Concavity as a diagnostic feature of visual scenes

Annie Cheng[a], **Dirk B. Walther**[b], **Soojin Park**[c,*], **Daniel D. Dilks**[a,*]

[a]Department of Psychology, Emory University, Atlanta, GA 30322, USA

[b]Department of Psychology, University of Toronto, Toronto, ON, Canada

[c]Department of Psychology, Yonsei University, Seoul, Republic of Korea

Abstract

Despite over two decades of research on the neural mechanisms underlying human visual scene, or place, processing, it remains unknown what exactly a "scene" is. Intuitively, we are always *inside* a scene, while interacting with the *outside* of objects. Hence, we hypothesize that one diagnostic feature of a scene may be concavity, portraying "inside", and predict that if concavity is a scene-diagnostic feature, then: 1) images that depict concavity, even non-scene images (e.g., the "inside" of an object – or concave object), will be behaviorally categorized as scenes more often than those that depict convexity, and 2) the cortical scene-processing system will respond more to concave images than to convex images. As predicted, participants categorized concave objects as scenes more often than convex objects, and, using functional magnetic resonance imaging (fMRI), two scene-selective cortical regions (the parahippocampal place area, PPA, and the occipital place area, OPA) responded significantly more to concave than convex objects. Surprisingly, we found no behavioral or neural differences between images of concave versus convex buildings. However, in a follow-up experiment, using tightly-controlled images, we unmasked a selective sensitivity to concavity over convexity of scene boundaries (i.e., walls) in PPA and OPA. Furthermore, we found that even highly impoverished line drawings of concave shapes are behaviorally categorized as scenes more often than convex shapes. Together, these results provide converging behavioral and neural evidence that concavity is a diagnostic feature of visual scenes.

Keywords

Parahippocampal place area; Occipital place area; Scene perception; Scene selectivity; High-level vision; Lateral occipital complex; fMRI

1. Introduction

Over the past two decades, cognitive neuroscientists have discovered dedicated regions of the human visual cortex that are selectively involved in visual scene, or place, processing. These cortical regions are said to be "scene selective" as they respond about two to four times more strongly to images of scenes (e.g., a kitchen) than to non-scene images, such as everyday objects (e.g., a cup) and faces. Furthermore, this selectivity persists across highly variable scene images that seem to share no common visual features, from images of empty rooms to landscapes, and, even more impressively, across scene images of different formats (e.g., color, grayscale, artificially rendered, and line-drawings) (Dillon et al., 2018; Epstein and Kanwisher, 1998; Kamps et al., 2016; Walther et al., 2011). Given the selective response to such a disparate set of scene images, how does the brain "know" what a scene is? In other words, what drives the scene-selective cortical regions to show a preferential response across highly variable scene stimuli over non-scene stimuli?

One proposed explanation is that the scene-selective cortical regions are tuned for low-level visual features that are more commonly found in scene stimuli than in non-scene stimuli (e.g., objects and faces). For example, the parahippocampal place area (PPA) (Epstein and Kanwisher, 1998) shows a preferential response to high-spatial frequency over low-spatial frequency information (Berman et al., 2017; Kauffmann et al., 2014; Rajimehr et al., 2011), cardinal orientations over oblique orientations (Nasr and Tootell, 2012), and rectilinear edges over curvilinear edges (Nasr et al., 2014) (for review, see Groen et al., 2017). However, a recent study (Bryan et al., 2016) found that one of these low-level features – rectilinearity – is not sufficient to explain the scene selectivity in PPA, casting doubt on the low-level explanation more generally. Why might the low-level visual features be insufficient to explain cortical scene selectivity? One possible reason is that these low-level features are not only common to scene stimuli, but also prominent in non-scene stimuli, especially objects, making them potentially unreliable diagnostic features of scenes. For a visual feature to be a reliable diagnostic feature of scenes, it should be not only more common, but also unique to scenes. Thus, what are those visual features that are unique to scene stimuli that may drive cortical scene selectivity?

Intuitively, we are always *inside* a scene, while interacting with the *outside* of objects; hence, we hypothesize that a potential diagnostic feature of a scene is concavity – a visual feature that conveys a viewer's state of being inside a space. Indeed, numerous neuroimaging studies highlight PPA's sensitivity to properties associated with concavity of space. For example, PPA i) responds more to images of empty rooms that depict a coherent spatial layout than to images of the same rooms but fractured and rearranged in which the spatial layout is disrupted (Epstein and Kanwisher, 1998; Kamps et al., 2016), ii) differentiates between open versus closed scenes (e.g. beach vs. urban street) (Kravitz et al., 2011; Park et al., 2011), iii) parametrically codes for the variation in size of indoor spaces (e.g. closet vs. auditorium) (Park et al., 2014), and iv) parametrically responds to the boundary height of a scene (Ferrara and Park, 2016). Here, in four experiments, using psychophysics and functional magnetic resonance imaging (fMRI), we directly test our hypothesis and predict that if concavity is a diagnostic feature of a scene, then 1) images of both scenes and objects that depict concavity will be behaviorally categorized as scenes more often than images of

scenes and objects that depict convexity (Experiment 1), 2) PPA will similarly respond more to concave scenes and objects than convex scenes and objects (Experiments 2 and 3), and 3) even highly impoverished line drawings of concave shapes will be behaviorally categorized as scenes more often than convex shapes (Experiment 4).

2. Materials and methods

2.1. Participants

Fifty participants were recruited from Amazon Mechanical Turk to participate in Experiment 1. No participants were excluded. Fifteen participants (Age: 19–32; 10 females) were recruited from the Emory University community to participate in Experiment 2, and no participants were excluded. A separate group of sixteen participants (Age: 18–31; 6 females) was recruited from the Johns Hopkins University community to participate in Experiment 3. In Experiment 3, two subjects were excluded: one due to excessive motion during scanning, and the other due to an inability to define any significant regions of interest (see Data Analysis). Thus, a total of fourteen subjects were analyzed in Experiment 3. A separate group of a hundred participants was recruited from Amazon Mechanical Turk to participate in Experiment 4; one subject was excluded from analysis due to incomplete response. The number of participants for each experiment was based on power analyses conducted before recruitment (see Supplementary Analysis 1). All participants gave informed consent and had normal or corrected-to-normal vision.

2.2. Visual stimuli

In Experiments 1 and 2, eleven real-world images of both concave and convex scenes (i.e., buildings) and objects (e.g., microwave, tissue box, etc.) were collected, totaling forty-four images (Fig. 1 A). During stimuli selection, we took three specific measures to eliminate potential confounds. First, given PPA's sensitivity to high spatial frequency information (Rajimehr et al., 2011), we matched the quantity of energy in the high spatial frequency range (qHF; >10 cycle/image) (Bainbridge and Oliva, 2015) across the conditions. A one-way ANOVA (Condition: Concave Buildings, Convex Buildings, Concave Objects, Convex Objects) revealed no main effect of Condition ($F_{(3,40)} = 0.57$, $p = .64$) in the qHF of the stimuli. Second, given PPA's preference for rectilinear over curvilinear features (Nasr et al., 2014), we included only images of rectilinear buildings and objects to eliminate the potential influence of a rectilinear bias. Third, we carefully selected stimuli with little to no words or symbols in an image to ensure participants are basing their judgment on the visual features of the stimuli.

In Experiment 3, artificial scene images of seven wall boundary conditions were created using Google SketchUp and Adobe Photoshop CS6 (Fig. 3 A). We first created three conditions of Concave boundaries (Concave 1, 2, 3) that vary parametrically in the angle at which two wall surfaces conjoin. To avoid the texture pattern on the boundary surfaces as a confounding variable, we applied the same set of 24 textures on the boundary surfaces in each condition. Furthermore, to ensure that the stimuli are scene like, we applied a light blue color to the portion of the image above the wall boundary and a green color to the portion of the image below the wall boundary to resemble the sky and field in an outdoor environment.

After creating the three Concave conditions, we swapped the left and right halves of each of the Concave condition stimulus to create the corresponding Convex conditions (Convex 1, 2, 3). As such, all low-level visual properties, including hue values, texture gradient of the wall pattern, and even the number of pixels that made up the boundary, were exactly equated between corresponding Concave and Convex conditions (e.g., Concave 1 vs. Convex 1). Finally, we also included a "flat" boundary condition as a baseline condition to estimate the neural response of a neural region to wall boundaries with no concavity nor convexity.

In Experiment 4, we created a stripped-down, line-drawing version of the stimuli conditions in Experiment 3, totaling seven conditions: Concave 1, 2, 3, Convex 1, 2, 3, and Flat. The shape and angle of the conditions exactly matched those in Experiment 3. To eliminate confounding visual features as much as possible, each condition is composed of only six connected lines, which are the minimal visual cues needed for constructing line drawings of concave and convex shapes.

2.3. Experimental design

In Experiment 1, each participant completed 4 practice trials (1 per condition: Concave Buildings, Convex Buildings, Concave Objects, Convex Objects) and 40 experimental trials (10 per condition). In each trial for both practice trials and experimental trials, an image was presented for 150 ms, followed by a blank screen for 400 ms, after which participants were asked to indicate whether the image they just saw was a "place" or an "object" (Fig. 1 B). Condition and image order were randomized. To ensure that participants had sufficient time to perceive and categorize the images, we decided on 150 ms as the image presentation time, since we know that humans take approximately 100 ms of image presentation to reach maximum performance on a basic-level scene categorization task (e.g., recognize an image as a forest, beach, etc.) (Greene and Oliva, 2009).

In Experiments 2 and 3, we used a region of interest (ROI) approach in which we localized the cortical regions of interest with the Localizer runs, and then used an independent set of Experimental runs to investigate the responses of these regions when viewing blocks of images from the stimulus categories of interest. Since PPA has been the center of the theoretical debates regarding scene selectivity, PPA is our primary ROI of interest. We, however, also examined the neural response in the two other known scene-selective regions —the occipital place area (OPA) (Dilks et al., 2013) and the retrosplenial complex (RSC) (Maguire, 2001) also known as the medial place area (MPA) (Silson et al., 2016). Finally, we also examined the neural response of an object-selective region (lateral occipital complex, LOC) and the primary visual cortex (V1; localized with independent Meridian Map runs) as control regions.

For the Localizer runs, we used a blocked design in which participants viewed images of faces, objects, scenes, and scrambled objects, as previously described (Epstein and Kanwisher, 1998), and widely used by multiple labs (e.g., Epstein et al., 2003; Cant and Goodale, 2011; Walther et al., 2011; Kravitz et al., 2011; Park et al., 2011; Golumb and Kanwisher, 2012; Ferrara and Park, 2016; Persichetti and Dilks, 2019). The scene stimuli include images from a wide variety of scene categories, such as indoor scenes, buildings, city streets, forest streams, and fields.

In both experiments, each participant completed two Localizer runs. In Experiment 2, each Localizer run was 336 s long. There were four blocks per stimulus category within each run, and 20 images from the same category within each block. Each image was presented for 300 ms, followed by a 500 ms ISI for a total of 16 s per block. Image order within each block was randomized. The order of the blocks in each run was palindromic, and the order of the blocks in the first half of the palindromic sequence was pseudo-randomized across runs. Five 16 s fixation blocks were included: one at the beginning, three in the middle interleaved between each set of stimulus blocks, and one at the end of each run. Participants performed a one-back repetition detection task, responding every time the same image was presented twice in a row.

In Experiment 3, each Localizer run was 426 s long. There were four blocks per stimulus category within each run, and 20 images from the same category within each block. Each image was presented for 600 ms, followed by a 200 ms interstimulus interval (ISI) for a total of 16 s per block. Image order within each block, as well as the order of blocks in each run, were randomized. Seventeen 10 s fixation blocks were included: one at the beginning, and one following each stimulus block. Participants performed a one-back repetition detection task, responding every time the same image was presented twice in a row.

For the Meridian Map runs, we used a blocked design in which participants viewed two flickering triangular wedges of checkerboards oriented either vertically (vertical meridians) or horizontally (horizontal meridians) to delineate borders of retinotopic areas, as previously described (Sereno et al., 1995; Spiridon and Kanwisher, 2002). In both experiments, each participant completed one run. Each run was 252 s long. There were five blocks per visual field meridians within a run. Each block flickered at 8Hz for a total of 12 s per block. Vertical meridians were always presented first, followed by horizontal meridians. Each block was preceded by a 12 s fixation period. Participants were instructed to fixate on a fixation dot during the display.

For the Experimental runs of Experiments 2 and 3, we used a block design in which participants viewed blocks of images from each condition of interest (see the Visual Stimuli section). In Experiment 2, participants completed four Experimental runs. Each run was 328 s long. There were four blocks per condition of interest within each run, and 12 images from the same condition within each block. Each image was presented for 500 ms, followed by a 500 ms ISI for a total of 12 s per block. Image order within each block, and the order of blocks in each run were randomized. Each block was preceded by an 8 s fixation block. Participants performed a one-back repetition detection task, responding every time the same image was presented twice in a row.

In Experiment 3, participants completed ten Experimental runs. Each run was 288 s long. There were two blocks per condition within each run, and 12 images from the same condition within each block. Each image was presented for 800 ms, followed by a 200 ms ISI for a total of 12 s per block. Image order within each block, and the order of blocks in each run were randomized. Each block was preceded by an 8 s fixation block. Participants performed a one-back repetition detection task, responding every time the same image was presented twice in a row.

In Experiment 4, each participant completed 7 trials. In each trial, an image was presented for an unlimited time, and participants were asked to indicate whether the image they saw was a "scene" or an "object". Condition order was randomized.

2.4. MRI scan parameters

In Experiment 2, scanning was done on a 3T Siemens Trio scanner at the Facility for Education and Research in Neuroscience (FERN) at Emory University (Atlanta, GA). Functional images were acquired using a 32-channel head matrix coil and a gradient echo single-shot echo planar imaging sequence. Twenty-eight slices were acquired for all runs: repetition time = 2 s; echo time = 30 ms; flip angle = 90°; voxel size = $1.5 \times 1.5 \times 2.5$ mm with a 0.2 mm interslice gap; and slices were oriented approximately between perpendicular and parallel to the calcarine sulcus, covering the occipital as well as the posterior portion of temporal lobes. Whole-brain, high-resolution T1-weighted anatomical images were also acquired with $1 \times 1 \times 1$ mm voxels.

In Experiment 3, scanning was done on a 3T Phillips fMRI scanner at the Kirby Research Center at Johns Hopkins University (Baltimore, MD). Functional images were acquired with a 32-channel head coil and a gradient echo-planar T2* sequence. Thirty-six slices were acquired for all runs: repetition time = 2 s; echo time = 30 ms; flip angle = 70°; voxel size = $2.5 \times 2.5 \times 2.5$ mm with 0.5 mm interslice gap; slices were oriented parallel to the anterior commissure-posterior commissure line. Whole-brain, high-resolution T1-weighted anatomical images were acquired using a magnetization-prepared rapid-acquisition gradient echo (MPRAGE) with $1 \times 1 \times 1$ mm voxels.

2.5. Data analysis

In Experiment 1, we calculated the proportion of place ratings for each condition for each subject and then we averaged the proportion of place-ratings for each condition across all subjects.

In Experiment 2, fMRI data were processed in FSL software (Smith et al., 2004) and the FreeSurfer Functional Analysis Stream (FS-FAST). Data were analyzed in each participant's native space. Preprocessing included skull-stripping (Smith, 2002), linear-trend removal, and three-dimensional motion correction using FSL's MCFLIRT tool. Data were then fit using a double gamma function, and spatially smoothed with a 4-mm kernel. No participant exhibited excessive head movement and hence no data were excluded.

In Experiment 3, fMRI data were processed in Brain Voyager QX software (Brain Innovation, Maastricht, Netherlands). Data were analyzed in each participant's ACPC-aligned space. Preprocessing included slice scan-time correction, linear trend removal, and three-dimensional motion correction. Data were then fit using a double gamma function, with no additional spatial or temporal smoothing performed. Experimental runs with excessive head movement (more than 3 mm deviation from the origin within a run and more than 8mm deviation from the origin across runs) were excluded from the analyses; one participant was excluded.

After preprocessing, the ROIs were bilaterally defined in each participant using data from the Localizer runs. PPA, OPA and RSC were defined as those regions that responded more strongly to scenes than objects ($p<10^{-4}$, uncorrected), whereas LOC was defined as those regions that responded more strongly to objects than scrambled objects, following the conventional method of Epstein and Kanwisher (1998) and Grill-Spector et al. (1998). To define V1, a contrast between the vertical and horizontal meridians from the Meridian Map runs were mapped onto a surface-rendered brain of each hemisphere, and the retinotopic borders of V1 were defined by the lower and upper vertical meridians.

In Experiment 2, PPA, LOC and V1 were defined in at least one hemisphere in all fifteen participants, while OPA was defined in at least one hemisphere in fourteen participants, and RSC in at least one hemisphere in twelve participants. For each ROI of each participant, the average response across voxels for each condition was extracted and converted to percent signal change (PSC) relative to fixation, and repeated-measures ANOVAs were performed. In Experiment 3, PPA was defined in at least one hemisphere in all fourteen participants; LOC and OPA were defined in at least one hemisphere in thirteen participants; V1 was defined in at least one hemisphere in twelve participants; and RSC was defined in at least one hemisphere in eight participants. The average response across voxels for each condition in an ROI was extracted and converted to PSC relative to fixation. To specifically examine the response to concavity and convexity in each ROI, we subtracted the PSC for each Concave and Convex condition from the Flat boundary condition to eliminate an ROI's baseline response to the presence of a boundary. To check whether having a smaller sample size for some ROIs and/or having different numbers of unilateral versus bilateral ROIs might affect the results, we defined the ROIs in the missing participants using a lower threshold (a minimum of $p<.01$) in the Localizer contrasts and examined the neural response of each ROI with these additional data—we found consistent response patterns (see Supplementary Analysis 2); thus, we proceeded with our data analyses using only the ROIs defined with the typical threshold for the Localizer runs. Finally, in Experiment 2, a 3 (ROI: PPA, LOC, V1) $\times$ 2 (Hemisphere: Left, Right) $\times$ 4 (Condition: Concave Buildings, Convex Buildings, Concave Objects, Convex Objects) repeated-measures ANOVA revealed no significant ROI $\times$ Hemisphere $\times$ Condition interaction ($F_{(6,78)} = 0.57$, $p = .75$, $\eta_p^2 = .04$). In Experiment 3, a 3 (ROI: PPA, LOC, V1) $\times$ 2 (Hemisphere: Left, Right) $\times$ 6 (Condition: Concave 1, 2, 3, Convex 1, 2, 3) repeated-measures ANOVA revealed no significant ROI $\times$ Hemisphere $\times$ Condition interaction ($F_{(10,40)} = 0.80$, $p = .63$, $\eta_p^2 = .17$). Thus, data from the left and right hemispheres of the same ROI for each Experiment were collapsed.

In addition to the ROI analysis described above, we also conducted a group-level, whole-brain analysis to explore the topography of activation over the entire visual cortex. In Experiment 2, these analyses were conducted using the same parameters as described above, except that the functional data were registered to the standard stereotaxic (MNI) space. We then contrasted the neural response for Concave versus Convex Objects across participants, and performed a voxel-wise nonparametric one-sample t-test using the FSL randomise program (Smith and Nichols, 2009) with default variance smoothing of 5 mm, which tests the t-value at each voxel against a null distribution generated from randomly flipping the sign of participants' t statistic maps 5000 random permutations of group membership. In Experiment 3, these analyses were also conducted using the same parameters as described in

the paragraphs above, except that the functional data was transformed into the standard Talairach (TAL) space, and spatially smoothed with a 4-mm kernel to enhance the spatial correspondence across subjects. We then contrasted the neural response for Concave versus Convex Boundary (combined across Angles) using a group-level fixed-effect analysis. Note that since our main goal was to examine the topography of activation and is exploratory in nature, we did not correct for multiple comparisons. Finally, to examine the topography of activation relative to the ROIs, we defined the group-level ROIs on the cortical surface using independent data from the Localizer runs (PPA, OPA, RSC: Scene – Object; LOC: Object – Scrambled).

In Experiment 4, we calculated the proportion of scene ratings for each condition and then tested the proportion against chance (i.e., 50%) using a Binomial Test. We also directly compared the proportion of scene ratings between the corresponding Concave and Convex Angle using the McNemar's test.

3. Results

3.1. Experiment 1

If concavity is a diagnostic feature of visual scenes, then images of both scenes and objects that depict concavity will be behaviorally categorized as scenes more often than images of scenes and objects that depict convexity. To test this prediction, we conducted a 2 (Category: Buildings, Objects) × 2 (Condition: Concave, Convex) repeated-measures ANOVA (Fig. 1 C). We found a significant main effect of Category, with a significantly greater proportion of place ratings for Buildings over Objects ($F_{(1,49)} = 1273.24$, $p<.001$, $\eta_p^2 = .96$), revealing that participants understood the task and had a clear sense of what is a "place" versus an "object". We also found a significant main effect of Condition ($F_{(1,49)} = 43.56$, $p<.001$, $\eta_p^2 = .47$), with a significantly greater proportion of place ratings for Concave images over Convex images, consistent with our hypothesis that images that depict concavity will be behaviorally categorized as scenes more often than images that depict convexity. Interestingly, however, we also found a significant Category × Condition interaction ($F_{(1,49)} = 102.35$, $p<.001$, $\eta_p^2 = .68$), with a significantly greater proportion of place ratings for Concave Objects than Convex Objects (post hoc comparison, $p<.001$), yet no difference between Concave Buildings and Convex Buildings ($p = .75$).

The above analysis suggests that while Concave Objects are indeed behaviorally categorized as scenes more often than Convex Objects, as hypothesized, participants however considered both Concave and Convex Buildings to be equally scene like, presenting seemingly contradictory results to our hypothesis. Why might there be no difference for place ratings of Concave versus Convex Buildings, when there is such a striking effect of concavity in biasing scene categorization of objects? A closer examination of the place ratings suggests that there may be a ceiling effect masking the effect of concavity in biasing scene categorization of Buildings. Specifically, participants reported 97.2% of Concave Buildings and 97.8% of Convex Buildings as places, demonstrating a nearly perfect categorization performance. Thus, it is plausible that, since we did not control for some visual features orthogonal to concavity and convexity (e.g., sky, horizon, depth), participants may have used these orthogonal cues in the Convex Buildings images for scene categorization, making the

task too easy and not sensitive enough to detect the effect of concavity in biasing scene categorization in these stimuli. As such, the lack of a difference between the place ratings of Concave and Convex Buildings is essentially a null effect and should not be taken as evidence against our hypothesis. Moreover, it is noteworthy that despite the easy task and loose control over the orthogonal visual cues, we nevertheless still find a striking difference in the place ratings of Concave Objects (33.6%) and Convex Objects (2.8%), highlighting the prominent effect of concavity in biasing scene categorization of object stimuli.

3.2. Experiment 2

Next, using fMRI, we hypothesized that if concavity is a diagnostic feature of visual scenes, then PPA will respond more to concave scenes and objects than convex scenes and objects. To test this prediction, we conducted a 2 (Category: Buildings, Objects) $\times$ 2 (Condition: Concave, Convex) repeated-measures ANOVA. We found a significant effect of Category ($F_{(1,14)} = 214.66$, $p<.001$, $\eta_p^2 = .94$), with an overall greater response to Buildings than to Objects, consistent with PPA's known selectivity for scenes (Epstein and Kanwisher, 1998; Kamps et al., 2016) (Fig. 1 D). We also found a significant main effect of Condition ($F_{(1,14)} = 18.11$, $p = .001$, $\eta_p^2 = .56$), with an overall greater response to Concave over Convex stimuli, consistent with our hypothesis. However, we also found a significant Category $\times$ Condition interaction ($F_{(1,14)} = 21.36$, $p<0.001$, $\eta_p^2 = .60$), with a significantly greater response to Concave Objects over Convex Objects (post-hoc comparison, $p<.001$), but no significant difference between Concave Buildings and Convex Buildings ($p = 0.73$), mirroring the behavioral results where Concave Objects are more scene like than Convex Objects, yet Concave and Convex Buildings are equally scene like.

Before further discussion of the results, it is important to reemphasize that, despite a lack of difference between PPA's response to Concave and Convex Buildings, we nevertheless found a strong effect of Concave Objects in driving PPA's response, supporting our hypothesis. However, regarding the lack of difference between Concave and Convex Buildings, these stimuli are the same stimuli used in Experiment 1, where we did not control for some orthogonal visual features; hence, it is possible that some of these orthogonal features may have driven PPA's response to "ceiling", making the effect of concavity in the PPA response to Buildings undetectable. As such, again, the null results for Concave versus Convex Buildings should not be taken as evidence against our hypothesis. Thus, to directly test whether PPA is sensitive to concavity in scene stimuli, we need to control for all visual features that are orthogonal to concavity and convexity, and we did just that in Experiment 3, discussed below.

But might PPA's selective response to Concave Objects be merely driven by participants paying more attention to Concave Objects that are perhaps more novel and interesting than Convex Objects? To address this question, we next examined the neural response in LOC. If PPA's selective response to Concave over Convex Objects is merely driven by participants paying more attention to Concave Objects, then LOC—an *object*-selective region—will show a similar, if not greater, attentional modulation in response. To test whether that is the case, we conducted a 2 (ROI: PPA, LOC) $\times$ 2 (Condition: Concave Objects, Convex Objects) repeated-measures ANOVA. We found a significant ROI $\times$ Condition interaction ($F_{(1,14)} =$

105.31, $p<.001$, $\eta_p^2 = .88$), revealing that only PPA, not LOC, responds selectively to objects depicting concavity, thus ruling out the alternative explanation that participants are merely paying more attention to Concave over Convex Objects.

Next, we further examined the neural response in LOC to test whether visual cues of convexity might selectively engage object processing, as proposed in a previous study (Haushofer et al., 2008). We conducted a 2 (Category: Buildings, Objects) × 2 (Condition: Concave, Convex) repeated-measures ANOVA. We found a significant effect of Category ($F_{(1,14)} = 25.56$, $p<.001$, $\eta_p^2 = .65$), with an overall greater response to Objects over Buildings, consistent with LOC's known selectivity for objects (Malach et al., 1995). We also found a significant effect of Condition ($F_{(1,14)} = 35.85$, $p<.001$, $\eta_p^2 = .72$), with LOC responding more to Convex over Concave stimuli. Crucially, we found a significant Category × Condition interaction ($F_{(1,14)} = 10.59$, p = .01, $\eta_p^2 = .43$), with LOC responding significantly more strongly to both buildings and objects depicting convexity than concavity (both $p<.05$), respectively, but with a greater PSC difference within Buildings than within Objects. Thus, LOC's selective sensitivity to convexity over concavity is consistent with previous findings that visual cues of convexity selectively engage the object processing system (Haushofer et al., 2008).

Next, to ensure that PPA's preferential response to objects depicting concavity is not merely driven by low-level visual information inherited from early visual cortex, we directly compared the neural responses in PPA and V1. A 2 (ROI: PPA, V1) × 2 (Condition: Concave Objects, Convex Objects) repeated-measures ANOVA revealed a significant ROI × Condition interaction ($F_{(1,14)} = 51.57$, $p<.001$, $\eta_p^2 = .79$), demonstrating a qualitatively different pattern of response in PPA than V1, and revealing that PPA's preferential response to Concave Objects is not merely driven by low-level visual information. One puzzling result, however, is that V1 seems to show a lower response to Concave Buildings relative to Convex Buildings—why might that be the case? One possibility is that, despite equating the overall quantity of energy in the high spatial frequency range across conditions to the best of our ability, the qHF in Concave Buildings is nevertheless numerically lower than Convex Buildings (mean difference = 5.90%; post-hoc comparison: $p = .40$). And with V1's known sensitivity to high spatial frequency information, V1 might be tracking such subtle, yet existing different high spatial frequency information, thus resulting in a relatively lower response to Concave Buildings.

Finally, we checked the reliability of our results by a split-half analysis (odd versus even runs). We found consistent results in both halves of the data (see Supplementary Analysis 3).

Having established the selective sensitivity to concavity in PPA, we next examined the neural response of OPA and RSC to test whether this sensitivity is specific to PPA or common across all scene-selective cortical regions. In OPA, a 2 (Category: Buildings, Objects) × 2 (Condition: Concave, Convex) repeated-measures ANOVA revealed a significant effect of Category ($F_{(1,13)} = 92.67$, $p<.001$, $\eta_p^2 = .88$), with an overall greater response to Buildings than Objects, consistent with OPA's known selectivity for scenes (Dilks et al., 2013; Kamps et al., 2016) (Fig. 1 D). We also found a significant main effect of Condition ($F_{(1,13)} = 31.51$, $p<.001$, $\eta_p^2 = .71$), with an overall greater response to Concave

over Convex stimuli, revealing a similar preference for concavity over convexity in OPA. Moreover, just like in PPA, there is a significant Category × Condition interaction ($F_{(1,13)}$ = 6.59, p = .02, η_p^2 = .34), with a significantly greater response to Concave over Convex Objects (post-hoc comparison, p<.001), and no significant difference between Concave and Convex Buildings (p = .13). Next, we directly compared the selective response for Concave over Convex Objects in PPA and OPA using a 2 (ROI: PPA, OPA) × 2 (Conditions: Concave Objects, Convex Objects) repeated-measures ANOVA. We found no significant ROI × Condition interaction ($F_{(1,13)}$ = 2.27, p = .16, η_p^2 = .15), indicating no difference in the preferential response to Concave Objects over Convex Objects in PPA and OPA.

To ensure that OPA's response is similar to PPA but not LOC and V1, we next directly compared the response in OPA with LOC and V1. A 2 (ROI: OPA, LOC) × 2 (Condition: Concave Objects, Convex Objects) repeated-measures ANOVA revealed a significant ROI × Condition interaction ($F_{(1,13)}$ = 50.94, p<.001, η_p^2 = .80), confirming that OPA, not LOC, responds selectively to objects depicting concavity. Similarly, using a 2 (ROI: OPA, V1) × 2 (Condition: Concave Objects, Convex Objects) repeated-measures ANOVA, we also found a significant ROI × Condition interaction ($F_{(1,13)}$ = 38.87, p<.001, η_p^2 = .75), demonstrating a qualitatively different pattern of response in OPA than V1. Together, these results suggest that OPA's selective response to Concave over Convex Objects is similar to PPA but not LOC and V1.

In RSC, a 2 (Category: Buildings, Objects) × 2 (Condition: Concave, Convex) repeated-measures ANOVA revealed a significant effect of Category ($F_{(1,11)}$ = 206.38, p<.001, η_p^2 = .95), with an overall greater response to Buildings than to Objects, consistent with RSC's known selectivity for scenes (Kamps et al., 2016; Maguire, 2001; Silson et al., 2016). Interestingly, however, despite a numerically greater response to Concave than Convex Conditions in both Buildings and Objects, we only found a marginally significant main effect of Condition ($F_{(1,11)}$ = 4.22, p = .07, η_p^2 = .28), with post-hoc comparison revealing no significant difference between Concave versus Convex Buildings (p = .12) and Objects (p = .38). Moreover, we also found no significant Category × Condition interaction ($F_{(1,11)}$ = 0.02, p = .88, η_p^2 = .002), unlike the response patterns in PPA and OPA.

Why do PPA and OPA show a selective sensitivity to Concave over Convex Objects whereas RSC does not? There are two possible reasons. The first possible reason is that we have less functionally defined RSCs in the participants, compared to PPA and OPA, thus less power to detect RSC's response pattern. To test this possibility, we defined RSC in the missing participants with a lower threshold for the Scene – Object contrast (a minimum of p<.01) in the independent Localizer runs and repeated the same statistical analyses with these additional data. We found consistent results (see Supplementary Analysis 4), indicating the response pattern is not merely a result of having less data for RSC. Another possible reason is that, relative to PPA and OPA, RSC has been proposed to be involved in processing more navigation- and memory-relevant information of a scene (e.g., relating a scene to the memory representation of a broader environment that is not immediately visible) (Aguirre et al., 1998; Maguire, 2001; Marchette et al., 2014; Park and Chun, 2009; Persichetti and Dilks, 2019) and also has stronger functional connectivity to the more anterior cortical regions, including the hippocampus and some cortical regions that are involved in processing

abstract navigation and mnemonic representation (Baldassano et al., 2016; Silson et al., 2019). Thus, RSC may be more susceptible to top-down influence and less sensitive to bottom-up, visual cues of concavity. What is the topography of Concave selectivity in relation PPA and OPA? Is it a preference common across the whole cortical region, or might there be subregions within PPA and OPA that are more sensitive for visual cues of concavity? To address this question, we next conducted a whole-brain analysis and examined the topography for Concave versus Convex Objects response on the cortical surface (Fig. 2). Note that since this is an exploratory analysis, we did not correct for multiple comparisons. We first verified whether the whole-brain analysis showed consistent results as the ROI analysis. Indeed, we found Concave-selective patches overlapping with PPA and OPA, and no Concave- nor Convex-selective clusters in RSC. In addition, we also found some small, but nevertheless existing Convex-selective clusters within LOC. Concerning the topography of Concave selectivity within PPA and OPA, we observed the strongest Concave selectivity localized in the more posterior parts of PPA and OPA, consistent with the findings that the posterior parts of the scene-selective regions are more involved in perceptual over memory-related processing (Baldassano et al., 2016; Bar and Aminoff, 2003; Silson et al., 2016, 2019).

3.3. Experiment 3

In Experiments 1 and 2, we found that Concave Objects are categorized as scenes more often than Convex Objects, both behaviorally and neurally, supporting our hypothesis that concavity is a diagnostic feature of a scene; however, we did not find any behavioral or neural distinction for Concave Buildings versus Convex Buildings. The lack of difference between both behavioral and neural categorization for Concave and Convex Buildings poses a challenge to our hypothesis: Is concavity truly a diagnostic cue for scenes, as hypothesized here, or are Concave Objects somehow a special class of stimuli that happen to be scene like, and drive PPA's response? As previously discussed, since we did not control for some additional orthogonal visual features in Experiments 1 and 2, we cannot determine whether PPA is sensitive to the concavity of scene stimuli. Hence, to directly test PPA's sensitivity to concavity of scene stimuli, in Experiment 3, we created three pairs of Concave versus Convex scene boundaries that are exactly equated in all visual properties orthogonal to concavity and convexity. If concavity is truly a diagnostic feature of a scene, then PPA will show a significantly greater response to Concave over Convex scene boundaries after all possible confounding variables are eliminated.

To test this prediction, we conducted a 2 (Boundary Type: Concave, Convex) × 3 (Angle: 1, 2, 3) repeated-measures ANOVA to directly compare Concave and Convex conditions (Fig. 3 B). We found a significant main effect of Boundary Type ($F_{(1,13)} = 7.24$, $p = .02$, $\eta_p^2 = .36$), with an overall greater response to the Concave conditions over the Convex conditions despite equated visual features, consistent with our hypothesis that PPA shows a significantly greater response to concavity than convexity of scene stimuli. We also found a significant effect of Angle ($F_{(2,26)} = 8.24$, $p = .002$, $\eta_p^2 = .39$), and a significant Boundary Type × Angle interaction ($F_{(2,26)} = 4.62$, $p = .02$, $\eta_p^2 = .26$), hinting at an increasing sensitivity to concavity over convexity as the angle between two surfaces increases. To further explore PPA's sensitivity to changes in concavity and convexity, we conducted a

three-level (Angle: 1, 2, 3) repeated-measures ANOVA for the Concave conditions and the Convex conditions, separately, and further examined the linear contrast for each Boundary Type. For the Concave conditions, we found a significant effect of Angle ($F_{(2,26)}$ = 8.84, p = .001, η_p^2 = .41), with a significant linear increase by Angle ($F_{(1,13)}$ = 15.21, p = .002, η_p^2 = .54), revealing PPA's sensitivity to changes in concavity. By contrast, for the Convex conditions, we did not find a significant effect of Angle ($F_{(2,26)}$ = 2.52, p = .10, η_p^2 = .16), nor a significant linear increase by Angle ($F_{(1,13)}$ = 1.28, p = .28, η_p^2 = .09). Finally, using a 2 (Boundary Type: Concave, Convex) × 3 (Angle: 1, 2, 3) repeated-measures ANOVA, we directly compared PPA's sensitivity to changes in concavity and convexity and found a significant linear trend interaction ($F_{(1,13)}$ = 9.44, p = .01, η_p^2 = .42), confirming that PPA is sensitive to changes in concavity, but not convexity.

Next, we examined LOC's response. To determine whether neural sensitivity to changes in concavity, but not convexity, is specific to PPA, we conducted a 2 (ROI: PPA, LOC) × 2 (Boundary Type: Concave, Convex) × 3 (Angle: 1, 2, 3) repeated-measures ANOVA, and found a significant ROI × Boundary Type × Angle interaction ($F_{(2,24)}$ = 4.95, p = .02, η_p^2 =.29), demonstrating a qualitatively different pattern of response in PPA than LOC (Fig. 3 B). Furthermore, we also found an ROI × Boundary Type interaction ($F_{(1,12)}$ = 42.47, p<.001, η_p^2 =.78), with PPA responding significantly more to concavity than convexity, and LOC responding significantly more to convexity than concavity (post-hoc comparisons, both p<.05), providing evidence that 1) PPA, but not LOC, shows a preferential response to concavity over convexity, and 2) LOC shows a preferential response to convexity over concavity of boundary cues. Next, to further characterize LOC's response to convexity versus concavity, we conducted a 2 (Boundary Type: Concave, Convex) × 3 (Angle: 1, 2, 3) repeated-measures ANOVA on LOC response alone. We found a significant effect of Boundary Type ($F_{(1,12)}$ = 21.23, p = .001, η_p^2 = .64), with LOC responding more to the Convex conditions than to the Concave conditions, consistent with results from Experiment 2 that visual cues of convexity selectively engage the object processing system. Furthermore, we also found a significant effect of Angle ($F_{(2,24)}$ = 39.81, p<.001, η_p^2 =.77) and significant linear increase by angle ($F_{(1,12)}$ = 100.34, p<.001, η_p^2 = .89), with LOC responding more to the most extreme angle (3), relative to the other two angles (1 and 2). Interestingly, we did not find a significant Boundary Type × Angle interaction ($F_{(2,24)}$ = 0.02, p =.98, η_p^2 =.002), and no significant linear trend interaction ($F_{(1,12)}$ = 0.05, p =.82, η_p^2 = .004), revealing LOC's similar sensitivity to both angle *changes* in concavity and convexity despite a preferential response for convexity over concavity, more generally.

Next, we tested whether the PPA response might be driven by low-level visual information directly inherited from early visual cortex, we compared the responses in PPA to V1. A 2 (ROI: PPA, V1) × 2 (Boundary Type: Concave, Convex) × 3 (Angle: 1, 2, 3) repeated-measures ANOVA revealed a significant ROI × Boundary Type × Angle interaction ($F_{(2,22)}$ = 60.72, p<.001, η_p^2 =.85), revealing a qualitatively different pattern of response in PPA than V1, confirming that PPA's preferential sensitivity to concavity over convexity is not driven by a mere response to low-level visual properties of the stimuli (Fig. 3 B). Note, however, like PPA, it seems as though V1 still tracks concavity in some sense; that is, decreasing in response as Angle increases, especially within the Concave conditions. What is driving this response pattern? One possibility is that since V1 is known to have a foveal

bias, and the amount of complex texture information on a boundary decreases in foveal vision as Angle becomes more extreme, especially within the Concave conditions, V1 may simply be tracking the changes of foveally-presented visual information. To test this possibility, we correlated the number of pixels that contain texture information in the center of an image (the central 50% of the image) with the V1 response across the conditions. Indeed, there is a strong and positive correlation between the number of textured pixels and V1 response ($r = .90$). By contrast, PPA and LOC show a weaker and negative correlation (PPA: $r = -.51$; LOC: $r = -.32$). Together, the contrast between the strength and direction of correlation in V1 and the other two regions suggests that V1's response is most likely driven by the low-level foveal visual information in the stimuli, and not the concavity and convexity of the boundary.

Finally, just like in Experiment 2, we checked the reliability of our results by a split-half analysis (odd versus even runs). We found consistent results in both halves of the data (see Supplementary Analysis 3).

Having established the selective sensitivity to concavity in PPA, we next tested whether OPA and RSC also show a similar sensitivity to concavity. In OPA, a 2 (Boundary Type: Concave, Convex) $\times$ 3 (Angle: 1, 2, 3) repeated-measures ANOVA revealed a significant main effect of Boundary Type ($F_{(1,12)} = 12.38$, $p = .004$, $\eta_p^2 = .51$), with an overall greater response to the Concave conditions over the Convex conditions, just like in PPA. We also found a significant effect of Angle ($F_{(2,24)} = 34.12$, $p<.001$, $\eta_p^2 = .74$), and a significant Boundary Type $\times$ Angle interaction ($F_{(2,24)} = 3.81$, $p = .04$, $\eta_p^2 = .24$), hinting at a greater sensitivity to changes in concavity over convexity. Thus, we next conducted a three-level (Angle: 1, 2, 3) repeated-measures ANOVA for the Concave conditions and the Convex conditions, separately, and further examined the linear contrast for each Boundary Type. For the Concave conditions, we found a significant effect of Angle ($F_{(2,24)} = 18.13$, $p<.001$, $\eta_p^2 = .60$), with a significant linear increase by Angle ($F_{(1,12)} = 22.94$, $p<.001$, $\eta_p^2 = .66$), revealing OPA's sensitivity to changes in concavity. Interestingly, for the Convex conditions, unlike in PPA, we also found a significant effect of Angle ($F_{(2,24)} = 10.91$, $p<.001$, $\eta_p^2 = .48$), with a significant linear increase by Angle ($F_{(1,12)} = 9.99$, $p = .01$, $\eta_p^2 = .45$), revealing OPA's sensitivity to changes in convexity. To directly compare OPA's sensitivity to concavity versus convexity, we conducted a 2 (Boundary Type: Concave, Convex) $\times$ 3 (Angle: 1, 2, 3) repeated-measures ANOVA. We found a significant linear trend interaction ($F_{(1,12)} = 5.53$, $p = .04$, $\eta_p^2 = .32$), indicating that, despite sensitivity to changes in convexity, OPA nevertheless shows a greater sensitivity to concavity over convexity. To directly compare PPA and OPA's selective sensitivity to concavity and convexity, we conducted a 2 (ROI: PPA, OPA) $\times$ 2 (Boundary Type: Concave, Convex) $\times$ 3 (Angle: 1, 2, 3) repeated-measures ANOVA. We found no significant ROI $\times$ Boundary Type interaction ($F_{(1,12)} = 3.56$, $p = .08$, $\eta_p^2 = .23$), and no significant ROI $\times$ Boundary Type $\times$ Angle interaction ($F_{(2,24)} = 1.18$, $p = .33$, $\eta_p^2 = .09$), indicating no significant difference in the selective sensitivity to concavity over convexity between PPA and OPA.

To ensure that the neural response in OPA is similar to PPA but not LOC and V1, we next directly compared OPA with LOC and V1. A 2 (ROI: OPA, LOC) $\times$ 2 (Boundary Type: Concave, Convex) $\times$ 3 (Angle: 1, 2, 3) repeated-measures ANOVA revealed a significant

ROI × Boundary Type × Angle interaction ($F_{(2,22)}$ = 5.34, p = .01, η_p^2 = .33), demonstrating a qualitatively different pattern of response in OPA than LOC. Similarly, using a 2 (ROI: OPA, V1) × 2 (Boundary Type: Concave, Convex) × 3 (Angle: 1, 2, 3) repeated-measures ANOVA, we also found a significant ROI × Boundary Type × Angle interaction ($F_{(2,20)}$ = 58.90, $p{<}.001$, η_p^2 =.86), demonstrating a qualitatively different pattern of response in OPA than V1. Thus, these results confirm that OPA shows a qualitatively similar neural response with PPA, but not LOC and V1.

In RSC, interestingly, unlike in PPA and OPA, a 2 (Boundary Type: Concave, Convex) × 3 (Angle: 1, 2, 3) repeated-measures ANOVA revealed no significant main effect of Boundary Type ($F_{(1,7)}$ = 1.64, p =.24, η_p^2 =.19), no significant main effect of Angle ($F_{(2,14)}$ =.25, p =.78, η_p^2 =.03), and no significant Boundary Type × Angle interaction ($F_{(2,14)}$ =.24, p =.79, η_p^2 =.03). Moreover, when we further tested for RSC's sensitivity to Angle using a planned, three-level (Angle: 1, 2, 3) repeated-measures ANOVA for the Concave conditions and the Convex conditions, separately, we consistently found no significant main effect of Angle (Concave: $F_{(2,14)}$ = 0.27, p =.77, η_p^2 =.04; Convex: $F_{(2,14)}$ = 0.20, p =.82, η_p^2 = .03) nor a significant linear increase by Angle (Concave: $F_{(1,7)}$ = 0.76, p =.41, η_p^2 = .10; Convex: $F_{(1,7)}$ = 0.03, p =.86, η_p^2 =.01). To ensure the lack of RSC response is not a result of having a smaller number of ROIs relative to PPA and OPA, we defined RSC in the missing participants with a lowered threshold for the Scene-Object contrast (a minimum of $p{<}.01$) in the independent Localizer runs and repeated the same statistical analyses with these additional data. We found consistent results (see Supplementary Analysis 4), indicating the response pattern is not merely a result of having fewer functionally defined RSC.

We next conducted a group-level, whole-brain analysis to examine the topography for selectivity for Concave versus Convex boundaries (averaged across Angle) (Fig. 4). Again, note that since this is an exploratory analysis, we did not correct for multiple comparisons.

Consistent with the ROI analysis, we found PPA and OPA overlap with Concave-selective regions, whereas LOC overlaps with Convex-selective regions. We next observed the activation topography in and around RSC. Interestingly, while we found no Concave- nor Convex-selective voxels within RSC, we observed some Concave-selective patches immediately posterior to RSC. Moreover, consistent with Experiment 2, we also found Concave selectivity localized at the more posterior parts of PPA. Together, these findings reveal a consistent localization of Concave selectivity in the posterior parts of the scene-selective regions even when all orthogonal visual features in the stimuli are tightly controlled.

In addition to Concave selectivity being localized in the posterior parts of scene-selective cortical regions, we also observed a stream of Concave-selective cortex leading up to PPA from the more posterior parts of the visual cortex, and a stream of Convex-selective cortex leading up to and extending beyond LOC. These separate Concave- and Convex-selective streams lie in juxtaposition and extend along the ventral surface of the occipitotemporal cortex, revealing a selectivity for concavity versus convexity beyond the scene- and object-selective cortical regions. To test whether these streams of Concave and Convex selectivity remain after controlling for multiple comparisons, we further thresholded the activation map

with a false discovery rate (FDR) of q<.05 and examined the activation map (see Supplementary Analysis 5)—we found a similar topography of activations. Together, these results are consistent with Experiment 2, and shed light onto a possible Concave-Convex organization in the ventral visual pathway.

3.4. Experiment 4

In Experiment 3, we found that, when using tightly controlled stimuli, concavity still drives cortical scene processing. Together with Experiment 2, then, we found neural evidence that concavity is a diagnostic feature for visual scenes, in both highly variable, naturalistic images and tightly controlled, artificial images. One remaining question, however, is whether concavity also biases *behavioral* categorization of visual inputs as scenes if all Concave-irrelevant features are tightly controlled. The strongest test then is to stripe away all visual features orthogonal to concavity and convexity, and test for behavioral categorization of highly impoverished line drawings of concave shapes as scenes versus objects. If concavity indeed drives behavioral scene categorization, then even highly impoverished line drawings of concave shapes should be categorized as scenes more often than convex shapes.

Indeed, we found that all the Concave shapes received over 70% of scene ratings as opposed to object ratings (Concave 1: 74.75%; Concave 2: 79.80%; Concave 3: 74.75%; Fig. 5). Moreover, using a Binomial Test, we found that the proportion of scene ratings for all the Concave shapes is significantly above chance (i.e., 50%; all p<.001), revealing a strong bias to behaviorally categorize even impoverished line drawings of concave shapes as scenes more often than objects. But might it simply due to participants not paying attention to the task and merely clicking the same button during the experiment, which happens to be "scene"? To test whether that might be the case, we next examined the response for the Flat and Convex shapes. Interestingly, unlike for the Concave shapes, participants did not have a preference towards whether these non-Concave shapes are categorized as scenes or objects —the proportions for scene versus object ratings for each of these conditions are split approximately in half (scene ratings for Flat: 51.52%; Convex 1: 50.51%; Convex 2: 58.59%; Convex 3: 44.44%), and a Binomial Test confirmed that none of the scene rating proportion for these non-Concave shapes is significantly different from chance (Flat: p = .84; Convex 1: p>.99; Convex 2: p = .11; Convex 3: p = .32). Thus, the contrast between participants' ratings for the Concave and non-Concave line drawings suggest that participants were indeed paying attention to the task and were not merely clicking buttons randomly.

But did participants actually rate the Concave shapes as scenes significantly more often than the Convex shapes, as we hypothesized? To answer this question, we directly compared the proportion of scene ratings between the Concave and Convex conditions with the same Angle (e.g., Concave 2 vs. Convex 2) using the McNemar's test. Consistent with our hypothesis, all the Concave shapes indeed showed a significantly greater proportion of scene ratings than the corresponding Convex shapes that shared the same Angle (p<.001 for all three pairs of comparisons), revealing that line drawings of the Concave shapes are indeed categorized as scenes more often than the Convex shapes across all Angles. Together, these results suggest even when all orthogonal stimulus features are eliminated, visual cues of

concavity still bias behavioral categorization of visual stimuli as scenes, providing convincing behavioral evidence for concavity as a diagnostic cue for visual scenes.

Finally, we checked the reliability of our results by repeating the experiment with a separate group of 50 participants. We found consistent results (see Supplementary Analysis 6). In addition, we also repeated the same experiment on the stimuli from Experiment 3 with a separate group of 100 participants to test whether the presence/absence of the boundary texture might have had affected participants' perception of these stimuli as scenes or objects, we found similar response patterns (see Supplementary Analysis 7).

4. Discussion

The current study asked whether concavity (versus convexity) may be a diagnostic visual feature enabling the human brain to "recognize" scenes. Consistent with this hypothesis, we present converging behavioral and neural evidence that 1) people behaviorally judged Concave Objects as scenes more often than Convex Objects, 2) PPA showed a selective response to Concave Objects over Convex Objects, 3) even though people behaviorally categorize both Concave and Convex Buildings to be both highly scene-like, and PPA shows a similar preferential response to Concave and Convex Buildings in naturalistic images, PPA nevertheless shows a selectively greater sensitivity to concavity over convexity of scene boundaries after all confounding visual features orthogonal to concavity and convexity are tightly controlled, and 4) consistently, people also categorized highly impoverished line drawings of concave shapes as scenes more often than convex shapes when all orthogonal visual features are eliminated. In addition to these main findings, we also found that LOC shows a preferential response to visual cues of convexity, over concavity, of both scene and object stimuli. Taken together, these results reveal that concavity is a diagnostic feature of a scene, and raise the intriguing hypothesis that concavity versus convexity may be a diagnostic visual feature enabling the human brain to differentiate scenes from objects.

Our finding that PPA responds selectively to visual cues of concavity, but not convexity, lends further support to previous findings that PPA represents the shape (or "spatial layout") of a scene in terms of the continuity of its spatial layout (e.g., PPA responds more strongly to intact spatial layouts – an empty apartment room – than to fragmented ones in which the walls, floors, and ceilings have been fractured and rear-ranged) (Epstein and Kanwisher, 1998; Kamps et al., 2016), the openness of the spatial layout (e.g., a desert vs. a forest) (Kravitz et al., 2011; Park et al., 2011), and the basic length and angle relations among the surfaces composing the spatial layout (Dillon et al., 2018). Note, however, that the previous work that investigated PPA's representation of spatial layout only tested the shape of concave indoor spaces. Thus, the present work extends the prior work by revealing that PPA is not responding merely to the "spatial layout" of a scene, more generally, but rather to the concavity that is defined by the conjoint surfaces of a scene, more specifically.

But how might PPA extract concavity (versus convexity) information from visual inputs? PPA represents many lower-level visual cues, including contour junctions (Choo and Walther, 2016), 3D orientation of surfaces (Lescroart and Gallant, 2019) and textures (Cant and Goodale, 2011; Cant and Xu, 2017), which when in certain configurations are diagnostic

of concavity and convexity. For example, concave spaces are usually made up of surfaces with contours conjoining at a Y-junction, whereas convex spaces are usually made up of surfaces with contours conjoining at an arrow-junction; 3D surfaces of concavity converge in depth, whereas 3D surfaces of convexity diverge in depth; and the surfaces that compose a concave space usually show an increasing density and distortion of texture patterns (i.e., texture gradient) as they converge, whereas surfaces of convexity usually show a decreasing density and distortion of texture patterns as they converge (see Supplementary Fig. S9 for a graphic illustration). Thus, it seems highly plausible that such representations of lower-level visual properties allow PPA to flexibly extract information of concavity across highly variable stimuli, as we propose here.

In addition to PPA, we also found a selective sensitivity to concavity over convexity of both naturalistic object images and also artificial scene boundaries in OPA. These results dovetail with OPA's sensitivity to the orientations and relationships among wall surfaces of an indoor scene (Dillon et al., 2018; Henriksson et al., 2019; Lescroart and Gallant, 2019). Interestingly, unlike PPA, OPA is also sensitive to the convexity of the wall boundaries in Experiment 3. Together with previous findings for a selective sensitivity to the egocentric distance (i.e., near vs. far) (Persichetti and Dilks, 2016) and sense information (i.e., left vs. right) (Dilks et al., 2011) of a scene in OPA, but not PPA, our findings lend further support to the dissociable roles of PPA and OPA in recognizing the kind of place (e.g., a kitchen vs. a bedroom) a scene is versus navigating through it (Persichetti and Dilks, 2018), as the angular relationship among wall surfaces is likely to be important for successful navigation through a scene but not as much for recognizing the kind of place it is.

We also examined RSC's response but found no significant response pattern in the ROI results; however, the whole-brain analysis in Experiment 3 revealed Concave selectivity in the cortical regions immediately posterior to RSC. Moreover, we also found a strong Concave selectivity in and around posterior PPA in both Experiment 2 and 3. One plausible reason for a similar posterior localization is that perceptual processing of visual scenes is known to occur in the more posterior parts of these scene-selective regions (Baldassano et al., 2016; Bar and Aminoff, 2003; Silson et al., 2016, 2019). Another plausible reason is that the brain is likely to recognize scene from non-scene stimuli *before* processing scene-specific information; thus, it makes sense that Concave selectivity would be localized at a more posterior region in which earlier stages of visual processing occur.

While our results point to concavity as a diagnostic feature of a scene, there are two caveats. The first caveat is that our results do not imply that concavity is the only diagnostic visual feature of a scene. In fact, as revealed by results in Experiment 1, despite more place ratings for Concave Objects over Convex Objects, participants nevertheless consider Building images in general to be more scene like than Concave Objects. Similarly, in Experiment 2, while PPA responds to Concave Objects more selectively than Convex Objects, PPA nevertheless responds most selectively to Concave and Convex Buildings, consistent with previous findings that PPA responds not only to concave indoor scenes, but also to "non-concave" images of outdoors scenes and buildings (Aguirre et al., 1998; Epstein and Kanwisher, 1998). Together, these results illustrate that, in addition to concavity, there exist other visual features that humans use to recognize scenes, and call for future work to explore

these other visual features. In particular, with previous studies highlighting a sensitivity to textures in PPA (Cant and Goodale; 2011; Cant and Xu, 2017; Lowe et al., 2017; Park and Park, 2017), and the textural differences between big versus small objects (Long et al., 2016; Long et al., 2018), we speculate that there exist certain textures that are common but unique to scenes, such as the sky and the grassy texture of an open field, and can be used by the cortical scene-processing system to recognize scene from non-scene stimuli. Moreover, it is also likely that, since a scene usually encompasses a large space and extends in depth, certain depth cues, such as texture gradients, might be common across scene but not non-scene stimuli, and thus could also be an indicator of visual scenes. Therefore, one fruitful future research direction is to investigate whether the cortical scene processing system may be driven by certain texture properties over others.

The second caveat is that despite matching for rectilinearity across conditions in all the experiments, the stimuli nevertheless contain rectilinear features, and we did not test for the effect of concavity in the absence of rectilinearity; thus, we cannot conclude whether concavity alone is diagnostic of visual scenes, or if it only plays a role in the presence of rectilinearity. However, considering how natural caves or the inside of an igloo—scene categories that are concave and curvilinear—is likely to be categorized as a scene, we speculate that even concavity alone, without rectilinearity, is still diagnostic of a scene. Nevertheless, more future research is needed to systematically disentangle the effect of concavity from rectilinearity in driving behavioral and neural scene categorization.

In addition to our main findings that concavity drives scene selectivity in PPA, we also found that LOC shows a selective response to convex, over concave, objects and scenes. This finding dovetails with several findings revealing that LOC is sensitive to the perceived 3-D structure (i.e., concave versus convex) of an object (Kourtzi et al., 2003), and that there is a privileged coding of convexity in LOC (Haushofer et al., 2008). Taken together, these studies, coupled with the findings here, shed light on the potential importance of convexity information in object processing.

Finally, in the whole-brain analysis conducted in Experiment 3, we found separate streams of cortical regions along the ventral occipitotemporal cortex that respond preferentially to concavity over convexity, and vice versa. Such findings are consistent with distinct channels for processing of landscape- and cave-like environmental shapes versus small, bounded object shapes in the inferotemporal cortex of non-human primates (Vaziri et al., 2014), and raise an intriguing possibility that Concave versus Convex might be an organizing principle of the ventral visual pathway that enables segregation of visual inputs into distinct scene- versus object-processing. However, since these findings are based on a fixed-effects analysis, which does not allow for generalization of inference to the population level, more future research is needed to further test for a large-scale Concave vs. Convex organization along the ventral visual pathway.

In conclusion, our results indicate that concavity is a diagnostic feature of a scene. Moreover, we have shown that convexity may also be a diagnostic feature of an object. These findings together shed light on the relevant features used in human scene and object

recognition, and demonstrate that the three-dimensional geometric relationships between surfaces enable the human brain to differentiate scenes from objects.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

Aguirre GK, Zarahn E, D'esposito M, 1998. An area within human ventral cortex sensitive to "building" stimuli: evidence and implications. Neuron 21 (2), 373–383. [PubMed: 9728918]

Bainbridge WA, Oliva A, 2015. A toolbox and sample object perception data for equalization of natural images. Data Brief 5, 846–851. [PubMed: 26693521]

Baldassano C, Esteva A, Fei-Fei L, Beck DM, 2016. Two Distinct Scene-Processing Networks Connecting Vision and Memory. Eneuro, p. 3.

Bar M, Aminoff E, 2003. Cortical analysis of visual context. Neuron 38 (2), 347–358. [PubMed: 12718867]

Berman D, Golomb JD, Walther DB, 2017. Scene content is predominantly conveyed by high spatial frequencies in scene-selective visual cortex. PLoS One 12 (12), e0189828. [PubMed: 29272283]

Bryan PB, Julian JB, Epstein RA, 2016. Rectilinear edge selectivity is insufficient to explain the category selectivity of the parahippocampal place area. Front. Human Neurosci 10, 137.

Cant JS, Goodale MA, 2011. Scratching beneath the surface: new insights into the functional properties of the lateral occipital area and parahippocampal place area. J. Neurosci 31 (22), 8248–8258. [PubMed: 21632946]

Cant JS, Xu Y, 2017. The contribution of object shape and surface properties to object ensemble representation in anterior-medial ventral visual cortex. J. Cogn. Neurosci 29 (2), 398–412. [PubMed: 27676614]

Choo H, Walther DB, 2016. Contour junctions underlie neural representations of scene categories in high-level human visual cortex. Neuroimage 135, 32–44. [PubMed: 27118087]

Epstein R, Kanwisher N, 1998. A cortical representation of the local visual environment. Nature 392 (6676), 598. [PubMed: 9560155]

Dilks DD, Julian JB, Kubilius J, Spelke ES, Kanwisher N, 2011. Mirror-image sensitivity and invariance in object and scene processing pathways. J. Neurosci 31 (31), 11305–11312. [PubMed: 21813690]

Dilks DD, Julian JB, Paunov AM, Kanwisher N, 2013. The occipital place area is causally and selectively involved in scene perception. J. Neurosci 33 (4), 1331–1336. [PubMed: 23345209]

Dillon MR, Persichetti AS, Spelke ES, Dilks DD, 2018. Places in the brain: bridging layout and object geometry in scene-selective cortex. Cereb. Cortex 28 (7), 2365–2374. [PubMed: 28633321]

Epstein R, Graham KS, Downing PE, 2003. specific scene representations in human parahippocampal cortex. Neuron 37 (5), 865–876. [PubMed: 12628176]

Ferrara K, Park S, 2016. Neural representation of scene boundaries. Neuropsychologia 89, 180–190. [PubMed: 27181883]

Golomb JD, Kanwisher N, 2012. Higher level visual cortex represents retinotopic, not spatiotopic, object location. Cereb. Cortex 22 (12), 2794–2810. [PubMed: 22190434]

Greene MR, Oliva A, 2009. The briefest of glances: The time course of natural scene understanding. Psychol. Sci 20 (4), 464–472. [PubMed: 19399976]

Grill-Spector K, Kushnir T, Hendler T, Edelman S, Itzchak Y, Malach R, 1998. A sequence of object-processing stages revealed by fMRI in the human occipital lobe. Hum. Brain Mapp 6 (4), 316–328. [PubMed: 9704268]

Groen II, Silson EH, Baker CI, 2017. Contributions of low-and high-level properties to neural processing of visual scenes in the human brain. Philosoph. Trans. R. Soc. B 372 (1714), 20160102.

Haushofer J, Baker CI, Livingstone MS, Kanwisher N, 2008. Privileged coding of convex shapes in human object-selective cortex. J. Neurophysiol 100 (2), 753–762. [PubMed: 18579661]

Henriksson L, Mur M, Kriegeskorte N, 2019. Rapid invariant encoding of scene layout in human OPA. Neuron 103 (1), 161–171. [PubMed: 31097360]

Kamps FS, Julian JB, Kubilius J, Kanwisher N, Dilks DD, 2016. The occipital place area represents the local elements of scenes. Neuroimage 132, 417–424. [PubMed: 26931815]

Kauffmann L, Ramanoël S, Peyrin C, 2014. The neural bases of spatial frequency processing during scene perception. Front. Integrat. Neurosci 8, 37.

Kourtzi Z, Erb M, Grodd W, Bülthoff HH, 2003. Representation of the perceived 3-D object shape in the human lateral occipital complex. Cereb. Cortex 13 (9), 911–920. [PubMed: 12902390]

Kravitz DJ, Peng CS, Baker CI, 2011. Real-world scene representations in high-level visual cortex: it's the spaces more than the places. J. Neurosci 31 (20), 7322–7333. [PubMed: 21593316]

Lescroart MD, Gallant JL, 2019. Human scene-selective areas represent 3D configurations of surfaces. Neuron 101 (1), 178–192. [PubMed: 30497771]

Long B, Konkle T, Cohen MA, Alvarez GA, 2016. Mid-level perceptual features distinguish objects of different real-world sizes. J. Exper. Psychol 145 (1), 95.

Long B, Yu CP, Konkle T, 2018. Mid-level visual features underlie the high-level categorical organization of the ventral stream. Proc. Natl. Acad. Sci 115 (38), E9015–E9024. [PubMed: 30171168]

Lowe MX, Rajsic J, Gallivan JP, Ferber S, Cant JS, 2017. Neural representation of geometry and surface properties in object and scene perception. Neuroimage 157, 586–597. [PubMed: 28647484]

Maguire E, 2001. The retrosplenial contribution to human navigation: a review of lesion and neuroimaging findings. Scand. J. Psychol 42 (3), 225–238. [PubMed: 11501737]

Malach R, Reppas JB, Benson RR, Kwong KK, Jiang H, Kennedy WA, … Tootell RB, 1995. Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. Proc. Natl. Acad. Sci 92 (18), 8135–8139. [PubMed: 7667258]

Marchette SA, Vass LK, Ryan J, Epstein RA, 2014. Anchoring the neural compass: coding of local spatial reference frames in human medial parietal lobe. Nat. Neurosci 17 (11), 1598. [PubMed: 25282616]

Nasr S, Tootell RB, 2012. A cardinal orientation bias in scene-selective visual cortex. J. Neurosci 32 (43), 14921–14926. [PubMed: 23100415]

Nasr S, Echavarria CE, Tootell RB, 2014. Thinking outside the box: rectilinear shapes selectively activate scene-selective cortex. J. Neurosci 34 (20), 6721–6735. [PubMed: 24828628]

Park J, Park S, 2017. Conjoint representation of texture ensemble and location in the parahippocampal place area. J. Neurophysiol 117 (4), 1595–1607. [PubMed: 28123006]

Park S, Chun MM, 2009. Different roles of the parahippocampal place area (PPA) and retrosplenial cortex (RSC) in panoramic scene perception. Neuroimage 47 (4), 1747–1756. [PubMed: 19398014]

Park S, Brady TF, Greene MR, 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 (4), 1333–1340. [PubMed: 21273418]

Park S, Konkle T, Oliva A, 2014. Parametric coding of the size and clutter of natural scenes in the human brain. Cereb. Cortex 25 (7), 1792–1805. [PubMed: 24436318]

Persichetti AS, Dilks DD, 2016. Perceived egocentric distance sensitivity and invariance across scene-selective cortex. Cortex 77, 155–163. [PubMed: 26963085]

Persichetti AS, Dilks DD, 2018. Dissociable neural systems for recognizing places and navigating through them. J. Neurosci 38 (48), 10295–10304. [PubMed: 30348675]

Persichetti AS, Dilks DD, 2019. Distinct representations of spatial and categorical relationships across human scene-selective cortex. Proc. Natl. Acad. Sci 116 (42), 21312–21317. [PubMed: 31570605]

Rajimehr R, Devaney KJ, Bilenko NY, Young JC, Tootell RB, 2011. The "parahippocampal place area" responds preferentially to high spatial frequencies in humans and monkeys. PLoS Biol 9 (4), e1000608. [PubMed: 21483719]

Sereno MI, Dale AM, Reppas JB, Kwong KK, Belliveau JW, Brady TJ, Rosen BR, Tootell RB, 1995. Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. Science 268, 889–893. [PubMed: 7754376]

Silson EH, Gilmore AW, Kalinowski SE, Steel A, Kidder A, Martin A, Baker CI, 2019. A posterior–anterior distinction between scene perception and scene construction in human medial parietal cortex. J. Neurosci 39 (4), 705–717. [PubMed: 30504281]

Silson EH, Steel AD, Baker CI, 2016. Scene-selectivity and retinotopy in medial parietal cortex. Front. Hum. Neurosci 10, 412. [PubMed: 27588001]

Smith SM, 2002. Fast robust automated brain extraction. Hum. Brain. Mapp 17 (3), 143–155. [PubMed: 12391568]

Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TE, Johansen-Berg H, Matthews PM, 2004. Advances in functional and structural MR image analysis and implementation as FSL. Neuroimage 23, S208–S219. [PubMed: 15501092]

Smith SM, Nichols TE, 2009. Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. Neuroimage 44 (1), 83–98. [PubMed: 18501637]

Spiridon M, Kanwisher N, 2002. How distributed is visual category information in human occipito-temporal cortex? An fMRI study. Neuron 35 (6), 1157–1165. [PubMed: 12354404]

Vaziri S, Carlson ET, Wang Z, Connor CE, 2014. A channel for 3D environmental shape in anterior inferotemporal cortex. Neuron 84 (1), 55–62. [PubMed: 25242216]

Walther DB, Chai B, Caddigan E, Beck DM, Fei-Fei L, 2011. Simple line drawings suffice for functional MRI decoding of natural scene categories. Proc. Natl. Acad. Sci 108 (23), 9661–9666. [PubMed: 21593417]
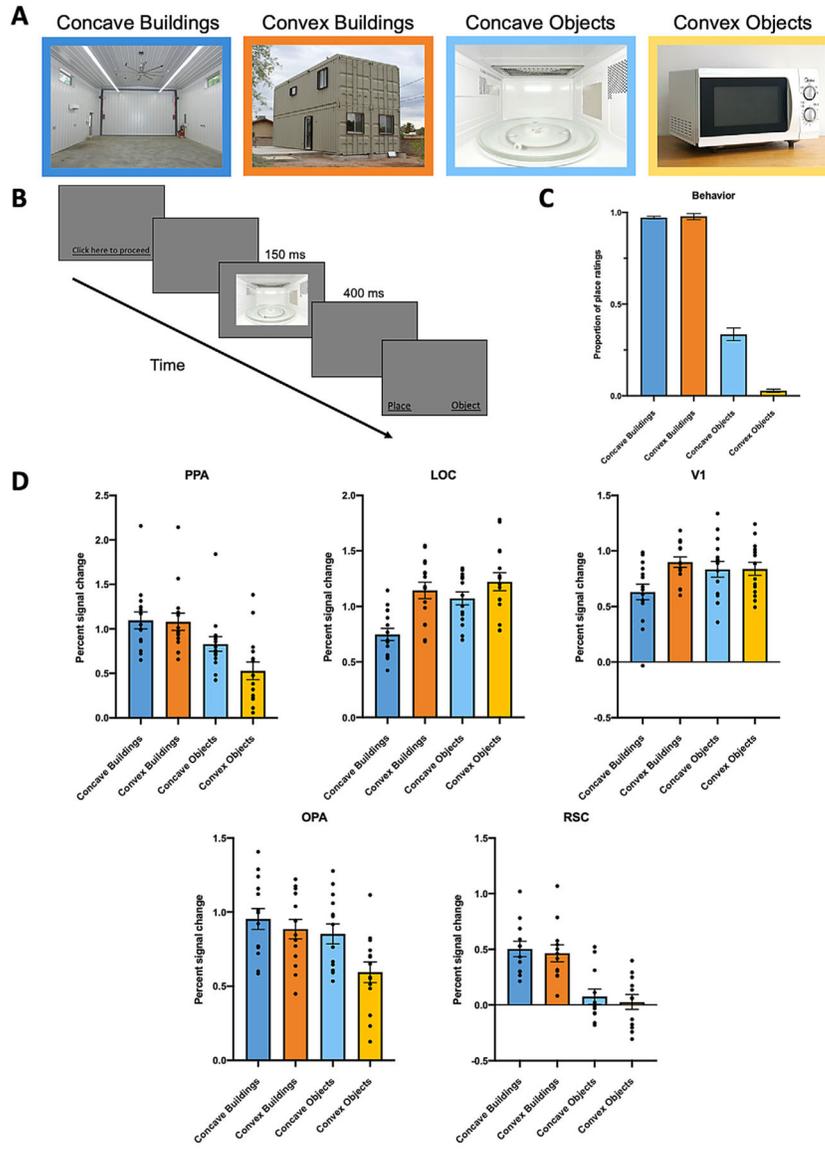
Fig. 1.

A, Example stimuli used in Experiments 1 and 2, varying in Category (Buildings, Objects) and Condition (Concave, Convex). **B,** Experimental procedure for Experiment 1. After a brief presentation of an image, participants were asked to indicate whether the image they just saw was a "place" or an "object". **C**, Participants' proportion of place ratings for all the stimuli conditions. There is a significantly greater proportion of place ratings for Concave Objects over Convex Objects, but not for Concave Buildings over Convex Buildings. **D,** Average percent signal change in each region of interest to the four conditions of interest. PPA and OPA responded significantly more to Concave Objects than Convex Objects, whereas LOC responded significantly more to Convex over Concave Buildings and Objects, respectively. The response patterns in PPA and OPA are qualitatively different from those in LOC and V1. We found no significant difference between RSC response to Concave versus

Convex Buildings and Objects, respectively. Error bars represent the standard error of the mean.
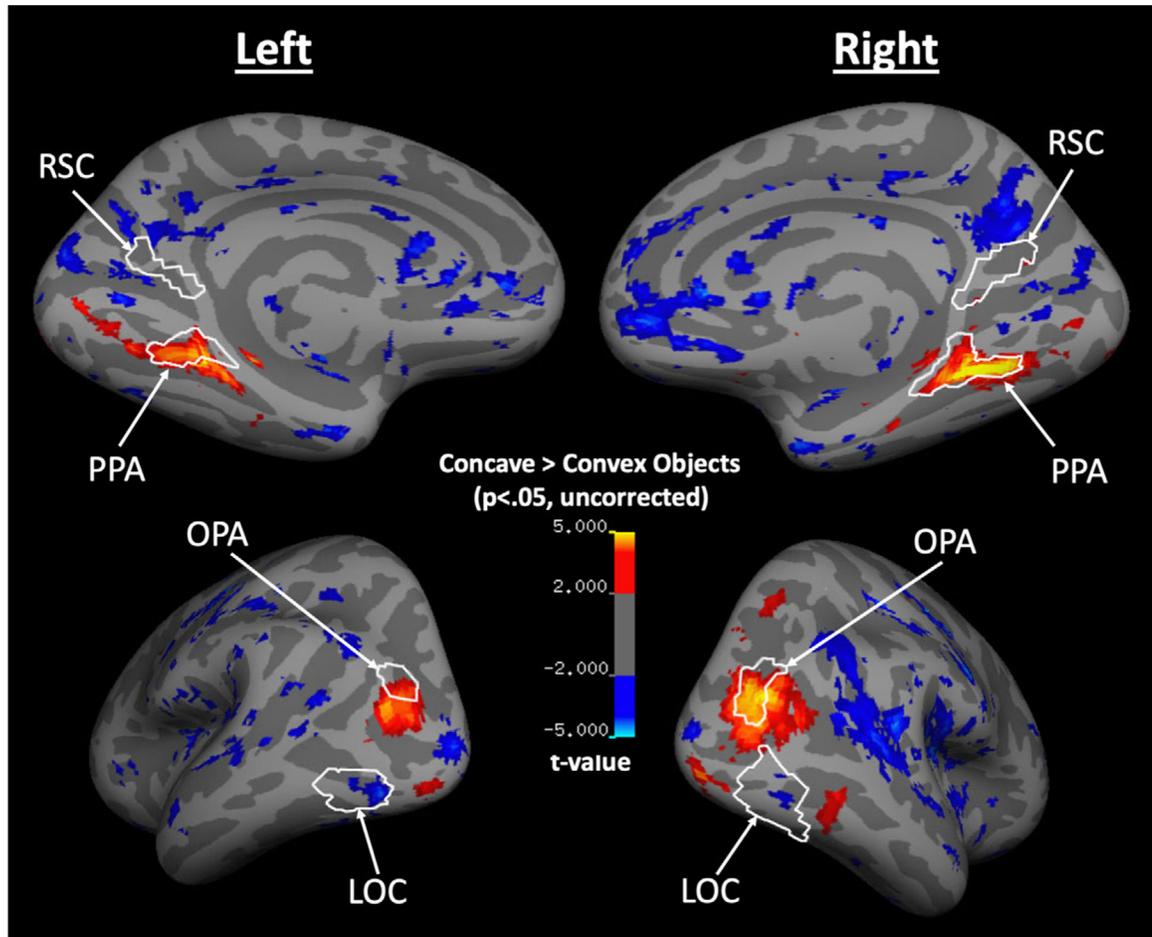
Fig. 2.
A group cortical surface map for regions that responded more to Concave Objects than Convex Objects. White lines indicate the ROIs that were functionally defined at the group level, using an independent set of Localizer runs. PPA and OPA overlap with cortical regions that respond more to Concave over Convex Objects, with a relatively stronger Concave selectivity at the more posterior parts of these cortical regions.
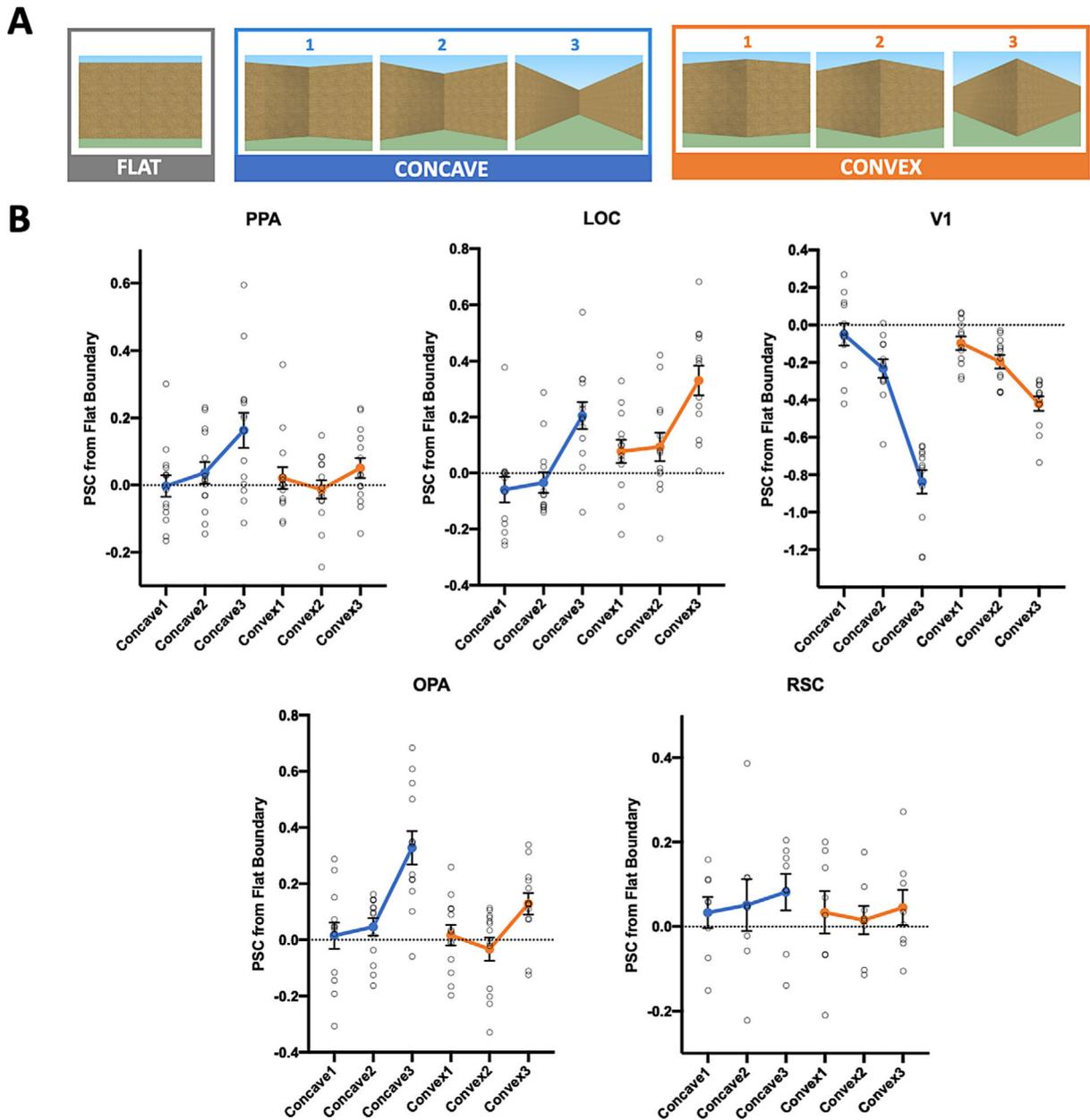
Fig. 3.

A, Example stimuli used in Experiment 3, varying in both Boundary Type (Concave, Convex) and Angle (1, 2, 3). All Concave and Convex conditions that share the same Angle are equated in the low-level visual information they contain. A "flat" boundary condition was also included, and used as a baseline condition to eliminate an ROI's baseline response to the presence of a boundary. **B,** Average percent signal change to the concave and convex boundaries relative to the flat boundary in each region of interest. PPA's response increases as the concavity of a boundary increases, but not with increasing convexity of a boundary. By contrast, LOC tracks changes in both concavity and convexity of a boundary, but shows a preferential response to convex boundaries. The pattern of response in PPA is qualitatively

different from those in V1 and LOC. OPA shows a greater sensitivity to changes in concavity over convexity. We found no significant response pattern in RSC. Error bars represent the standard error of the mean.
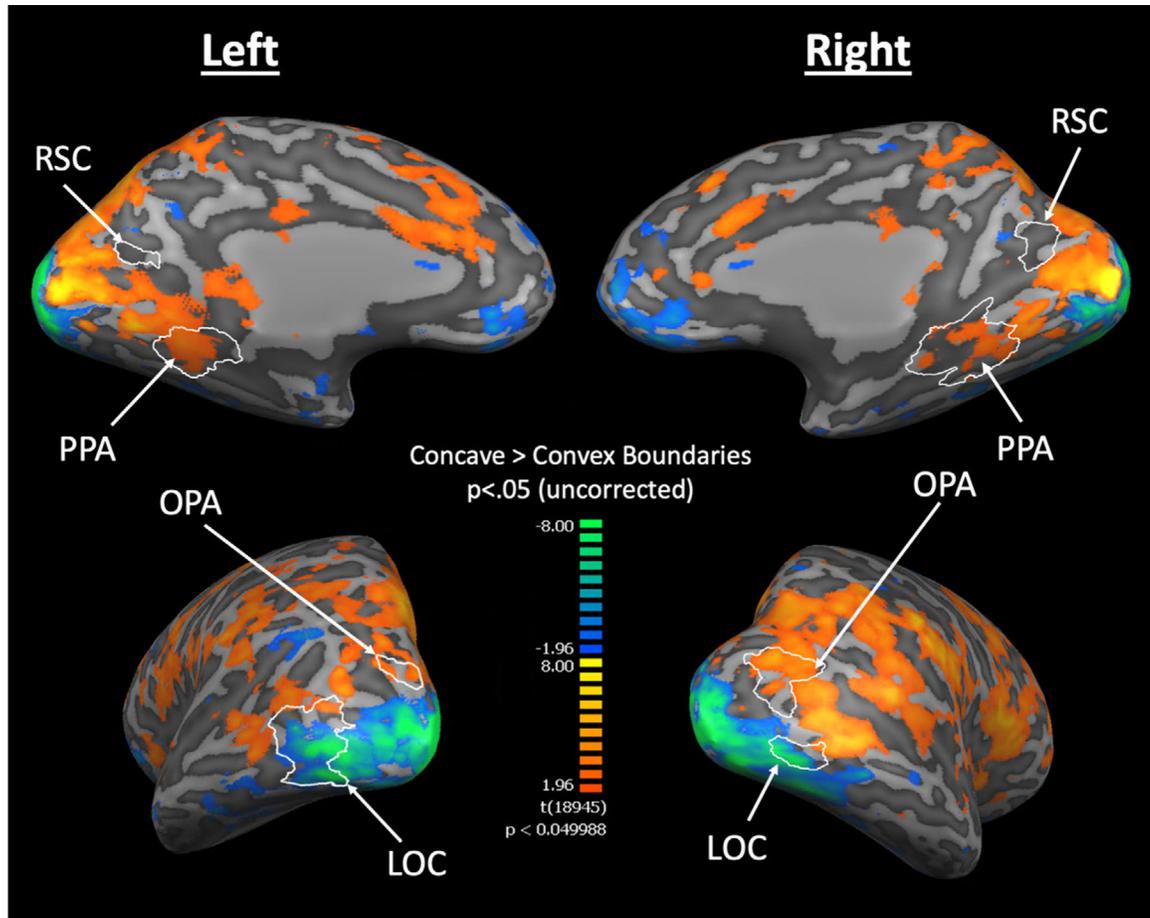
Fig. 4.
A group cortical surface map for regions that responded more to Concave than Convex Boundaries (averaged across Angles 1, 2, 3). White lines indicate the ROIs that are functionally defined at the group level using an independent set of Localizer runs. PPA and OPA overlap with the Concave-selective cortical regions, with Concave selectivity localized at the posterior parts of PPA and RSC. We also observed distinct streams of Concave and Convex selectivity along the ventral occipitotemporal cortex.
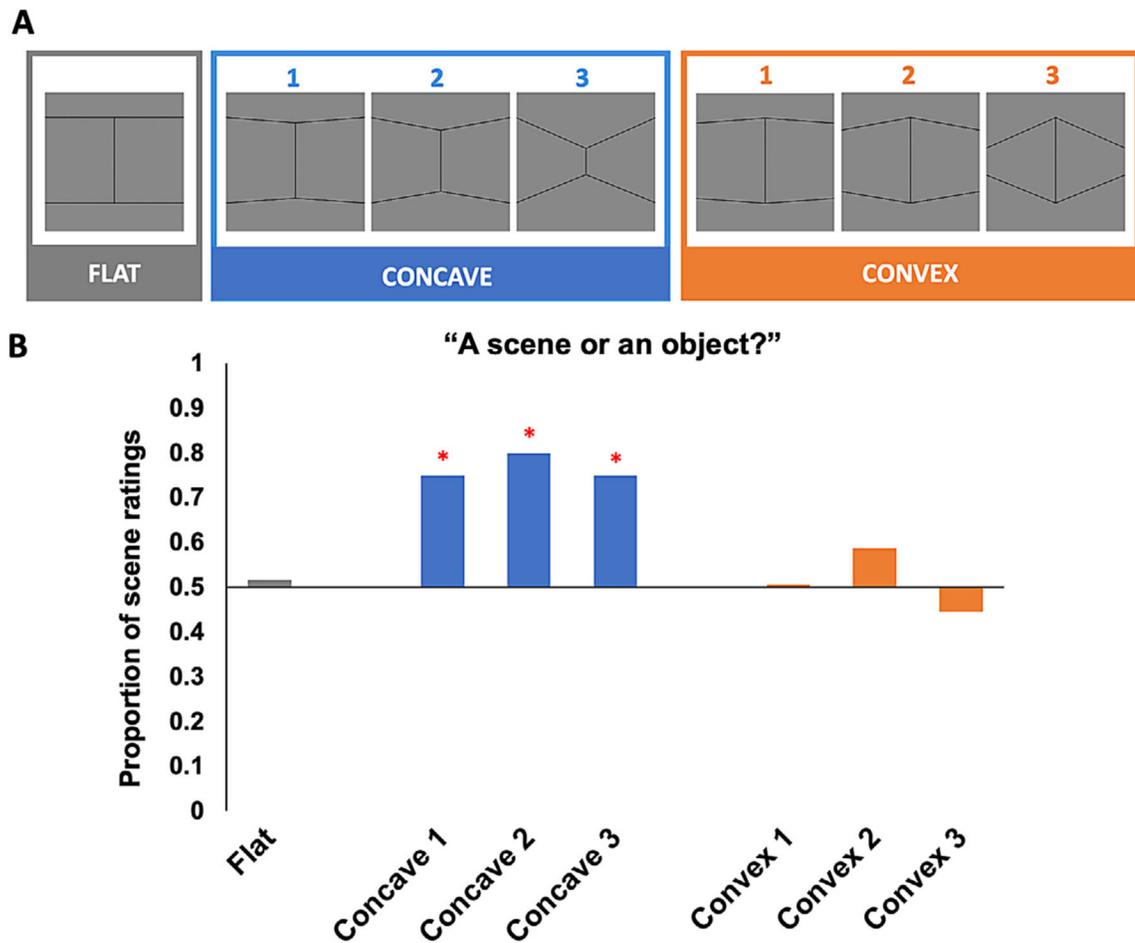
Fig. 5.
A, Stimuli tested in Experiment 4. **B,** Participants' proportion of scene ratings when asked to indicate whether each of these line-drawing images was a "scene" or an "object". The proportion of scene ratings for all the Concave conditions is significantly above chance, but not for the Flat nor the Convex conditions. Participants also rated the Concave shapes as scenes significantly more often than the corresponding Convex shapes with the same Angle.