Processing multiple visual objects is limited by overlap in neural channels

Michael A. Cohen[a,1], Talia Konkle[a], Juliana Y. Rhee[b], Ken Nakayama[a], and George A. Alvarez[a]

Departments of [a]Psychology and [b]Molecular and Cellular Biology, Harvard University, Cambridge, MA 02140

Edited by Randolph Blake, Vanderbilt University, Nashville, TN, and approved May 2, 2014 (received for review September 23, 2013)

High-level visual categories (e.g., faces, bodies, scenes, and objects) have separable neural representations across the visual cortex. Here, we show that this division of neural resources affects the ability to simultaneously process multiple items. In a behavioral task, we found that performance was superior when items were drawn from different categories (e.g., two faces/two scenes) compared to when items were drawn from one category (e.g., four faces). The magnitude of this mixed-category benefit depended on which stimulus categories were paired together (e.g., faces and scenes showed a greater behavioral benefit than objects and scenes). Using functional neuroimaging (i.e., functional MRI), we showed that the size of the mixed-category benefit was predicted by the amount of separation between neural response patterns, particularly within occipitotemporal cortex. These results suggest that the ability to process multiple items at once is limited by the extent to which those items are represented by separate neural populations.

working memory | capacity limitations | representational similarity | competition | visual cognition

An influential idea in neuroscience is that there is an intrinsic relationship between cognitive capacity and neural organization. For example, seminal cognitive models claim there are distinct resources devoted to perceiving and remembering auditory and visual information (1, 2). This cognitive distinction is reflected in the separate cortical regions devoted to processing sensory information from each modality (3). Similarly, within the domain of vision, when items are placed near each other, they interfere more than when they are spaced farther apart (4, 5). These behavioral effects have been linked to receptive fields and the retinotopic organization of early visual areas, in which items that are farther apart activate more separable neural populations (6–8). Thus, there are multiple cognitive domains in which it has been proposed that capacity limitations in behavior are intrinsically driven by competition for representation at the neural level (4, 7–10).

However, in the realm of high-level vision, evidence linking neural organization to behavioral capacities is sparse, although neural findings suggest there may be opportunities for such a link. For example, results from functional MRI (fMRI) and single-unit recording have found distinct clusters of neurons that selectively respond to categories such as faces, bodies, scenes, and objects (11, 12). These categories also elicit distinctive activation patterns across the ventral stream as measured with fMRI (13, 14). Together, these results raise the interesting possibility that there are partially separate cognitive resources available for processing different object categories.

In contrast, many prominent theories of visual cognition do not consider the possibility that different categories are processed by different representational mechanisms. For example, most models of attention and working memory assume or imply that these processes are limited by content-independent mechanisms such as the number of items that can be represented (15–18), the amount of information that can be processed (19–21), or the degree of spatial interference between items (4, 22–24). Similarly, classical accounts of object recognition are intended to apply equally to all object categories (25, 26). These approaches implicitly assume that visual cognition is limited by mechanisms that are not dependent on any major distinctions between objects.

Here, we examined (*i*) how high-level visual categories (faces, bodies, scenes, and objects) compete for representational resources in a change-detection task, and (*ii*) whether this competition is related to the separation of neural patterns across the cortex. To estimate the degree of competition between different categories, participants performed a task that required encoding multiple items at once from the same category (e.g., four faces) or different categories (e.g., two faces and two scenes). Any benefit in behavioral performance for mixed-category conditions relative to same-category conditions would suggest that different object categories draw on partially separable representational resources. To relate these behavioral measures to neural organization, we used fMRI to measure the neural responses of these categories individually and quantified the extent to which these categories activate different cortical regions.

Overall, we found evidence for separate representational resources for different object categories: performance with mixed-category displays was systematically better than performance with same-category displays. Critically, we also observed that the size of this mixed-category benefit was correlated with the degree to which items elicited distinct neural patterns, particularly within occipitotemporal cortex. These results support the view that a key limitation to simultaneously processing multiple high-level items is the extent to which those items are represented by nonoverlapping neural channels within occipitotemporal cortex.

Results

Behavioral Paradigm and Results. To measure how items from different categories compete for representation, participants performed a task that required encoding multiple items at once. The stimulus set included images of faces, bodies, scenes, and objects matched for luminance and contrast (Fig. S1 shows the full stimulus set). On each trial, four different items were presented

Significance

Human cognition is inherently limited: only a finite amount of visual information can be processed at a given instant. What determines those limits? Here, we show that more objects can be processed when they are from different stimulus categories than when they are from the same category. This behavioral benefit maps directly onto the functional organization of the ventral visual pathway. These results suggest that our ability to process multiple items at once is limited by the extent to which those items compete with one another for neural representation. Broadly, these results provide strong evidence that the capacities and limitations of human behavior can be inferred from our functional neural architecture.

Author contributions: M.A.C., T.K., K.N., and G.A.A. designed research; M.A.C. and J.Y.R. performed research; M.A.C., T.K., J.Y.R., and G.A.A. analyzed data; and M.A.C., T.K., J.Y.R., K.N., and G.A.A. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. E-mail: michaelthecohen@gmail.com.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1317860111/-/DCSupplemental.

for 800 ms with one item in each visual quadrant. Following a blank display (1,000 ms), the items reappeared with one item cued by a red frame, and participants reported if that item had changed (Fig. 1*A*). Changes occurred on half the trials and could occur only at the cued location.

The critical manipulation was that half the trials were same-category displays (e.g., four faces or four scenes), whereas the other half of trials were mixed-category displays (e.g., two faces and two scenes). Whenever an item changed, it changed to another item from the same category (e.g., a face could change into another face, but not a scene). Each participant was assigned two categories, such as faces and scenes, or bodies and objects, to obtain within-subject measures for the same-category and mixed-category conditions. Across six groups of participants, all six pairwise combinations of category pairings were tested (*SI Materials and Methods*).

If items from one category are easier to process, participants might pay more attention to the easier category in mixed-category displays. To address this concern, we averaged performance across the tested categories (e.g., for the face–scene pair, we averaged over whether a face or scene was tested). Thus, any differences in overall performance for the mixed-category and same-category conditions cannot be explained by attentional bias toward one particular category. We also took several steps to ensure that performance was approximately matched on the same-category conditions for all categories. First, we carefully selected our stimulus set based on a series of pilot studies (*SI Materials and Methods*). Second, before testing each participant, we used an adaptive calibration procedure to match performance on the same-category conditions, by adjusting the transparency of the items (*Materials and Methods*). Finally, we adopted a conservative exclusion criterion: any participants whose performance on the same-category displays (e.g., four faces compared with four scenes) differed by more than 10% were not included in the main analysis (*SI Materials and Methods*). This exclusion procedure ensured that there were no differences in difficulty between the same-category conditions for any pair of categories ($P > 0.16$ for all pairings; Fig. S2). Although these exclusion criteria were chosen to isolate competition between items, our overall behavioral pattern and its relationship to neural activation is similar with and without behavioral subjects excluded (Fig. S3).

Overall, we observed a mixed-category benefit: performance on mixed-category displays was superior to performance on same-category displays ($F_{1,9} = 19.85$; $P < 0.01$; Fig. 1*B*). This suggests that different categories draw on separate resources, improving processing of mixed-category displays. Moreover, although there was a general benefit for mixing any two categories, a closer examination suggested that the effect size depended on which categories were paired together (regression model comparison, $F_{5,54} = 2.29$; $P = 0.059$; *SI Materials and Methods*). The mixed-category benefit for each pairing, in order from largest to smallest, was: bodies–scenes, 5.6%, SEM = 1.5%; faces–scenes, 5.2%, SEM = 1.3%; bodies–faces, 3.3%, SEM = 1.1%; bodies–objects, 3.3%, SEM = 1.2%; faces–objects, 2.4%, SEM = 1.9%; scenes–objects, −0.8%, SEM = 1.9% (Fig. 1*B*). The variation in the size of the mixed-category benefit suggests that categories do not compete equally for representation and that there are graded benefits depending on the particular combination of categories.

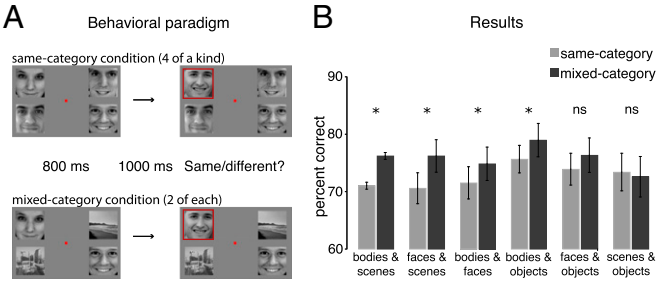

Fig. 1. (*A*) Behavioral paradigm. Participants were shown two successively presented displays with four items in each display (*Materials and Methods*). On the second display, one item was cued (red frame) and participants reported if that item had changed. In the same-category condition, items came from the same stimulus category (e.g., four faces or four scenes). In the mixed-category conditions, items came from two different categories (e.g., two faces and two scenes). (*B*) Behavioral experiment results. Same-category (light gray) and mixed-category (dark gray) performance is plotted in terms of percent correct for all possible category pairings. Error bars reflect within-subject SEM (48) (*$P < 0.05$).

What is the source of the variability in the size of the mixed-category benefit? We hypothesize that visual object information is represented by a set of broadly tuned neural channels in the visual system, and that each stimulus category activates a subset of those channels (7–10, 27, 28). Under this view, items compete for representation to the extent that they activate overlapping channels. The differences in the size of the mixed-category benefit may thus result from the extent to which the channels representing different categories are separated.

Importantly, although this representational-competition framework explains why varying degrees of mixed-category benefits occur, it cannot make a priori predictions about why particular categories (e.g., faces and scenes) interfere with one another less than other categories (e.g., objects and scenes). Thus, we sought to (*i*) directly measure the neural responses to each stimulus category and (*ii*) use these neural responses to predict the size of the mixed-category benefit between categories. Furthermore, by assuming a model of representational competition in the brain, we can leverage the graded pattern of behavioral mixed-category benefits to gain insight into the plausible sites of competition at the neural level.

Measuring Neural Separation Among Category Responses. Six new participants who did not participate in the behavioral experiment were recruited for the fMRI experiment. Participants viewed stimuli presented in a blocked design, with each block composed of images from a single category presented in a single quadrant (*Materials and Methods*). The same image set was used in the behavioral and fMRI experiments. Neural response patterns were measured separately for each category in each quadrant of the visual field. There are two key properties of this fMRI design. First, any successful brain/behavior relationship requires that behavioral interference between two categories can be predicted from the neural responses to those categories measured in isolation and across separate locations. Second, by using two groups of participants, one for behavioral measurements and another for neural measurements, any brain/behavior relationship cannot rely on individual idiosyncrasies in object processing and instead reflects a more general property of object representation in behavior and neural coding.

To probe how different neural regions predict behavioral interference, we divided the set of visually active voxels into four sectors in each participant: occipitotemporal, occipitoparietal, early visual (V1–V3), and prefrontal cortex (PFC). These sectors were defined from independent functional localizers (*SI Materials and Methods* and Fig. S4). Performing the analysis in these sectors allowed us to examine the neural response patterns across the major divisions of the visual system: early retinotopic cortex, the what pathway, the where/how pathway (29, 30), as well as in a frontal lobe region associated with working memory (31).

We defined the neural separation between any two categories as the extent to which the stimuli activate different voxels. To quantify this, we first identified the most active voxels in each sector for each of the categories, at a particular threshold [e.g., the 10% most active voxels for objects (objects > rest) and the top 10% most active voxels for scenes (scenes > rest); Fig. 2]. Next, we calculated the percentage of those voxels that were shared by each category pairing (i.e., percent overlap). This

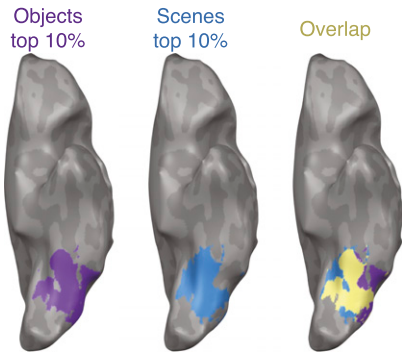

Fig. 2. Visualization of the neural separation procedure. Activation and overlap among the 10% most active voxels for objects and scenes in the occipitotemporal sector is shown in a representative subject. The 10% most active voxels for each category are colored as follows: objects purple, scenes blue. The overlap among these active voxels are shown in yellow. For visualization purposes, this figure shows the most active voxels and overlapping voxels combined across all locations; for the main analysis, overlap was computed separately for each pair of locations.

percent overlap measure was then converted to a neural separation measure as 1 minus percent overlap. The amount of neural separation for every category pairing was then calculated at every activation threshold (from 1% to 99%; *Materials and Methods*). Varying the percent of active voxels under consideration allows us to probe whether a sparse or more distributed pool of representational channels best predicts the behavioral effect. This was done in every sector of each fMRI participant. In addition, we also used an area under the curve (AUC) analysis, which integrates over all possible activation thresholds, to compute an aggregate neural separation measure for each category pairing. Finally, we performed a standard representational similarity analysis in which the neural patterns of each category pairing were compared by using a pattern dissimilarity measure [1 minus the Pearson correlation between two response patterns across the entire sector (14); *SI Materials and Methods*].

Neural Separation Predicts the Mixed-Category Benefit. To assess the degree to which neural separation predicted the mixed-category benefit, we correlated the amount of neural separation for every category pairing in each individual fMRI participant with the size of the mixed-category benefits from the behavioral experiments (i.e., a random effects analysis of the distribution of brain/behavior correlations; *SI Materials and Methods*). An illustration of this analysis procedure using data obtained from occipitotemporal cortex is shown in Fig. 3. We chose this analysis because it allows for a stronger inference about the generality of our results relative to a fixed effects analysis on the neural and behavioral data (14). In addition, we were confident in our ability to analyze each fMRI participant individually given the highly reliable nature of our neural data (average within-subject split-half reliability in occipitotemporal, $r = 0.82$; occipitoparietal, $r = 0.79$; early visual, $r = 0.86$; and prefrontal, $r = 0.65$; *SI Materials and Methods* and Figs. S5 and S6).

To determine whether the most active voxels alone could predict the mixed-category benefit, we correlated the amount of neural separation in the 10% most active voxels with the size of the mixed-category benefit. In this case, we found a significant correlation in occipitotemporal cortex of each participant (average $r = 0.59$, $P < 0.01$) with a smaller, but still significant, correlation in occipitoparietal cortex (average $r = 0.30$, $P < 0.01$) and no correlation in early visual (average $r = -0.03$, $P = 0.82$) or prefrontal cortex (average $r = -0.06$, $P = 0.63$; Fig. 4*B*). A leave-one-category-out analysis confirmed that the correlations in each of these sectors were not driven by any particular category (*SI Materials and Methods*). It should be noted that, given the fine-grained retinotopy in early visual cortex, objects presented across visual quadrants activate nearly completely separate regions, and this is reflected in the neural separation measure (ranging from 93% to 96% separation). Thus, by design, we anticipated that the neural separation of these patterns in the early visual cortex would not correlate with the behavioral results.

To compare correlations between any two sectors, a paired *t* test was performed on the Fisher *z*-transformed correlations. In this case, the correlation in occipitotemporal cortex was significantly greater than the correlations in all other sectors (occipitoparietal, $t_5 = 5.14$, $P < 0.01$; early visual, $t_5 = 4.68$, $P < 0.01$; PFC, $t_5 = 4.67$, $P < 0.01$). Together, these results show that the degree of neural overlap between stimulus categories, particularly within occipitotemporal cortex, strongly predicts the variation in the size of the behavioral mixed-category benefit for different categories. Moreover, because this analysis considers only 10% of voxels in a given sector, these results indicate that a relatively sparse set of representational channels predicts the behavioral effect.

Next, we varied the activation threshold to test whether a more restricted or expansive pool of neural channels could best predict the graded patterns of the behavioral mix effect. The percentage of most-active voxels used for the separation analysis was systematically increased from 1% to 99% (at 100%, there is complete overlap between all pairs of categories because every voxel is included for every category). The brain/behavior correlation of every subject as a function of percentile for each sector is shown in Fig. 5. This analysis revealed that the behavioral effect is well-predicted by the amount of neural separation across a broad range of occipitotemporal cortex, regardless of the percentile chosen for the neural separation analysis. Put another way, for any two categories, the degree of overlap among the most active voxels is similarly reflected across the majority of the entire activation profile.

To assess the statistical reliability of this result in a way that does not depend on a particular activation threshold, we used an AUC analysis to compute the aggregate neural separation between categories for all subjects in all sectors. These values were then correlated with the behavioral results, and the results were similar to those observed when considering only the top 10% most active voxels (Fig. 4*C*). There was a significant correlation

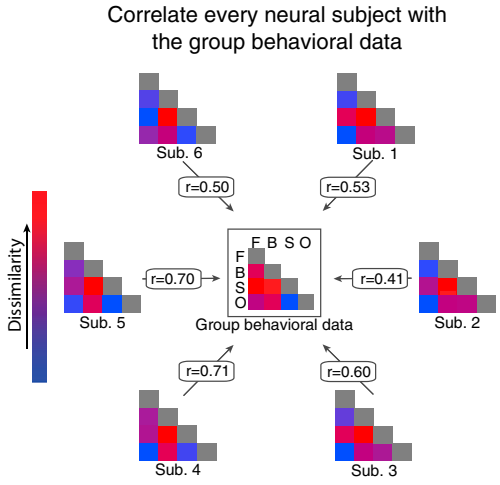

Fig. 3. Visualization of our analysis procedure. The center matrix represents the group data from the behavioral experiment with the color of each square corresponding to the size of the mixed-category benefit for that category pairing (Fig. 1). The six remaining matrices correspond to each fMRI participant, with the color of each square corresponding to the amount of neural separation between two categories in occipitotemporal cortex at the 10% activation threshold (Fig. 2). The correlations (*r*) are shown for each fMRI participant. Note that the *r* values shown here are the same as those shown in Fig. 4*B* for occipitotemporal cortex.

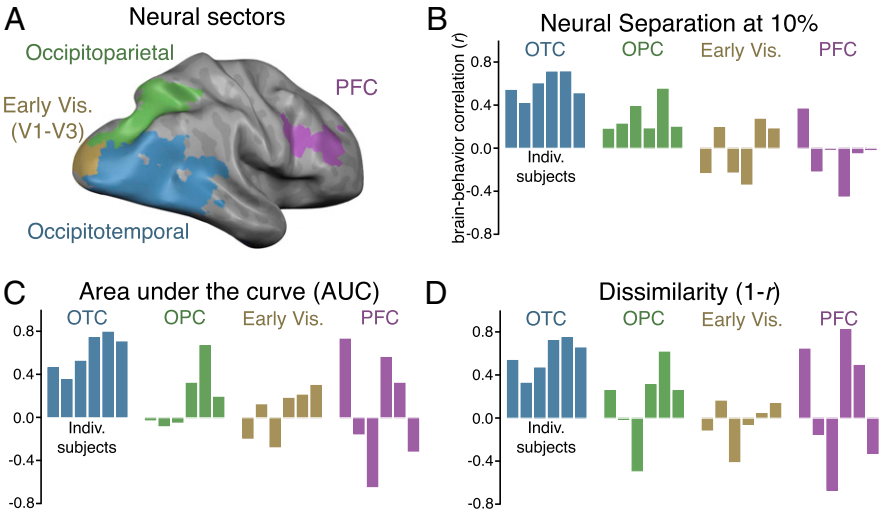

Fig. 4. (*A*) Visualization of the four sectors from a representative subject. (*B*) Brain/behavior correlations in every sector for each fMRI participant at the 10% activation threshold, with *r* values plotted on the *y* axis. Each bar corresponds to an individual participant. (*C*) Brain/behavior correlations in every sector for each participant when using the AUC analysis and (*D*) representational dissimilarity ($1 - r$).

in occipitotemporal cortex (average $r = 0.62$, $t_5 = 6.28$, $P < 0.01$), whereas little to no correlation was observed in the remaining sectors (occipitoparietal, $r = 0.20$; early visual, $r = 0.06$; prefrontal, $r = 0.11$; $P > 0.21$ in all cases; Fig. 4*C*). In addition, the correlation in occipitotemporal cortex was significantly greater than the correlations in all other sectors (occipitoparietal, $t_5 = 8.36$, $P < 0.001$; early visual, $t_5 = 6.86$, $P < 0.01$), except PFC, in which the effect was marginal ($t_5 = 2.30$, $P = 0.07$), likely because the brain/behavior correlation measures in PFC were highly inconsistent across participants. It is worth noting that, in the dorsal stream sector of occipitoparietal cortex, although there was a significant correlation at the 10% cutoff, the AUC analysis did not show a significant correlation ($t_5 = 1.45$, $P > 0.21$), suggesting that only the most active voxels in occipitoparietal cortex predict the behavioral data.

A convergent pattern of results was found by using a representational similarity analysis (14). That is, in occipitotemporal cortex, pattern dissimilarity (i.e., $1 - r$) across all pairs of categories strongly predicted the magnitude of the mixed-category benefit (average $r = 0.60$, $t_5 = 6.77$, $P < 0.01$; Fig. 4*D*). None of the other sectors showed a significant brain/behavior correlation using this neural measure ($P > 0.37$ in all cases), and direct comparisons between sectors show that the brain/behavior correlation in occipitotemporal cortex was significantly greater than those in the other sectors [$P < 0.05$ in all cases except in PFC, in which the difference was not significant ($t_5 = 1.87$, $P = 0.12$), again likely because of the brain/behavior correlations being highly variable in PFC].

To what extent do the category-selective regions for faces, bodies, scenes, and objects found in the occipitotemporal sector (11) drive these results? To address this question, we calculated the brain/behavior correlation in occipitotemporal cortex when considering only category selective regions (e.g., fusiform face area (FFA)/occipital face area (OFA) and parahippocampal place area (PPA)/retrosplenial cortex (RSC) when comparing faces and scenes) or only the cortex outside the category selective regions by using pattern dissimilarity as our measure of representational similarity (*SI Materials and Methods*). This analysis revealed a strong brain/behavior correlation within the category-selective regions (average $r = 0.62$, $P < 0.01$) and outside the category selective regions ($r = 0.60$, $P < 0.01$; Fig. S7), with no difference between these two correlations ($t_5 = 0.10$, $P = 0.92$).

Different assumptions about neural coding are tested by our two analyses: neural separation tests the idea that information is conveyed by maximal neural responses; neural similarity assumes that information is conveyed over the full distribution of responses within some circumscribed cortical territory. These measures dissociate in the occipitoparietal sector. The neural overlap analysis revealed that only the most active voxels have systematic differences among categories, suggesting that there is reliable object category information along portions of this sector (32). This observation was missed by the representational similarity analysis, presumably because many of the dorsal stream voxels are not as informative, making the full neural patterns subject to more noise. This result also highlights that the selection of voxels over which pattern analysis is performed can be critical to the outcomes. In contrast, in the occipitotemporal cortex, the separation and similarity metrics strongly correlated with behavior, and thus cannot distinguish between the functional roles of strong overall responses and distributed patterns. Nevertheless, this convergence strongly demonstrates that neural responses across the entire occipitotemporal cortex have the requisite representational structure to predict object-processing capacity in behavior.

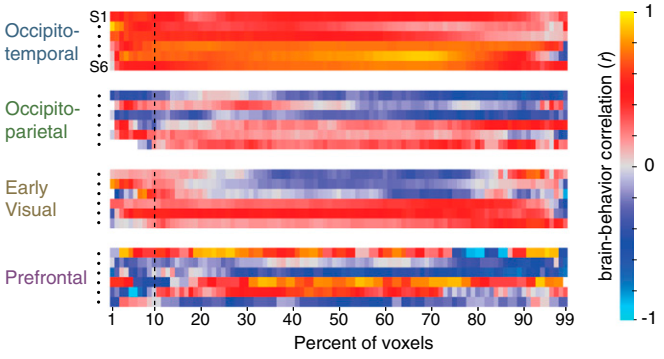

Fig. 5. Brain/behavior correlation in each subject for every sector for all possible activation thresholds. Each row shows the results for one of the six individual subjects (in the same order as shown in Fig. 4 for each sector). The *x* axis corresponds to the percentage of active voxels considered for the neural overlap analysis. The dashed vertical line marks the brain/behavior correlation when considering the 10% most active voxels, corresponding to the data plotted in Fig. 4*B*.

Discussion

Here we characterized participants' ability to simultaneously process multiple high-level items and linked this behavioral capacity to the underlying neural representations of these items. Participants performed better in a change-detection task when items were from different categories than when items were from the

same category. This suggests that, within the domain of high-level vision, items do not compete equally for representation. By using fMRI to independently measure the pattern of activity evoked by each category, we found that the magnitude of the mixed-category benefit for any category pairing was strongly predicted by the amount of neural separation between those categories in occipitotemporal cortex. These data suggest that processing multiple items at once is limited by the extent to which those items draw on the same underlying representational resources.

The present behavioral results challenge many influential models of attention and working memory. These models are typically derived from studies that use simple stimuli (e.g., colorful dots or geometric shapes), and tend to posit general limits that are assumed to apply to all items equally. For example, some models propose that processing capacity is set by a fixed number of pointers (15–18) or fixed resource limits (19–21), or from spatial interference between items (4, 22–24), none of which are assumed to depend on the particular items being processed. However, the present results demonstrate that the ability to process multiple items at once is greater when the items are from different categories. We interpret this finding in terms of partially separate representational resources available for processing different types of high-level stimuli. However, an alternative possibility is that these effects depend on processing overlap instead of representational overlap. For example, it has been argued that car experts show greater interference between cars and faces in a perceptual task than nonexperts because only experts use holistic processing to recognize both cars and faces (33, 34). The present behavioral data cannot distinguish between these possibilities. Future work will be required to determine which stages of perceptual processing show interference and whether this interference is best characterized in terms of representational or processing overlap.

Given that the size of the behavioral mixed-category benefit varied as a function of which categories were being processed, what is the source of this variability? We found that neural responses patterns, particularly in occipitotemporal cortex, strongly predicted the pattern of behavioral interference. This relationship between object processing and the structure of occipitotemporal cortex is intuitive because occipitotemporal cortex is known to respond to high-level object and shape information (11–14) and has receptive fields large enough to encompass multiple items in our experimental design (35). However, some aspects of the correlation between behavior and occipitotemporal cortex were somewhat surprising. In particular, we found that the relative separation between stimulus categories was consistent across the entire response profile along occipitotemporal cortex. That is, the brain/behavioral relationship held when considering the most active voxels, the most selective voxels (e.g., FFA/PPA), or those voxels outside of classical category selective regions.

The fact that the brain/behavior correlation can be seen across a large majority of occipitotemporal cortex is not predicted by expertise-based (36) or modular (11) models of object representation. If this correlation was caused by differences in expertise between the categories, one might expect to see a significant correlation only in FFA (37). Similarly, if competition within the most category-selective voxels drove the behavioral result, we would expect to only find a significant brain/behavior correlation within these regions. Of course, it is important to emphasize that the present approach is correlational, so we do not know whether all or some subset of occipitotemporal cortex is the underlying site of interference. Future work with the use of causal methods (38), or that explores individual differences in capacity and neural organization (39), will be important to explore these hypotheses.

In light of the relationship between behavioral performance and neural separation, it is important to consider the level of representation at which the competition occurs. For example, items might interfere with each other within a semantic (40), categorical (11), or perceptual (41, 42) space. Variants of the current task could be used to isolate the levels of representation involved in the mixed-category benefit and its relationship to neural responses. For example, the use of exemplars with significant perceptual variation (e.g., caricatures, Mooney faces, and photographs) would better isolate a category level of representation. Conversely, examining the same type of brain/behavioral relationship within a single category would minimize the variation in semantic space and would target a more perceptual space. Although the present data cannot isolate the level of representation at which competition occurs, it is possible that neural competition could occur at all levels of representation, and that behavioral performance will ultimately be limited by competition at the level of representation that is most relevant for the current task.

The idea of competition for representation is a prominent component of the biased competition model of visual attention, which was originally developed based on neurophysiological studies in monkeys (7), and has been expanded to explain certain human neuroimaging and behavioral results (43). These previous studies have shown that, if two items are presented close enough to land within a single receptive field, they compete for neural representation, such that the neural response to both items matches a weighted average of the response to each individual item alone (44). When attention is directed to one of the items, neural responses are biased toward the attended item, causing the neuron to fire as if only the attended item were present (7–10, 44). In the present study, we did not measure neural competition directly. Instead, we measured neural responses to items presented in isolation and used similarity in those responses to predict performance in a behavioral task. We suggest that the cost for neural similarity reflects a form of competition, but we cannot say how that competition manifests itself (e.g., as suppression of overall responses or a disruption in the pattern of responses across the population) or if these mechanisms are the same from early to high-level visual cortex (45–47). Thus, parameterizing neural similarity and measuring neural responses to items presented simultaneously will be essential for addressing the relationship between neural similarity and neural competition.

Overall, the present findings support a framework in which visual processing relies on a large set of broadly tuned coding channels, and perceptual interference between items depends on the degree to which they activate overlapping channels. This proposal predicts that a behavioral mixed-category benefit will be obtained for tasks that require processing multiple items at once, to the extent that the items rely on separate channels. It is widely known that there are severe high-level constraints on our ability to attend to, keep track of, and remember information. The present work adds a structural constraint on information processing and perceptual representation, based on how high-level object information is represented within the visual system.

Materials and Methods

Behavioral Task. Participants ($N = 55$) viewed four items for 800 ms, followed by a fixation screen for 1,000 ms, followed by a probe display of four items, one of which was cued with a red frame. For any display, images were randomly chosen from the stimulus set with the constraint that all images on a given display were unique. There were no changes on half the trials. On the other trials, the cued item changed to a different item from the same category (e.g., from one face to another). Participants reported if the cued item had changed. The probe display remained on screen until participants gave a response by using the keyboard.

On same-category trials, all four items came from the same category (e.g., four faces or four scenes), with each category appearing equally often across trials. On mixed-category trials, there were two items from each category (e.g., two faces and two scenes). Items were arranged such that one item from each category appeared in each visual hemifield. Across the mixed-category trials, both categories were tested equally often. The location of the cued item was chosen in a pseudorandom fashion.

fMRI Task. Six participants (none of whom performed the behavioral task) completed the fMRI experiment. There were eight runs of the main experiment, one run of meridian mapping to localize early visual areas (V1–V3), one run of a working memory task used to localize PFC, and two localizer runs used to define the occipitotemporal and occipitoparietal sectors as well as

category-selective regions (FFA/OFA, PPA/RSC, extrastriate body area/fusiform body area, and lateral occipital). Participants were shown the same faces, bodies, scenes, and objects that were used in the behavioral experiments (Fig. S1), presented one at a time in each of the four quadrants [top left (TL), top right (TR), bottom left (BL), bottom right (BR)], following a standard blocked design. In each 16-s stimulus block, images from one category were presented in isolation (one at a time) at one of the four locations; 10-s fixation blocks intervened between each stimulus block. A total of 11 items were presented per block for 1 s with a 0.45-s intervening blank. Participants were instructed to maintain fixation on a central cross and to press a button indicating when the same item appeared twice in a row, which happened once per block. For any given run, all four stimulus categories were presented in two of the four possible locations, for two separate blocks per category × location pair, yielding 16 blocks per run. Across eight runs, the categories were presented at each pair of locations (TL–TR; TL–BR; BL–BR; BL–TR), yielding eight blocks for each of the 16 category × location conditions (*SI Materials and Methods* provides information on localizer runs).

Neural Separation Analysis. The logic of this analysis is to compute the proportion of voxels that are activated by any two categories (e.g., faces and scenes): if no voxels are coactivated, there is 100% neural separation, whereas, if all voxels are coactivated, there is 0% separation. This analysis relies on one free parameter, which sets the percent of the most active voxels to consider as the available representational resources of each object category. In addition, we take into account location by considering the overlap between the two categories at all pairs of locations, and then averaging across location pairs.

To compute the neural separation between two categories within a sector (e.g., faces and scenes in occipitotemporal cortex), we used the following procedure. (*i*) The responses (β) for each category–location pair were sorted and the top *n* (as a percentage) was selected for analysis, where *n* was varied from 1% to 99%. (*ii*) Percent overlap at a particular threshold was computed as the number of voxels that were shared between any two conditions at that threshold, divided by the number of selected voxels for each condition (e.g., if 10 voxels overlap among the top 100 face voxels and the top 100 scene voxels, the face–scene overlap would be 10/100 = 10%). To take into account location, percent overlap was calculated separately for all 12 possible location pairs, e.g., faces-TL–scenes-TR, faces-TL–scenes-BL, faces-TL–scenes-BR. Fig. 2 shows an example whereby $n = 10\%$ for the activation patterns of objects (purple) scenes(blue) and shared (yellow); and (*iii*) Finally, we averaged across these 12 overlap estimates to compute the final overall estimate of overlap between a pair of categories. This measure can be interpreted as the degree to which two different categories in two different locations will recruit similar cortical territory. We computed percent overlap at each percentile (i.e., 1%–99% of the most active voxels), generating a neural overlap curve, and converted this to a percentage separation measure by taking 1 minus the percent overlap. This procedure was conducted for all pairs of categories, for all sectors, for all subjects.

ACKNOWLEDGMENTS. We thank Marvin Chun, James DiCarlo, Nancy Kanwisher, Rebecca Saxe, Amy Skerry, Maryam Vaziri-Pashkam, and Jeremy Wolfe for comments on the project. This work was supported by National Science Foundation (NSF) Graduate Research Fellowship (to M.A.C.), National Research Service Award Grant F32EY022863 (to T.K.), National Institute of Health National Eye Institute-R01 Grant EY01362 (to K.N.), and NSF CAREER Grant BCS-0953730 (to G.A.A.).

1. Baddeley AD, Hitch G (1974) Working memory. *The Psychology of Learning and Motivation: Advances in Research and Theory*, ed Bower GH (Academic, New York), pp 47–89.
2. Duncan J, Martens S, Ward R (1997) Restricted attentional capacity within but not between sensory modalities. *Nature* 387(6635):808–810.
3. Gazzaniga M, Ivry RB, Mangun GR (2008) *Cognitive Neuroscience* (Norton, New York).
4. Franconeri SL, Alvarez GA, Cavanagh P (2013) Flexible cognitive resources: Competitive content maps for attention and memory. *Trends Cogn Sci* 17(3):134–141.
5. Pelli DG, Tillman KA (2008) The uncrowded window of object recognition. *Nat Neurosci* 11(10):1129–1135.
6. Wandell BA, Dumoulin SO, Brewer AA (2007) Visual field maps in human cortex. *Neuron* 56(2):366–383.
7. Desimone R, Duncan J (1995) Neural mechanisms of selective visual attention. *Annu Rev Neurosci* 18:193–222.
8. Kastner S, et al. (2001) Modulation of sensory suppression: Implications for receptive field sizes in the human visual cortex. *J Neurophysiol* 86(3):1398–1411.
9. Kastner S, De Weerd P, Desimone R, Ungerleider LG (1998) Mechanisms of directed attention in the human extrastriate cortex as revealed by functional MRI. *Science* 282(5386):108–111.
10. Beck DM, Kastner S (2005) Stimulus context modulates competition in human extrastriate cortex. *Nat Neurosci* 8(8):1110–1116.
11. Kanwisher N (2010) Functional specificity in the human brain: A window into the functional architecture of the mind. *Proc Natl Acad Sci USA* 107(25):11163–11170.
12. Bell AH, et al. (2011) Relationship between functional magnetic resonance imaging-identified regions and neuronal category selectivity. *J Neurosci* 31(34):12229–12240.
13. Haxby JV, et al. (2001) Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293(5539):2425–2430.
14. Kriegeskorte N, et al. (2008) Matching categorical object representations in inferior temporal cortex of man and monkey. *Neuron* 60(6):1126–1141.
15. Pylyshyn ZW, Storm RW (1988) Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spat Vis* 3(3):179–197.
16. Drew T, Vogel EK (2008) Neural measures of individual differences in selecting and tracking multiple moving objects. *J Neurosci* 28(16):4183–4191.
17. Awh E, Barton B, Vogel EK (2007) Visual working memory represents a fixed number of items regardless of complexity. *Psychol Sci* 18(7):622–628.
18. Zhang W, Luck SJ (2008) Discrete fixed-resolution representations in visual working memory. *Nature* 453(7192):233–235.
19. Alvarez GA, Cavanagh P (2004) The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychol Sci* 15(2):106–111.
20. Ma WJ, Husain M, Bays PM (2014) Changing concepts of working memory. *Nat Neurosci* 17(3):347–356.
21. Franconeri SL, Alvarez GA, Enns JT (2007) How many locations can be selected at once? *J Exp Psychol Hum Percept Perform* 33(5):1003–1012.
22. Delvenne J-F (2005) The capacity of visual short-term memory within and between hemifields. *Cognition* 96(3):B79–B88.
23. Delvenne J-F, Holt JL (2012) Splitting attention across the two visual fields in visual short-term memory. *Cognition* 122:258–263.
24. Franconeri SL, Jonathan SV, Scimeca JM (2010) Tracking multiple objects is limited only by object spacing, not by speed, time, or capacity. *Psychol Sci* 21(7):920–925.
25. Biederman I (1987) Recognition-by-components: A theory of human image understanding. *Psychol Rev* 94(2):115–147.
26. Tarr MJ, Bülthoff HH (1998) Image-based object recognition in man, monkey and machine. *Cognition* 67(1-2):1–20.
27. Wong JH, Peterson MS, Thompson JC (2008) Visual working memory capacity for objects from different categories: A face-specific maintenance effect. *Cognition* 108(3):719–731.
28. Olsson H, Poom L (2005) Visual memory needs categories. *Proc Natl Acad Sci USA* 102(24):8776–8780.
29. Ungerleider LG, Mishkin M (1982) Two cortical visual systems. *Analysis of Visual Behavior*, eds Ingle DJ, Goodale MA, Mansifield RJW (MIT Press, Cambridge, MA), pp 549–583.
30. Goodale MA, Milner AD (1992) Separate visual pathways for perception and action. *Trends Neurosci* 15(1):20–25.
31. Curtis CE, D'Esposito M (2003) Persistent activity in the prefrontal cortex during working memory. *Trends Cogn Sci* 7(9):415–423.
32. Konen CS, Kastner S (2008) Two hierarchically organized neural systems for object information in human visual cortex. *Nat Neurosci* 11(2):224–231.
33. McKeeff TJ, McGugin RW, Tong F, Gauthier I (2010) Expertise increases the functional overlap between face and object perception. *Cognition* 117(3):355–360.
34. McGugin RW, McKeeff TJ, Tong F, Gauthier I (2011) Irrelevant objects of expertise compete with faces during visual search. *Atten Percept Psychophys* 73(2):309–317.
35. Kravitz DJ, Vinson LD, Baker CI (2008) How position dependent is visual object recognition? *Trends Cogn Sci* 12(3):114–122.
36. Gauthier I, Skudlarski P, Gore JC, Anderson AW (2000) Expertise for cars and birds recruits brain areas involved in face recognition. *Nat Neurosci* 3(2):191–197.
37. Tarr MJ, Gauthier I (2000) FFA: A flexible fusiform area for subordinate-level visual processing automatized by expertise. *Nat Neurosci* 3(8):764–769.
38. Afraz SR, Kiani R, Esteky H (2006) Microstimulation of inferotemporal cortex influences face categorization. *Nature* 442(7103):692–695.
39. Park J, Carp J, Hebrank A, Park DC, Polk TA (2010) Neural specificity predicts fluid processing ability in older adults. *J Neurosci* 30(27):9253–9259.
40. Moores E, Laiti L, Chelazzi L (2003) Associative knowledge controls deployment of visual selective attention. *Nat Neurosci* 6(2):182–189.
41. Tanaka K (2003) Columns for complex visual object features in the inferotemporal cortex: Clustering of cells with similar but slightly different stimulus selectivities. *Cereb Cortex* 13(1):90–99.
42. Ullman S (2007) Object recognition and segmentation by a fragment-based hierarchy. *Trends Cogn Sci* 11(2):58–64.
43. Scalf PE, Torralbo A, Tapia E, Beck DM (2013) Competition explains limited attention and perceptual resources: Implications for perceptual load and dilution theories. *Front Psychol* 4:243.
44. Reynolds JH, Desimone R (1999) The role of neural mechanisms of attention in solving the binding problem. *Neuron* 24(1):19–29, 111–125.
45. Treue S, Hol K, Rauber H-J (2000) Seeing multiple directions of motion-physiology and psychophysics. *Nat Neurosci* 3(3):270–276.
46. Carandini M, Movshon JA, Ferster D (1998) Pattern adaptation and cross-orientation interactions in the primary visual cortex. *Neuropharmacology* 37(4-5):501–511.
47. MacEvoy SP, Tucker TR, Fitzpatrick D (2009) A precise form of divisive suppression supports population coding in the primary visual cortex. *Nat Neurosci* 12(5):637–645.
48. Loftus GR, Masson MEJ (1994) Using confidence intervals in within-subject designs. *Psychon Bull Rev* 1(4):476–490.

Supporting Information

Cohen et al. 10.1073/pnas.1317860111

SI Materials and Methods

Behavioral Materials and Methods. ***Stimuli.*** For all behavioral experiments, stimuli were presented on a 24-inch LCD monitor with a 60-Hz refresh rate. Participants sat ∼57 cm away from the monitor such that 1° of visual angle subtended 1 cm on the display. Experiments were created and controlled on a computer running MATLAB with the Psychophysics Toolbox (1, 2). Images were presented at 6° × 6° of visual angle, with a different image appearing in each quadrant of the visual field, 8.4° away from fixation. Within a hemifield, the center-to-center distance of items was ∼7.5°, whereas the center-to-center distance of two items that were in different hemifields but on the same horizontal plane was 15.4°. A red fixation dot (0.55°) was presented in the middle of the display. The background of the display had an average luminance of 73.8 cd/m^2.

Stimulus selection. Stimuli were chosen in an attempt to minimize the possibility that participants could perform the task by focusing on low- or midlevel features. All items were matched for luminance and contrast. All faces were vertically oriented, looking at the camera, smiling, Caucasian, and approximately between the ages of 20–50 y, and each image was cropped so the outer contours of the head/hair were not showing. Thus, participants would have to represent the "face" rather than notice a change in outer contour or a difference in hair texture. Body images were of a single individual (M.A.C.) wearing the same outfit in a variety of action poses (e.g., jumping, throwing). Scenes were a variety of natural scene categories that never contained faces, bodies, or any of the objects from the object category. Objects were specifically selected to have a similar, round outer contour. This was done in an attempt to force participants to focus on the object itself (e.g., an apple or a clock), rather than just focus on the bounding contour alone.

Procedure. Each participant initially performed a calibration block of same-category trials with their assigned categories (e.g., four faces, four scenes). In this block, there were 30 practice trials and 120 experimental trials. Half the experimental trials were allocated for each of the two assigned categories. QUEST, a Bayesian adaptive staircase (3) procedure, was used to adaptively change the transparency of the items until performance on same-category displays for both categories was ∼70% (actual same-category performance across all six conditions was 73.83%). These transparency thresholds were then used for the stimuli in the main experiment. In the main experiment, participants performed 20 practice trials followed by four blocks of 80 experimental trials (320 total experimental trials). Within a block, all trial types (e.g., same-category/mixed-category, different types of mixed-category display configurations) were shuffled and appeared in a random order.

Participants and exclusion criterion. To obtain an $N = 10$ for each of the six category pairings (e.g., faces–scenes, faces–bodies), 55 participants completed one or more of the category-pairings experiments. The visibility matching procedure described earlier was designed to ensure that performance was equal on both the same-category conditions. However, in some cases, individuals showed large differences in performance across the two categories that could not be eliminated with this procedure. Thus, participants whose performance on the two same-category conditions differed by more than 10% were excluded. To ensure that this procedure did not result in an atypical sample, we correlated the size of the mixed-category benefit when the exclusion criterion was and was not applied. There was a very strong correlation between these data sets ($r = 0.93$) with the different mixed-category benefits as follows: faces and scenes, 5.2% with exclusion, 5.1% without exclusion; bodies and scenes, 5.6% with exclusion, 5.2% without exclusion; bodies and faces, 3.3% with exclusion, 3.2% without exclusion; bodies and objects, 3.3% with exclusion, 4.2% without exclusion; faces and objects 2.4% with exclusion, 2.6% without exclusion; objects and scenes −0.8% with exclusion, 2.1% without exclusion. Fig. S5 shows the different brain/behavior correlations in occipitotemporal cortex using the top 10% and area under the curve measurements with vs. without excluded participants.

Functional MRI Acquisition. Structural and functional imaging data were collected on a 3-T Siemens Trio scanner at the Harvard University Center for Brain Sciences. Structural data were obtained in 176 axial slices with 1 × 1 × 1 mm voxel resolution (repetition time 2,200 ms). Functional blood oxygenation level-dependent data were obtained by using a gradient-echo echo-planar pulse sequence (33 axial slices parallel to the anterior commissure–posterior commissure line; 70 × 70 matrix; field of view 256 × 256 mm; 3.1 × 3.1 × 3.1 mm voxel resolution; gap thickness 0.62 mm; repetition time 2,000 ms; echo time 60 ms; flip angle = 90°). A 32-channel phased-array head coil was used. Stimuli were generated using the Psychophysics toolbox for MATLAB and displayed with an LCD projector onto a screen in the scanner that subjects viewed via a mirror attached to the head coil.

Functional MRI Experiment Localizer Runs. ***Meridian map runs.*** Participants were instructed to maintain fixation and were shown blocks of flickering black-and-white checkerboard wedge stimuli, oriented along the vertical or horizontal meridian. The apex of each wedge was at fixation and the base extended to 8° of visual angle in the periphery, with a width of 4.42°. The checkerboard pattern flickered at 8 Hz. The run consisted of four vertical meridian and four horizontal meridian blocks. Each stimulus block was 12 s with a 12-s intervening blank period. The orientation of the stimuli (vertical vs. horizontal) alternated from one block to the other.

Prefrontal cortex runs. Participants performed a change detection task in which they had to say if an item changed between displays. Each display consisted of four items with each item placed in one of the four visual quadrants. The configuration of items on these displays was not identical to the configuration used in the behavioral experiment (Fig. 1). In this case, all items were equidistant from fixation and each item was equally close to the two adjacent items within and across the visual hemifields. On every trial, all items came from the same category (e.g., four faces, four bodies), with the images and locations selected randomly. The first display appeared for 1 s, followed by a 0.7-s blank interval with a fixation cross, and then a second display was presented for 1 s. Participants responded during a 1.8-s intertrial interval. Immediately before each trial, the black fixation dot turned red to alert participants that the next trial was about to begin. A change detection block consisted of eight trials, in which each of the four categories was used on two trials. Changes occurred on half of the trials such that a change occurred once with every category in every block. Each run was composed of three change detection blocks of 32 s each, with fixation periods of 32 s following each block.

Localizer runs. Participants performed a one-back repetition detection task with blocks of faces, bodies, scenes, objects, and scrambled objects. Stimuli in these runs were different from those in the experimental runs. Each run consisted of 10 stimulus blocks

of 16 s, with intervening 12-s blank periods. Each category presented twice per run, with the order of the stimulus blocks counterbalanced in a mirror reverse manner (e.g., face, body, scene, object, scrambled, scrambled, objects, scene, body, face). Within a block, each item was presented for 1 s followed by a 0.33-s blank. Additionally, these localizer runs contained an orthogonal motion manipulation: In half of the blocks, the items were presented statically at fixation; in the remaining half of the blocks, items moved from the center of the screen toward either one of the four quadrants or along the horizontal and vertical meridians at 2.05 °/s. Each category was presented in a moving and stationary block.

Mixed-Category Benefit Analysis. The graded nature of the mixed category benefit was evaluated by comparing regression models whereby category pair was included as a factor (i.e., a separate parameter was fit for each category pair) vs. not included as a factor (i.e., a null model with an intercept term only). Model comparison was conducted by ANOVA to test whether the model that estimated the mix effect separately for each category pair fit the data reliably better than a model that fit a single intercept across all category pairs.

Functional MRI Data Analysis. All functional MRI (fMRI) data were processed by using Brain Voyager QX software (Brain Innovation). Preprocessing steps included 3D motion correction, slice scan-time correction, linear trend removal, temporal high-pass filtering (0.01-Hz cutoff), spatial smoothing (4 mm full width at half-maximum kernel), and transformation into Talairach space. Statistical analyses were based on the general linear model. All general linear model (GLM) analyses included boxcar regressors for each stimulus block convolved with a γ-function to approximate the idealized hemodynamic response. Motion correction regressors were included as regressors of no interest. For each experimental protocol, separate GLMs were computed for each participant and run, yielding β-maps for each condition.

Defining Neural Sectors. Sectors were defined in each participant using the following procedure. By using the localizer runs, a set of visually active voxels was defined based on the contrast of [faces + bodies + scenes + objects] vs. rest (false discovery rate < 0.05, cluster threshold 150 contiguous 1 × 1 × 1 voxels) within a gray matter mask. To divide these visually responsive voxels into sectors, the "EarlyV" sector included all active voxels within V1, V2, and V3, which were defined by hand on an inflated surface representation based on the horizontal vs. vertical contrasts of the meridian mapping experiment. The occipitotemporal and occipitoparietal sectors were then defined as all remaining active voxels (outside of EarlyV), where the division between the dorsal and ventral streams was drawn by hand in each participant starting at the edge of EarlyV and determined by the spatial profile of active voxels along the surface. Finally, the prefrontal cortex runs were used to identify the prefrontal cortex sector from the contrast of working memory vs. rest; this region was masked by the gray matter but not by visually active voxels (because, for some subjects, no frontal voxels were significantly visually active).

Neural Dissimilarity Analysis. This analysis follows the representational similarity methods outlined in the work of Kriegeskorte et al. (4). For each pair of categories, we computed the Pearson correlation between the response patterns (β) across the entire sector (e.g., the occipitotemporal cortex), and converted this to a dissimilarity measure (1 − r). Location was taken into account in the same way as described in the channel separation analysis (*Methods and Materials*).

Category-Selective Region of Interest Analysis. To define category-selective regions, we computed standard contrasts for face selectivity [faces > (bodies scenes objects)], scene selectivity [scenes > (bodies faces objects)], and body selectivity [bodies > (faces scenes objects)] based on independent localizer runs. For object-selective areas, the contrast of objects > scrambled was used. In each participant, face-, body-, scene-, and object-selective regions were defined by using a semiautomated procedure that selects all significant voxels within a 9-mm radius spherical region of interest (ROI) around the weighted center of category-selective clusters (5), where the cluster is selected based on proximity to the typical anatomical location of each region based on a meta-analysis. All ROIs for all participants were verified by eye and adjusted if necessary. Category-selective regions included fusiform face area (FFA) and fusiform body area (FBA) (faces), parahippocampal place area (PPA) and retrosplenial cortex (RSC) (scenes), extrastriate body area (EBA) and FBA (bodies), and lateral occipital (LO) (objects).

We next computed neural dissimilarity inside the category-selective ROIs for all pairs of categories, considering only the ROIs specific for the two categories in each pair. For example, the correlation between the response patterns to scenes and the response patterns to faces was computed considering only the voxels within scene- and face-selective regions, and converted to a dissimilarity measures (1 − r). This dissimilarity measure for faces and scenes was then compared with the size of the mixed-category benefit with faces and scenes. Location was taken into account following the same procedure described in the channel separation analysis. To compute neural dissimilarity outside the category-selective ROIs, the same procedure was followed, but the voxels under consideration for scene and face competition, for example, were all voxels in the occipitotemporal cortex that were not in any of the scene- or face-selective regions.

Within-Subject Pattern Reliability. We computed the reliability of neural response patterns separately for each fMRI participant and each brain sector by using the following procedure. The data were split into two halves, such that each half contained one run of every possible location–location pairing [i.e., each half had a (top left) − (bottom right) run, a (bottom left) − (bottom right) run]. Next, the pattern for each category in each location was correlated between the two halves of the data (e.g., activity for faces in the top left for half 1 was correlated with activity for faces in the top left for half 2). This correlation was calculated for each category in each location. Averaged across categories, locations, and subjects, the overall correlations were quite high in each sector (Fig. S3), indicating that these neural patterns were highly reliable within subjects.

Between-Subject Consistency in Neural Overlap. In every sector of each participant, the amount of overlap (top 10% and area under the curve) or dissimilarity ($1 - r$) was calculated for each category pairing. Those measurements across all category pairings ($n = 6$ total) were then compared between all subjects. The correlation between each subject in every sector by using all three similarity measures was quite high overall, indicating a high degree of between-subject consistency in neural overlap between categories (Fig. S4).

Assessing the Statistical Significance of the Brain/Behavior Correlations. All brain/behavior correlations were computed for each individual fMRI subject against the group behavioral data. For each fMRI subject, an r value (Pearson correlation value) was calculated by using a particular neural similarity measure (i.e., neural separation, pattern dissimilarity). The r values obtained for a given brain/behavior correlation were Fisher z-transformed and then tested for statistical significance by using a random effects analysis. To compare the relative strength of two given correlations (e.g., is the

brain/behavior correlation stronger in occipitotemporal cortex compared with early visual areas), a within-subject t test was performed on the transformed correlation coefficients obtained from two regions.

It should be noted that, in every subject, a brain/behavior correlation was obtained by using the same group average behavioral data. However, despite the common behavioral measures, the brain/behavior correlation values across brain participants are still independent from one another and are not intrinsically correlated. Simulations were run to ensure the independence of such correlation values and more generally to validate the statistical procedure used. In these simulations, the behavioral data were fixed and the brain data were randomly set for each subject, the brain/behavior correlation was obtained for each subject, and the distribution of these brain/behavior correlations was assessed. The results of this simulation showed a null distribution of correlation values centered around 0 that was well-approximated by a normal distribution following the Fisher z-transformation.

Do Outlier Category Pairings Drive These Brain/Behavior Correlations? Given that there were only six category pairings used in this experiment, it is possible that one particular pairing (e.g., scenes and bodies) could be an outlier that drives the observed correlation. This possibility challenges our claim that the magnitude of the mixed-category benefit for any category pairing is strongly predicted by the amount of neural separation. To address this concern, we conducted two supplemental analyses.

First, we conducted a bootstrap analysis, in which we sampled, with replacement, six random categories pairings from the set, and conducted the same analysis on that random sample. Thus, on some iterations, a particular category pair (e.g., scenes and bodies) will be left out completely, whereas, on other iterations, that pair could be included multiple times (giving more weight to that pairing). For each iteration, the random set of six category pairings were used to calculate a new, subject-averaged brain/behavior correlation. This process was repeated 10,000 times to obtain a distribution of correlation values. We found that in, occipitotemporal cortex, a brain/behavior correlation of zero was not within the 95% interval when using the 10% overlap, area under the curve, or pattern dissimilarity analysis ($P < 0.05$ in all cases). As points of comparison, zero was within the 95% interval in occipitoparietal, early visual, and prefrontal cortex for all neural measure we used ($P > 0.20$ in all cases).

Second, we also analyzed the data after having excluded each of the different category pairings one at a time (e.g., leave out faces and scenes and keep the rest, leave out faces and bodies and keep the rest). In this case, the brain/behavior correlations in occipitotemporal remained significant with all three neural measures regardless of which of category pairings were left out ($P < 0.01$ in all cases except when bodies and scenes were left out and we used the pattern dissimilarity measure, in which $P = 0.59$). Taken together, these two results suggest that the significant brain/behavior correlations we find in occipitotemporal cortex are not driven by outlier category pairings.

1. Brainard DH (1997) The psychophysics toolbox. *Spat Vis* 10(4):433–436.
2. Pelli DG (1997) The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spat Vis* 10(4):437–442.
3. Watson AB, Pelli DG (1983) QUEST: A Bayesian adaptive psychometric method. *Percept Psychophys* 33(2):113–120.
4. Kriegeskorte N, et al. (2008) Matching categorical object representations in inferior temporal cortex of man and monkey. *Neuron* 60(6):1126–1141.
5. Peelen MV, Downing PE (2005) Within-subject reproducibility of category-specific visual activation with functional MRI. *Hum Brain Mapp* 25(4):402–408.

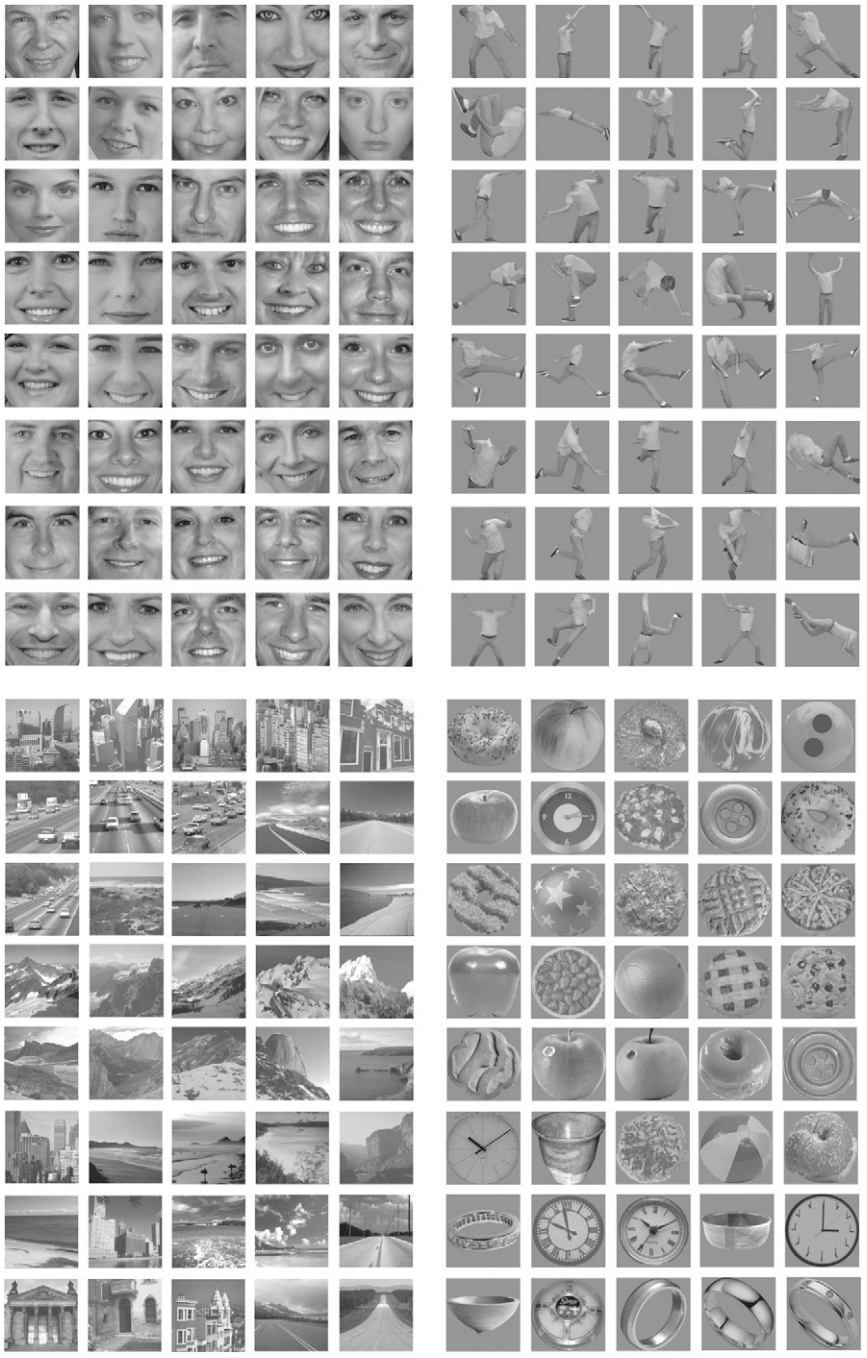

Fig. S1. Stimuli used in all experiments. *SI Materials and Methods* provides a description on the stimulus selection criteria.

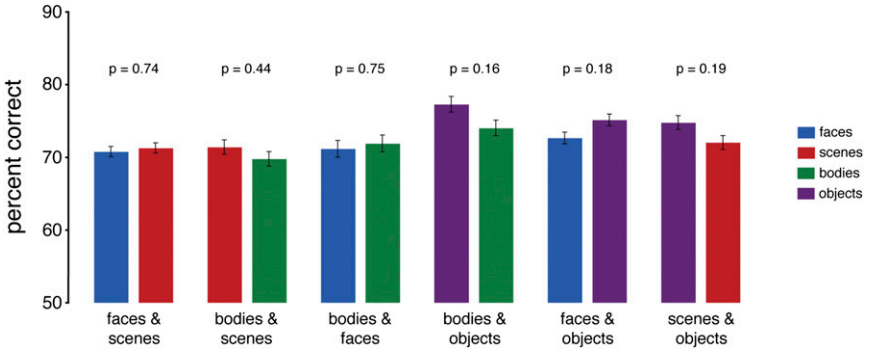

Fig. S2. Performance for same-category conditions across all category pairings. Each group of bars reflects one of the six possible category pairings (e.g., faces and scenes). Each bar reflects the percent correct on the displays in which all of the items were from the same category (e.g., all four faces or all four scenes). Error bars reflect within-subject SEM. Note that, for each category pairing, the data here were averaged together to compute the same-category performance data presented in Fig. 1.

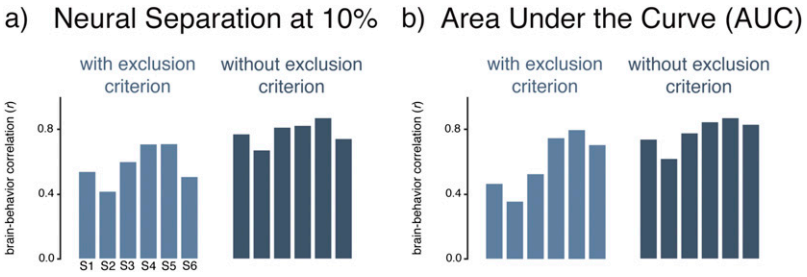

Fig. S3. Brain/behavior correlations in occipitotemporal cortex for each fMRI participant, with behavioral subjects excluded (light blue) and without behavioral subjects excluded (dark blue). The brain/behavior correlation (*r*) is plotted on the *y* axis. Each bar represents an individual fMRI participant. Correlations in occipitotemporal cortex when using (*A*) the 10% activation threshold and (*B*) the area under the AUC analysis.

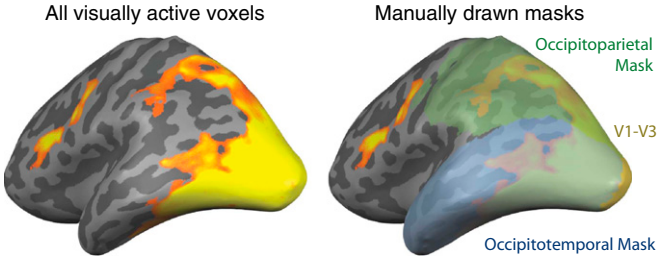

Fig. S4. All visually active voxels as obtained by the contrast [faces + bodies + scenes + objects] vs. rest (false discovery rate < 0.05, cluster threshold 150 contiguous 1 × 1 × 1 voxels) and shown on the inflated left hemisphere of a representative participant. For each participant, early visual cortex (V1–V3) was first defined with meridian map on the inflated cortical surface. Once defined, masks for the occipitotemporal (light blue) and occipitoparietal (light green) cortices were manually drawn starting at the edge of V1–V3, up through the division between the ventral and dorsal pathways on the lateral surface, and continued to include all voxels within each pathway. It was from these masks that all active voxels were selected, which in turn became the ROIs used in the analysis.

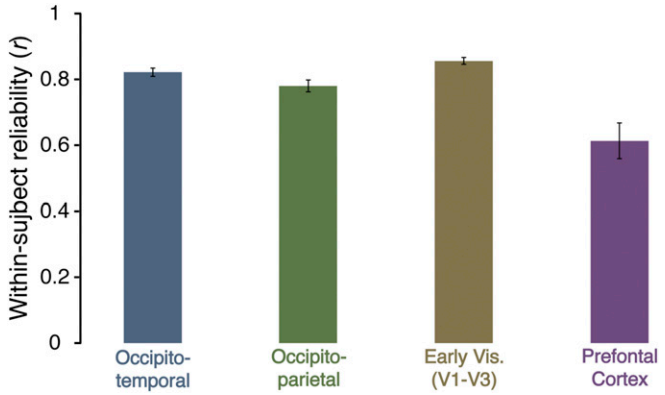

Fig. S5. Average within-subject split-half reliability of activation patterns, averaged across category, location, and subject. Error bars denote ±1 SEM. *SI Materials and Methods* provides a description of how within-subject reliability was calculated.

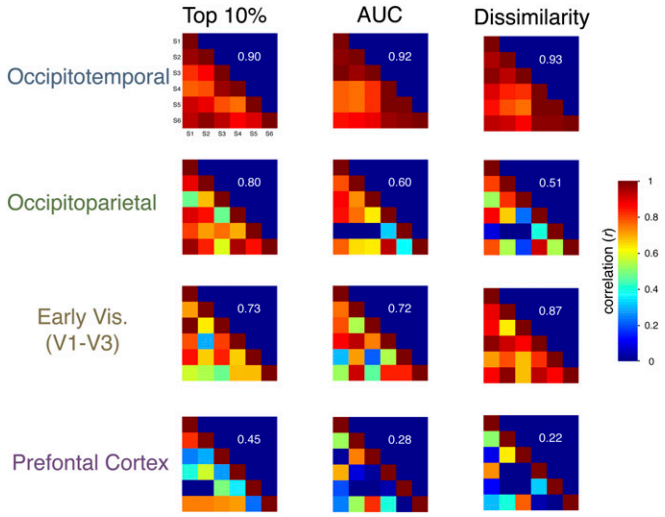

Fig. S6. Between-subject consistency in neural separation. Each plot shows the correlation between subjects in the amount of neural separation for each category pairing, with separate plots for each sector (rows) and for each neural separation measure (columns). Warmer colors denote higher correlations, and cooler colors denote lower correlations.

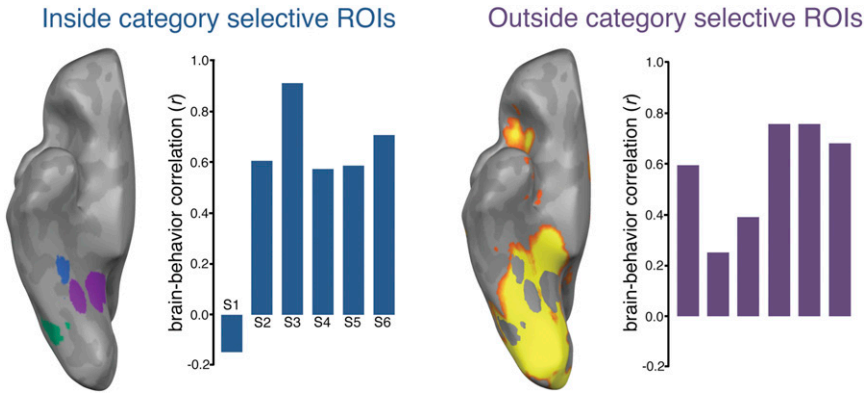

Fig. S7. Brain/behavior correlation considering the category-selective ROIs within occipitotemporal cortex (*Left*) or excluding the category-selective ROIs within occipitotemporal cortex (*Right*). For each category pairing, only the voxels selective for either of those two categories was included (faces, FFA/occipital face area; bodies, EBA/FBA; scenes, PPA/RSC; and objects, LO). Each individual fMRI participant is plotted on the *x* axis. The *y* axis shows the brain/behavior correlation (*r*) (*SI Materials and Methods*).