Task alters category representations in prefrontal but not high-level visual cortex

Lior Bugatus\textsuperscript{a,*}, Kevin S. Weiner\textsuperscript{b}, Kalanit Grill-Spector\textsuperscript{a,b}

\textsuperscript{a} Department of Psychology, Stanford University, Stanford, CA, 94305, USA
\textsuperscript{b} Stanford Neuroscience Institute, Stanford University, Stanford, CA 94305, USA

A central question in neuroscience is how cognitive tasks affect category representations across the human brain. Regions in lateral occipito-temporal cortex (LOTC), ventral temporal cortex (VTC), and ventro-lateral prefrontal cortex (VLPFC) constitute the extended “what” pathway, which is considered instrumental for visual category processing. However, it is unknown (1) whether distributed responses across LOTC, VTC, and VLPFC explicitly represent category, task, or some combination of both, and (2) in what way representations across these subdivisions of the extended ‘what’ pathway may differ. To fill these gaps in knowledge, we scanned 12 participants using fMRI to test the effect of category and task on distributed responses across LOTC, VTC, and VLPFC. Results reveal that task and category modulate responses in both high-level visual regions, as well as prefrontal cortex. However, we found fundamentally different types of representations across the brain. Distributed responses in high-level visual regions are more strongly driven by category than task, and exhibit task-independent category representations. In contrast, distributed responses in prefrontal cortex are more strongly driven by task than category, and contain task-dependent category representations. Together, these findings of differential representations across the brain support a new idea that LOTC and VTC maintain stable category representations allowing efficient processing of visual information, while prefrontal cortex contains flexible representations in which category information may emerge only when relevant to the task.

Introduction

The human ventral stream, also referred to as the “what” pathway, extends from occipital cortex to high-level visual regions in lateral occipito-temporal cortex (LOTC) and ventral temporal cortex (VTC), and is involved in object recognition and categorization (Goodale and Milner, 1992; Grill-Spector and Weiner, 2014; Mishkin and Ungerleider, 1982). LOTC and VTC have both clustered (Cohen et al., 2000; Downing et al., 2001; Epstein and Kanwisher, 1998; Kanwisher, 1997) and distributed representations of object categories (Grill-Spector and Weiner, 2014; Haxby et al., 2001), which are organizational features believed to contribute to the speed of categorization (Thorpe et al., 1996). Further, a large body of research has documented that distributed responses across LOTC and VTC contain linearly separable category representations (Cox and Savoy, 2003; Grill-Spector and Weiner, 2014; Jacques et al., 2016; Kravitz et al., 2011; Kriegeskorte et al., 2008; Sayres and Grill-Spector, 2008; Spiridon and Kanwisher, 2002; Walther et al., 2009; Weiner and Grill-Spector, 2010). Linear separability is computationally advantageous because it enables combinatorial coding of distributed responses resulting in representations of tens of thousands of categories (Haxby et al., 2001, 2011), as well as rapid and efficient readout of category information (DiCarlo and Cox, 2007; Grill-Spector and Weiner, 2014).

The “what” pathway is thought to extend from VTC to ventro-lateral prefrontal cortex (VLPFC) for two reasons. First, VTC is anatomically connected to VLPFC via white matter connections such as the uncinate fasciculus (Catani et al., 2002; De Schotten et al., 2012; Gerbella et al., 2010; Schmahmann and Pandya, 2006; Ungerleider et al., 1989). Second, VLPFC is sensitive to the content of visual stimuli, rather than the location of visual stimuli, particularly under selective attention (Baldauf and Desimone, 2014; Çukur et al., 2013; Harel et al., 2014; Peelen et al., 2009; Rushworth et al., 2005) and working memory tasks (Courtney et al., 1997; Goldman-Rakic, 1996; Muller et al., 2002; Volle et al., 2008; Demb et al., 1995). Additional evidence suggests that “what” responses related to stimulus attributes emerge in VLPFC only when they are task-relevant (Asaad et al., 2000; Johnston and Everling, 2006; Lee and Baker, 2016; McKee et al., 2014; Miller and Cohen, 2000; Waskom et al., 2014). Thus, while VLPFC is considered part of...
Since a key computational goal of the ventral “what” pathway is categorization (Haxby et al., 2001; Kiani et al., 2007; Kriegeskorte et al., 2008; Grill-Spector and Weiner 2014), it is natural to ask whether VLPFC responses also contain category representations. Previous studies provide some evidence for visual category representations in VLPFC. Indeed, both electrophysiology experiments in macaques and fMRI measurements in humans observed face-selective responses (Scalaidhe et al. 1997; Tsao et al., 2008; Chan, 2013), as well as representation of categorical boundaries between pairs of stimuli (Freedman et al., 2001; Miller et al., 2002; Jiang et al., 2007; Meyer et al., 2011). For example, Jiang et al. (2007) showed that human VLPFC responses are sensitive to categorical boundaries, rather than physical differences among exemplars of the category. However, the nature of distributed responses in VLPFC has not been extensively investigated. Thus, it remains unknown if distributed responses in VLPFC contain category representations for multiple categories, and if these representations are robust and linearly separable as found in high-level visual cortex.

Notably, research in both human and non-human primates suggests that category-selective responses in VLPFC may be task-dependent (Jiang et al., 2007; McKee et al., 2014; Rigotti et al., 2013; Romanski, 2004). For example, McKee et al. (2014) reported that category-selective responses in VLPFC were observed only when category distinctions were task-relevant, but not during passive viewing. Additionally, researchers have observed that selective attention to categories alters distributed responses in VLPFC (Çukur et al., 2013; Peelen et al., 2009). However, prior studies differ on how task and attention affect category-selective responses in high-level visual cortex. While some studies report task independent category-selective responses (McKee et al. 2014, Jiang et al. 2007), other studies indicate that selective attention can alter distributed responses in high-level visual cortex (Çukur et al., 2013; Peelen et al., 2009). Thus, it remains unresolved whether task effects are equivalent or different from the effects of selective attention, and if these effects are region-specific or common across the extended “what” pathway.

Given these unresolved issues, we asked: (1) Do distributed responses across LOTC, VTC, and VLPFC explicitly represent category, task, or some combination of both? (2) Do category representations differ across cortical regions constituting the extended “what” pathway? (3) Are category representations similarly affected by task and selective attention?

To address these questions, we measured distributed responses in LOTC, VTC, and VLPFC while participants viewed images from five object categories and performed three types of cognitive tasks: oddball (OB), working memory (WM), and selective attention (SA, Fig. 1). We chose these tasks as they (1) have been widely used in studies of high-level visual cortex and VLPFC, (2) vary in the cognitive operations required, and (3) vary in the relevancy of the stimulus to the task. We first investigated how task and visual category affect amplitudes of distributed responses across the extended “what” pathway. We then examined how task and selective attention affect distributed category-selective representations. Critically, for each of these analyses, we examined whether effects were uniform or differential across LOTC, VTC, and VLPFC.

**Fig. 1. Experimental design.** (A) Example images from each category used in the experiment. This base set of images was used for all experiments. Superimposed stimuli from these images were generated for the selective attention (SA) task, see Fig. 1D. (B–D) Example stimuli for a block for each task in the experiment. Block lengths were equated for the number and duration of stimuli. (B) Oddball (OB): subjects indicated via a button when a phase-scrambled image appeared. (C) Working memory (WM): subjects indicated when a stimulus repeated after an intervening stimulus. (D) Selective Attention (SA): subjects viewed superimposed stimuli of two categories (in this example faces and cars) and were asked to attend to one category (as indicated by the cue preceding each block). Subjects indicated when a stimulus of the attended category was presented upside down.
Materials and methods

Participants

Twelve subjects (5 female, ages 23–44) from Stanford University participated in three experiments. All subjects were right-handed and had normal or corrected-to-normal vision. Participants gave written informed consent. All procedures were approved by the Stanford Internal Review Board on Human Subjects Research.

Experimental design

Stimuli

In all experiments, subjects viewed grayscale images of stimuli from five categories: faces, bodies, cars, houses, and pseudowords (Fig. 1A, example stimuli). We chose these categories because we wanted to include a wide range of categories that have also been widely used in the field to examine distributed responses to categories in high-level visual cortex (e.g. Haxby et al., 2001; Spiridon 2005; Peelen and Downing, 2007; Op De Beeck, 2008; Stligiani, Weiner and Grill-Spector, 2016). Face stimuli consisted of male and female faces presented at various poses and positions. Cars were of various makes and models, presented at different views, sizes, and positions. House stimuli consisted of San Francisco houses and buildings. Finally, pronounceable pseudowords (courtesy of Laurie Glezer, see full description in Glezer et al. (2009)) were presented with different textures and viewing orientations to match the variable viewing conditions of other categories. A total of 1440 pictures composed this base stimulus set, which is a subset of the image set used in (Stiglani et al., 2015).

In order to minimize low-level difference among categories, images were embedded in a 10.5° phase-scrambled background generated from a different randomly selected image from the set. The retinal position of the image was jittered around fixation. All tasks used the same base stimulus set.

Tasks and design

Each experiment contained images from the same five stimulus categories, but the task varied for each experiment. The tasks were: oddball, working memory, and selective attention. Task order was counterbalanced across participants. In all tasks, images were presented in 8 s blocks, in which 8 images of a single category were displayed at a rate of 1 Hz. Blocks containing images were interleaved with blocks of a blank screen (1/3 of blocks; 8 s each). Block order was counter-balanced across categories. Subjects were instructed to fixate on a central dot and perform one of the following tasks:

**Oddball (OB) task** (Fig. 1B): Subjects viewed the images while fixating and responded via button press when a phase-scrambled image without an object appeared. 0, 1, or 2 phase-scrambled targets occurred randomly in each block. This task has the lowest attentional demands and depth of encoding, as the performance of the task does not depend on the presented category. We chose this task as this task keeps participants engaged during the experiment and is frequently used in studies of high-level visual cortex (Gauthier, 2000; Glezer et al., 2009; Haxby et al., 2000; Kanwisher et al., 1997; Stligiani et al., 2015; Weiner and Grill-Spector, 2010; Weiner and Grill-Spector, 2011).

**Working Memory (WM) task** (Fig. 1C): Subjects viewed images while fixating and responded via button press when an image repeated after an intervening stimulus (2-back task). 0, 1, or 2 repeats occurred randomly in each block. This task not only requires active engagement of working memory mechanisms, but also encoding of specific exemplars of a category. WM tasks are also frequently used in studies of high-level vision (e.g. Golari et al., 2007; Kanwisher et al., 1997), and WM of exemplar identity is thought to engage the VLPFC (Goldman-Rakic, 1996; Muller et al., 2002; Volle et al., 2008).

Selective Attention (SA) task (Fig. 1D): Each image consisted of superimposed exemplars from two categories. Each block contained a superimposition of images of the same two categories (e.g. faces and cars). Across blocks, all possible pairings of the 5 categories occurred. Before each block, a cue indicating the name of a category appeared for 1 s. The cue indicated which category was to be attended during the block. Subjects were instructed to selectively attend to the cued category and report when an item of that category was presented upside down. 0, 1, or 2 images of either the attended or unattended category were presented upside down at random in each block. While the SA task requires active filtering of category-irrelevant information in images containing superimposed items from two categories, it does not require deep encoding of specific exemplars or maintaining items in memory. We chose this task as selective attention is thought to involve the VLPFC (Baldauf and Desimone, 2014) and prior research of high-level visual cortex suggests that selective attention to attributes of the visual stimuli alters responses in regions involved in processing of the stimuli (Cukur et al., 2013; Harel et al., 2014; Koida and Komatsu, 2007; Peelen et al., 2009).

Subjects performed three runs of each of the tasks. Different runs of the same task did not repeat any images. Within each task, image order presentation was randomized. Each run contained 40 blocks. Due to the inclusion of cues, runs of the SA task were 40 s longer than WM or OB, totaling 360 s for SA and 320 s for WM and OB. Task conditions were matched in the number of blocks and target trials.

MRI

Data acquisition

Subjects were scanned using a General Electric Sigma MR750 3T scanner located in the Center for Cognitive and Neurobiological Imaging (CNI) at Stanford University using a custom-built 32-channel head coil. Using an EPI sequence with level 3 multi-slice acceleration, we acquired 48 slices at 2.4 mm isotropic resolution, FOV=192 mm, TE=30 ms, TR=1 s, and flip angle=62°. The slice prescription covered the entire brain, except the very superior portion of the cortex, roughly corresponding to superior motor and somatosensory cortices. Whole-brain anatomical images were acquired in the same session using a T1 SPGR sequence with a resolution of 2.4 mm isotropic voxels, FOV=192 mm, flip angle=15°.

A whole brain high resolution MRI volume of each subject was acquired in a separate session. Scans were acquired at 1mm isotropic resolution, FOV=240 mm, flip angle=12°. Data from fMRI scans were aligned to this volume, which was used to create a cortical surface reconstruction of each subject's brain.

Software

fMRI data analysis was performed in MATLAB (www.mathworks.com) and using the mrVista analysis software (http://github.com/vistalah) developed at Stanford University. We used the hemodynamic response function as implemented in the Statistical Parametric Mapping (SPM) package (http://www.fil.ion.ucl.ac.uk/spm). Automated anatomical segmentation of white and gray matter was performed with FreeSurfer (http://freesurfer.net), and manual corrections were performed using ITKGray (http://web.stanford.edu/group/vista/cgi-bin/wiki/index.php/ItkGray). Group-level statistical analysis was done in the R statistical analysis software suite (https://www.r-project.org/) using RStudio (https://www.rstudio.com/) as well as statistical packages ‘afex’ (https://cran.r-project.org/web/packages/afex/index.html), ‘ez’ (https://cran.r-project.org/web/packages/ez/index.html), and ‘lmPerm’ (https://cran.r-project.org/web/packages/lmPerm/index.html). Plotting used functions from the ‘ggplot2’ library (http://ggplot2.org/).
Behavioral analysis

We estimated subjects’ performance across tasks and categories by calculating the hit rate, false alarm rate, and response times (RT) for each task. Due to false alarm rates being near floor (< 2.1% for all tasks), we only report the hit rate and RT. We used repeated-measures Friedman rank sum test to evaluate significant differences in performance across tasks. This was followed by a pairwise comparisons using Wilcoxon signed-rank test. Performance data from the oddball task from one subject was not collected due to technical error and that subject was omitted from the analysis of behavioral data.

Analysis of behavioral data during scan reveal that the oddball task was easier than the other tasks, as subjects were faster (oddball (OB): 551 ms ± 45 ms, mean ± standard deviation; working memory (WM): 604 ms ± 36 ms; selective attention (SA): 662 ms ± 20 ms) and more accurate (hit rate OB: 0.93 ± 0.06; WM: 0.69 ± 0.16; SA: 0.80 ± 0.1) in the OB task than the other two tasks. A Friedman rank-sum test revealed significant differences in hit rate (χ²(2)=9.45, p=0.009) across tasks. A Wilcoxon signed-rank test showed that hit rates were significantly higher during the OB task than WM and SA (Zs > 104, Ps < 0.003). Likewise, response times (RT) were fastest for the OB, and slowest for SA (Friedman rank-sum test, χ²(2)=18.182, p < 0.001). While response times were slower for SA than WM (Z=109, p = 0.001), accuracy was marginally higher for SA than WM (Z=90, p=0.0556). This may be a result of different response time-accuracy tradeoffs across these two tasks. Overall, analysis of behavioral data suggests that participants had faster reactions and higher accuracy in the OB task than the other two tasks and suggests that subjects likely used different cognitive resources or strategies across these tasks.

Anatomical data

Each T1 anatomy was segmented into gray and white matter using FreeSurfer automatic segmentation tools. The resulting segmentations were manually validated and corrected for artifacts and mislabeling using ITKsag. Anatomical and functional data were registered and aligned using mrVista alignment tools. For purposes of visualization and ROI tracing, 3D meshes of each subject’s brain were generated using mrMesh from the mrVista software suite (Wandell et al., 2000).

Meshes were inflated in order to expose the cortex within sulci.

Anatomical regions of interest (ROIs)

Anatomical ROIs were defined for each individual subject based on anatomical landmarks. The following ROIs were defined in each hemisphere: lateral occipito-temporal cortex (LOTC), ventral temporal cortex (VTC), and ventro-lateral prefrontal cortex (VLPFC). See Fig. 2 for illustration of these anatomical ROIs.

The boundaries of LOTC were defined in as in Weiner and Grill-Spector (2013; Fig. 2-top). Posterior: the convergence of the intraparietal sulcus (IPS) and the descending limb of the superior temporal sulcus (STS); anterior: anterior tip of the mid-fusiform sulcus (MFS; Weiner et al., 2014); superior: the dorsal lip of the STS; and inferior: the occipitotemporal sulcus (OTS). The sizes of the LOTC ROIs (mean ± standard error of the mean) were 879 ± 54.27 voxels for the left hemisphere, and 1084 ± 57.44 voxels for the right.

The boundaries for VTC were defined in as in Weiner and Grill-Spector (2013; Fig. 2-top). Posterior: posterior transverse collateral sulcus (ptCoS); anterior: anterior tip of the MFS; the STS; and inferior: the medial lip of the collateral sulcus (CoS). The sizes of the VTC ROIs were 1086 ± 69.28 voxels for the left hemisphere, and 1023 ± 43.87 voxels for the right.

The boundaries of VLPFC were defined in as in Levy and Wagner (2011; Fig. 1). Posterior: caudal lip of the precentral sulcus; anterior: anterior edge of pars orbitalis; superior: the dorsal lip of the inferior frontal sulcus (IFS); and inferior: the dorsal lip of the lateral sulcus. The sizes of the

VLPFC ROIs were 1394 ± 91.22 voxels for the left hemisphere, and 1373 ± 56.29 voxels for the right.

Functional data

All data were analyzed in the native brain space of each subject, without spatial smoothing. FMRI data were aligned to the whole brain anatomy of each subject. Automatic quality control algorithms employed in the CNR screened for excessive subject motion and scanner-related artifacts. Four scan sessions exhibited excessive subject motion (> 3 voxels movement) or scanner artifacts and were discarded. These subjects were scanned again at a later date using different images and task-order than the original sessions. All scans were motion corrected, both within-scan and between-scans, using mrVista motion correction algorithms. A 20 Hz high-pass temporal filter was applied to all time courses to remove slow undulations in scanner signal, including drift. The time series of each voxel was transformed from arbitrary scanner units to percentage signal change.

General Linear Model (GLM)

We convolved the design matrix of each experiment with the hemodynamic response function implemented in SPM to generate predictors of the BOLD response for each condition. A GLM was applied to fit each voxel’s data and estimate regularized response coefficients (betas) for each predictor. Betas are in units of percentage signal change. GLM betas were generated separately for each run to estimate the response for each possible category and task combination in each voxel.

Multivoxel pattern analysis (MVPA) of response amplitudes to visual categories and tasks

We examined the distributed patterns of neural activity elicited by each visual category and task by conducting a multivoxel pattern analysis (MVPA; Haxby et al., 2001) of responses in each anatomical ROI defined in each subject, as many studies have shown that distributed responses across LOTC and VTC contain linearly separable category information (Cox and Savoy, 2003; Jacques et al., 2016; Kravitz et al., 2011; Kriegeskorte et al., 2008; Sayres and Grill-Spector, 2008; Spiridon and Kanwisher, 2002; Walthier et al., 2009; Weiner and
Grill-Spector, 2010). For every voxel, we estimated the response to each condition and task from the GLM of each run (conditions in the SA task are labeled by the attended category). The relative response to each condition was calculated by subtracting the mean response across all conditions. This procedure reduces between-voxel and between-region differences in the mean response amplitude (see Sayres and Grill-Spector (2008)). Multivoxel patterns (MVPs) - vectors of responses across all voxels in an anatomical ROI – were generated for each category and task, and separately for each run.

For each subject, and for each anatomical ROI, we calculated correlations between pairs of MVPs to a category and task combination in each run. This was done for all unique pairwise combination of runs, then averaged across combinations. This resulted in 15×15 representational similarity matrix (RSM) for each subject and ROI. Each cell in the matrix represents the average correlation between two MVPs averaged across all runs as described above. We then averaged these matrices across subjects for each hemisphere and anatomical ROI, resulting in the RSMs shown in Fig. 3B–G.

The average RSM across subjects for each ROI were converted into a dissimilarity matrix that was used for hierarchical clustering (1-r of the RSMs). As there were no significant hemispherical differences in the RSMs, the matrices were averaged across hemispheres prior to conversion. Hierarchical clustering was run on the dissimilarity matrix using the ‘hclust’ method in R and visualized as dendrograms in Fig. 3H.

We used ideal representational similarity matrices (shown in Fig. 3A) as factors in a multiple linear regression analysis to assess the size and significance of the category and task factors used to explain the measured RSMs in each anatomical ROI. The regression was done for each subject, hemisphere, and anatomical ROI, generating the model fit and predictor coefficients per regression model. Results were then averaged across subjects for each anatomical ROI in each hemisphere. Data are shown in Fig. 4.

Statistical analysis
Repeated-measures ANOVAs were used to determine statistical significance of effects throughout the paper. For all ANOVAs in this paper, the assumptions of sphericity are tested. If sphericity was violated, we report Greenhouse-Geisser corrected values. Assumption of normality was tested on model residuals after accounting for all effects. If violations of normality were found, factorial permutation testing (using 10000 permutations) was used to assess the significance of effects, as permutation testing makes no assumption about the distribution of the data.

Assessing the null distribution of the regression analysis
To test the statistical significance of the beta values from the regression analysis, we first evaluated the null distribution using a nonparametric permutation test. For each subject, hemisphere, and anatomical ROI, we permuted the RSM in two ways, one to estimate the null distribution of the category regressor, and the second to estimate the null distribution of task regressor. For category, we permuted the RSM cells within a given task, e.g. permuted category cells within the WM condition. For task, we permuted the RSM cells between tasks given the same category condition combination, e.g. permuted the cells between OB, WM, and SA for any particular combination of category and condition. This permutation was performed 1000 times for each subject, resulting in 12000 iterations for each anatomical ROI across subjects. Each permuted RSM was subjected to the regression analysis and the betas of task and category were derived for the permuted RSM. For each anatomical ROI, the average betas for the permuted task and category effect were estimated from the 12000 iterations.

Assessing the source of similarity among distributed responses during the selective attention task
During the SA task, superimposed stimuli from two categories are displayed at any moment, simultaneously driving visual responses for both categories. While the RSM in Fig. 3 is sorted by the attended category in the SA task, all other conditions also contained blocks in which that category is presented but unattended (25% of each condition). As MVPs of an image containing two categories may be a sum of the MVPs of each category presented in isolation (MacEvo and Epstein, 2011), it is possible that the elevated similarity among all MVPs during the SA task is driven by the similarity between MVPs containing a common category irrespective of attention.

To test this hypothesis, we generated MVPs for all trial types in the SA condition. As before, we subtracted the mean response across all SA conditions from each voxel, such that the MVP reflects the response relative to the mean response of the voxel during the SA task. This results in a 20×20 RSM per anatomical ROI and hemisphere, where each condition reflects the MVP for each attended and unattended stimulus pair (e.g. face attended, body unattended). We then proceeded to calculate the correlations between MVPs based on the number of shared categories. For example, 2 shared categories: e.g. face attended + body unattended correlated with body attended + face unattended; 1 shared category: e.g., car attended + body unattended correlated with house attended + car unattended; and 0 shared categories: e.g. word attended + house unattended correlated with face attended + car unattended. We calculated the mean correlation across these combinations of common categories for each subject, anatomical ROI, and hemisphere, and then averaged across subjects. Fig. 5 shows the results of this analysis averaged across hemispheres (as hemispheric differences were not significant).

Reproducibility of category selectivity within and across tasks
To evaluate whether category selectivity was stable within and across tasks, we generated maps of category selectivity separately for each task, as well as for attentional state within the selective attention task (category-attended vs. category unattended).

While the exemplars and the categories are the same across experiments, the visual presentation varies across tasks. To derive the contribution of task and stimulus to distributed category-selective responses in an unbiased way, we analyzed the data using only conditions that had identical visual content. This was done in two ways depending on the task. In the OB and WM, the stimuli in each trial were identical except for the task stimuli: subjects responded to the occurrence of phase scrambled images in the OB task, and to repeated images in the WM task. Thus, to equate stimuli across these tasks, we excluded only the image trials that contained the phase scrambled/repeated images since these were not shared between the two tasks. These were coded as a nuisance regressor in the GLM. This resulted in removal of approximately 8% of the trials – or 264 out 3000 trials across all experiments (88 from each). In the SA experiment, the stimuli were identical across attended and unattended conditions. Here we excluded the images in which subjects made a behavioral response to equate both the stimulus and response across attended and unattended conditions.

Within task reproducibility (OB/WM): We calculated the mean correlation between category selectivity maps for each category using all pairings of independent runs for a particular task. For each subject, we then calculated the mean reproducibility across categories for each anatomical ROI in both hemispheres. Data were then averaged across hemispheres, as there were no significant differences between hemispheres. The overall within-task reproducibility was calculated for each ROI and subject by averaging the reproducibility across tasks. Fig. 7A shows the average reproducibility across subjects across both tasks. Fig. 7B shows the within-task reproducibility of each task alone. Mauchly’s test of sphericity revealed that the assumption of sphericity has been violated (W=0.53, p=0.04), thus Greenhouse-Geisser corrected values are reported.

Between-task reproducibility (OB/WM): We calculated the mean correlation among category selectivity maps for each category across
OB and WM tasks using all pairings of runs. In each subject and ROI, data were averaged across categories and hemispheres. Results in Fig. 7A show the average correlation across subjects. We used a repeated-measures ANOVA with factors of ROI and task to test if reproducibility varies across task and ROIs. Mauchly’s test of sphericity revealed that the assumption of sphericity has been violated (W=0.232, p < 0.001), thus Greenhouse-Geisser corrected values are reported.

Reproducibility of unattended and attended categories in the SA task:
For each attentional state (attended/unattended), we calculated the mean correlation between category selectivity maps for each category using all pairings of independent runs of the SA task on matching stimuli. For each subject, we then calculated the mean reproducibility across categories for each anatomical ROI in both hemispheres. Data were then averaged across hemispheres, as there were no significant differences between hemispheres. Results in Fig. 8A show the average reproducibility (correlation) across categories, subjects, and hemispheres for the attended and unattended conditions. We used a repeated-measures ANOVA with factors of ROI and attentional state to test if reproducibility varies across attentional conditions and ROIs.

Classification of categories during attended and unattended conditions in the SA tasks

To test how attention affects category information we used a winner-take-all classifier (WTA, Sayres and Grill-Spector, 2008; Weiner and Grill-Spector, 2010) to test if the category can be decoded from the functional selectivity map from a given ROI within and across attention conditions. Specifically, we used a WTA classifier trained on one run to test how well it classified the other two runs, either within an attention condition (training and testing the same attention condition) or across attention conditions (training and testing different attention conditions). Classification was done on all permutations of training and testing runs. Fig. 8B shows the results of this analysis averaged across permutations, categories, hemispheres, and subjects. We employed a repeated-measures ANOVA with factors of anatomical ROI, classification type (within/between), task, and hemisphere to determine if classification performance differed significantly across any of these factors. Mauchly’s test of sphericity revealed that the assumption of sphericity has been violated (W=0.232, p < 0.001), thus Greenhouse-Geisser corrected values are reported.

Classification performance of visual category from distributed responses across tasks

Finally, we used the full set of data from all three tasks regardless of differences in visual stimuli, to test for category information content. For the SA task, we considered the category-attended condition. We reasoned that if we can decode category information not only within a task, but also across tasks, and despite differences in visual presenta-
tion, it will provide strong support for the robustness of category information despite changes in task context and stimuli.

For each subject and ROI, we performed the following process. Of the 3 runs available for each task, we used one run as a training set and the other two as test sets. The WTA classifier determines the category based on the highest correlation between the test category selectivity map and each of the trained category selectivity maps. The process is repeated with all runs being used as either training or test sets, then averaged across runs and categories. As 5 categories were tested, the chance level is 20%. We separately calculated the WTA classifier performance when training and testing category selectivity maps within the same task and from different tasks. Fig. 9 shows the results of this analysis averaged across categories and subjects. As both violations of sphericity and normality were detected, data was tested for statistical significance using permutation testing and non-parametric rank test.

**Results**

**How do task and visual category affect distributed responses in LOTC, VTC, and VLPFC?**

To address this question, we examined the information in distributed multivoxel patterns (MVPs) of response amplitudes across VTC, LOTC, and VLPFC for each of the five visual categories and the three cognitive tasks. To test the relationship among MVPs for these 15 experimental conditions, we generated representational similarity matrices (RSMs) for each anatomical ROI within each subject. Each cell in the RSM indicates the similarity (Pearson correlation) between a pair of MVPs for a specific category and task. We reasoned that if distributed responses are primarily driven by category, MVPs of the same category under different tasks would be similar while MVPs of different categories in the same task would be dissimilar, resulting in the ideal RSM illustrated in Fig. 3A-top. Alternatively, if distributed responses are primarily driven by task, MVPs of different categories under the same task would be similar, while MVPs of the same category under different tasks would be dissimilar, resulting in the RSM illustrated in Fig. 3A-bottom.

Fig. 3 shows the average RSM across 12 subjects for LOTC, VTC, and VLPFC. In both LOTC (Fig. 3B/E) and VTC (Fig. 3C/F) the main diagonal shows positive correlations, indicating that MVPs for different categories in the same task are positively correlated, but MVPs to different categories in the same task are zero or negatively correlated (Fig. 3D/G). The representational structure of distributed responses in LOTC, VTC, and VLPFC is also illustrated in the dendrograms generated from hierarchical clustering (Fig. 3H).

The relative contribution of task and visual category was estimated via a multiple linear regression analysis of RSMs from individual subjects using the ideal RSMs for category and task as predictors. Summary results are shown in Fig. 4A. Across subjects, the variance explained by the ideal model with factors of task and category was higher than 45% for all subjects, across all ROIs (LOTC: 53% ± 2%, mean ± standard error; VTC: 50% ± 2%; VLPFC: 48% ± 2%) and the contribution of both task and category predictors was statistically significant (ps < .001). In VLPFC, the effect of task was larger than the effect of category (Fig. 4A), consistent with the results illustrated by the average RSMs in Fig. 3. However, in LOTC and VTC the contribution of task and category were surprisingly similar. An examination of the LOTC and VTC average RSMs in Fig. 3 suggests that the task effect may be driven by the higher similarity among MVPs of different categories during the selective attention (SA) task, but not during the working memory (WM) or oddball (OB) tasks.

**Is the putative task effect during SA due to visual presentation of superimposed stimuli from two categories?**

During the SA task (Fig. 1D), each block contains images from two superimposed categories. The MVPs during the SA in the analyses above were labeled by the attended category. However, if the MVP of response to superimposed images is the average of MVPs of responses to each category in isolation, as suggested by MacEvoy and Epstein (2011), then the positive correlations between MVPs during the SA task might be driven by the common categories occurring across conditions. In other words, if conditions containing the same categories under different attention conditions generate similar distributed responses (e.g. the MVP of face attended + body unattended may be similar to the MVP of body attended + face unattended), their positive correlation would result in overall increase in similarity between MVPs during the SA task compared to the other tasks, in turn, inflating the apparent task effect in the RSM.

To test this hypothesis, we analyzed the similarity among MVPs when they contained 2 shared categories in different attention conditions (e.g. car attended + body unattended vs. car unattended + body attended), 1 shared category (e.g. car attended + body unattended vs. car unattended + face attended), or 0 shared categories (e.g. car attended + body unattended vs. face attended + words unattended). As predicted by our hypothesis, in bilateral LOTC and VTC, the correlations between MVPs containing 2 shared categories in different attention conditions were significantly positive, zero when they contained 1 common category, and negative when there were no shared categories (Fig. 5). Conversely, in VLPFC the correlations among MVPs containing 2, 1, or 0 shared categories were all negative (Fig. 5), indicating that MVPs containing overlapping categories were dissimilar when the attended category was different. Differences across ROIs in the correlation between MVPs as a function of shared categories were statistically significant (ROI by number of shared categories interac-
Largely driven by task, not category.

MVPs in LOTC and VTC are driven by both category and task, with significantly larger category than task effect. The task effect was nearly double that of the category effect, by a factor of about 2. The category effect was greater than the task effect, by a factor of about 4.47, p < 0.001, permutation testing). Importantly, in both LOTC and VTC, we find a significant effect of number of shared categories (p < 0.001, permutation testing), suggesting that the positive correlations among SA MVPs are driven by blocks containing the same categories, as well as in different attention conditions. Considering these results, the apparent task effect on correlations among MVPs in LOTC and VTC during the SA task is likely driven by visual responses to the shared categories.

Reevaluating the contribution of task and visual category to distributed responses across anatomical regions

To obtain an unbiased estimate of task and category effects, we (1) repeated the multiple linear regression analysis of LOTC, VTC, and VLPFC RSMs on the OB and WM tasks, excluding the SA task and (2) matching the visual stimuli across WM and OB tasks by excluding task trials (see Materials and Methods). The variance explained by the regression model increased when considering these two matched tasks as compared to all three tasks, and particularly so in the VLPFC (LOTC: 55% ± 2%, mean ± standard error; VTC: 50% ± 3%; VLPFC: 65% ± 3%). A repeated-measures ANOVA on the multiple regression weights (β), with anatomical ROI (LOTC, VTC, VLPFC) and effect type (task/category) as factors, revealed a significant anatomical ROI by effect type interaction (F(2, 44)=33.72, p < 0.001), indicating that the contribution of task and category varies across anatomical ROIs. Notably, results show a significantly larger category than task effect in LOTC and VTC (Fig. 4B; t(22) > 4.47, p < 0.001), whereby the category effect was nearly double that of the task effect. In VLPFC, we observed an opposite pattern in which the task effect was significantly larger than the category effect, by a factor of about five (Fig. 4B; t(22)=6.08, p < 0.001). Together, these results reveal that (1) MVPs in LOTC and VTC are driven by both category and task, with category effects larger than tasks effects, and (2) MVPs in VLPFC are largely driven by task, not category.

Does task or attention affect the spatial topology of category selectivity maps?

Our results show that response amplitudes are independently driven by both category and task, with differing contributions across cortex. However, it remains to be determined how distributed patterns of category selectivity are affected by task and attention. This question is important for two reasons. First, it is unknown if task or attention not only modulate response amplitudes, but also the relative responses across categories, which would affect the spatial topography of category selectivity across cortex. Second, it is unknown if the task-relevancy of the stimulus affects distributed category selectivity.

Therefore, we examined distributed category selectivity for the stimulus matched oddball (OB) and working memory (WM) as well as for the attended and unattended conditions of each category appearance in the SA task (SA-A and SA-UA, respectively). In each pairing, stimuli were identical, and in one condition the category was relevant for the task (WM; SA-A), and in the other it was irrelevant for the task (OB; category-unattended SA-U). Category selectivity was defined in each voxel as the t-value indicating the degree to which its response to one category was higher compared to all other categories (see Materials and Methods). Given the widespread use of category selectivity to both define and characterize functional regions, we asked: (1) Are the spatial patterns of category selectivity reproducible across tasks and attention conditions? (2) Do task or attention alter the distributed patterns of category selectivity across anatomical ROIs?

Qualitative examination of category selectivity maps revealed similar patterns in LOTC and VTC across tasks. In the example subject shown in Fig. 6, the topography of face selectivity in LOTC is similar across conditions, and especially between OB, WM, and SA-A: voxels showing preference for faces (hot colors) are consistently found in the ventral and posterior aspect of the map and voxels preferring nonfaces (cool colors) are located more superiorly. Similarly, in VTC, the lateral fusiform gyrus shows consistent face selectivity across tasks, and medial VTC shows preference for non-faces (Fig. 6-middle). Note that in both LOTC and VTC, the SA-U seems to display lowered category selectivity and slightly different topology compared to the other three conditions. In contrast, qualitative inspection of VLPFC (Fig. 6-bottom) revealed considerably more variability of face selectivity across tasks and attentional states, both in terms of the magnitude of face selectivity and in the location of the voxels illustrating the highest selectivity.

Does task affect the reproducibility of category selectivity maps?

To quantify these observations across subjects, we evaluated the reproducibility of each of the five category selectivity maps for faces, bodies, cars, houses, and pseudowords across tasks and attention conditions. We calculated the correlation between the selectivity map for each category across independent runs, within and across OB and WM tasks, as well as within and across attended and unattended conditions in the SA task. These pairwise comparisons enable us to compare the effects of task and attention, respectively, on identical stimuli in each paired comparison. We reasoned that if the task (or attention) changes the spatial patterns of category selectivity, within-task (within-attention) correlations will be positive and higher than across-task (across-attention) correlations. Fig. 7A summarizes the average correlations among category selectivity maps for OB and WM tasks across LOTC, VTC, and VLPFC, illustrating two main findings. First, correlations of category selectivity maps were significantly positive both within and across tasks in all three ROI (ts(11) > 6.1; ps < 0.001, Fig. 7A), indicating that, in general, category selectivity is reproducible. Second, the reproducibility of category selectivity significantly varied across anatomical ROIs. A 2-way repeated-measures ANOVA with factors of ROI (LOTC/VTC/VLPFC) and comparison type (within-task/between-task), revealed a significant main effect of ROI (F(1.94, 21.38)=160.59, p < 0.001;
Fig. 6. Example face selectivity maps across tasks in a representative subject. Each panel shows a zoomed view on an inflated right hemisphere, where each row is cropped around an anatomical ROI. From left to right are the different cognitive tasks under which each selectivity map was produced, as well as the two attentional states within SA. OB: Oddball; WM: working memory; SA-A: selective attention, category-attended, and SA-UA: selective attention, category unattended. Top: face selectivity in LOTC. Middle: face selectivity in VTC. Bottom: face selectivity in VLPFC. Color bar indicates degree of face selectivity (t-value of contrast faces > non-faces) in each voxel.

Fig. 7A) and no other significant main effects or interactions. Consistent with the single subject example shown in Fig. 6, category selectivity maps in LOTC and VTC were significantly more reproducible both within and across OB and WM tasks than those in VLPFC (F(22) > 9.52, p < 0.001). The difference among the anatomical ROIs is striking. Correlations between category selectivity maps in LOTC and VTC using different stimuli across runs and different tasks were higher than 0.4 (Mean ± SD averaged across both within and between task; LOTC: 0.46 ± 0.093; VTC: 0.41 ± 0.1), whereas in VLPFC they were below 0.1 (VLPFC: 0.11 ± 0.045).

We next examined if the within-task reproducibility of category selectivity maps varied across the OB and WM tasks. We calculated the reproducibility of category selectivity maps separately for each task. Results of this analysis reveal three findings, shown in Fig. 7B. First, as in the prior analysis, within-task reproducibility of category selectivity maps varied across ROIs (repeated-measures ANOVA with anatomical ROI and task as factors, F(1.99, 21.87) = 176.78, p < 0.001), with LOTC and VTC yielding higher reproducibility than VLPFC. Second, there was no significant main effect of task across anatomical ROIs. Third, we found a significant interaction between ROI and task (F(1.36, 14.97) = 13.25, p = 0.002). In LOTC and VTC, reproducibility of category selectivity was similar during WM compared to OB. However, in VLPFC the reproducibility of category selectivity maps was significantly higher for WM compared to OB (t(22) > 4.33, p < 0.001). Further, during the OB task, the reproducibility of category selectivity maps in VLPFC was not significantly different than zero. This suggests that in VLPFC, category representations are reproducible when the category content of the stimulus is task-relevant.

Does attention affect the reproducibility and information of category selectivity maps?

We then calculated the reproducibility of category selectivity maps under the category-attended and the category-unattended conditions during the SA task. Even though stimuli are identical and only attention conditions varied, results show that different from task, selective attention has a profound effect on reproducibility of category selectivity in all regions. Specifically, in all ROIs, reproducibility of category selectivity maps was significantly higher when the category was attended than when it was not (main effect of attention, 2-way repeated-measures ANOVA with factors of anatomical ROI and attention condition, F(1, 11) = 86.18, p < 0.001, Fig. 8A). Notably, reproducibility of category selectivity varied across ROIs (main effect of anatomical ROI, F(2, 22) = 118.34, p < 0.001) and attention differentially affected category selectivity across ROIs (anatomical ROI by attention interaction, F(2, 66) = 35.93, p < 0.001, Fig. 8A). In LOTC and VTC, category selectivity maps were positive and reproducible both when the category was attended and when it was unattended (correlation significantly different than zero, ts(11) > 8.86, p < 0.001). By contrast, in VLPFC, category selectivity maps were reproducible only when the category was attended (r(11) = 0.64, p < 0.001), but not when the category was unattended (r(11) = 0.4, p = 0.69).

To further examine the effect of selective attention on category information, we used a winner-take-all (WTA) classifier to test whether we can decode the category from distributed responses within and across attention conditions. Classification results show that selective
attention has a substantial effect on category information whereby classification performance was significantly lower when either training or testing on unattended categories (Fig. 8A). Indeed, a 3-way repeated-measures ANOVA with factors of attention during training (attended/unattended), attention during testing (attended/unattended), and ROI (LOTC, VTC, VLPFC) revealed both main effect of attentional state during training ($F(1,11)=17.87, p=0.001$) and main effect of attention during testing ($F(1,11)=14.27, p=0.003$). Overall, the highest performance was when training and testing used the attended category, and was poorest when training and testing used the unattended category. As in the reproducibility analyses, category information varied across ROIs: classification was significantly higher in LOTC and VTC than VLPFC (main effect of ROI, $F(1,76,19.32)=177.222, p<0.001$). In particular, in LOTC and VTC, classification was above 70% and significantly above chance in all conditions, even when both training and testing used the unattended category ($t(11)>12.87; ps<0.001$). In contrast, in VLPFC, there was significant above chance classification ($t(11)=6.99; p<0.001$) only when both the training and testing used the category attended condition. That is, when the category was unattended, either during training or testing, category information could not be classified from VLPFC (not significantly different than chance, $t(23)<1.26; ps>0.23$). Together these results suggest that category information is available both when it is attended and unattended in high-level visual cortex, but category information in VLPFC is present only when it is attended.

How does task affect category information content?

The prior analyses reveal that classification performance was more robust to changing attention conditions than reproducibility of category selectivity. This suggests that classification performance may be a more direct metric to determine category information across tasks. Therefore, as a final analysis, we examined if category information changed across tasks by evaluating classification performance within and across all three tasks, finding three main results.

First, classification within and between tasks was significantly above chance in all ROIs and conditions (Wilcoxon rank-sum tests; Zs > 2.8, ps < 0.002). Second, classification performance significantly varied across anatomical ROIs (main effect, $p<0.001$, factorial permutation testing), in which average classification performance was higher in LOTC and VTC compared to VLPFC. Strikingly, in both LOTC and VTC, classification performance was above 85% irrespective of training task and classification type (LOTC: 93% ± 1.3% within task, 86% ± 2.2% between tasks, mean ± standard error; VTC: 94% ± 1.4% within, 91% ± 1.3% between, Fig. 9). However, in VLPFC, the mean classification performance was substantially lower (averaging 52% ± 2.7% within task and 45% ± 2.8% between tasks, Fig. 9). Third, classification performance showed a significant three-way interaction between ROI, training task, and classification type ($p=0.029$; permutation testing). Specifically, in LOTC and VTC, both the training task and classification type had minimal effect on classification performance (< 7.5% across all pairwise comparisons), while in VLPFC, both training task and classification type had a substantial effect on classification performance (10–18%). Finally, in VLPFC, we observed that when training with the WM and SA tasks, between-task classification was lower than within-task classification, but when training with the OB task, between-task classification was higher than within-task classification. Since the OB yielded the lowest within-task classification (Zs > 2.84, ps < 0.002), we further examined if inclusion of this task generated the drop in classification performance between tasks. Results from a control analysis, using only WM and SA tasks, indicate that between-task classification in VLPFC was still lower than within-task classification (between-task: 45.4% ± 3% mean ± SEM; within-task: 60.8% ± 3%). Together, these analyses suggest that the changing nature of VLPFC category information across tasks is not merely driven by differences in overall performance.

Discussion

In the present study, we measured distributed responses in the extended “what” pathway including high-level visual cortex (LOTC and VTC) and prefrontal cortex (VLPFC) while participants viewed exemplars from multiple categories and performed different tasks. Our results revealed four main findings. First, the amplitudes of distributed responses throughout the extended “what” pathway are driven by both visual category and cognitive task. Second, distributed responses in high-level visual cortex are primarily driven by category, while distributed responses in VLPFC are dominated by task. Third, category-selective representations are consistent within and across tasks in high-level visual cortex, but vary across tasks in VLPFC. Fourth, selective attention affects distributed category representations both in high-level visual cortex and VLPFC. In the sections below, we first elaborate on these findings and then propose that efficient visual categorization may be achieved by a combination of stable and flexible aspects of distributed representations within the extended “what” pathway.
Our findings reveal differential effects of task and category across high-level visual and prefrontal cortices: representations in LOTC and VTC are largely task-independent, while those in VLPFC are task-dependent. These findings are consistent with accumulating evidence that representations within VLPFC are more sensitive to changes in stimulus compared to high-level visual cortex (Jiang et al., 2007; Lee and Baker, 2016; Bracci et al., 2017). Consistent findings of task sensitivity in VLPFC are likely due to the fact that neural responses within VLPFC are not only sensitive to visual stimuli, but also to a combination of factors including task rules and response types (Rigotti et al., 2013).

Concomitant with this idea, our data suggest that category representations in VLPFC are drastically impacted by the relevance of the category for the task. For example, VLPFC showed the lowest reproducibility of category selectivity and approached chance classification performance during (1) the oddball task, which required no category information and (2) when categories were task-irrelevant and unattended during the selective attention task. In contrast, VLPFC showed reproducible category selectivity and high classification performance during (1) the working memory task, which required fine-grained categorical representations at the exemplar-level and (2) attended conditions during the selective attention task. Together these data suggest that VLPFC may allocate neural resources to maintain category information only when it is task-relevant. This interpretation is supported by prior empirical evidence showing that (1) neurons in macaque VLPFC show selective responses when category information is task-relevant (Asaad et al., 2000; Freedman et al., 2001; McKee et al., 2014) and (2) neural responses in human VLPFC are modulated by task-relevancy (Jiang, 2007; Lee and Baker, 2016; Bracci et al., 2017).

While we did not observe task-independent distributed category representations in VLPFC, prior studies have shown evidence for task-independent category responses at the level of single neurons in VLPFC (Freedman et al., 2003; Meyer et al., 2011; Miller et al., 2002). Disparities across studies may be due to resolution differences between fMRI measurements and single neuron electrophysiology. For example, if task-independent neurons are sparse and scattered throughout VLPFC, they may not modulate distributed responses sufficiently to consistently be captured using fMRI. Future experiments combining single neuron recordings and fMRI (e.g., Dubois et al., 2015) can test if (1) different kinds of category information in VLPFC are present at different spatial scales and (2) task-dependent category-selective neurons are more abundant or more clustered than task-independent neurons in VLPFC.

Implications for understanding category representations in human prefrontal cortex

We interpret the task-dependent responses of VLPFC as a type of flexibility, which may function to accomplish the goals of a cognitive task through an optimization procedure that involves accessing category representations at different spatial scales. For example, category representations in high-level visual cortex have a nested structure spanning several spatial scales: large spatial scales of large-scale maps contain more abstract, superordinate category information (e.g. animate vs. inanimate) while smaller spatial scales of fine-scale clusters (e.g. faces vs. bodyparts) contain basic-level information (Grill-Spector and Weiner, 2014). If VLPFC reads out category information from VTC to use in a flexible, task-relevant manner, an open question for future research is whether categorical judgments at different levels of abstraction (e.g. superordinate vs. basic) affect the spatial scale of task-dependent category representation in VLPFC.

Our results also have important implications for defining category-selective regions (Downing et al., 2001; Epstein and Kanwisher, 1998; Kanwisher et al., 1997; Kanwisher, 2017; Saxe et al., 2006). Category-selective regions are typically defined as a set of contiguous voxels in a predictable anatomical location that respond significantly more strongly to exemplars of one category compared to other categories. Fundamental to this “functional localization” approach is the assumption that the anatomical location of a functional ROI is stable and does not vary with task. Our data show that this assumption is valid for high-level visual cortex, but not for VLPFC. Consequently, the spatial location of category-selective voxels in VLPFC may change with task-context and attentional states. While recent studies of VLPFC have utilized functional localizers to identify regions associated with processing faces (Chan, 2013; Ishai et al., 2005; Rajimehr et al., 2009; Tsao et al., 2008) or language (Fedorenko et al., 2012) in this cortical expanse, our data indicate that a combination of task and stimulus may be necessary for a robust localization of functional regions in VLPFC.

Selective attention effects are region-general

In addition to the region-specific effects of task on category representations in VLPFC that we covered in the first section of the Discussion, selective attention had a region-general effect across LOTC, VTC, and VLPFC. Specifically, both reproducibility and information content decreased across high-level visual cortex and VLPFC between category-attended and category-unattended conditions in the SA task. This result is in agreement with studies showing that selective attention alters distributed category representations across cortex, particularly in high-level visual cortex and VLPFC (Çukur et al., 2013; Peelen et al., 2009). In these studies, similar to ours, effects of attention on category representations in high-level visual cortex occurred when the task included an element of visual competition. We believe that these converging findings suggest there may be a fundamental difference between task engagement and selective attention, in which the latter specifically requires restricted processing to the attended stimulus in the face of competition. Future work can elucidate the effects of competition and degree of spatial overlap on distributed responses.
A computational advantage of stable category representations in high-level visual cortex and flexible category representations in prefrontal cortex

We hypothesize that multiple types of representations across cortical regions functionally complement one another for efficient and flexible categorization. In terms of stability, both LOTC and VTC exhibit high reproducibility of category selectivity, which may be mediated by a hierarchical feed-forward processing (Riesenhuber and Poggio, 1999; Jiang et al., 2006; Yamins et al., 2014). Additionally, LOTC and VTC exhibit linearly separable distributed representations across tasks, which enables fast and efficient read-out of category information that is task-independent. In terms of flexibility, category representations within VLPFC are task-dependent and display different degrees of linear separability depending on task, as evident from the differences in classification performance across tasks. We propose that task-independent category representations in high-level visual cortex provide stable visual representations, while task-dependent representations in VLPFC provide flexible representations that may enhance task-relevant information. In turn, a key aspect of the flexibility in VLPFC may be to recruit neural resources for visual categorization only when these representations are necessary to achieve the cognitive task at hand.

An alternative explanation of VLPFC responses stems from the multiple demand (Duncan, 2010) framework. Under this view, the differences in VLPFC category representations across tasks may be related to differences in cognitive demand. One aspect of this hypothesis suggests that additional cortical resources are recruited as task difficulty increases (Crittenden and Duncan, 2014), resulting in more distinct category representations. This hypothesis could not be directly tested under the current design as the qualitative differences across our tasks do not enable equating task difficulty under a single metric (e.g., accuracy or response time). We wish to underscore that the tasks in the present study were purposely chosen to be as distinct as possible to maximize the effect size of task, while also testing for robustness of category representations. Future work can differentiate between our relevancy hypothesis and the multiple demand account by concurrently parameterizing both task difficulty and stimulus relevance within a single task.

Conclusions

In sum, we suggest that differential representations of category and task across high-level visual and prefrontal cortices reflect the brain’s need for a balance between stable and flexible category representations to accommodate task goals. Too much stability would generate rigid neural representations unable to accommodate new information, while too much flexibility would generate a confusable and unstable representation of the visual world that is functionally unreliable. The present data indicate that high-level visual cortex contains stable, task-independent categorical representations, while VLPFC contains flexible category representations that can be modified to optimize task-relevant information.

Acknowledgements

This work was supported by the National Eye Institute (NEI) Grant 1R01EY02391501A1 and the Stanford CNI seed grant.

References


