

The occipital place area represents the local elements of scenes



Frederik S. Kamps^a, Joshua B. Julian^b, Jonas Kubilius^c, Nancy Kanwisher^d, Daniel D. Dilks^{a,*}

^a Department of Psychology, Emory University, Atlanta, GA 30322, USA

^b Department of Psychology, University of Pennsylvania, Philadelphia, PA 19104, USA

^c Laboratories of Biological and Experimental Psychology, Faculty of Psychology and Educational Sciences, KU Leuven, Leuven, Belgium

^d Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

ARTICLE INFO

Article history:

Received 1 May 2015

Accepted 21 February 2016

Available online 27 February 2016

Keywords:

OPA

Transverse Occipital Sulcus (TOS)

Parahippocampal Place Area (PPA)

Retrosplenial Complex

Scene perception

fMRI

ABSTRACT

Neuroimaging studies have identified three scene-selective regions in human cortex: parahippocampal place area (PPA), retrosplenial complex (RSC), and occipital place area (OPA). However, precisely what scene information each region represents is not clear, especially for the least studied, more posterior OPA. Here we hypothesized that OPA represents local elements of scenes within two independent, yet complementary scene descriptors: spatial boundary (i.e., the layout of external surfaces) and scene content (e.g., internal objects). If OPA processes the local elements of spatial boundary information, then it should respond to these local elements (e.g., walls) themselves, regardless of their spatial arrangement. Indeed, we found that OPA, but not PPA or RSC, responded similarly to images of intact rooms and these same rooms in which the surfaces were fractured and rearranged, disrupting the spatial boundary. Next, if OPA represents the local elements of scene content information, then it should respond more when more such local elements (e.g., furniture) are present. Indeed, we found that OPA, but not PPA or RSC, responded more to multiple than single pieces of furniture. Taken together, these findings reveal that OPA analyzes local scene elements – both in spatial boundary and scene content representation – while PPA and RSC represent global scene properties.

© 2016 Elsevier Inc. All rights reserved.

Introduction

Functional magnetic resonance imaging (fMRI) studies have reliably identified three scene-selective regions in human cortex: the parahippocampal place area (PPA) (Epstein and Kanwisher, 1998), the retrosplenial complex (RSC) (Maguire, 2001), and the occipital place area (OPA) (Dilks et al., 2013), also known as transverse occipital sulcus (TOS) (Grill-Spector, 2003) (Fig. 1). However, precisely what information each region extracts from scenes is far from clear—particularly for the least studied OPA.

More is known about information processing in face- and body-selective cortical systems. In face processing, the more posterior occipital face area (OFA) responds strongly to face parts (i.e., eyes, nose, mouth) regardless of their spatial arrangement, whereas the more anterior fusiform face area (FFA) represents face parts and the typical spatial arrangement of these parts (e.g., two eyes above a nose above a mouth) (Yovel and Kanwisher, 2004; Pitcher et al., 2007; Liu et al., 2009). Similarly, among body-selective regions, the more posterior extrastriate body area (EBA) represents body parts, with the response to body parts rising gradually as more of the body is visible (e.g., a single finger versus a hand with five fingers). By contrast, the more anterior

fusiform body area (FBA) is sensitive to the whole body, not the amount of body shown (Taylor et al., 2007). Here we ask whether the scene processing system exhibits a similar functional division of labor. In particular, we hypothesize that the more posterior OPA represents scenes at the level of local elements, while the more anterior PPA and RSC represent the global properties of scenes.

But what are the local elements of a scene? Initial clues can be found in behavioral and computational work suggesting that scenes are represented by two independent, yet complementary descriptors: i) *spatial boundary*, or the external shape, size, and scope of the space, and ii) *scene content*, or the internal features of the scene encompassing objects, textures, colors, and materials (Oliva and Torralba, 2001, 2002). Within spatial boundary representation, local scene elements may be the major surfaces and planes that together comprise the spatial boundary (i.e., walls, floors, and ceilings). Evidence for this possibility comes from the finding that PPA responds significantly more to images of intact, empty apartment rooms than to these same rooms in which the walls, floors, and ceilings were fractured and rearranged, such that they no longer defined a coherent space (Epstein and Kanwisher, 1998). This possibility also dovetails with one approach in robotic mapping that assumes that elements of the environment consist of large, flat surfaces (e.g., ceiling and walls) (Thrun, 2002). Next, within scene content representation, local scene elements may be the individual objects, textures, colors, and materials that make up the internal content of a scene (Oliva and Torralba, 2001, 2002). For example, a piece of furniture

* Corresponding author at: Department of Psychology, Emory University, 36 Eagle Row, Atlanta, GA 30322, USA.

E-mail address: dilks@emory.edu (D.D. Dilks).

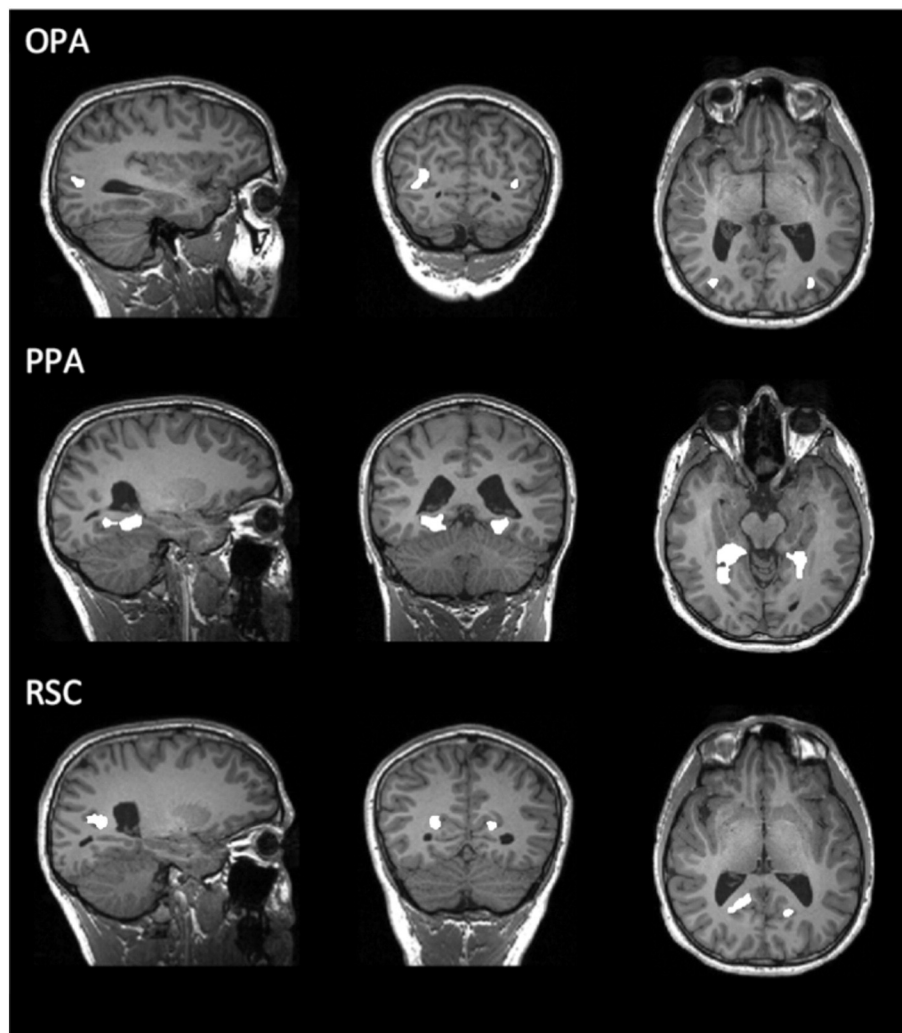


Fig. 1. Scene-selective regions of interest (ROIs) in a sample participant. Occipital Place Area (OPA), Parahippocampal Place Area (PPA), and Retrosplenial Complex (RSC), labeled accordingly. Using 'Localizer' runs, these ROIs were selected as those regions responding significantly more to scenes than objects ($p < 0.0001$, uncorrected). Responses of these ROIs to the experimental conditions were then tested using an independent set of data (Experimental runs).

might be considered a local scene element insofar as furniture is an object that is typically associated with particular places or contexts (e.g., a sofa is typically found in a living room) (Bar and Aminoff, 2003), and is different from other 'objects' because it is generally large and not portable (Mullaly and Maguire, 2011; Konkle and Oliva, 2012; Troiani et al., 2014).

To test whether the more posterior OPA represents local scene elements within both spatial boundary and scene content representation, we examined responses in OPA (as well as PPA and RSC) to images of 1) empty rooms; 2) these same rooms 'fractured' and rearranged such that the walls, floors, and ceilings no longer defined a coherent space; 3) single, nonfurniture objects; 4) single pieces of furniture; and 5) multiple pieces of furniture (Fig. 2). Within spatial boundary representation, if the more posterior OPA processes scenes at the level of local elements, then it should not represent the coherent spatial arrangement of the elements, but rather the local elements themselves. As such, we predicted that OPA would respond similarly to the empty and the fractured rooms. By contrast, if the more anterior PPA and RSC encode global representations of the spatial boundary, then they should respond more to images of empty rooms that depict a coherent layout than to images of fractured and rearranged rooms in which the spatial boundary is disrupted. Within scene content representation, if OPA is sensitive to the local elements of scenes (i.e., furniture), then it should respond more when more such elements are presented. As such, we predicted that OPA

would respond more to images of multiple pieces of furniture than to images of single pieces of furniture. By contrast, if PPA and RSC represent global properties of scene content, then their responses should be independent of the amount of content (i.e., furniture) presented.

Methods

Participants

Twenty-five participants (Age: 18–25; 12 from Emory University, 13 from MIT; 13 females, 12 males) were recruited for this experiment. Two participants were excluded from further analyses because of nonsignificant localizer results, and one for excessive motion during scanning, yielding a total of 22 participants reported here. All participants gave informed consent and had normal or corrected-to-normal vision.

Design

We used a region of interest (ROI) approach in which we localized category-selective regions (Localizer runs) and then used an independent set of runs to investigate their responses to a variety of stimulus categories (Experimental runs). For both Localizer and Experimental



Fig. 2. Example stimuli used in the Experimental runs. From top row to bottom row: 1) intact, empty apartment rooms (intact rooms); 2) indoor rooms whose walls, floors, and ceilings were fractured and rearranged such that they no longer defined a coherent space (fractured rooms); 3) single, non-furniture objects (single objects); 4) single pieces of furniture (single furniture); and 5) multiple pieces of furniture (multiple furniture).

runs, participants performed a one-back task, responding every time the same image was presented twice in a row.

For the Localizer runs, ROIs were identified using a standard method described previously (Epstein and Kanwisher, 1998). Specifically, a blocked design was used in which participants viewed images of scenes, objects, faces, and scrambled objects. Each participant completed 2–3 localizer runs. Each run was 336 s long and consisted of 4 blocks per stimulus category. The order of blocks in each run was palindromic (e.g., faces, objects, scenes, scrambled objects, scrambled objects, scenes, objects, faces, etc.) and the order of blocks in the first half of the palindromic sequence was pseudorandomized across runs. Each block contained 20 images from the same category for a total of 16 s blocks. Each image was presented for 300 ms, followed by a 500 ms interstimulus interval, and subtended $8 \times 8^\circ$ of visual angle. We also included five 16 s fixation blocks: one at the beginning, three in the middle interleaved between each palindrome, and one at the end of each run.

For the Experimental runs, participants viewed runs during which 16 s blocks (20 stimuli per block) of either 8 (at Emory) or 12 (at MIT) categories of images were presented. Five of the categories were common between Emory and MIT, and tested the central hypotheses described here; the additional categories tested unrelated hypotheses. Each image was presented for 300 ms, followed by a 500 ms interstimulus interval, and subtended $8 \times 8^\circ$ of visual angle. At Emory, participants viewed 8 runs, and each run contained 21 blocks (2 blocks of each condition, plus 5 blocks of fixation), totaling 344 s. At MIT, participants viewed 12 runs, and each run contained 16 blocks (one block of

each of the 12 different stimulus categories, and 4 blocks of fixation), totaling 256 s.

For the five categories of interest, we used the same stimuli presented in Epstein and Kanwisher (1998; indicated with an asterisk), as well as one other category (Fig. 2): (1*) photographs of apartment rooms with all furniture and objects removed (intact rooms); (2*) the same rooms but fractured into their component surfaces and rearranged such that they no longer defined a coherent space (fractured rooms); (3*) single non-furniture objects (single objects); (4) single items of furniture (single furniture); and (5*) arrays of all of the objects from one of the furnished rooms cut out from the original background and rearranged in a random configuration (multiple furniture).

fMRI scanning

All scanning was performed on a 3T Siemens Trio scanner. At Emory, scans were conducted in the Facility for Education and Research in Neuroscience. Functional images were acquired using a 32-channel head matrix coil and a gradient-echo single-shot echoplanar imaging sequence (28 slices, TR = 2 s, TE = 30 ms, voxel size = $1.5 \times 1.5 \times 2.5$ mm, and a 0.25 interslice gap). At MIT, scans were conducted at the Athinoula A. Martinos Imaging Center at the McGovern Institute for Brain Research. Functional images were acquired using a 32-channel head matrix coil and a gradient-echo single-shot echoplanar imaging sequence (28 slices, TR = 2 s, TE = 30 ms, voxel size = $1.4 \times 1.4 \times 2.0$ mm, and a 0.2 interslice gap). For all scans, slices were oriented approximately between perpendicular and parallel to the

calcarine sulcus, covering the occipital and temporal lobes. Whole-brain, high-resolution anatomical images were also acquired for each participant for purposes of registration and anatomical localization (see Data analysis).

Data analysis

fMRI data analysis was conducted using the FSL software (FMRIB's Software Library; www.fmrib.ox.ac.uk/fsl) (Smith et al., 2004) and the FreeSurfer Functional Analysis Stream (FS-FAST; <http://surfer.nmr.mgh.harvard.edu/>). ROI analysis was conducted using the FS-FAST ROI toolbox. Before statistical analysis, images were motion corrected (Cox and Jesmanowicz, 1999). Data were then detrended and fit using a double gamma function. Localizer data, but not experimental data, were spatially smoothed (5-mm kernel). After preprocessing, scene-selective regions OPA, PPA, and RSC were bilaterally defined in each participant (using data from the independent localizer scans) as those regions that responded more strongly to scenes than objects ($p < 10^{-4}$, uncorrected), as described previously (Epstein and Kanwisher, 1998) (Fig. 1). OPA, PPA, and RSC were identified in at least one hemisphere in all participants. We further defined two additional ROIs as control regions (again, using data from the localizer scans). First, we functionally defined foveal confluence (FC) – the region of cortex responding to foveal stimulation (Dougherty et al., 2003) – as the region that responded more strongly to scrambled objects than to intact objects ($p < 10^{-4}$, uncorrected), as described previously (MacEvoy and Yang, 2012; Linsley and MacEvoy, 2014; Persichetti et al., 2015). FC was identified in at least one hemisphere in all participants. Second, we functionally defined the object-selective lateral occipital complex (LOC) as the region responding more strongly to objects than scrambled objects ($p < 10^{-4}$, uncorrected), as described previously (Grill-Spector et al., 1998). LOC was identified in at least one hemisphere in all participants. Within each ROI, we then calculated the magnitude of response (percent signal change, or PSC) to the five categories of interest, using the data from the experimental runs. A 2 (hemisphere: Left, Right) \times 5 (condition: intact rooms, fractured rooms, single objects, single furniture, multiple furniture) repeated-measures ANOVA for each scene ROI was conducted. We found no significant hemisphere \times condition interaction in OPA ($p = 0.62$), PPA ($p = 0.08$), or RSC ($p = 0.43$). Thus, both hemispheres were collapsed for further analyses.

In addition to the ROI analysis described above, we also performed a group-level analysis to explore responses to the experimental conditions across the entire slice prescription. This analysis was conducted using the same parameters as were used in the ROI analysis, with the exceptions that the experimental data were spatially smoothed with a 5-mm kernel, and registered to standard stereotaxic (MNI) space. For each contrast, we performed a nonparametric one-sample t-test using the FSL randomize program (Winkler et al., 2014) with default variance smoothing of 5 mm, which tests the t value at each voxel against a null distribution generated from 5000 random permutations of group membership. The resultant statistical maps were then corrected for multiple comparisons ($p < 0.05$, FWE) using threshold-free cluster enhancement (TFCE) (Smith and Nichols, 2009).

Results

Local elements of spatial boundary

If OPA represents local scene elements (i.e., walls, floors, and ceilings) within spatial boundary representation, then it should respond to those elements regardless of how they are arranged relative to one another. To test this prediction, we first compared responses to the intact and fractured rooms in each scene-selective region individually (Fig. 3A). For OPA, a paired t-test revealed no significant difference between intact and fractured rooms ($t_{(21)} = 0.89$, $p = 0.39$), consistent with the local scene elements hypothesis for OPA. For PPA,

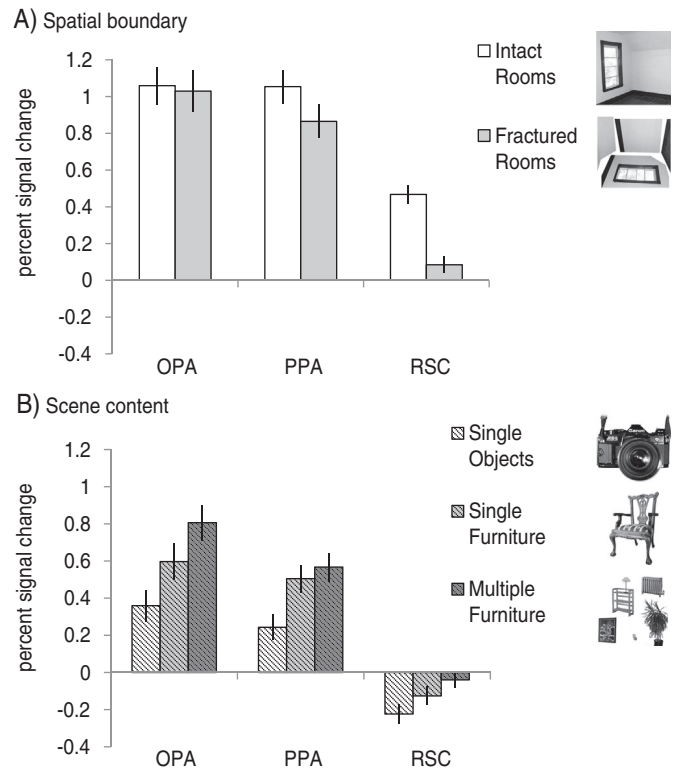


Fig. 3. Average percent signal change in OPA, PPA, and RSC to the five conditions. Error bars indicate the standard error of the mean. (A) Spatial boundary representation in OPA, PPA, and RSC. A 3 (ROI: OPA, PPA, RSC) \times 2 (room type: intact, fractured) repeated measures ANOVA revealed a significant interaction ($F_{(2, 42)} = 33.66$, $p < 0.001$), with OPA responding significantly more to fractured rooms relative to intact rooms, compared with both PPA and RSC (interaction contrasts, both p values < 0.001). (B) Scene content representation in OPA, PPA, and RSC. A 3 (ROI: OPA, PPA, RSC) \times 3 (condition: single objects, single furniture, multiple furniture) repeated-measures ANOVA revealed a significant interaction ($F_{(4, 84)} = 14.58$, $p < 0.001$), with both OPA and PPA responding more to single furniture than to single objects, relative to RSC, and OPA responding more to multiple furniture than to single furniture, relative to both PPA and RSC (interaction contrasts, all p values < 0.01).

a paired t-test revealed a significantly greater response to intact than fractured rooms ($t_{(21)} = 5.68$, $p < 0.001$), replicating previous findings (Epstein and Kanwisher, 1998). Finally, for RSC, a paired t-test revealed a significantly greater response to intact rooms than fractured rooms ($t_{(21)} = 8.81$, $p < 0.001$), consistent with previous reports of spatial boundary representation in RSC (Harel et al., 2012).

The analyses above suggest that the three scene-selective regions represent spatial boundary information differently, so next we directly compared their response profiles. A 3 (ROI: OPA, PPA, RSC) \times 2 (room type: intact, fractured) repeated measures ANOVA revealed a significant interaction ($F_{(2, 42)} = 33.66$, $p < 0.001$), with OPA responding significantly more to fractured rooms compared to intact rooms, relative to both PPA and RSC (interaction contrasts, both p values < 0.001). This result shows that the scene-selective regions represent spatial boundary information differently, with OPA representing the local elements (e.g., walls, floors, ceilings) that compose the spatial boundary, and PPA and RSC representing the global spatial arrangement of these elements relative to one another.

Local elements of scene content

Next, if OPA represents local scene elements (i.e., furniture) within scene content representation, then it should be sensitive to the number of such elements in a scene. To test this prediction, we first examined responses to images of single objects, single furniture, and multiple furniture in each scene-selective region individually (Fig. 3B). For OPA, a

three-level repeated-measures ANOVA revealed a significant main effect ($F_{(2, 42)} = 75.80, p < 0.001$), with OPA responding significantly more to both single and multiple furniture than single objects – suggesting that furniture may indeed be considered a local element of a scene – and, as predicted, significantly more to multiple furniture than single furniture (Bonferroni corrected post hoc comparisons, all p 's < 0.001).

But might OPA be sensitive to the amount of any object information more generally, rather than the amount of scene content information in particular? To address this question, in ten of our participants, we included an additional condition – images of multiple non-furniture objects (multiple objects). Importantly, the number and position of objects in the multiple object images was matched with the number and position of pieces of furniture in the multiple furniture images. A paired t-test revealed a significantly greater response to multiple furniture than multiple objects ($t_{(9)} = 4.34, p < 0.01$) in OPA, suggesting that OPA does not simply represent the number of any sort of objects, but rather scene-related objects in particular. This finding dovetails with a recent study that also found a significantly greater response to images of multiple furniture than multiple objects in OPA (referred to as TOS in that study) (Bettencourt and Xu, 2013).

For PPA, a three-level repeated-measures ANOVA revealed a significant main effect ($F_{(2, 42)} = 52.93, p < 0.001$), with PPA responding significantly more to both multiple and single furniture than single objects (Bonferroni corrected post hoc comparisons, both p 's < 0.001) – consistent with previous findings (Bar and Aminoff, 2003; Mullally and Maguire, 2011; Harel et al., 2012) – but similarly to multiple and single furniture (Bonferroni corrected post hoc comparison, $p = 0.19$). Finally, RSC did not respond above baseline to any object or furniture condition, consistent with previous findings that RSC is not sensitive to scene content (Harel et al., 2012).

The above analyses suggest that the three scene-selective regions encode scene content information differently, so next we directly tested this suggestion. A 3 (ROI: OPA, PPA, RSC) \times 3 (condition: single objects, single furniture, multiple furniture) repeated measures ANOVA revealed a significant interaction of ROI and condition ($F_{(4, 84)} = 14.58, p < 0.001$), with both OPA and PPA responding significantly more to single furniture than to single objects, relative to RSC, and critically with OPA responding more to multiple furniture than to single furniture relative to both PPA and RSC (interaction contrasts, all p values < 0.01). These results suggest that the three scene-selective regions represent scene content information differently, with OPA more sensitive than both PPA and RSC to the amount of content.

Local scene element representation in OPA does not reflect low-level visual information representation

But might the more posterior OPA – adjacent to early retinotopic regions – be responding to low-level visual information in our stimuli, like an early retinotopic region? We addressed this concern in two ways. First, given studies showing increased sensitivity to high versus low spatial frequency information in scene selective cortex (Rajimehr et al., 2011; Kauffmann et al., 2015), we performed a power spectrum analysis on our stimuli to explore differences between the five conditions across the range of spatial frequencies (Fig. 4B) (Park et al., 2014). At low spectral energy levels (10% power), a one-way ANOVA revealed significant differences between the five conditions ($F_{(4,95)} = 4.39, p < 0.005$). Post hoc analyses indicated that this effect was driven by the difference in low spatial frequencies between the intact and fractured rooms (Tukey HSD, $p = 0.001$). No other conditions significantly differed from one another (all p values > 0.05). At high spectral energy levels (90% power), a one-way ANOVA also revealed significant differences

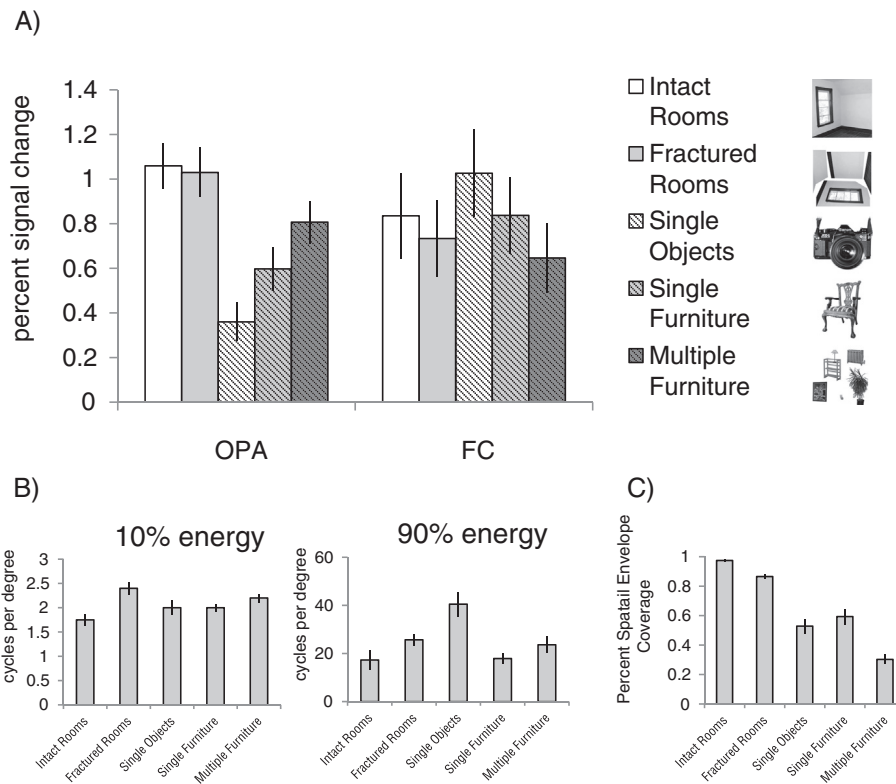


Fig. 4. (A) Average percent signal change in OPA and FC to the five conditions. Error bars indicate the standard error of the mean. A 2 (ROI: OPA, FC) \times 5 (condition: intact rooms, fractured rooms, single objects, single furniture, multiple furniture) repeated measures ANOVA revealed a significant interaction ($F_{(4,84)} = 62.03, p < 0.001$), indicating that the pattern of responses in OPA was not as expected for a retinotopic region. (B) Distribution of low (10% energy) and high (90% energy) spatial frequency information across the five conditions. (C) Average percent spatial envelope coverage in each condition. Importantly, while our stimuli were not precisely matched for spatial frequency content or spatial envelope coverage, none of these sources of low-level visual information predicts the pattern of responses observed in OPA.

between the five conditions ($F_{(4,95)} = 6.84, p < 0.001$). Post hoc analyses indicated that this effect was driven by the difference in high spatial frequency between images of single objects and all other conditions (Tukey HSD, all p values < 0.05). No other conditions significantly differed from one another (all p values > 0.45). While this analysis indicates that our stimuli were not precisely matched for spatial frequency information, crucially, neither of these patterns of results can explain the pattern observed in OPA (or PPA or RSC), ruling out the possibility that differential sensitivity to spatial frequency information can explain our findings. Second, we compared responses in OPA to those in FC, an early retinotopic region sensitive to such low level differences (Fig. 4A). A 2 (ROI: OPA, FC) \times 5 (condition: intact rooms, fractured rooms, single objects, single furniture, multiple furniture) repeated measures ANOVA revealed a significant interaction of ROI and condition ($F_{(4,84)} = 62.03, p < 0.001$), indicating that the pattern of responses in OPA was not as expected for an early retinotopic region. In particular, whereas OPA responded as predicted for a region sensitive to local elements of scenes, FC responded as expected for a low-level visual region sensitive to foveal information in two ways. First, FC should respond more to conditions with more information in the higher spatial frequency range, given that foveal cells are tuned to higher spatial frequencies than parafoveal cells (De Valois et al., 1982). Analysis of high spatial frequency information across our stimuli revealed that the single object condition had more high spatial frequency information than any other condition (Fig. 4B). Not surprisingly then, FC responded more to the single object condition than any other condition, consistent with differences in higher spatial frequency information across the conditions. Second, FC should respond more to conditions with more information presented at the fovea compared to those conditions with less information. To quantify the amount of visual information presented at the fovea in each condition, we calculated the percentage of non-white pixels in each image (in other words, how much of the image's spatial envelope was covered) (Fig. 4C). A one-way ANOVA revealed significant differences in spatial envelope coverage across the five conditions ($F_{(4,95)} = 56.78, p < 0.001$), with both the intact and fractured rooms, followed by both the single object and single furniture conditions, covering significantly more of the spatial envelope than the multiple furniture condition (Tukey's HSD, both p 's < 0.001). Not surprisingly then, FC responded more to single objects and single furniture than it did to multiple furniture, consistent with the differing amounts of information contained in the scene content conditions. Intriguingly, however, FC did not respond more to the intact and fractured rooms than single objects, as might be expected if spatial envelope explained the full response profile of FC. Taken together, the above analyses demonstrate that OPA is indeed scene-selective, rather than retinotopic.

Local scene element representation in OPA does not reflect general object representation

How does the observed sensitivity to furniture information in OPA compare to that in an object-selective region? To address this question, we compared the magnitude of responses in OPA with those in object-selective lateral occipital complex (LOC) across all five conditions (Fig. 5). A 2 (ROI: OPA, LOC) \times 5 (object type: intact rooms, fractured rooms, single object, single furniture, multiple furniture) repeated-measures ANOVA revealed a significant interaction, ($F_{(4, 84)} = 151.59, p < 0.001$), suggesting that OPA and LOC are functionally distinct, and consistent with previous reports that OPA is selectively involved in scene, not object processing (Dilks et al., 2013). The differential information processing in these regions was particularly clear in the difference between OPA and LOC responses to single objects and single furniture, where OPA responses were significantly lower to single objects than to single furniture, compared with LOC (interaction contrast, $p < 0.002$). This finding indicates that whereas LOC shows a general sensitivity to object information (i.e., responding similarly to the single object and single furniture conditions), OPA responds selectively to

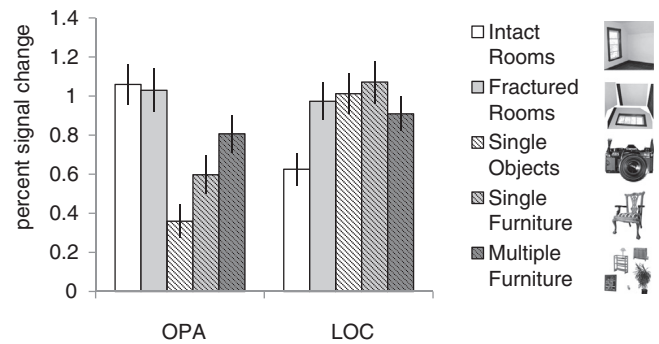


Fig. 5. Average percent signal change in OPA and object-selective LOC for all five conditions. Error bars depict the standard error of the mean. A 2 (ROI: OPA, LOC) \times 5 (condition: intact rooms, fractured rooms, single objects, single furniture, multiple furniture) repeated-measures ANOVA revealed a significant interaction ($F_{(4, 84)} = 151.59, p < 0.001$), indicating that the pattern of activity in LOC across these conditions was indeed distinct from that in OPA.

scene-related objects in particular (i.e., furniture). Next, OPA responded significantly more to images of multiple furniture than single furniture, relative to LOC (interaction contrast, $p < 0.001$). The reduced response to multiple furniture versus single furniture in LOC is consistent with an fMRI study showing that LOC responds less to arrays of multiple objects than single objects (Nasr et al., 2015), as well as with neurophysiological studies suggesting that responses in monkey temporal cortex to multiple objects are generally equal to or less than the strongest response to one stimulus (Rousselet et al., 2004). Finally, LOC responded significantly more to fractured rooms than empty rooms, relative to OPA (interaction contrast, $p < 0.001$). This finding suggests that LOC treats the major surfaces as “objects” when they are fractured from the background that defines the scene, but not when they are properly configured as part of the spatial boundary. By contrast, OPA responds similarly to local scene elements regardless of their coherent spatial arrangement. Taken together, the above analyses reveal that local scene element representation in OPA is distinct from general object representation in LOC.

Do other regions beyond OPA represent local elements of scenes?

Finally, to investigate whether cortical regions beyond the functionally defined OPA might represent local elements of scenes, we performed a group-level analysis exploring responses to our experimental conditions across the entire slice prescription. This analysis focused on two contrasts testing the central claims of the local elements hypothesis: “intact vs. fractured rooms” and “multiple vs. single furniture”. If a region represents local scene elements, then it should respond similarly to intact and fractured rooms, and more to multiple than single furniture. We found three regions exhibiting this pattern of results, including i) a region in right lateral superior occipital lobe that overlapped with a group-defined right OPA (contrast = “intact rooms vs. single objects”), consistent with the ROI analysis; ii) a set of regions in the peripheral portion of the calcarine sulcus, which likely reflect low-level visual differences between the stimuli; and iii) a region in superior parietal lobe (see Supplemental Fig. 1, Supplemental Table 1). Intriguingly, this superior parietal region showed overlapping activation with a region responding more to “intact rooms vs. single objects”, suggesting that it may further be scene-selective. We did not observe increased responses to “intact vs. fractured rooms” within the group-defined PPA or RSC; however, this lack of findings in the group analysis may reflect anatomical variability in these regions across subjects (Fedorenko et al., 2010). Thus, the present group analysis is consistent with our ROI analysis showing that OPA represents local scene elements, and provides evidence for another region in superior parietal lobe that may likewise represent local elements of scenes. Importantly, our slice prescription primarily targeted the ventral visual pathway,

leaving open the possibility that regions outside of this slice prescription may likewise represent local elements of scenes. Further, it is possible that other regions even within this slice prescription represent local scene elements, but are too anatomically variable to be detected in the present group analysis (Saxe et al., 2006; Fedorenko et al., 2010).

Discussion

Here we asked whether the more posterior OPA represents local elements of scenes, while the more anterior PPA and RSC represent global scene properties. We tested this hypothesis within two independent, yet complementary scene descriptors: spatial boundary information and scene content information. For spatial boundary representation, OPA responded similarly to images of intact rooms and these same rooms fractured and rearranged, such that they no longer define a coherent space, suggesting that OPA is sensitive to the presence of these surfaces independent of their coherent spatial arrangement. By contrast, PPA and RSC responded significantly more to the empty than the fractured rooms, indicating that these regions encode global representations of spatial boundary. For scene content representation, OPA and PPA, but not RSC, responded more strongly to furniture than other objects. Only OPA, however, was sensitive to the amount of furniture, suggesting that OPA represents local elements of scene content individually, while PPA encodes global representations of scene content that are independent of the number of such elements that make it up. Importantly, the pattern of responses in OPA was distinct from that in an early visual region (FC), was not explained by low-level visual information (i.e., high or low spatial frequency information, spatial envelope coverage), and further was distinct from the pattern of responses in object selective cortex (LOC). These findings are consistent with the hypothesis that OPA analyzes scenes at the level of local elements—both in spatial boundary and scene content representation—while PPA and RSC represent global properties of scenes. This work may provide a missing piece of evidence for an overarching functional organization of information processing in ventral visual cortex, where more posterior regions (OPA, OFA, EBA) are sensitive to local elements, while more anterior regions (PPA, FFA, FBA) represent global configurations (Yovel and Kanwisher, 2004; Pitcher et al., 2007; Taylor et al., 2007; Liu et al., 2009). Note that this functional division of labor does not necessarily imply a particular hierarchy of processing (i.e., parts processed before wholes), but rather simply reflects a common division of labor across category-selective cortex in the ventral visual pathway.

Our findings showing that PPA represents both spatial boundary and scene content are consistent with many other studies (Epstein and Kanwisher, 1998; Janzen and van Turenout, 2004; Kravitz et al., 2011; MacEvoy and Epstein, 2011; Mullally and Maguire, 2011; Park et al., 2011; Harel et al., 2012; Bettencourt and Xu, 2013). Further, our finding that RSC responded to intact over fractured rooms, and not at all to objects or furniture, is consistent with previous reports of spatial boundary sensitivity, but not scene content sensitivity, in RSC (Maguire, 2001; Ino et al., 2002; Epstein et al., 2007; Park and Chun, 2009; Harel et al., 2012). Finally, while some recent studies have focused on retinotopic and low-level visual functions in these regions (Rajimehr et al., 2011; Nasr et al., 2014; Watson et al., 2014; Kauffmann et al., 2015; Silson et al., 2015; Watson et al., 2016), our findings suggest that scene-selective cortex represents properties of scenes independent of such lower-level visual representations—even in the case of OPA, which lies immediately adjacent to the retinotopically defined area V3A (Grill-Spector, 2003; Nasr et al., 2011).

In contrast to PPA and RSC, little is known about information processing in OPA. For example, this is the first study to our knowledge to explore spatial boundary processing in OPA. Another study has explored scene content processing in OPA, and found that OPA was sensitive to multiple properties of objects presented in isolation, including physical size, fixedness, “placeness”, and the degree to which the object is “space defining” (Troiani et al., 2014), consistent with our finding that OPA

responds more to furniture than non-furniture objects. Note that the increased response to furniture in OPA does not necessarily indicate that OPA represents the object category of “furniture” per se, nor the precise identity of the local elements (e.g., a “chair”); rather, OPA may be sensitive to multiple mid-level object properties commonly captured by the category of furniture (e.g., furniture tends to be larger and more fixed compared with other non-furniture objects). Finally, while the present study tested the local elements hypothesis in the context of indoor scenes only, this hypothesis is readily extended to outdoor scenes as well, where OPA may likewise represent major surfaces that make up the spatial boundary (e.g., the ground plane, the side of a building) and large, fixed objects that make up the scene content (e.g., a parked car, a bench).

Beyond spatial boundary and content information, two recent studies have reported that OPA encodes ‘sense’ (left/right) (Dilks et al., 2011) and egocentric distance information (Persichetti and Dilks, in press), suggesting a role for OPA in navigation. At first glance, our finding that OPA is relatively insensitive to spatial boundary appears inconsistent with the role of OPA in navigation, insofar as fractured rooms do not imply a navigable space. However, we propose that while OPA may not represent allocentric spatial relationships between local components of scenes, such as how floors are arranged relative to walls, it may nevertheless represent egocentric spatial information about local scene elements, such as the distance and direction of boundaries (e.g., walls) and obstacles (e.g., furniture) relative to the viewer. Thus, our hypothesis that OPA represents the local elements of scenes is compatible with the proposed role of OPA in navigation, and may point to a role for OPA in visually-guided navigation and obstacle avoidance.

In conclusion, we found differential representation of spatial boundary and scene content information across scene-selective cortex. Unlike PPA and RSC, OPA does not represent the spatial boundary of a scene per se, but rather responds to the surfaces that make up that spatial boundary regardless of their arrangement. Further OPA represents scene content and the amount of such content, unlike PPA and RSC, suggesting that OPA encodes local, individual elements of scene content. Together, these findings support the hypothesis that OPA represents the local elements of scenes, while PPA and RSC represent global scene properties.

Acknowledgments

We would like to thank the Facility for Education and Research in Neuroscience (FERN) Imaging Center in the Department of Psychology, Emory University, Atlanta, GA, as well as the Athinoula A. Martinos Imaging Center at the McGovern Institute for Brain Research, MIT, Cambridge, MA. We would also like to thank Samuel Weiller and Andrew Persichetti for their technical support and insightful comments. The work was supported by Emory College, Emory University (DD), National Institute of Child Health and Human Development grant T32HD071845 (FK), and National Institutes of Health grant EY013455 (NK). Jonas Kubilius currently is a research assistant of the Research Foundation—Flanders (FWO). The authors declare no competing financial interests.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.neuroimage.2016.02.062>.

References

- Bar, M., Aminoff, E., 2003. Cortical analysis of visual context. *Neuron* 38, 347–358.
- Bettencourt, K.C., Xu, Y.D., 2013. The role of transverse occipital sulcus in scene perception and its relationship to object individuation in inferior intraparietal sulcus. *J. Cogn. Neurosci.* 25, 1711–1722.
- Cox, R.W., Jesmanowicz, A., 1999. Real-time 3D image registration for functional MRI. *Magn. Reson. Med.* 42, 1014–1018.
- De Valois, R.L., Albrecht, D.G., Thorell, L.G., 1982. Spatial frequency selectivity of cells in macaque visual cortex. *Vis. Res.* 22, 545–559.

- Dilks, D.D., Julian, J.B., Kubilius, J., Spelke, E.S., Kanwisher, N., 2011. Mirror-image sensitivity and invariance in object and scene processing pathways. *J. Neurosci.* 31, 11305–11312.
- 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, 1331–1336a.
- Dougherty, R.F., Koch, V.M., Brewer, A.A., Fischer, B., Modersitzki, J., Wandell, B.A., 2003. Visual field representations and locations of visual areas V1/2/3 in human visual cortex. *J. Vis.* 3, 586–598.
- Epstein, Kanwisher, 1998. A cortical representation of the local visual environment. *Nature* 392, 598–601.
- 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.
- Fedorenko, E., Hsieh, P.J., Nieto-Castanon, A., Whitfield-Gabrieli, S., Kanwisher, N., 2010. New method for fMRI investigations of language: defining ROIs functionally in individual subjects. *J. Neurophysiol.* 104, 1177–1194.
- Grill-Spector, K., 2003. The neural basis of object perception. *Curr. Opin. Neurobiol.* 13, 159–166.
- Grill-Spector, K., Kushnir, T., Hendler, T., Edelman, S., Itzhak, Y., Malach, R., 1998. A sequence of object-processing stages revealed by fMRI in the human occipital lobe. *Hum. Brain Mapp.* 6, 316–328.
- Harel, A., Kravitz, D.J., Baker, C.I., 2012. Deconstructing visual scenes in cortex: gradients of object and spatial layout information. *J. Mol. Neurosci.* 48, S50–S50.
- Ino, T., Inoue, Y., Kage, M., Hirose, S., Kimura, T., Fukuyama, H., 2002. Mental navigation in humans is processed in the anterior bank of the parieto-occipital sulcus. *Neurosci. Lett.* 322, 182–186.
- Janzen, G., van Turenout, M., 2004. Selective neural representation of objects relevant for navigation. *Nat. Neurosci.* 7, 673–677.
- Kauffmann, L., Ramanan, S., Guyader, N., Chauvin, A., Peyrin, C., 2015. Spatial frequency processing in scene-selective cortical regions. *NeuroImage* 112, 86–95.
- Konkle, T., Oliva, A., 2012. A real-world size organization of object responses in occipitotemporal cortex. *Neuron* 74, 1114–1124.
- 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.
- Linsley, D., MacEvoy, S.P., 2014. Evidence for participation by object-selective visual cortex in scene category judgments. *J. Vis.* 14.
- Liu, J., Harris, A., Kanwisher, N., 2009. Perception of face parts and face configurations: an fMRI study. *J. Cogn. Neurosci.* 22, 203–211.
- MacEvoy, S.P., Epstein, R.A., 2011. Constructing scenes from objects in human occipitotemporal cortex. *Nat. Neurosci.* 14, 1323–1329.
- MacEvoy, S.P., Yang, Z., 2012. Joint neuronal tuning for object form and position in the human lateral occipital complex. *NeuroImage* 63, 1901–1908.
- Maguire, E.A., 2001. The retrosplenial contribution to human navigation: a review of lesion and neuroimaging findings. *Scand. J. Psychol.* 42, 225–238.
- Mullally, S.L., Maguire, E.A., 2011. A new role for the parahippocampal cortex in representing space. *J. Neurosci.* 31, 7441–7449.
- Nasr, S., Liu, N., Devaney, K.J., Yue, X., Rajimehr, R., Ungerleider, L.G., Tootell, R.B., 2011. Scene-selective cortical regions in human and nonhuman primates. *J. Neurosci.* 31, 13771–13785.
- 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., Stemmann, H., Vanduffel, W., Tootell, R.B., 2015. Increased visual stimulation systematically decreases activity in lateral intermediate cortex. *Cereb. Cortex* 25, 4009–4028.
- Oliva, A., Torralba, A., 2001. Modeling the shape of a scene: a holistic representation of the spatial envelope. *Int. J. Comput. Vis.* 42, 145–175.
- Oliva, A., Torralba, A., 2002. Scene-centered description from spatial envelope properties. *Lect. Notes Comput. Sci.* 2525, 263–272.
- 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.
- 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., Konkle, T., Oliva, A., 2014. Parametric coding of the size and clutter of natural scenes in the human brain. *Cereb. Cortex.*
- Persichetti, A.S., Dilks, D.D., 2016. Perceived egocentric distance sensitivity and invariance across scene-selective cortex. *Cortex* (in press).
- Persichetti, A.S., Aguirre, G.K., Thompson-Schill, S.L., 2015. Value is in the eye of the beholder: early visual cortex codes monetary value of objects during a diverted attention task. *J. Cogn. Neurosci.* 27, 893–901.
- Pitcher, D., Walsh, V., Yovel, G., Duchaine, B., 2007. TMS evidence for the involvement of the right occipital face area in early face processing. *Curr. Biol.* 17, 1568–1573.
- Rajimehr, R., Devaney, K.J., Bilenko, N.Y., Young, J.C., Tootell, R.B., 2011. The "parahippocampal place area" responds preferentially to high spatial frequencies in humans and monkeys. *PLoS Biol.* 9, e1000608.
- Rousselet, G.A., Thorpe, S.J., Fabre-Thorpe, M., 2004. How parallel is visual processing in the ventral pathway? *Trends Cogn. Sci.* 8, 363–370.
- Saxe, R., Brett, M., Kanwisher, N., 2006. Divide and conquer: a defense of functional localizers. *NeuroImage* 30, 1088–1096 (discussion 1097–1089).
- Silson, E.H., Chan, A.W., Reynolds, R.C., Kravitz, D.J., Baker, C.I., 2015. A retinotopic basis for the division of high-level scene processing between lateral and ventral human occipitotemporal cortex. *J. Neurosci.* 35, 11921–11935.
- Smith, S.M., Nichols, T.E., 2009. Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. *NeuroImage* 44, 83–98.
- Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E.J., Johansen-Berg, H., Bannister, P.R., De Luca, M., Drobnjak, I., Flitney, D.E., Niazy, R.K., Saunders, J., Vickers, J., Zhang, Y.Y., De Stefano, N., Brady, J.M., Matthews, P.M., 2004. Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage* 23, S208–S219.
- Taylor, J.C., Wiggett, A.J., Downing, P.E., 2007. Functional MRI analysis of body and body part representations in the extrastriate and fusiform body areas. *J. Neurophysiol.* 98, 1626–1633.
- Thrun, S., 2002. Robotic Mapping: A Survey. In: Lakemeyer, G., Nebel, B. (Eds.), *Exploring Artificial Intelligence in the New Millennium*. Morgan Kaufmann.
- Troiani, V., Stigliani, A., Smith, M.E., Epstein, R.A., 2014. Multiple object properties drive scene-selective regions. *Cereb. Cortex* 24, 883–897.
- 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.
- Watson, D.M., Hymers, M., Hartley, T., Andrews, T.J., 2016. Patterns of neural response in scene-selective regions of the human brain are affected by low-level manipulations of spatial frequency. *NeuroImage* 124, 107–117.
- Winkler, A.M., Ridgway, G.R., Webster, M.A., Smith, S.M., Nichols, T.E., 2014. Permutation inference for the general linear model. *NeuroImage* 92, 381–397.
- Yovel, G., Kanwisher, N., 2004. Face perception: domain specific, not process specific. *Neuron* 44, 889–898.