pLOC patterns distinguished between single object positions, they did not discriminate between different spatial configurations of a pair, suggesting functional segregation of category and position information. Position/Configuration classification accuracy in the PPA was marginally above chance for object pairs, but not for single objects. Error bars are s.e.m.

complicating factor in any such analysis was the fact that object pairs necessarily contained objects at two positions in the visual field. If category-selective neurons were also position-selective, the relationship between the pattern evoked by a given object pair and the patterns evoked by its constituent objects might vary depending on whether the single objects were presented at the same locations as they appeared in the pair or at opposite locations. It was therefore important for us to understand the degree to which multi-voxel patterns varied with object position.

To assess the position-specificity of activity patterns in LOC, we measured our ability to classify the positions of single objects based on multi-voxel patterns. In other words, we assessed whether (for example) brush-on-top/nothing-on-bottom could be distinguished from nothing-on-top/brush-on-bottom. This was done by examining the patterns evoked by each kind of object (brushes, cars, chairs, and shoes) at each screen position (top, bottom) and assessing the extent to which the patterns elicited by an object of a given category in the two halves of the data were more similar when stimuli were presented at the same position than when they were presented in opposite positions. For these analyses, we divided LOC in each subject into anterior (aLOC) and posterior (pLOC) regions which corresponded to regions that have been previously termed posterior fusiform and LO, respectively [S4]. A recent study by Schwarzlose et al. [S5] showed that activity patterns evoked by single objects in pLOC carried significant information about object position, while activity patterns in aLOC were considerably less dependent upon objects' screen positions. In keeping with these results, we were able to classify single object positions in pLOC ($t(11) = 3.95$, $p = 0.002$) but not in aLOC ($t(11) = 1.04$, $p = 0.32$), nor in the PPA ($t(11) = 1.04$, $p = 0.23$), as shown in Supplemental Figure 2 (orange bars).

These data suggest a convergence between information about object category and position in pLOC. However, a more nuanced picture emerged when we examined the spatial-specificity of patterns evoked by pairs. Classification accuracy was not significantly above chance in either aLOC ($t(11) = 0.58$, $p= 0.62$) or pLOC ($t(11) = 0.75$, $p = 0.47$) when we attempted to distinguish between the two spatial configurations of each pair, for example between brush-on-top/chair-on-bottom and

chair-on-top/brush-on bottom (Supplemental Figure 2, green bars). These data indicate that, to the extent that neurons in pLOC encode information about object position, this information is carried by a population of neurons different from that which carries information about category. That is, while the responses of some neurons may discriminate between a stimulus at the top or bottom stimulus position, and other neurons may discriminate between a chair and shoe, neurons tuned jointly for position and object category do not appear to compose a significant fraction of the pLOC population.

In the PPA, we did observe a trend towards above-chance classification accuracy for spatial configurations of object pairs ($t(11) = 2.02$, $p = 0.065$). Although falling short of significance, this result is particularly interesting insofar as it suggests the possibility that some neurons in the PPA may be tuned to the spatial relationship between objects.

We also used the results of the searchlight classification analysis to identify any regions in the posterior half of the brain that reliably differentiated between spatial configurations of object pairs. Searchlight positions matching this criterion were relatively rare: while subjects had an average of 169 searchlights that could classify pair identity above an arbitrary threshold of 75%, they only had an average of 62 searchlights that could classify pair configuration at the same threshold. By comparison, an average of 2,356 searchlights in each subject could correctly identify the position of a single object at a rate greater than 75%. Group random-effects analysis showed no notable regions of high pair configuration accuracy outside LOC.

Relationship between pair and single object responses in other ROIs

We analyzed the relationship between responses to single objects and pairs for each voxel in the PPA, the FFA, and the non-brain ROI, using the same regression analysis detailed for LOC in the main text. In contrast to the findings in LOC, we saw no positive relationship between searchlight classification accuracy and R^2 in the FFA (mean correlation = 0.06, $t(11) = 0.77$, $p = 0.45$) or non-brain region (mean correlation = 0.095, $t(11) = 1.23$, $p = 0.24$). Taking note of the fact that some non-brain searchlights had, by chance, classification accuracies above 50%, the absence of any significant positive correlation demonstrates that higher R^2 values are not a trivial accompaniment to high classification scores, providing a control against which to judge the positive correlations we observed in LOC.

Interestingly, we did observe a significant positive correlation between classification rank and R^2 in the PPA (mean = 0.21, $t(11) = 3.25$, $p = 0.0077$). This result may seem surprising given the fact that whole-PPA patterns did not discriminate reliably among object pairs (Figure 1B). However, patterns derived from voxels in the top 30 searchlight positions in the PPA did exhibit pair classification accuracies significantly above chance ($t(11) = 5.17$, $p = 0.0002$). As in LOC, the slope of linear regressions between pair and summed single-object responses for the top 30 searchlight clusters in the PPA was 0.51, with upper and lower 95% confidence limits of 0.73 and 0.27, respectively. As in LOC, the average regression intercept term in the PPA did not differ significantly from zero ($t(11) = 1.59$, $p = 0.14$). By analogy to LOC, these results suggest that at least a subset of PPA voxels do contain information about object pairs,

and that responses of these voxels to pairs are linearly related to their responses to constituent objects.

With this in mind, we tested whether patterns evoked by object pairs in the PPA could be classified based on synthetic patterns created by averaging the patterns evoked by the constituent single objects. As in LOC, the classifier was able to correctly identify patterns evoked by object pairs based on these synthetic patterns at rates well above chance (PPA, $t(9) = 4.85$, $p = 0.0007$) when PPA data were restricted to the top 30 searchlight clusters.

We also performed regression analysis on retinotopic cortex, which was defined from functional localizer scans as regions with significantly higher responses to scrambled objects than intact objects. Classification accuracy in retinotopic cortex was not above chance for either single objects ($t(11) = 1.31$, $p = 0.22$) or pairs ($t(11) = 0.005$, $p = 0.99$). This indicates that the responses of voxels in early visual cortex did not vary reliably across single objects or pairs. Consistent with this, when we calculated regressions between voxel responses to single objects and object pairs there was no trend toward higher R^2 as a function of searchlight accuracy; across subjects, the mean correlation between R^2 and pair accuracy rank did not differ significantly from zero ($t(11) = 0.8693$, $p(11) = 0.40$).

Validation of Regression Analysis

Our regression analysis was based on the responses of each voxel to stimuli averaged across all six scans in our experiment. We used data from these same six scans to measure searchlight classification accuracy. In standard pattern classification studies, using the same data for both voxel selection and pattern classification ("peeking") can produce artificially elevated performance. It was therefore necessary to ensure that the positive relationship we observed between accuracy and regression R^2 (and slope) was not a trivial outcome of voxel selection.

Such a confound was unlikely in our analysis because voxels were selected entirely on the basis of searchlight classification performance for pairs without any consideration whatsoever of their classification performance for single objects. Because single and pair responses constitute different data sets collected at different time points, the fact that voxels within a given searchlight carry information about pairs should not bias us towards finding a linear (or any other) relationship between their responses to object pairs and responses to single objects if none really exists. In contrast, if such a relationship did exist, then we would expect that it should be most evident among voxel clusters with the highest signal-to-noise ratios (see "Voxel Selection" in Supplemental Discussion).

Nevertheless, we undertook several additional analyses to ensure that the positive correlations we observed between pair classification performance and both R^2 and slope in LOC were not artifacts of voxel selection. First, we recomputed both searchlight classification accuracies and regression analyses using data from separate scans. Classification accuracy was measured for data drawn from four of the six total scans, and regression analysis was performed on the remaining two. This process was repeated for each possible draw of 4 scans for each subject (e.g., scans 1,2,3,4 versus 5 and 6, then scans 1, 3, 4, 5 versus 2, 6, etc.). Data were accumulated across all 15

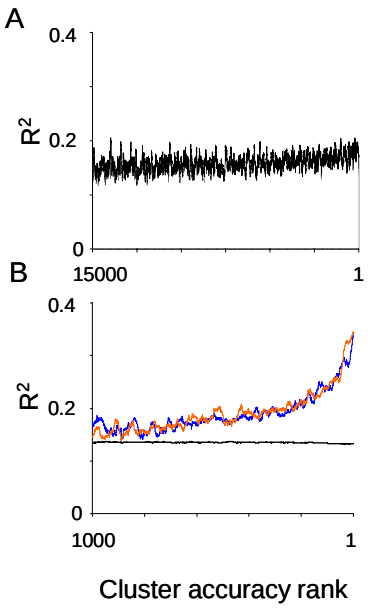

Figure S3. Validation of regression analysis. A) R^2 as a function of searchlight cluster pair classification rank when R^2 and classification accuracy are computed from separate data sets, averaged across subjects. Best-performing clusters are those with low rank numbers, at the right end of the x-axis. As in Figure 3 of the main text, R^2 showed a significant positive correlation to cluster accuracy. Ranking extends to higher numbers than in Figure 2 in the main text due to the 15 possible ways in which non-overlapping data sets could be drawn from 6 scans (see Supplemental Results). B) Comparison of relationship between R^2 and pair classification accuracy rank when R^2 is computed from original data (red) and after permuting single object condition labels (black). Permuting single-object labels completely abolishes the upward trend in R^2 as classification accuracy improves, indicating that this trend is not a trivial outcome of high classification accuracy. Also shown is R^2 as a function of single-object classification rank (blue), which does not differ meaningfully from R^2 based on pair rank.

possible draws for each subject, but searchlight classification accuracies from each draw were associated exclusively with the results of regression analysis from the corresponding two remaining scans.

Regression R^2 values for all searchlight positions across all draws were placed in a single pool for each subject and then ranked according to their associated classification accuracy, and the resulting sorted values were averaged across subjects (Supplemental Figure 3A). If the positive correlation we observed between R^2 and searchlight accuracy shown in Figure 2 were a result of peeking, we should see no such correlation when classification scores and regression analysis were performed on non-overlapping data sets. Instead, we observed a small but highly significant ($p < 10^{-6}$) positive correlation, qualitatively similar to the data shown in Figure 2. Further, 10 of 12 subjects still had correlation coefficients significantly greater than 0 ($p < 0.05$).

Though this result indicates that the positive correlation between searchlight accuracy and R^2 is not *solely* attributable to voxel selection, it does not eliminate the possibility that voxel selection made *some* contribution to the R^2 trend in Figure 2. To assess what bias, if any, was contributed by voxel selection, we recomputed our original regression analysis (which included data from all 6 scans for each subject) after randomly permuting the condition labels for single-object responses independently for each LOC voxel. It is critical to note that this permutation step in no way altered each searchlight's classification score, which was based solely on data for *pairs*. If the trend observed in Figure 2 were even in part a result of voxel selection, we should see a positive correlation between searchlight rank and R^2 even after label permutations. Instead, we find that R^2 from permuted regressions remains constant as a function of classification accuracy, showing no bias towards higher values as classification

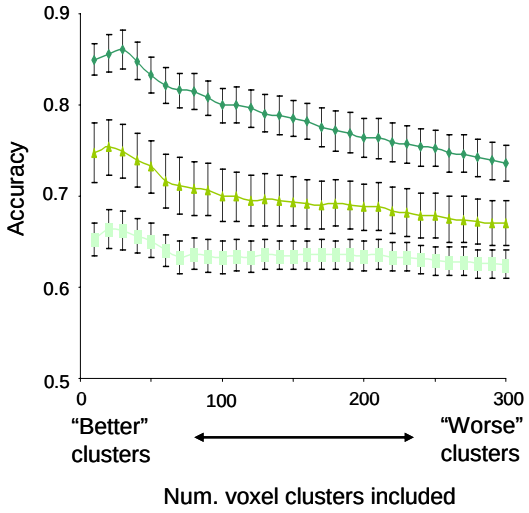

Figure S4. LOC classification accuracy for pairs and synthetic pairs as a function of pattern voxel count, averaged across subjects. Data are accuracy in classifying pair patterns based actual pair patterns (dark green), patterns derived from the means of single object patterns (light green), and patterns derived from a MAX-function combination of single object patterns (cyan). X-axis denotes the size of the pattern, expressed in terms of the number of searchlight positions included, binned in 10-searchlight increments. Searchlights were added in descending order of classification accuracy. Error bars are s.e.m.

improves. This is shown by the black line in Supplemental Figure 3B, which represents subject-averaged data after 1000 permutation cycles for each subject. Also plotted in Supplemental Figure 3B is R^2 as a function of pair classification accuracy rank (repeated from Figure 2B), as well as R^2 as a function of single object classification accuracy rank. These two functions are almost identical, further demonstrating that voxel selection did not bias our assessment of a linear relationship between single objects and object pairs.

We also wished to confirm that our regression analysis truly reflected pair responses that were the simple averages of the responses to their constituent objects, rather than weighted averages. To do so, we began by performing simulations that allowed us to examine the expected outcome of the linear regression analysis for a set of artificial voxels. In the first simulation, the responses of each voxel to pairs were the weighted averages (rather than simple average) of its single object responses, with the response to the more-preferred object in a pair weighted more strongly (i.e. weights for each object were proportional to their single-object responses). Linear regressions between these pair responses and the sum of their constituent object responses produced a unique pattern of predictive accuracy. Specifically, we found that residual error between pair response predicted by linear fits and actual pair responses was distributed unevenly across pairs, with the largest residuals attaching to pairs composed of objects with the largest differences in their single-object responses. This contrasted with the results of a a second simulation, in which pair responses were constructed from the simple averages of component responses (with noise added). In this case, regressions between pairs and the sum of their constituent objects had residuals that were uniformly distributed across pairs.

With this in mind, we next examined the distribution of residuals from linear regressions performed on real fMRI data from voxels within the 30 highest accuracy searchlights within each subject. For each voxel, we computed the squared error between the actual response to each object pair and the response predicted for that pair by linear regression. We then ranked these error terms by the magnitudes of the response differences between each pair's constituent objects. Based on the results of our simulations, we expected to see the largest residual error among pairs that ranked

high by this measure. However, a one-way repeated measures ANOVA showed no significant effect of this ranking on the magnitude of residual error for paired objects across subjects ($F(5,11) = 1.17$, $p = 0.34$). Thus we find no evidence that responses to pairs were weighted, as opposed to simple, averages of their constituent object responses. However, it is unclear whether this finding would generalize to a situation in which one of the two objects of a pair were "preferred" in terms of cortical specialization (e.g. face or house). In this case, we might expect the mean rule to break down, perhaps reflecting an absence of multistimulus normalization across separate domain-specific systems.

Supplemental Discussion

Voxel Selection

One challenge in assessing the ability of any model to characterize the relationship between single and pair responses is that of voxel selection. For at least two reasons it would be unreasonable to expect the responses of every voxel to vary reliably enough across the stimuli to allow us to discern a relationship between responses to different stimuli. The first reason is the extreme sparseness of any experimental stimulus set. Given the vast number of objects which the visual system is forced to encode, the odds are high that the four object categories we chose lie outside the tuning envelope of many neurons. Response variability among voxels corresponding to populations of such neurons will necessarily be dominated by noise, and will therefore defy any model. (It is worth noting that this challenge is not unique to fMRI. In single-unit studies of macaque IT, neurons are often excluded from further study if they do not respond to at least one element in an experimental stimulus set [S6, S7].) A second reason is partial-volume effects. Each voxel represents neural activity among a large and potentially heterogeneous population of neurons. Even though individual elements of these populations may differentiate reliably between stimuli, their aggregate activity may show considerably less selectivity. As such, response variability among these voxels will also tend to be dominated by noise.

In light of these factors, it is not surprising that, on average, 54% of LOC searchlight clusters and 69% of PPA clusters had pair classification accuracies below the 95th percentile of searchlight accuracies in the non-brain ROI. In other words, these LOC and PPA clusters had accuracies indistinguishable from chance, indicating that the responses of voxels within them were not strongly related to pair identity. Therefore, their pair responses could not be expected have any relationship, linear or otherwise, to responses evoked by single objects, and consideration of these voxels will tend to artificially reduce the predictive ability of any model. Given these factors, we used searchlight classification performance as an independent estimate of the signal-to-noise ratio of voxels within each searchlight, as explained in the Results.

Classification performance in ROIs outside LOC

We found that voxel patterns from regions outside the LOC often showed poor classification performance for single objects and object pairs. While patterns in PPA

differentiated between single objects, they did not differentiate between pair patterns at a rate above chance, and FFA patterns did not differentiate among either single objects or pairs.

These results might at first appear to contradict previous work, most notably by Haxby et al. [S8], showing that information about object identity could be gleaned from patterns in the FFA and PPA. Similar results were reported by Reddy and Kanwisher [S2]. However, both of these studies reported average classification accuracy across a stimulus set that included stimuli preferred by FFA and PPA as well; thus, above-chance accuracy for non-preferred objects may have been driven to a large extent by correct classifications between preferred and non-preferred objects (e.g., house versus brush), rather than among non-preferred objects (e.g., brush versus chair). Although Reddy and Kanwisher confirmed that patterns in both FFA and PPA did discriminate between shoes and cars, these were the only non-preferred objects that were used, making the generality of the result unclear. In contrast, Spiridon and Kanwisher [S9], using a stimulus set similar to ours (including objects in four categories not preferred by either FFA or PPA), found that patterns in the FFA and PPA did not reliably discriminate among four categories of single objects.

Given the fact that PPA patterns did reliably discriminate among single objects, it might seem somewhat surprising that they did not discriminate among object pairs (at least when whole-PPA patterns were examined; as noted above, subsets of voxels within the PPA were more discriminative). After all, object pairs might be taken to form a minimal scene, a stimulus class which robustly activates the PPA. It is worth noting, however, that there are at least three important differences between the real-world scenes that robustly activate the PPA and the object pairs shown here. First, real-world scenes contain fixed background elements such as walls and ground planes, which are believed to be critical for eliciting a strong PPA response [S10]. Second, the objects in the current study were not presented in their typical spatial relationships. Finally, the stimuli in the current experiment were equated for angular subtense even though they possessed very different real-world sizes. The close apposition of a brush and car of similar angular subtense, in an unnatural spatial relationship to each other, and in the absence of surrounding background information is unlikely to convey any meaningful sense of a scene.

Supplemental References

S1. Epstein, R.A., and Higgins, J.S. (2006). Differential parahippocampal and retrosplenial involvement in three types of visual scene recognition. Cereb. Cortex.

S2. Reddy, L., and Kanwisher, N. (2007). Category selectivity in the ventral visual pathway confers robustness to clutter and diverted attention. Curr Biol *17*, 2067-2072.

S3. Kriegeskorte, N., Goebel, R., and Bandettini, P. (2006). Information-based functional brain mapping. Proc Natl Acad Sci U S A *103*, 3863-3868.

S4. Malach, R., Levy, I., and Hasson, U. (2002). The topography of high-order human object areas. Trends in Cognitive Sciences *6*, 176-184.

S5. Schwarzlose, R.F., Swisher, J.D., Dang, S., and Kanwisher, N. (2008). The distribution of category and location information across object-selective regions in human visual cortex. Proc Natl Acad Sci U S A *105*, 4447-4452.

S6. Zoccolan, D., Cox, D.D., and DiCarlo, J.J. (2005). Multiple object response normalization in monkey inferotemporal cortex. Journal of Neuroscience *25*, 8150-8164.

S7. Sheinberg, D.L., and Logothetis, N.K. (2001). Noticing familiar objects in real world scenes: the role of temporal cortical neurons in natural vision. Journal of Neuroscience *21*, 1340-1350.

S8. Haxby, J.V., Gobbini, M.I., Furey, M.L., Ishai, A., Schouten, J.L., and Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. Science *293*, 2425-2430.

S9. Spiridon, M., and Kanwisher, N. (2002). How distributed is visual category information in human occipito-temporal cortex? An fMRI study. Neuron *35*, 1157-1165.

S10. Epstein, R., and Kanwisher, N. (1998). A cortical representation of the local visual environment. Nature *392*, 598-601.