
Article:
Watson, David Mark, Hymers, Mark, Hartley, Tom et al. (2016) Patterns of neural response in scene-selective regions of the human brain are affected by low-level manipulations of spatial frequency. Neuroimage. pp. 107-117. ISSN: 1053-8119

https://doi.org/10.1016/j.neuroimage.2015.08.058

PATTERNS OF NEURAL RESPONSE IN SCENE-SELECTIVE REGIONS OF THE HUMAN BRAIN ARE AFFECTED BY LOW-LEVEL MANIPULATIONS OF SPATIAL FREQUENCY

David M. Watson, Mark Hymers, Tom Hartley, Timothy J. Andrews*

Department of Psychology and York Neuroimaging Centre,

University of York, York, YO10 5DD, United Kingdom

* Corresponding author: timothy.andrews@york.ac.uk

Tables: 4

Figures: 10

Supplementary Figures: 4

Abstract: 222

Introduction: 630

Discussion: 1748

Acknowledgments

We would like to thank Andre Gouws and Sam Johnson for their help at various stages of this project.

ABSTRACT

Neuroimaging studies have found distinct patterns of response to different categories of scenes. However, the relative importance of low-level image properties in generating these response patterns is not fully understood. To address this issue, we directly manipulated the low level properties of scenes in a way that preserved the ability to perceive the category. We then measured the effect of these manipulations on category-selective patterns of fMRI response in the PPA, RSC and OPA. In Experiment 1, a horizontal-pass or vertical-pass orientation filter was applied to images of indoor and natural scenes. The image filter did not have a large effect on the patterns of response. For example, vertical- and horizontal-pass filtered indoor images generated similar patterns of response. Similarly, vertical- and horizontal-pass filtered natural scenes generated similar patterns of response. In Experiment 2, low-pass or high-pass spatial frequency filters were applied to the images. We found that image filter had a marked effect on the patterns of response in scene-selective regions. For example, low-pass indoor images generated similar patterns of response to low-pass natural images. The effect of filter varied across different scene-selective regions, suggesting differences in the way that scenes are represented in these regions. These results indicate that patterns of response in scene-selective regions are sensitive to the low-level properties of the image, particularly the spatial frequency content.

INTRODUCTION

Despite their spatial complexity and heterogeneity, human observers are able to reliably categorise real world scenes even when images are presented rapidly (Greene and Oliva, 2009; Potter, 1975) or visually degraded (Torralba, 2009; Walther et al., 2011). This capacity is thought to be based on neural activity in regions of human visual cortex that are selectively responsive to visual scenes (Aguirre and D'Esposito, 1997; Dilks et al., 2013; Epstein and Kanwisher, 1998; Maguire, 2001; Nasr et al., 2011). While studies using univariate fMRI analyses have reported comparable levels of response within these regions to different images of scenes (Epstein and Kanwisher, 1998), more recent reports employing multivariate techniques have shown that there are distinct patterns of response to different categories of scene (Walther et al., 2011, 2009) suggesting a finer-grained organisation that might underpin perceptual discriminations. However, the functional dimensions that shape these patterns have not been fully resolved.

Some reports have argued that patterns of response reflect high-level, categorical differences amongst scenes (Walther et al., 2011, 2009). For example, Walther and colleagues (2011) showed that the ability to decode scene categories from fMRI data was similar for photographs and line drawings, suggesting some level of invariance to the low level properties of images. However, other studies have suggested that patterns of response in scene-selective regions may be better explained in terms of visual properties of scenes such as spatial layout (e.g. Kravitz et al., 2011; Park et al., 2011; Watson et al., 2014). This latter account is consistent with the sensitivity of the amplitude of response in these regions for orientation (Nasr and Tootell, 2012), spatial frequency (Musel et al., 2014; Rajimehr et al., 2011), visual contrast (Kauffmann et al., 2015), rectilinearity (Nasr et al., 2014), and visual field location (Arcaro et al., 2009; Golomb and Kanwisher, 2012; Levy et

al., 2001). Nevertheless, these studies employed univariate analyses, so it remains unclear whether these modulations in the amplitude of response also affect the pattern of response.

In a recent study, we demonstrated that low-level properties of visual scenes, (defined by the GIST descriptor; Oliva and Torralba 2001), predicted patterns of neural response in scene-selective regions (Watson et al., 2014). However, images drawn from the same scene category are likely to have similar low-level properties (Oliva and Torralba, 2001). So, reliable category-specific patterns of response are expected under both categorical and image-based accounts. Therefore, it remains unclear whether patterns are determined primarily by membership of a common category or by the shared low-level image statistics characteristic of that category.

In the current study, we provide a direct comparison of the relative importance of image properties and category in determining patterns of response in scene-selective regions. Participants viewed images from two different categories of scene (indoor and natural) that are known to have distinct image properties (Oliva and Torralba, 2001) and to elicit different patterns of response in scene-selective regions (Walther et al., 2009; Watson et al., 2014). Low-level visual properties of the scenes were manipulated by filtering the images by orientation (Experiment 1) and spatial frequency (Experiment 2) as previous reports have suggested functional biases for these properties (Nasr and Tootell, 2012; Rajimehr et al., 2011). Using multi-voxel pattern analysis (MVPA), we compared the similarity of the patterns of neural response to each condition across the core scene regions (PPA, RSC, OPA). Our prediction was that if scene-selective regions are sensitive to image properties, then some degree of similarity should be seen between conditions sharing the same filter. If scene-selective regions are solely sensitive to category, then conditions

sharing the same category should elicit similar patterns of response regardless of the low-level manipulation. The use of pattern analysis allows us to determine whether image properties are an important organizing factor in the topography of this region of the brain.

METHODS

Participants

25 participants (8 males; mean age, 25.52; age standard deviation, 4.28; age range, 19-33) took part in Experiment 1 and 24 (8 males; mean age, 25.46; age standard deviation, 3.27; age range, 20-32) took part in Experiment 2. All participants were neurologically healthy, right-handed, and had normal or corrected-to-normal vision. Written consent was obtained for all participants and the study was approved by the York Neuroimaging Centre Ethics Committee.

Stimuli

Visual stimuli were back-projected onto a custom in-bore acrylic screen at a distance of approximately 57 cm from the participant with all images subtending approximately 10.7° of visual angle. Images presented in the main experiment runs were taken from the LabelMe scene database (http://cvcl.mit.edu/database.htm; Oliva & Torralba, 2001) and presented in greyscale. The image set comprised 128 images; 64 indoor and 64 natural scenes. These categories were selected on the basis of their inclusion in previous studies of scene processing (Oliva and Torralba, 2001; Walther et al., 2009). Images were first converted to greyscale – this is important as the filtering process can produce undesirable artifacts in colour images. For instance, high-pass filtering a colour image is likely to introduce false colour into areas of the image not passed by the filter, which will now appear a colour given by the mean luminance of each colour channel. Next, luminance histograms were equated across all images using the MATLAB SHINE toolbox (Willenbockel et al., 2010) prior to any filtering. The full sets of indoor and natural images are shown in Supplementary Figures 1 and 2 respectively.

Filtering was performed by weighting the Fourier spectrum of each image to preserve either horizontal or vertical orientations (Experiment 1), or high or low spatial frequencies (Experiment 2). In Experiment 1, filters were wrapped Gaussian profiles, with a wide angle cut-off (FWHM = 75°) that ensured images remained recognisable after filtering. In Experiment 2 filters were Gaussian profiles with cut-offs set at less than 2 cycles/degree and greater than 6 cycles/degree at FWHM for the low- and high-pass filters respectively. Filter cut-offs for Experiment 2 were based upon those used in previous literature (Oliva and Schyns, 1997; Schyns and Oliva, 1999, 1994). A soft window was applied around the edges of all images to reduce wrap-around edge artifacts associated with the filtering process. Figure 1 shows examples of the images used in each experiment.

[Figure 1 near here]

For each experiment, an additional localiser scan was performed. An independent set of 64 scene images were drawn from the SUN database (http://groups.csail.mit.edu/vision/SUN/; Xiao et al., 2010) and presented in full colour. The SUN database is hierarchically organised into manmade-indoor, manmade-outdoor and natural-outdoor scenes, and stimuli were drawn in approximately equal numbers from each of these 3 classifications. Fourier-scrambled images were created by applying the same set of random phases to each 2-dimensional frequency component in each colour channel of the original image while keeping the magnitude constant. Intact and scrambled images were then rescaled to have a mean luminance equal to that of the images used in the experimental scan. Figure 2a shows examples of the images used in the localiser scan.

Experimental Design

During the localiser scan, participants viewed images from 2 stimulus conditions: (1) intact scene images and (2) phase scrambled versions of the same images in condition 1. During the experimental scan participants viewed images from 4 stimulus conditions comprising 2 scene categories (indoor and natural) across 2 levels of filtering (Experiment 1: horizontal-pass, vertical-pass; Experiment 2: low-pass, high-pass).

In both the localiser and experimental scans, images from each condition were presented in a blocked fMRI design with 9 images per block (8 unique and 1 repeated). Each image was presented for 750ms followed by a 250ms grey screen that was equal in mean luminance to the scene images. Each stimulus block was separated by a 9s period in which the same grey screen as used in the inter-stimulus interval was presented. In order to minimise eye movements a central fixation cross was superimposed on all images and the grey screen and participants were instructed to maintain fixation for the duration of both scans. Each condition was repeated 8 times in a counterbalanced block design giving a total of 16 and 32 blocks in the localiser and experimental scans respectively. To maintain attention throughout the scan sessions participants performed a one-back task in which they were required to detect the repeated presentation of one image in each block, responding to the repeated image with a button press. By using a passive task we avoid biasing neural responses towards either one of our experimental manipulation; for instance, a categorisation task might bias responses towards the category manipulation, whereas an image-based task might bias responses towards the filter manipulation.

Imaging Parameters

All scanning was conducted at the York Neuroimaging Centre (YNiC) using a GE 3 Tesla HDx Excite MRI scanner. A Magnex head-dedicated gradient insert coil was used in conjunction with a birdcage, radiofrequency coil tuned to 127.7MHz. Data were collected from 38 contigual axial slices via a gradient-echo EPI sequence (TR = 3s, TE = 32.5ms, FOV = 288x288mm, matrix size = 128x128, voxel dimensions = 2.25x2.25 mm, slice thickness = 3mm, flip angle = 90°).

fMRI Analysis

Univariate analyses of the fMRI data were performed with FEAT v5.98 (http://www.fmrib.ox.ac.uk/fsl). In all scans the initial 9s of data were removed to reduce the effects of magnetic stimulation. Motion correction (MCFLIRT, FSL, Jenkinson et al., 2002) was applied followed by temporal high-pass filtering (Gaussian-weighted least-squared straight line fittings, sigma=50s). Spatial smoothing (Gaussian) was applied at 6mm FWHM to both the localiser and experiment runs, in line with previous studies employing smoothing in conjunction with MVPA (Op de Beeck, 2010; Watson et al., 2014). Parameter estimates were generated for each condition by regressing the hemodynamic response of each voxel against a box-car regressor convolved with a single-gamma HRF. Next, individual participant data were entered into higher-level group analyses using a mixed-effects design (FLAME, FSL). Functional data were first registered to a high-resolution T1-anatomical image and then onto the standard MNI brain (ICBM152).

A scene-selective region of interest was defined from the localiser data of both experiments using the contrast of intact scenes > scrambled scenes (Figure 2b). The intact scenes share the same amplitude spectra with their phase scrambled counterparts, thus

such a contrast provides a clearer control for low-level visual differences than other commonly used contrasts such as scenes > objects or scenes > faces. For instance, although scenes and objects / faces differ in their category membership, they also differ in a large number of image properties (e.g. spatial frequency, orientation, retinotopic eccentricity, etc.). Given that this experiment aimed to investigate the neural representation of image properties, it was important to use the contrast that provided a stronger control for such visual differences. This ROI therefore provides a definition including scene-selective voxels across a wide extent of cortex – this enables us to test the distributed neural representations of the images as originally described by Haxby et al. (2001). This scene-selective ROI was used for subsequent MVPA across both experiments.

[Figure 2 near here]

We also generated more restrictive ROIs constrained to the classical scene-selective regions (parahippocampal place area (PPA), retrosplenial complex (RSC), occipital place area (OPA)) that have been reported in previous fMRI studies (Dilks et al., 2013; Epstein and Kanwisher, 1998; Maguire, 2001). Within the MNI-2x2x2mm space, group intact>scrambled statistical maps were first averaged across the experiments. Next, seed points were defined at the peak voxels within the average intact>scrambled statistical map for each region (PPA, RSC, OPA) in each hemisphere. For a given seed, a flood fill algorithm was used to identify a cluster of spatially contiguous voxels around that seed which exceeded a given threshold. This threshold was then iteratively adjusted till a cluster size of approximately 500 voxels was achieved (corresponding to a volume of 4000mm^3); actual cluster sizes ranged from 499-501 voxels as an optimal solution to the algorithm was not always achievable. This step ensures that estimates of multi-voxel pattern similarity are not biased by the different sizes

of ROIs being compared. Clusters were combined across hemispheres to yield 3 ROIs, each comprising approximately 1000 voxels. These regions are shown in Figure 3. MNI co-ordinates of the seeds are given in Table 1. These seed points had similar locations to those reported in previous literature (see Watson et al. 2014 – Supplementary Table 1). To ensure clusters were appropriately sized we additionally repeated our analyses across using clusters across a range of sizes from 200-500 voxels. We found that the cluster size made little to no difference upon the main results (Supplementary Figure 3). An additional early visual control ROI was defined from the V1 region of the Jülich histological atlas (Amunts et al., 2000; Eickhoff et al., 2005). We also tested for possible differences in response within the PPA region by splitting this region precisely halfway along its posterior-anterior extent into a posterior PPA and an anterior PPA region.

[Figure 3 and Table 1 near here]

Next, we measured patterns of response to different stimulus conditions in each experiment. Parameter estimates were generated for each condition in the experimental scans. The reliability of response patterns was tested using a leave-one-participant-out (LOPO) cross-validation paradigm (Poldrack et al., 2009; Shinkareva et al., 2008) in which parameter estimates were determined using a group analysis of all participants except one (Supplementary Figure 4). This generated parameter estimates for each scene condition in each voxel. This LOPO process was repeated such that every participant was left out of a group analysis once. These data were then submitted to correlation-based pattern analyses (Haxby et al., 2014, 2001) implemented using the PyMVPA toolbox (http://www.pymvpa.org/; Hanke et al., 2009). Parameter estimates were normalised by subtracting the mean response per voxel across all experimental conditions (see Haxby et

al., 2001). For each iteration of the LOPO cross-validation, the normalized patterns of response to each stimulus condition were correlated between the group and the left-out participant. This allowed us to determine whether there are reliable patterns of response that are consistent across individual participants. A Fisher's z-transformation was then applied to the correlations prior to further statistical analyses.

We next used a representational similarity analysis (RSA; Kriegeskorte et al. 2008) utilising multiple regression to assess the relative contributions of category information and image properties to the neural response patterns. For each factor (category and filter type) a binary regressor was generated representing a model correlations matrix whereby ones were placed on those elements where the relevant factor was shared and zeroes on all other elements. The regressors therefore represent the extreme cases where the patterns of response are entirely predicted by either the scene category or by the filtering; these regressors are illustrated for Experiments 1 and 2 in Figure 5a-b and Figure 8a-b respectively. Each regressor was then repeated and tiled across LOPO iterations. The outcomes measure was defined as the MVPA correlation matrices concatenated across LOPO iterations. All regressors and outcomes were then Z-scored such that all outputs of the regression model are given in standardised units. These regressors and outcomes were then entered into the multiple regression model. This analysis yielded a beta value and associated standard error for each regressor which would be expected to differ significantly from zero if that regressor were able to explain a significant amount of the variance in the MVPA correlations. A t-contrast was used to assess the significance of the differences between the betas.

Behavioural Experiment

In order to ensure that the filtering process did not disrupt the ability of participants to perceive the scenes categorically, we conducted an additional behavioural experiment. A new set of 20 participants (5 males; mean age, 26.80; age standard deviation, 3.32; age range, 23-34) were presented with the images used in the fMRI experiments plus their unfiltered equivalents. This produced 10 conditions across 2 categories (indoor, natural) and 5 levels of filtering (horizontal-pass, vertical-pass, high-pass, low-pass, unfiltered). For each participant, images were divided into 5 subsets and then each subset randomly assigned to a different filtering condition such that participants only saw each image once across all filtering conditions. A chin rest was used to maintain viewing distance across participants. Images subtended a visual angle of approximately 10.7°. In each trial a fixation screen was presented for 1000ms, followed by an image for 750ms. Importantly, both visual angle and stimulus duration were set to match those of the fMRI experiment. Following this, a blank screen was presented for 2250ms or until the participant made a response. Participants were required to indicate, with a button press, whether the image was of an indoor or natural scene as quickly and as accurately as possible, and were able to respond immediately after stimulus onset.

RESULTS

Experiment 1

In Experiment 1, we measured patterns of neural response to different categories of scene (indoor and natural) filtered by orientation (horizontal-pass and vertical-pass). Figure 4 shows the normalised group responses to each condition across the scene-selective ROI. Responses above the mean are shown in red and responses below the mean are shown in blue.

[Figure 4 near here]

A correlation based MVPA (Haxby et al., 2001) was conducted to measure the similarity of the neural responses to different conditions (Figure 5c). To test the contribution of category and image factors to the neural responses, we used a representational similarity analysis (Kriegeskorte et al., 2008). Model correlation matrices were generated representing the extreme cases where the patterns of response are entirely predicted by the scene category (Figure 5a) or by the orientation filter (Figure 5b). These were then used as regressors in a multiple regression analysis of the fMRI data. Figure 5d shows the resulting coefficients for each regressor. Both the category ($\beta = 0.82$, $p < .001$) and filter regressors ($\beta = 0.17$, $p < .001$) explained a significant amount of the variance in the MVPA correlation matrix. However, a t-contrast revealed that the category regressor explained significantly more variance than the filter regressor ($t = 12.84$, $p < .001$). A series of post-hoc paired-sample t-tests were used to compare the critical elements of the correlations matrix representing the same-category, different-filter and different-category, same-filter correlations. In all cases, same-category/different-filter correlations were found to be significantly greater than different-category/same-filter correlations (indoor-

horizontal-pass/indoor-vertical-pass > indoor-horizontal-pass/natural-horizontal-pass: $t(24) = 13.32$, $p < .001$; natural-horizontal-pass/natural-vertical-pass > indoor-horizontal-pass/natural-horizontal-pass: $t(24) = 7.07$, $p < .001$; indoor-horizontal-pass/indoor-vertical-pass > indoor-vertical-pass/natural-vertical-pass: $t(24) = 14.68$, $p < .001$; natural-horizontal-pass/natural-vertical-pass > indoor-vertical-pass/natural-vertical-pass: $t(24) = 8.64$, $p < .001$). An additional post-hoc test did not find a significant difference between correlations in the indoor-horizontal-pass/natural-horizontal-pass and the indoor-vertical-pass/natural-vertical-pass comparison ($t(24) = 1.13$, $p = .271$). Thus, patterns were no more of less similar for horizontal-pass than vertical-pass filtered images.

[Figure 5 near here]

Restricting the regression analysis to the standard scene-selective regions (PPA, RSC, OPA) revealed a similar pattern of results (Figure 6). Responses in the PPA were significantly predicted by the category ($\beta = 0.85$, $p < .001$) but not the filter regressor ($\beta = 0.04$, $p = .204$), with significantly more variance explained by the category than the filter regressor ($t = 16.34$, $p < .001$). Responses in the RSC were significantly predicted by the category ($\beta = 0.77$, $p < .001$) but not the filter regressor ($\beta = 0.02$, $p = .529$), with significantly more variance explained by the category than the filter regressor ($t = 12.01$, $p < .001$). Responses in the OPA were significantly predicted by the category ($\beta = 0.73$, $p < .001$) but not the filter regressor ($\beta = 0.07$, $p = .095$), with significantly more variance was explained by the category than the filter regressor ($t = 10.21$, $p < .001$). In contrast to the scene regions, responses in the early visual (V1) control region were significantly predicted by both the category ($\beta = 0.36$, $p < .001$) and filter regressors ($\beta = 0.25$, $p < .001$). There was no

significant difference between the effect of category and filter ($t = 1.28$, $p = .203$). Results of post-hoc t-tests for these regions are given in Table 2.

[Figure 6 and Table 2 near here]

Experiment 2

In Experiment 2, we measured patterns of neural response to different categories of scene (indoor and natural) filtered by spatial frequency (high-pass and low-pass). Figure 7 shows the normalised group responses to each condition across the scene-selective ROI. Responses above the mean are shown in red and responses below the mean are shown in blue.

[Figure 7 near here]

Correlation based MVPA was used to assess the similarity of the neural responses across different conditions. The influence of category and image factors on the fMRI data was assessed using a representational similarity analysis. Model correlation matrices representing the cases where responses are entirely predicted by the scene category (Figure 8a) or by the spatial frequency filtering (Figure 8b) were entered as regressors in a multiple regression analysis of the fMRI data (Figure 8c). Figure 8d shows the resulting coefficients for each regressor. Both the category ($\beta = 0.23$, $p < .001$) and filter regressors ($\beta = 0.86$, $p < .001$) explained a significant amount of the variance in the MVPA data. However, in contrast to Experiment 1, the filter regressor explained significantly more variance than the category regressor ($t = 16.93$, $p < .001$). Post-hoc tests revealed greater different-category/same-filter than same-category/different-filter correlations in all cases (indoor-high-pass/natural-

high-pass > indoor-high-pass/indoor-low-pass: $t(23) = 17.56$, $p < .001$; indoor-high-pass/natural-high-pass > natural-high-pass/natural-low-pass: $t(23) = 10.29$, $p < .001$; indoor-low-pass/natural-low-pass > indoor-high-pass/indoor-low-pass: $t(23) = 20.26$, $p < .001$; indoor-low-pass/natural-low-pass > natural-high-pass/natural-low-pass: $t(23) = 15.95$, $p < .001$). An additional post-hoc test revealed significantly higher correlations in the indoor-low-pass/natural-low-pass than the indoor-high-pass/natural-high-pass comparison ($t(23) = 10.51$, $p < .001$), indicating greater similarity in the neural response patterns across low-pass than high-pass filtered images.

[Figure 8 near here]

Restricting the regression analyses to the standard scene-selective regions (PPA, RSC, OPA) revealed a more variable pattern of results (Figure 9). Responses in the PPA were significantly predicted by both the category ($\beta = 0.66$, $p < .001$) and filter regressors ($\beta = 0.43$, $p < .001$). However, in contrast to the scene-selective region as a whole, more variance was explained by the category than the filter regressor ($t = 4.33$, $p < .001$) in this subregion. Responses in the RSC were significantly predicted by both the category ($\beta = 0.35$, $p < .001$) and filter regressors ($\beta = 0.53$, $p < .001$) but in this case slightly more variance was explained by the filter than the category regressor ($t = 2.41$, $p = .017$). Responses in the OPA were significantly predicted by both the category ($\beta = 0.22$, $p < .001$) and filter regressors ($\beta = 0.66$, $p < .001$), but again significantly more variance was explained by the filter than the category regressor ($t = 6.25$, $p < .001$). Responses in the V1 control region were significantly predicted by the filter ($\beta = 0.95$, $p < .001$) but not the category regressor ($\beta = 0.03$, $p = .213$), with significantly more variance explained by the filter than the category regressor ($t = 29.96$, $p < .001$). Results of post-hoc t-tests for these regions are given in Table 3.

[Figure 9 and Table 3 near here]

Previous experiments have suggested a possible division of labour between anterior and posterior regions of the PPA (Aminoff et al., 2007; Arcaro et al., 2009; Baldassano et al., 2013; Epstein, 2008). Accordingly, we re-analysed our data by splitting the PPA region halfway along its posterior-anterior extent and repeating the pattern analyses within each division. Responses in the posterior PPA region were significantly predicted by both the category ($\beta = 0.19$, $p < .001$) and filter regressors ($\beta = 0.63$, $p < .001$), with significantly more variance explained by the filter regressor ($t = 5.90$, $p < .001$). Representations in the anterior PPA appeared more similar to the overall PPA region, with responses significantly predicted by both the category ($\beta = 0.75$, $p < .001$) and filter regressors ($\beta = 0.31$, $p < .001$), but with significantly more variance explained by the category regressor ($t = 8.51$, $p < .001$). These results are shown in Figure 10. Our results therefore show a change in selectivity within the PPA, with a shift from more image-based to more category-based representations along a posterior-to-anterior axis.

[Figure 10 near here]

Behavioural Experiment

In order to ensure that the filtering process did not disrupt the ability of participants to perceive the scenes categorically, we conducted an additional behavioural experiment. Participants were presented with the images from the fMRI experiments plus their unfiltered equivalents whilst performing a scene categorisation task. Percentage accuracy scores and median RTs were calculated for each condition within each participant (Table 4).

Mean accuracy across all conditions was 95.63 ± 1.34% (range 89.17 – 97.92%). Mean RT across all conditions was 598 ± 26 msec (range: 566 - 611). These results show that participants were able to categorize all stimulus conditions well above chance levels.

[Table 4 near here]

DISCUSSION

The aim of this study was to compare the relative effect of low-level image properties and high-level categorical factors on the patterns of fMRI response in scene-selective regions. Participants viewed images from indoor and natural scene categories that were filtered by orientation and spatial frequency. These manipulations had a marked effect on the low level image properties. Nevertheless, a behavioural experiment using stimulus presentation parameters matched to those of the fMRI experiments revealed that these manipulations preserved the ability to accurately categorize the images. We then measured the patterns of response in scene-selective regions. We found that orientation filtering had a significantly smaller effect on patterns of response than category. In contrast, spatial frequency filtering had a significantly greater effect on patterns of response compared to category. These results show that patterns of neural response in scene-selective cortices revealed by fMRI are sensitive to low-level properties of the image, particularly the spatial frequency content.

Previous studies have established that distinct patterns of neural response are elicited by viewing different categories of scene (Walther et al., 2011, 2009). These findings have been taken to suggest a categorical organisation of scene-selective cortices in which response properties are linked to the semantic properties of the image. It has also been shown that the semantic content of scene images can be used to predict neural responses during viewing of natural scenes (Huth et al., 2012; Stansbury et al., 2013) and to reconstruct scene images from neural responses in higher visual areas (Naselaris et al., 2009). However, other studies suggest that categorical factors may not provide a complete account of the organization of scene-selective regions. For instance, reports by both Kravitz et al. (2011) and Park et al. (2011) suggest that responses in PPA are better predicted by

image properties (open versus closed) than by the categorical content (indoor versus natural) of scenes. It has also been shown that visual properties can be used to discriminate between different categories of scenes (Torralba and Oliva, 2003). These findings suggest that a fuller understanding of the principles governing organization of ventral visual cortex will hinge on determining the way in which patterns of brain activity reflecting semantic, spatial and functional properties of scenes are derived from their lower level visual properties.

Recently, we showed that the statistical properties of visual images can be used to predict patterns of response in high-level visual cortex (Andrews et al., 2015; Rice et al., 2014; Watson et al., 2014). These results provide an alternative framework for understanding the topographic organization of the ventral visual pathway in which the appearance of category-selective patterns of response may emerge from the combinations of low-level image properties that typically co-occur in different image categories (see also Hanson et al., 2004; Op de Beeck et al., 2008). To directly test the role of image properties, we measured the effect of low-level image manipulations on patterns of response in scene-selective regions. We found a significant effect of spatial frequency filter on patterns of response in scene-selective cortex. For example, indoor low-pass images generated similar patterns of response to natural low-pass images. Similarly, indoor high-pass images generated similar patterns to natural high-pass images. These results show that patterns of response to scenes are sensitive to the low-level properties of the image. Previous univariate fMRI studies have shown that there are biases in the magnitude of the response to different spatial frequencies in scene-selective regions (Kauffmann et al., 2014; Rajimehr et al., 2011). However, changes in the amplitude of response can occur without a change in the pattern of response. Our findings fundamentally extend these earlier studies by showing

that the spatial frequency of the image can also influence the pattern of response in scene-selective regions. This suggests that this property of the image is a key feature underlying the functional organisation of scene-selective regions.

How do we explain the category-specific patterns of response found in scene-selective regions (Walther et al., 2011, 2009)? Rather than reflecting an organization based on categorical properties of the stimulus, we propose that scene-selective regions have a topographic organization that is based on image properties (Andrews et al., 2015). We suggest that the appearance of category selectivity may reflect the characteristic combinations of low-level image properties that co-occur in different types of scenes. Because images from different scene categories have distinct image properties (Watson et al., 2014), images from a particular scene category will activate spatially-selective patterns of response. Although patterns of response in scene-selective regions may be dominated by the features characteristic of specific natural categories, they may remain sensitive to low-level manipulations.

Our findings appear to contrast with a previous study that reported scene category can be decoded from photographs and line drawings of scenes, and that decoding generalises between these visual representations (Walther et al., 2011). As line drawings represent a visually impoverished version of photographic images, it is argued that these results are indicative of image-invariant, categorical representations in scene-selective regions. Our results suggest that such effects could alternatively be understood in terms of the low-level visual properties of images, such as spatial frequency. Line-drawings reduce an image to a subsample of its edge boundaries, and thus represent an extreme high-pass representation of the original image. Consequently, despite being visually impoverished line drawings will nevertheless maintain similar high spatial frequency content to their original

images. Thus, generalisation between each visual representation could reflect sensitivity within the neural patterns to the high spatial frequency content of the image.

Despite showing that manipulations of spatial frequency did affect the patterns of response in scene-selective regions, we also found a smaller but significant effect of scene category across the whole scene-selective ROI. When the scene-selective ROI was subdivided into different sub-divisions (PPA, RSC, TOS/OPA), we found that, although filter and category influenced the patterns of response, the relative contribution of category and filter varied between regions. For instance, the effect of the spatial frequency filter was greater than that of the category in both the OPA and RSC, while in the PPA the effect of category was greater than filter. This suggests that while all scene-selective regions remain sensitive to the low-level visual properties of scenes, there may be a shift towards a more categorical representation in some regions. Presumably, these differences in selectivity reflect the different computational processes that are thought to occur in different scene-selective regions. For instance, it has been proposed that the PPA and RSC may form distinct but complimentary roles within the scene processing network, with the PPA primarily focussed on representing the spatial components of the immediately visible scene, whilst the RSC is more concerned with representing the scene within the wider spatial environment (Epstein and Higgins, 2007; Epstein, 2008; Epstein et al., 2007; Park and Chun, 2009). Meanwhile, the more posterior OPA has been proposed to be a lower-level component of a hierarchical scene processing network (Dilks et al., 2013), perhaps analogous to proposed roles for the occipital face area within the face processing network (Haxby et al., 2002). We additionally observed a shift from more image-based to more category-based representations along a posterior-to-anterior axis within the PPA. This suggests an organisation in which representations become less dependent on the individual

visual components of images in more anterior regions of parahippocampal cortex, consistent with previous studies suggesting a division of labour along this axis (Baldassano et al., 2013; Epstein, 2008).

In contrast to spatial frequency, we found that manipulating the orientation content of the image had a much smaller effect on the patterns of response across scene-selective cortex. For example, indoor vertical-pass images generated similar patterns of response to indoor horizontal-pass images and natural vertical-pass images generated similar patterns to natural horizontal-pass images. Our results suggest that not all low level properties exert the same degree of influence on large scale patterns of response in scene-selective cortex. This result may seem at odds with a previous study that reported orientation biases in scene-selective regions (Nasr and Tootell, 2012). However, this study differed from our study in two important ways. First, our filters only included the cardinal orientations (horizontal and vertical) and so did not coincide with the cardinal versus oblique orientation bias shown by Nasr and Tootell (2012). Indeed, they did not report any significant differences between cardinal orientations. Second, they used a univariate analysis in which the magnitude of response to cardinal orientations was compared to oblique orientations. In contrast, we investigated the pattern of response across the cortical surface. It is possible to find overall differences in the magnitude of the response between conditions that are not reflected in the pattern of response. So, the finding that the current analyses did not show a significant effect of orientation filtering upon the pattern of response should not be taken as meaning that the regions do not have low-level orientation biases. Rather, it simply means that (horizontal vs. vertical) orientation biases are not found in the pattern of response detected by fMRI.

To understand how the neural representation of scenes changes through the processing hierarchy, we measured the patterns of response in V1. We found that the pattern of response in V1 showed some differences to the patterns found in the scene-selective regions. For instance, while the orientation filters had little effect on the responses in the scene selective regions, a significant effect of both orientation filter and category was found in V1. Furthermore, although a significant effect of both spatial frequency filters and category was observed in scene-selective regions, there was only an effect of spatial frequency filters on the pattern of response in V1. It is important to note, however, that although image filtering techniques do preserve categorical information, they also preserve other visual dimensions that are not influenced by the filtering manipulation. So, the observed effects of the category manipulation may be attributable not only to categorical factors, but also to visual properties that were not affected by the filtering. For example, the effect of category in V1 in Experiment 1 is unlikely to reflect a higher-level representation of scenes in this region, but it is more likely to be driven by differences in the remaining non-orientation-sensitive visual information (such as spatial frequency). Nevertheless, our results indicate a gradual transition in responses to low-level properties such that later processing regions (e.g., PPA) are increasingly sensitive to those features which serve to distinguish behaviourally distinct environments.

In conclusion, in this study we directly determined the effect of low-level image manipulations on the patterns of neural response to different scene categories. We found clear evidence that scene-selective regions were sensitive to the low-level visual content of the image, and that spatial frequency was more influential than orientation content in determining the coarse-scale patterns measured by the MVPA. The sensitivity to image properties shown in this study fundamentally extends previous univariate reports of image

biases in the magnitude of response in scene-selective regions. By showing that the pattern of response to scenes can be influenced by the spatial frequency content of the image, our results suggest that this image property is an important organizing factor in the topographic organization of scene-selective regions of the brain.

REFERENCES

Aguirre, G.K., D'Esposito, M., 1997. Environmental knowledge is subserved by separable dorsal/ventral neural areas. J. Neurosci. 17, 2512–2518.

Aminoff, E., Gronau, N., Bar, M., 2007. The parahippocampal cortex mediates spatial and nonspatial associations. Cereb. Cortex 17, 1493–503.

Amunts, K., Malikovic, A., Mohlberg, H., Schormann, T., Zilles, K., 2000. Brodmann's areas 17 and 18 brought into stereotaxic space - where and how variable? Neuroimage 11, 66–84.

Andrews, T.J., Watson, D.M., Rice, G.E., Hartley, T., 2015. Low-level properties of natural images predict topographic patterns of neural response in the ventral visual pathway visual pathway. J. Vis. 15, 1–12.

Arcaro, M.J., McMains, S.A., Singer, B.D., Kastner, S., 2009. Retinotopic Organization of Human Ventral Visual Cortex. J. Neurosci. 29, 10638–10652.

Baldassano, C., Beck, D.M., Fei-Fei, L., 2013. Differential connectivity within the Parahippocampal Place Area. Neuroimage 75, 228–37.

Dilks, D.D., Julian, J.B., Paunov, A.M., Kanwisher, N., 2013. The Occipital Place Area Is Causally and Selectively Involved in Scene Perception. J. Neurosci. 33, 1331–1336.

Eickhoff, S.B., Stephan, K.E., Mohlberg, H., Grefkes, C., Fink, G.R., Amunts, K., Zilles, K., 2005. A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. Neuroimage 25, 1325–35.

Epstein, R., Higgins, J.S., 2007. Differential parahippocampal and retrosplenial involvement in three types of visual scene recognition. Cereb. Cortex 17, 1680–1693.

Epstein, R., Kanwisher, N., 1998. A cortical representation of the local visual environment. Nature 392, 598–601.

Epstein, R.A., 2008. Parahippocampal and retrosplenial contributions to human spatial navigation. Trends Cogn. Sci. 12, 388–396.

Epstein, R.A., Parker, W.E., Feiler, A.M., 2007. Where am I now? Distinct roles for parahippocampal and retrosplenial cortices in place recognition. J. Neurosci. 27, 6141–6149.

Golomb, J.D., Kanwisher, N., 2012. Higher Level Visual Cortex Represents Retinotopic, Not Spatiotopic, Object Location. Cereb. Cortex 22, 2794–2810.

Greene, M.R., Oliva, A., 2009. The Briefest of Glances: The Time Course of Natural Scene Understanding. Psychol. Sci. 20, 464–472.

Hanke, M., Halchenko, Y.O., Sederberg, P.B., Hanson, S.J., Haxby, J. V, Pollmann, S., 2009. PyMVPA: a Python Toolbox for Multivariate Pattern Analysis of fMRI Data. Neuroinformatics 7, 37–53.

Hanson, S.J., Matsuka, T., Haxby, J. V, 2004. Combinatorial codes in ventral temporal lobe for object recognition: Haxby (2001) revisited: is there a "face" area? Neuroimage 23, 156–166.

Haxby, J. V, Connolly, A.C., Guntupalli, J.S., 2014. Decoding Neural Representational Spaces Using Multivariate Pattern Analysis. Annu. Rev. Neurosci. 37, 435–456.

Haxby, J. V, Gobbini, M., Furey, M., Ishai, A., Schouten, J., Pietrini, P., 2001. Distributed and overlapping representations of faces and objects in ventral temporal cortex. Science 293, 2425–2430.

Haxby, J. V, Haxby, J. V, Hoffman, E. a, Hoffman, E. a, Gobbini, M.I., Gobbini, M.I., 2002. Human neural systems for face recognition and social communication. Biol. Psychiatry 51, 59–67.

Huth, A.G., Nishimoto, S., Vu, A.T., Gallant, J.L., 2012. A Continuous Semantic Space Describes the Representation of Thousands of Object and Action Categories across the Human Brain. Neuron 76, 1210–1224.

Jenkinson, M., Bannister, P., Brady, M., Smith, S., 2002. Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images. Neuroimage 17, 825–841.

Kauffmann, L., Ramanoël, S., Guyader, N., Chauvin, A., Peyrin, C., 2015. Spatial frequency processing in scene-selective cortical regions. Neuroimage 112, 86–95.

Kauffmann, L., Ramanoël, S., Peyrin, C., 2014. The neural bases of spatial frequency processing during scene perception. Front. Intergrative Neurosci. 8, 1–14.

Kravitz, D.J., Peng, C.S., Baker, C.I., 2011. Real-World Scene Representations in High-Level Visual Cortex: It's the Spaces More Than the Places. J. Neurosci. 31, 7322–7333.

Kriegeskorte, N., Mur, M., Bandettini, P.A., 2008. Representational similarity analysis - connecting the branches of systems neuroscience. Front. Syst. Neurosci. 2, 1–28.

Levy, I., Hasson, U., Avidan, G., Hendler, T., Malach, R., 2001. Center – periphery organization of human object areas. Nat. Neurosci. 4, 533–539.

Maguire, E., 2001. The retrosplenial contribution to human navigation: A review of lesion and neuroimaging findings. Scand. J. Psychol. 42, 225–238.

Musel, B., Kauffmann, L., Ramanoël, S., Giavarini, C., Guyader, N., Chauvin, A., Peyrin, C., 2014. Coarse-to-fine Categorization of Visual Scenes in Scene-selective Cortex. J. Cogn. Neurosci. 26, 2287–2297.

Naselaris, T., Prenger, R.J., Kay, K.N., Oliver, M., Gallant, J.L., 2009. Bayesian Reconstruction of Natural Images from Human Brain Activity. Neuron 63, 902–915.

Nasr, S., Echavarria, C.E., Tootell, R.B.H., 2014. Thinking Outside the Box: Rectilinear Shapes Selectively Activate Scene-Selective Cortex. J. Neurosci. 34, 6721–6735.

Nasr, S., Liu, N., Devaney, K.J., Yue, X., Rajimehr, R., Ungerleider, L.G., Tootell, R.B.H., 2011. Scene-Selective Cortical Regions in Human and Nonhuman Primates. J. Neurosci. 31, 13771–13785.

Nasr, S., Tootell, R.B.H., 2012. A cardinal orientation bias in scene-selective visual cortex. J. Neurosci. 32, 14921–6.

Oliva, A., Schyns, P.G., 1997. Coarse Blobs or Fine Edges ? Evidence That Information Diagnosticity Changes the Perception of Complex Visual Stimuli. Cogn. Psychol. 34, 72–107.

Oliva, A., Torralba, A., 2001. Modeling the Shape of the Scene: A Holistic Representation of the Spatial Envelope. Int. J. Comput. Vis. 42, 145–175.

Op de Beeck, H.P., 2010. Against hyperacuity in brain reading: spatial smoothing does not hurt multivariate fMRI analyses? Neuroimage 49, 1943–8.

Op de Beeck, H.P., Haushofer, J., Kanwisher, N.G., 2008. Interpreting fMRI data: maps, modules and dimensions. Nat. Rev. Neurosci. 9, 123–135.

Park, S., Brady, T.F., Greene, M.R., Oliva, A., 2011. Disentangling Scene Content from Spatial Boundary: Complementary Roles for the Parahippocampal Place Area and Lateral Occipital Complex in Representing Real-World Scenes. J. Neurosci. 31, 1333–1340.

Park, S., Chun, M.M., 2009. Different roles of the parahippocampal place area (PPA) and retrosplenial cortex (RSC) in panoramic scene perception. Neuroimage 47, 1747–1756.

Poldrack, R.A., Halchenko, Y.O., Hanson, S.J., 2009. Decoding the large-scale structure of brain function by classifying mental states across individuals. Psychol. Sci. 20, 1364–72.

Potter, M.C., 1975. Meaning in visual search. Science 187, 965–966.

Rajimehr, R., Devaney, K.J., Bilenko, N.Y., Young, J.C., Tootell, R.B.H., 2011. The "Parahippocampal Place Area" Responds Preferentially to High Spatial Frequencies in Humans and Monkeys. PLoS Biol 9, e1000608.

Rice, G.E., Watson, D.M., Hartley, T., Andrews, T.J., 2014. Low-level image properties of visual objects predict patterns of neural response across category-selective regions of the ventral visual pathway. J. Neurosci. 34, 8837–8844.

Schyns, P.G., Oliva, A., 1994. From Blobs to Boundary Edges: Evidence for Time- and Spatial-Scale-Dependent Scene Recognition. Psychol. Sci. 5, 195–200.

Schyns, P.G., Oliva, A., 1999. Dr. Angry and Mr. Smile: when categorization flexibly modifies the perception of faces in rapid visual presentations. Cognition 69, 243–65.

Shinkareva, S. V, Mason, R.A., Malave, V.L., Wang, W., Mitchell, T.M., Just, M.A., 2008. Using FMRI brain activation to identify cognitive states associated with perception of tools and dwellings. PLoS One 3, e1394.

Stansbury, D.E., Naselaris, T., Gallant, J.L., 2013. Natural Scene Statistics Account for the Representation of Scene Categories in Human Visual Cortex. Neuron 79, 1025–1034.

Torralba, A., 2009. How many pixels make an image? Vis. Neurosci. 26, 123–131.

Torralba, A., Oliva, A., 2003. Statistics of natural image categories. Netw. Comput. Neural Syst. 14, 391–412.

Walther, D.B., Caddigan, E., Fei-Fei, L., Beck, D.M., 2009. Natural Scene Categories Revealed in Distributed Patterns of Activity in the Human Brain. J. Neurosci. 29, 10573–10581.

Walther, D.B., Chai, B., Caddigan, E., Beck, D.M., Fei-Fei, L., 2011. Simple line drawings suffice for functional MRI decoding of natural scene categories. Proc. Natl. Acad. Sci. 108, 9661–9666.

Watson, D.M., Hartley, T., Andrews, T.J., 2014. Patterns of response to visual scenes are linked to the low-level properties of the image. Neuroimage 99, 402–410.

Willenbockel, V., Sadr, J., Fiset, D., Horne, G.O., Gosselin, F., Tanaka, J.W., 2010. Controlling low-level image properties: The SHINE toolbox. Behav. Res. Methods 42, 671–684.

Xiao, J.X., Hays, J., Ehinger, K.A., Oliva, A., Torralba, A., 2010. SUN Database: Large-scale Scene Recognition from Abbey to Zoo, in: IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Soc, Los Alamitos, pp. 3485–3492.

Table 1. Peak MNI mm co-ordinates and thresholds of standard scene-selective clusters (PPA, RSC, OPA).

Region	Hemisphere	*x*	*y*	*z*	Threshold (Z)
PPA	L	-24	-52	-14	5.21
	R	26	-50	-16	5.68
RSC	L	-18	-62	4	4.24
	R	16	-54	-2	4.92
OPA	L	-36	-88	4	5.23
	R	36	-82	4	5.54

Table 2. Experiment 1: t-statistics and significance of post-hoc pairwise t-tests for standard scene selective regions (PPA, RSC, OPA) and V1 (* p < .05, ** p < .01, *** p < .001).

	PPA	RSC	OPA	V1
IHo/IVe > IHo/NHo	9.35***	9.30***	9.25***	2.49(ns)
NHo/NVe > IHo/NHo	8.68***	8.77***	7.05***	-2.13(ns)
IHo/IVe > IVe/NVe	9.84***	7.14***	6.97***	3.83**
NHo/NVe > IVe/NVe	9.26***	7.05***	5.09***	-0.18(ns)

Table 3. Experiment 2: t-statistics and significance of post-hoc pairwise t-tests for standard scene selective regions (PPA, RSC, OPA) and V1 (* p < .05, ** p < .01, *** p < .001).

	PPA	RSC	OPA	V1
IHi/NHi > IHi/ILo	-3.42**	2.48(ns)	6.12***	18.23***
IHi/NHi > NHi/NLo	-5.55***	1.06(ns)	1.63(ns)	17.33***
ILo/NLo > IHi/ILo	-0.80(ns)	4.10**	7.89***	17.94***
ILo/NLo > NHi/NLo	-2.89*	1.89(ns)	5.49***	16.79***

Table 4. Behavioural experiment: average accuracy and response times (± 1 SEM).

Category	Filter	Accuracy (% correct)	Response Time (ms)
Indoor	Horizontal-pass	95.83 ± 1.28	597 ± 18
	Vertical-pass	95.42 ± 1.41	611 ± 19
	High-pass	97.08 ± 1.09	569 ± 16
	Low-pass	94.58 ± 1.51	611 ± 18
	Unfiltered	95.42 ± 1.54	566 ± 17
Natural	Horizontal-pass	97.50 ± 1.06	604 ± 23
	Vertical-pass	95.83 ± 1.54	595 ± 21
	High-pass	97.50 ± 1.06	589 ± 16
	Low-pass	89.17 ± 2.10	664 ± 24
	Unfiltered	97.92 ± 0.83	581 ± 20

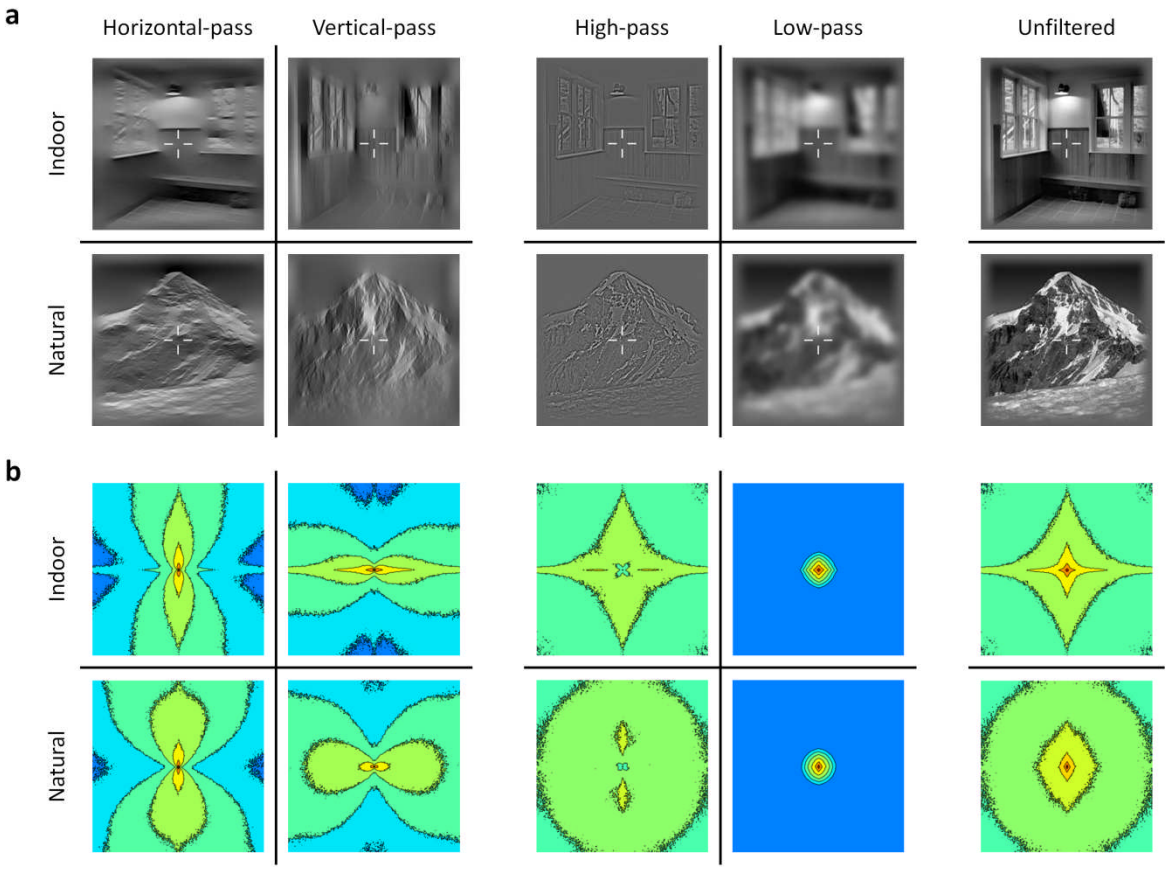


Fig 1. (a) Examples of images from conditions in Experiment 1 (left panels) and Experiment 2 (middle panels). For comparison, equivalent unfiltered images are shown (right panels). (b) Average Fourier amplitude spectra across all images in each condition.

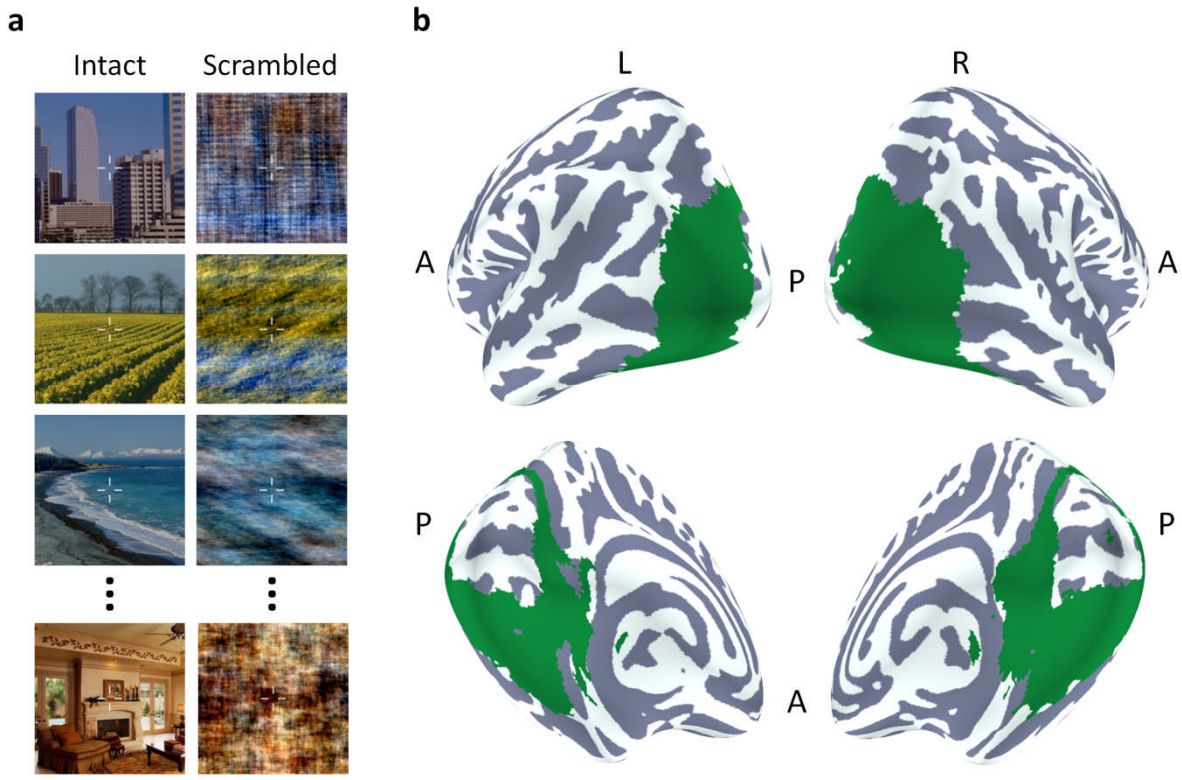


Fig. 2. (a) Examples of images presented in the localiser scan. (b) Mask used for ROI analyses defined by the contrast of intact > scrambled.

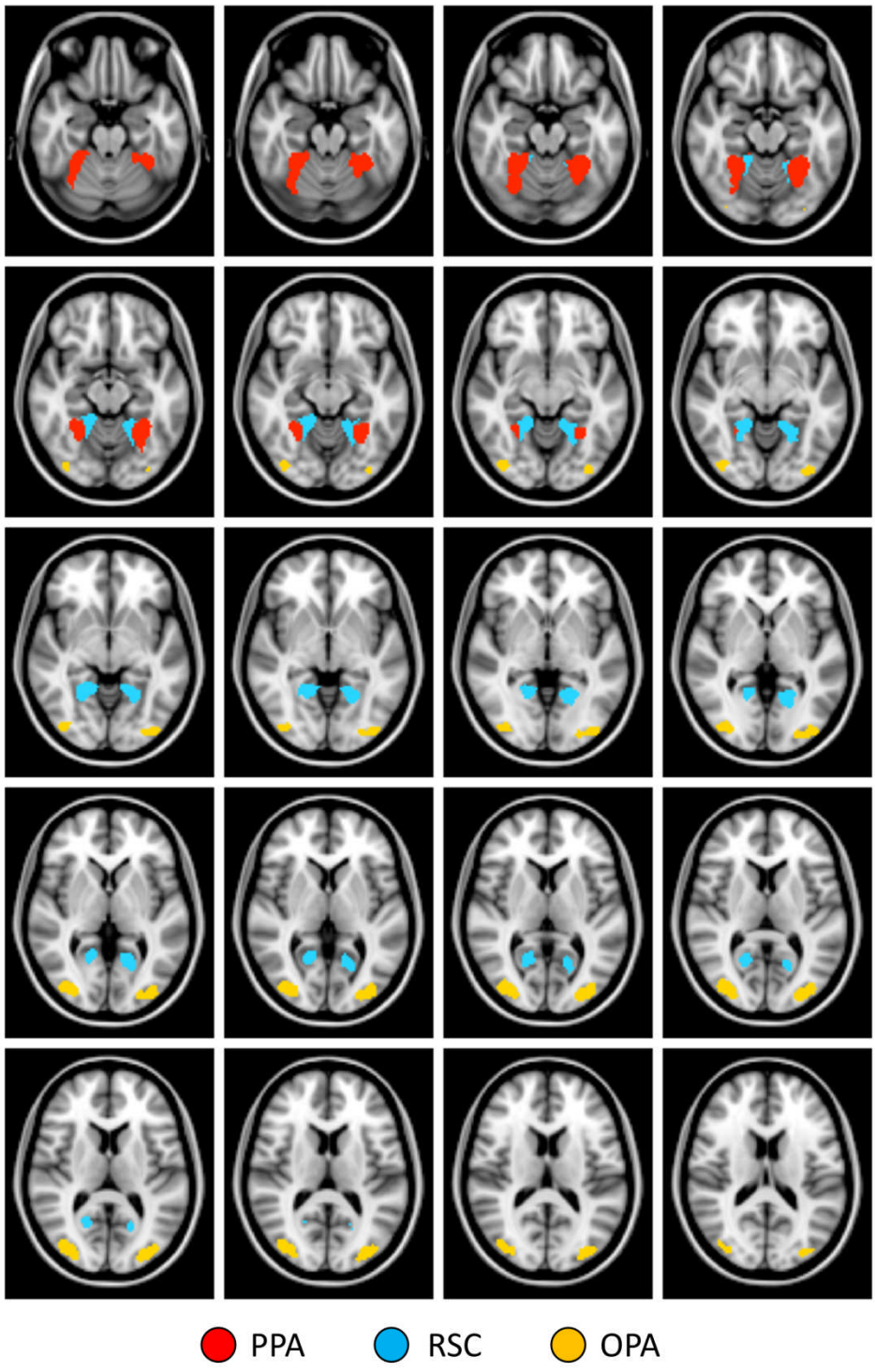


Fig. 3. Masks used for ROI analyses of core scene regions. Each mask comprises approximately 500 voxels (4000mm^3) in each hemisphere. Slices of MNI brain span the range from Z = -22 to Z = 16 in 2mm increments.

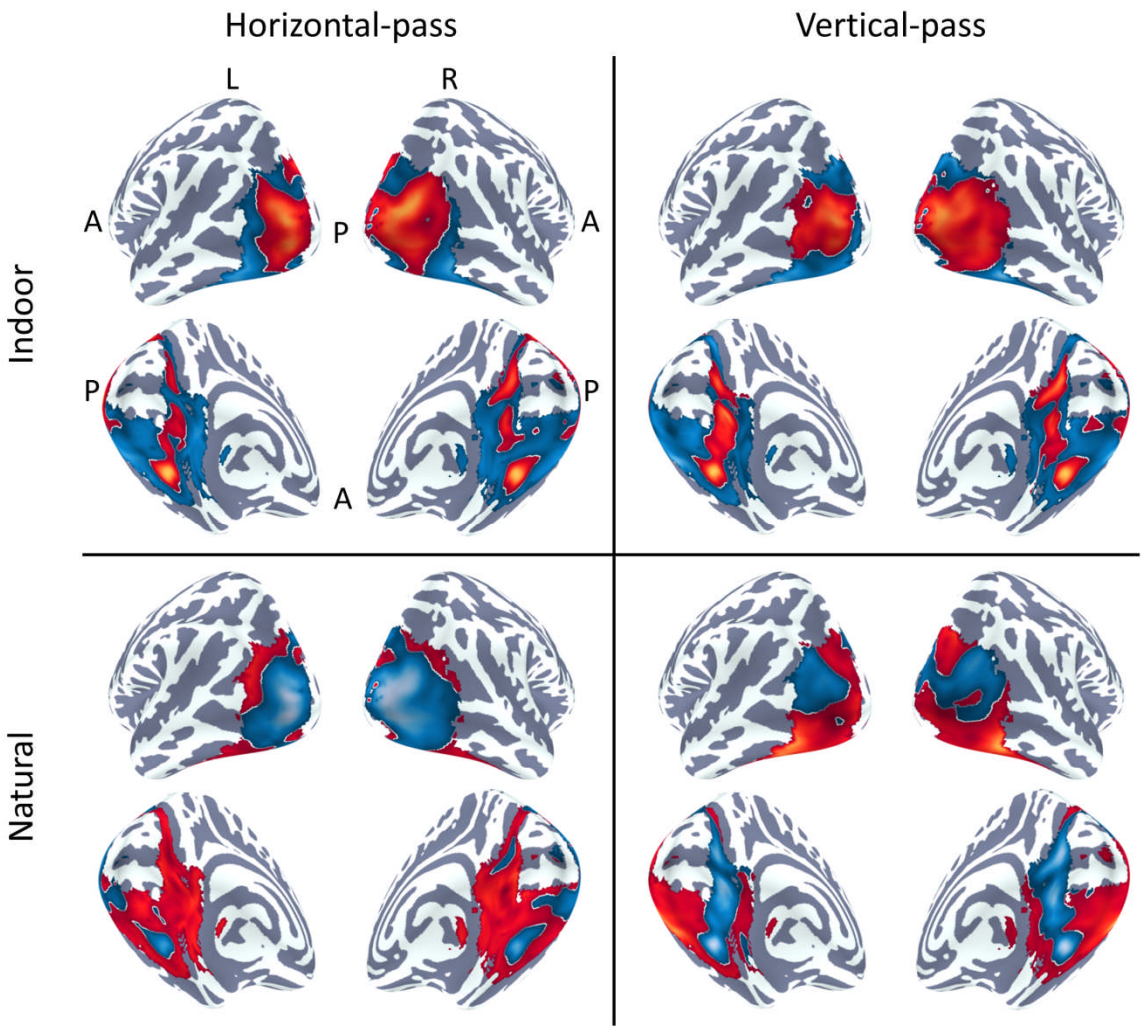


Fig. 4. Group patterns of response to conditions in Experiment 1. Patterns are restricted to regions defined by the response of intact scenes > scrambled scenes. Red and blue colours indicate normalized values above and below the mean respectively.

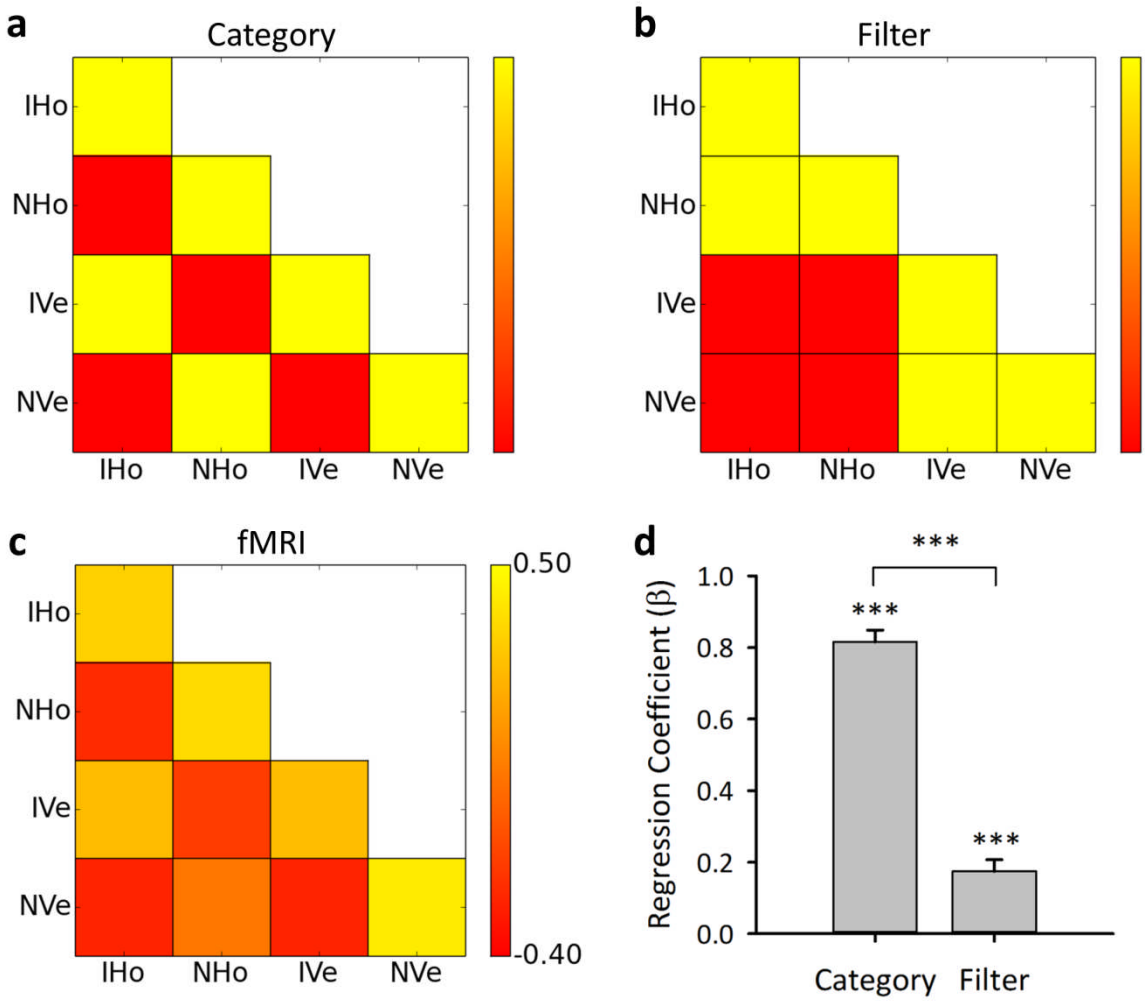


Fig. 5. Experiment 1 analysis. Condition labels: indoor horizontal-pass (IHo), natural horizontal-pass (NHo), indoor vertical-pass (IVe), natural vertical-pass (NVe). Binary models were defined representing the cases where the patterns of response are entirely predicted by either the category (a) or the filter type (b). These were entered into a multiple regression analysis as regressors, while the fMRI MVPA correlations (c) were entered as outcomes. The resulting regression coefficients are shown in (d). Error bars represent 1 SEM. (* $p < .05$, ** $p < .01$, *** $p < .001$).

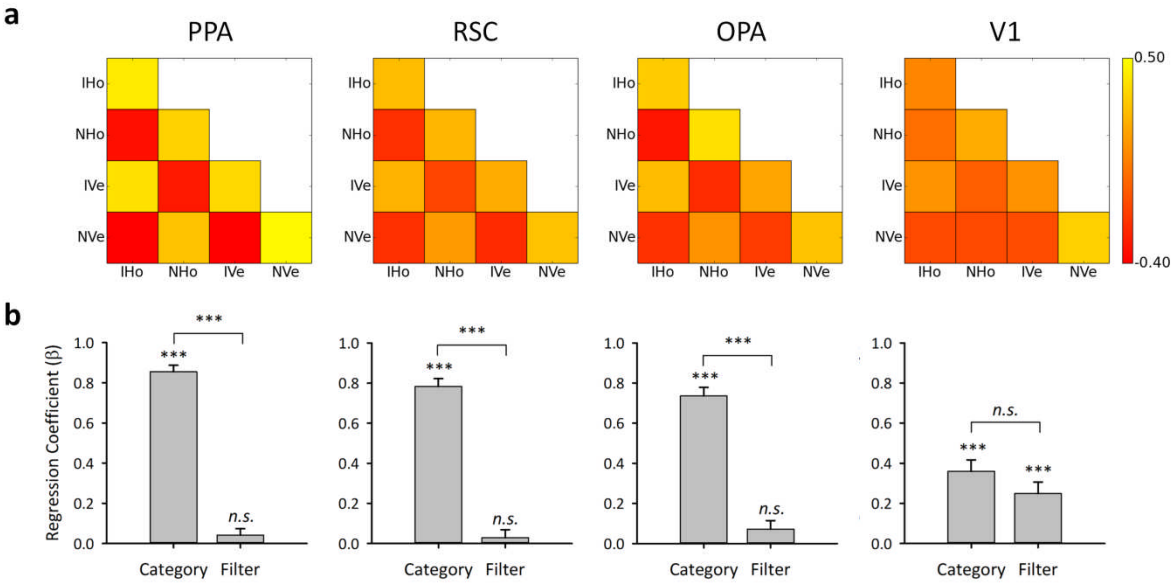


Fig. 6. Experiment 1: standard scene-selective regions and V1. (a) MVPA correlation matrices. (b) These matrices were compared against binary regressors of category and filter effects using a multiple regression analysis; resulting beta coefficients are shown for each regressor. Error bars represent 1 SEM. (* $p < .05$, ** $p < .01$, *** $p < .001$).

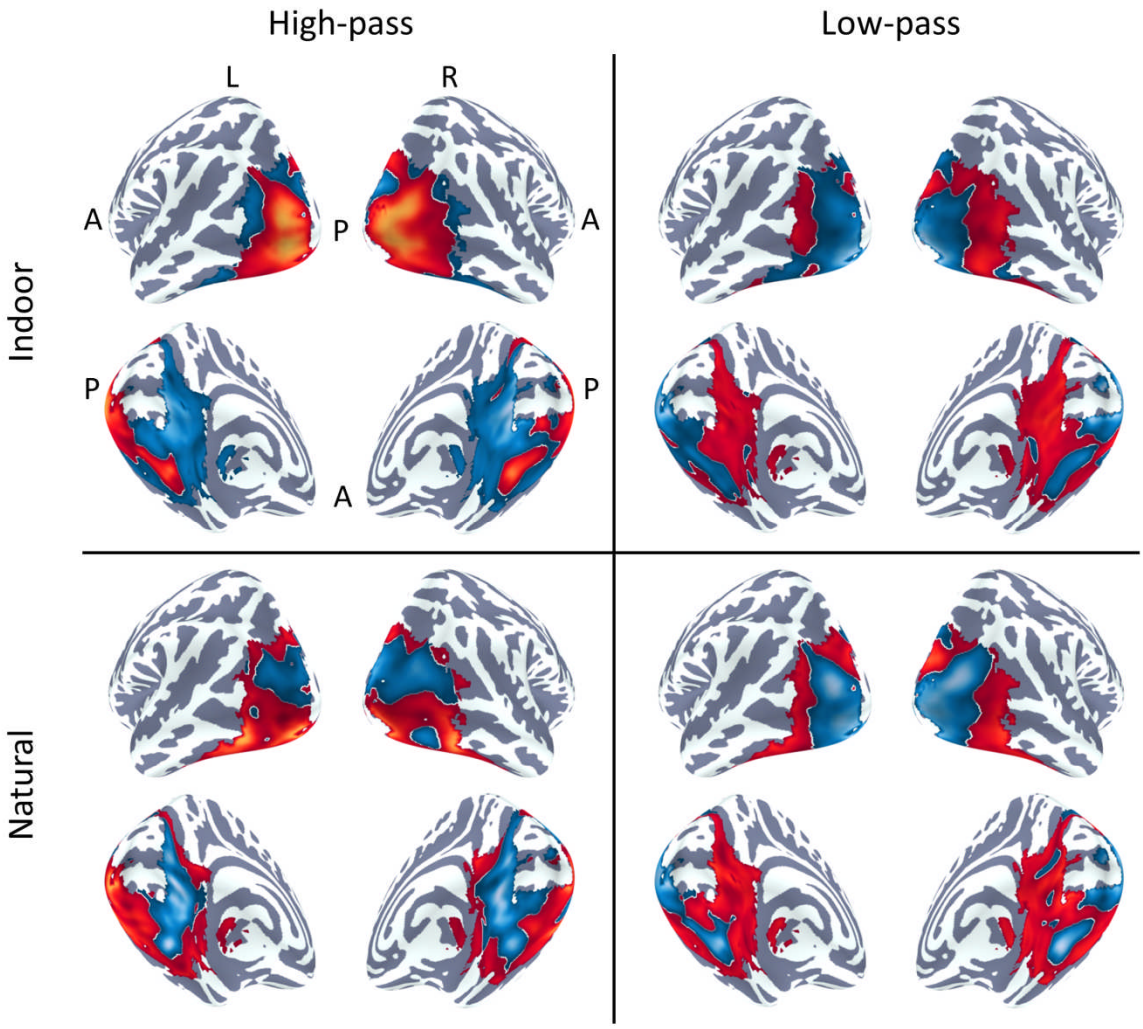


Fig. 7. Group patterns of response to conditions in Experiment 2. Patterns are restricted to regions defined by the response of mixed scenes > scrambled scenes. Red and blue colours indicate normalized values above and below the mean respectively.

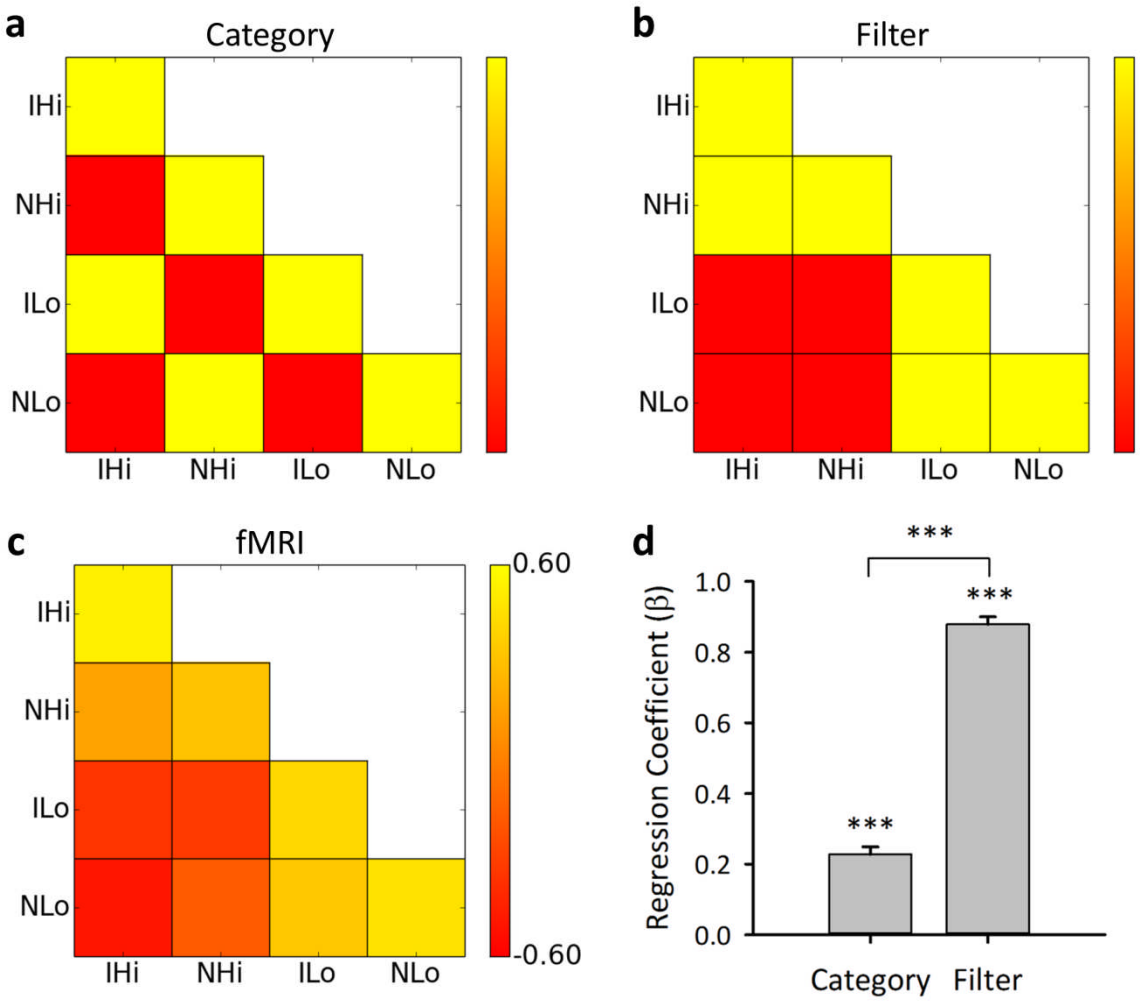


Fig. 8. Experiment 2 analysis. Condition labels: indoor high-pass (IHi), natural high-pass (NHi), indoor low-pass (ILo), natural low-pass (NLo). Binary models were defined representing the cases where the patterns of response are entirely predicted by either the category (a) or the filter type (b). These were entered into a multiple regression analysis as regressors, while the fMRI MVPA correlations (c) were entered as outcomes. The resulting regression coefficients are shown in (d). Error bars represent 1 SEM. (* $p < .05$, ** $p < .01$, *** $p < .001$).

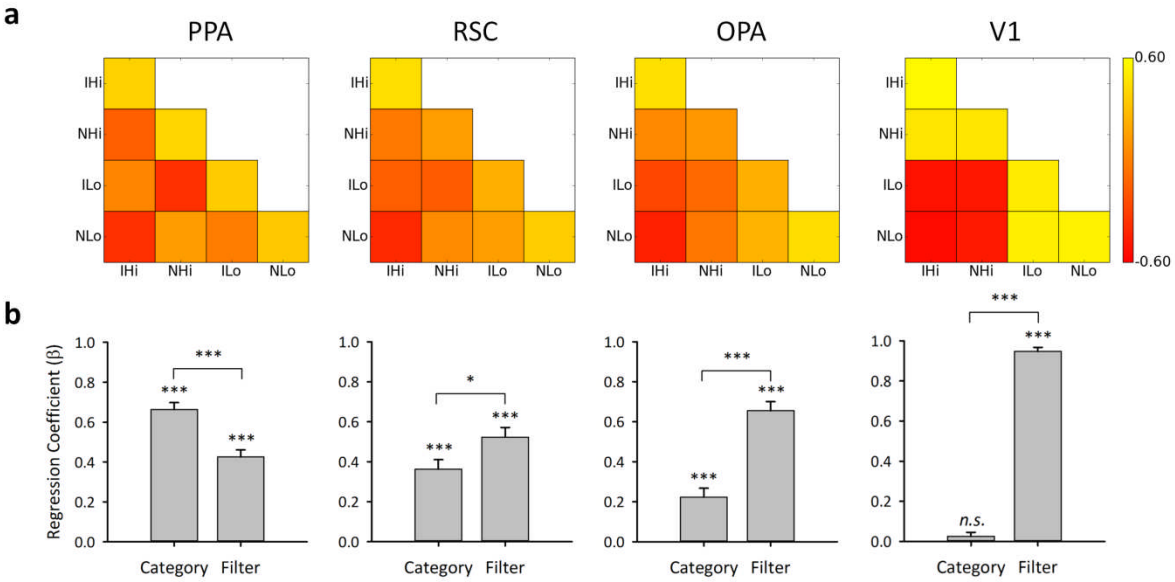


Fig. 9. Experiment 2: standard scene-selective regions. (a) MVPA correlation matrices. (b) These matrices were compared against binary regressors of category and filter effects using a multiple regression analysis; resulting beta coefficients are shown for each regressor. Error bars represent 1 SEM. (* p < .05, ** p < .01, *** p < .001).

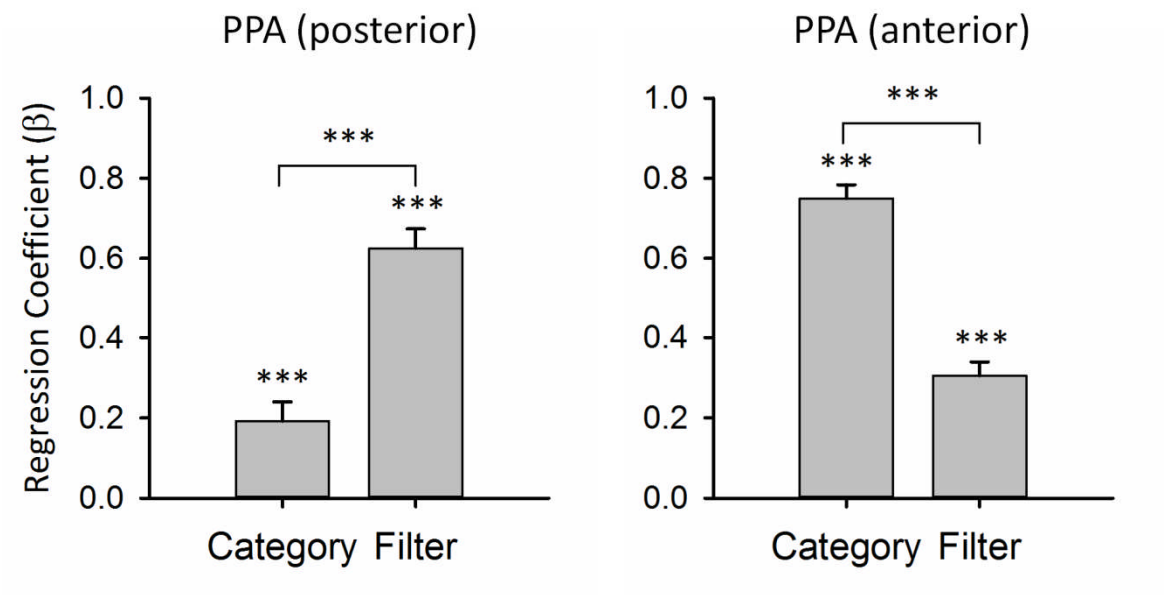
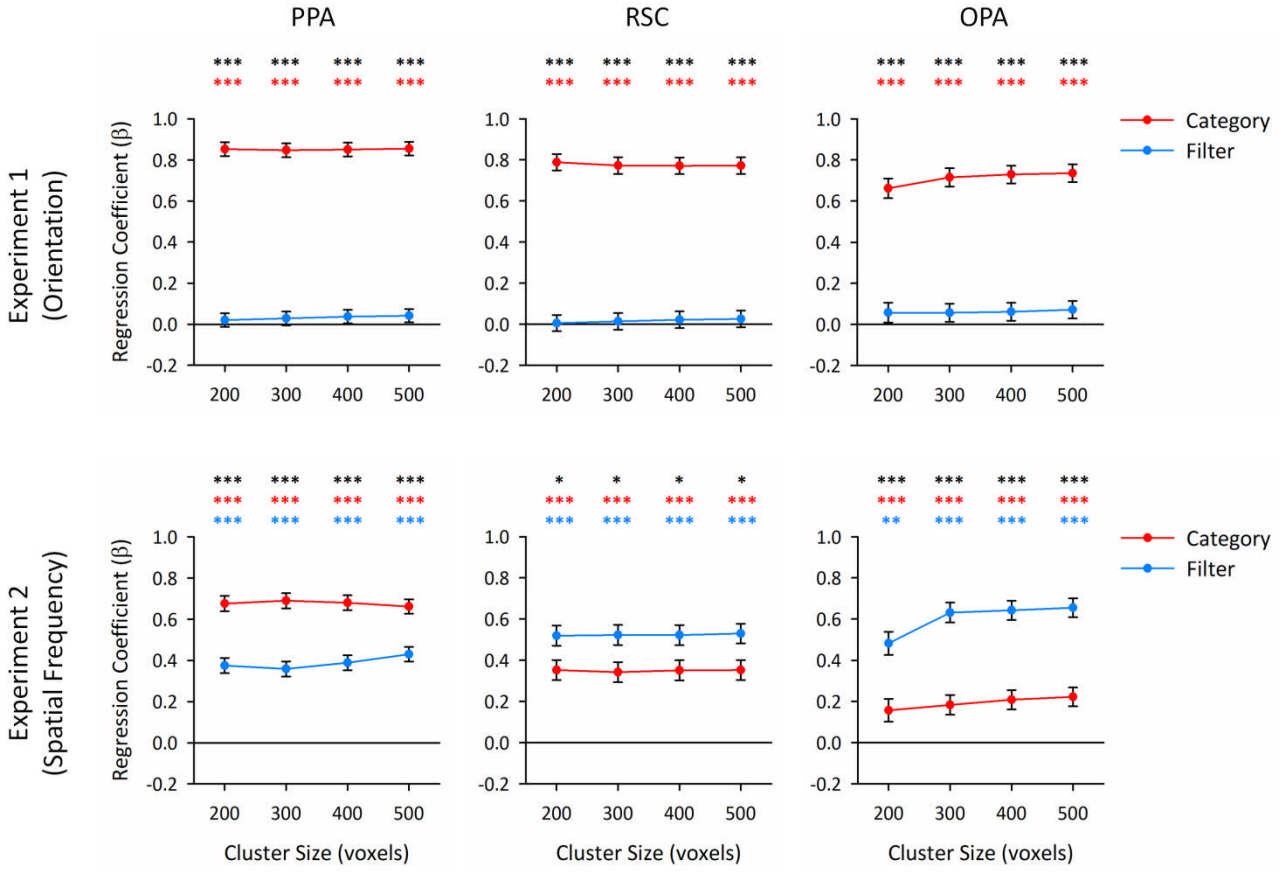


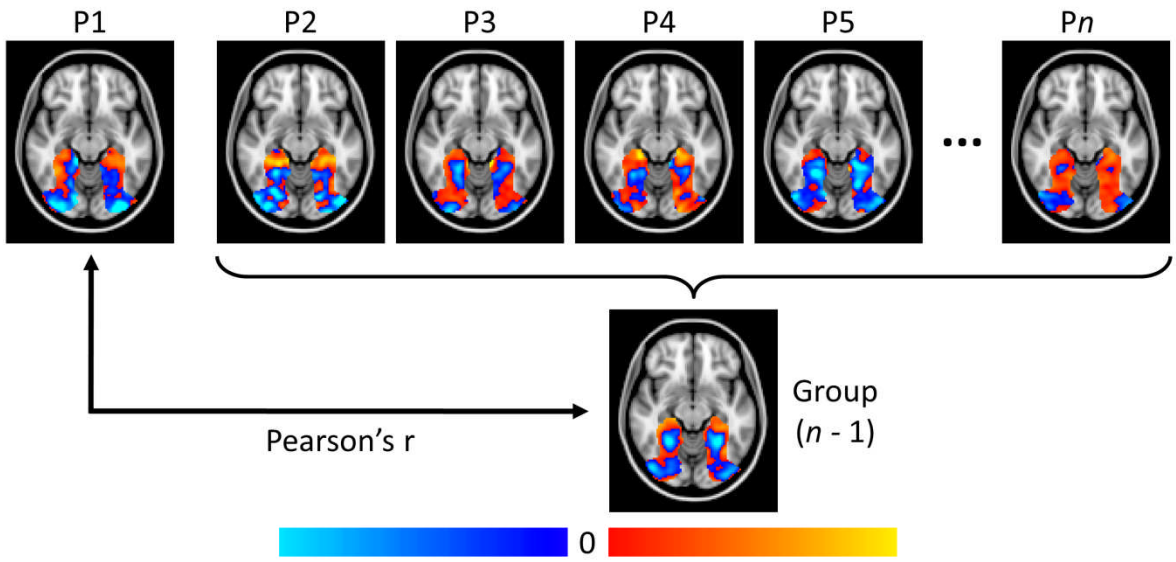
Fig. 10. Experiment 2: Analysis of anterior and posterior PPA divisions. The PPA region was divided halfway along its posterior-anterior extent, and the pattern analyses and representational analyses repeated for each division separately. The resulting regression coefficients are displayed above. Error bars represent 1 SEM. (* $p < .05$, ** $p < .01$, *** $p < .001$).

Suppl. Fig. 1. Full unfiltered indoor scene image set.

Suppl. Fig 2. Full unfiltered natural scene image set.



Suppl. Fig. 3. A flood-fill algorithm was used to identify ROIs for each of the scene-selective regions (PPA, RSC, OPA) in each hemisphere. Clusters were defined to comprise 200, 300, 400, or 500 contiguous voxels, and then combined across hemispheres for each region to yield final ROIs comprising 400, 600, 800, or 1000 voxels respectively. The multi-voxel pattern analyses and representational similarity analyses were conducted for each ROI independently (see Methods for full details). The resulting regression coefficients are displayed above; coloured asterisks indicate the significance of the corresponding regressors, whilst black asterisks indicate the significance of the contrast between the regressors (*** $p < .001$, ** $p < .01$, * $p < .05$). Error bar represent 1 SEM. In all cases, cluster size is seen to have little effect upon the results of the regression analyses.



Suppl. Fig. 4. Schematic diagram of pattern analysis and leave-one-participant-out (LOPO) cross-validation scheme. Analyses correlated individual patterns of response with the group pattern of response derived from all participants except that individual. This process was then repeated across all LOPO iterations.