OXFORD

Cerebral Cortex, 2025, **35**, bhaf082
https://doi.org/10.1093/cercor/bhaf082
Advance access publication date 23 April 2025
Original Article

A scene-selective region in the superior parietal lobule for visually guided navigation

Hee Kyung Yoon[1], Yaelan Jung[1], Andrew S. Persichetti[2], and Daniel D. Dilks[1,*]

[1]Department of Psychology, Emory University, 36 Eagle Row, Atlanta, GA, 30322, United States
[2]Section on Cognitive Neuropsychology, Laboratory of Brain and Cognition, National Institute of Mental Health, National Institutes of Health, Bethesda, MD 20892, United States

*Corresponding author: Daniel D. Dilks, Department of Psychology, Emory University, 36 Eagle Row, Atlanta, GA 30322, United States. Email: dilks@emory.edu

Growing evidence indicates that the occipital place area (OPA) is involved in "visually guided navigation." Here, we propose that a recently uncovered scene-selective region in the superior parietal lobule is also involved in visually guided navigation. First, using functional magnetic resonance imaging (fMRI), we found that the superior parietal lobule (SPL) responds significantly more to scene stimuli than to face and object stimuli across two sets of stimuli (i.e. dynamic and static), confirming its scene selectivity. Second, we found that the SPL, like the OPA, processes two kinds of information necessary for visually guided navigation: first-person perspective motion and sense (left/right) information in scenes. Third, resting-state fMRI data revealed that SPL is preferentially connected to OPA, compared to other scene-selective regions, indicating that SPL and OPA are part of the same system. Fourth, analysis of previously published fMRI data showed that SPL, like OPA, responds significantly more while participants perform a visually guided navigation task compared to both a scene categorization task and a baseline task, further supporting our hypothesis in an independent dataset. Taken together, these findings indicate the existence of a new scene-selective region for visually guided navigation and raise interesting questions about the precise role that SPL, compared to OPA, may play within visually guided navigation.

Keywords: connectivity; fMRI; navigation; parietal cortex; place processing.

Introduction

Cognitive neuroscience of the past three decades has revealed a set of three cortical regions that together make up human visual place (or "scene") processing: the parahippocampal place area (PPA; Epstein and Kanwisher 1998), the occipital place area (OPA; Dilks et al. 2013), and the medial place area (MPA; or retrosplenial complex, RSC; Silson et al. 2016; Maguire 2001). These three regions are so-called "scene selective" because they each respond about two to four times more to images of scenes than to images of faces and objects in adult human functional magnetic resonance imaging (fMRI) studies. Beyond scene selectivity, however, a considerable body of evidence suggests that each of these three scene-selective regions plays a distinct role within the broader domain of human visual scene processing (Walther et al. 2009; Dilks et al. 2011; Walther et al. 2011; Persichetti and Dilks 2016; Kamps et al. 2016b; Bonner and Epstein 2017; Dillon et al. 2018; Suzuki et al. 2021; Jones et al. 2023). For example, we have proposed that PPA is involved in "scene categorization" (i.e. our ability to recognize a scene as a particular kind of place), OPA is involved in "visually guided navigation" (i.e. our ability to move about the immediately visible environment, avoiding boundaries and obstacles), and MPA is involved in "map-based navigation" (i.e. our ability to find our way from a specific place to some distant, out-of-sight place) (Dilks et al. 2022).

However, is each of these scene-selective regions the only one involved in its proposed distinct function? Probably not. Indeed, several previous fMRI studies have hinted at regions within the superior parietal lobule (SPL) that, like OPA, may be involved in visually guided navigation (van Assche et al. 2016; Kamps et al. 2016a; Kamps et al. 2016b; Persichetti and Dilks 2018; Suzuki et al. 2021; Sulpizio et al. 2023; Kennedy et al. 2024). However, whether these regions (i) are indeed scene selective, as expected if a region is involved in visually guided navigation and (ii) represent information specific to visually guided navigation remains unknown, as none of the above studies showed both. For example, Persichetti and Dilks (2018) reported a scene-selective region (responding more to images of scenes than objects) within the SPL that overlapped with a swath of cortex extending from the inferior to superior parietal lobule that responded more while participants performed a "visually guided navigation" task than when they performed a "scene categorization" task—suggesting a scene-selective region in the SPL involved in visually guided navigation. Note, however, that since this finding was not the focus of the paper, the authors never directly tested this hypothesis, revealing that this scene-selective region was indeed involved in visually guided navigation and not scene categorization. Similarly, another paper by Kennedy et al. (2024) showed that a region within the posterior intraparietal gyrus, which they termed the "posterior intraparietal gyrus scene-selective (PIGS) site," is scene selective—responding significantly more to images of scenes than to faces or objects. They further found that this region processes egocentric motion through scenes, showing that it elicits a greater response to coherent motion through scenes compared to incoherent motion through scenes. However, despite the intriguing possibility that this region then may be involved in visually guided navigation—since egocentric motion through scenes is one kind of information necessary for visually guided navigation—this study did not include any nonscene motion conditions, such as faces and objects in motion, to conclude that this region is selectively responding to egocentric motion through scenes versus motion more generally.

Received: September 23, 2024. **Revised:** March 10, 2025. **Accepted:** March 14, 2025

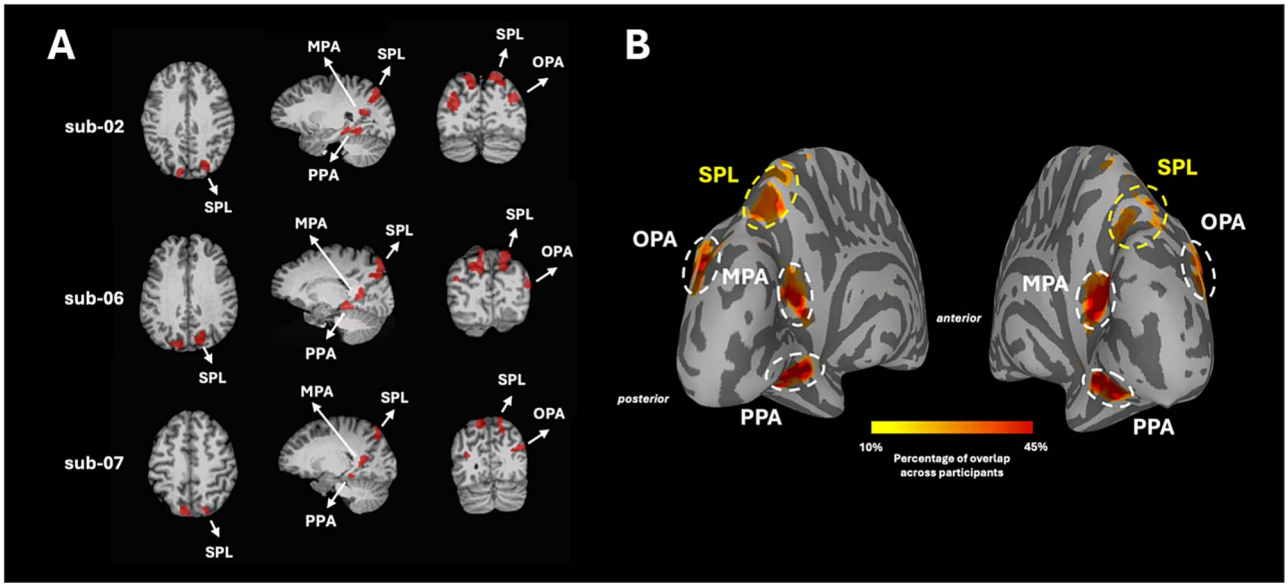

Fig. 1. A) Scene-selective regions (Dynamic Scenes > Dynamic Objects) shown in the axial, sagittal, and coronal imaging planes (from left to right) in 3 representative participants, from Experiment 1, in a volumetric presentation. Note in the sagittal plane that SPL is aligned diagonally with MPA and PPA, allowing for a convenient identification of this region. Likewise, in the coronal plane, SPL can be found in the same plane as OPA. See Supplementary Fig. 1 for the identification of SPL in all participants. B) Probabilistic maps of scene-selective regions (left and right hemispheres) derived from the Dynamic Scenes > Dynamic Objects contrast (across 16 participants from Experiment 1) projected onto a 3D reconstruction of the brain. Mean MNI coordinates (x, y, z) for SPL were as follows: left SPL (−10.75, −74.81, 53.75) and right SPL (16.56, −74.31, 50.38). Created by the authors.

Thus, here we ask whether a part of the superior parietal lobule is indeed scene selective *and* is involved in visually guided navigation across four experiments. In Experiments 1 and 2, we confirmed a scene-selective region in SPL (Fig. 1) and then tested whether SPL processes two kinds of information necessary for visually guided navigation specifically—that is, first-person perspective motion information through scenes (i.e. information mimicking the actual visual experience of walking through a place) and "sense" (left/right) information in scenes (Kamps et al. 2016b; Jones et al. 2023; Jung et al 2024). In Experiment 3, using resting-state fMRI, we asked whether SPL is preferentially connected to OPA, compared to both PPA and MPA, to determine whether SPL and OPA are part of the same system, both supporting visually guided navigation. In Experiment 4, using a preexisting fMRI dataset (Persichetti and Dilks 2018), we asked whether the SPL responds significantly more while participants performed a "visually guided navigation" task compared to both a "scene categorization" and a "baseline" task, further supporting our hypothesis in an independent dataset.

Materials and methods

Experiment 1

In Experiment 1, using fMRI, we examined (i) whether a region within the SPL is scene selective, as proposed in previous work (Persichetti and Dilks 2018; Kennedy et al. 2024), and (ii) whether SPL responds specifically to first-person perspective motion through scenes—again, one kind of information necessary for visually guided navigation—compared to two other types of nonscene motion (i.e. faces in motion and objects in motion).

Participants

Seventeen healthy participants were recruited for this experiment. All participants gave informed consent. All participants had normal or corrected-to-normal vision and had no history of neurological or psychiatric conditions. One participant was excluded for excessive motion. Thus, we report the results from 16 participants (ages 20 to 40; $M = 26.6$, 8 females) for the task-based fMRI analysis. This sample size was determined based on previous work employing a similar paradigm (Kamps et al. 2016b; Jung et al. 2018; Jones et al. 2023). Participants were compensated for their participation upon completing the study. All procedures were approved by the Emory University Institutional Review Board.

Design and stimuli

We used a region of interest (ROI) approach in which we used one set of runs to define SPL, PPA, OPA, and MPA and then used another independent set of runs to investigate the responses in each of these ROIs, as detailed in the Data Analysis section. For all runs, we used a blocked design in which participants viewed stimuli from six conditions: Dynamic Scenes, Static Scenes, Dynamic Faces, Static Faces, Dynamic Objects, and Static Objects. For each of the Dynamic Scene, Dynamic Face, and Dynamic Object conditions, there were 48, 3-s video clips, subtending 23 × 15 degrees of visual angle. The Dynamic Scene stimuli portrayed first-person perspective motion, as one would experience from walking through a place (Kamps et al. 2016b; Kamps et al. 2020b) (Fig. 2). Indeed, the video clips were filmed by walking through 8 different places (e.g. a parking garage, a hallway) at a typical pace with a camera (a Sony HDR XC260V HandyCam with a field of view of 90.3 × 58.9 degrees) held at eye level. The Static Scenes stimuli were created by taking stills from each Dynamic Scene video clip at 1-, 2-, and 3-s time points, resulting in 144 images. These still images were presented in groups of three images from the same video clip, and each image was presented for 1 s with no interstimulus interval (ISI), thus equating the presentation time of the static images with the duration of the video clips from which they were made (still totaling 3 s). Importantly, the still images were presented in a quasi-random sequence such that first-person perspective motion could not be inferred and

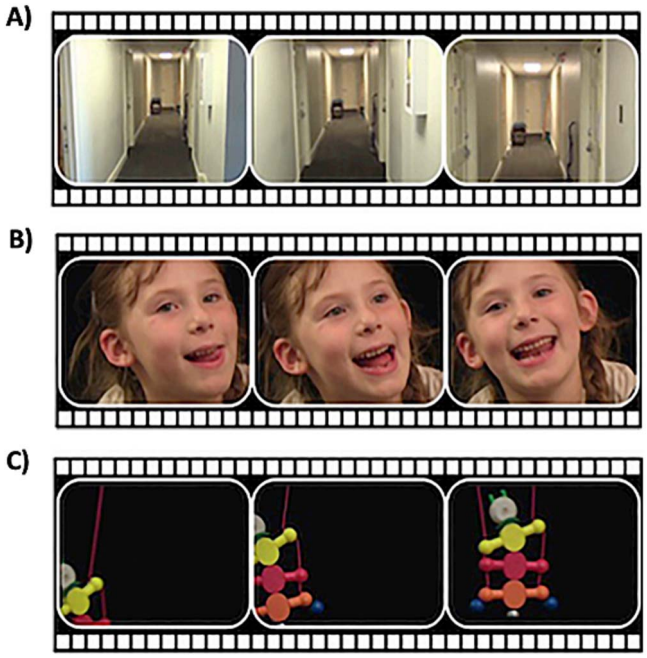

Fig. 2. Example dynamic stimuli, replicated from Kamps et al. (2020b). The conditions included A) Dynamic Scenes, which consisted of 3-s video clips of first-person perspective motion through a scene; Static Scenes, which consisted of three 1-s stills taken from the Dynamic Scenes condition and presented in a quasi-random order, such that first-person perspective motion could not be inferred; B) Dynamic Faces, which consisted of 3-s video clips of only the faces of children against a black background as they laughed, smiled, and looked around; and Static Faces, which consisted of three 1-s stills taken from the Dynamic Faces and presented in quasi-random order; C) Dynamic Objects, which consisted of 3-s video clips of moving toys against a black background; Static Objects, which consisted of three 1-s stills taken from Dynamic Objects and presented in a quasi-random order. Replicated from Kamps et al. 2020b. Used with permission from Elsevier B.V. Permission granted under a Formal Permission License (FPL).

quasi-random such that the still images were never presented in their original temporal order. Like the video clips, the still images also subtended 23 × 15 degrees of visual angle. The Dynamic Face stimuli showed only the faces of 7 children against a black background as they smiled, laughed, and looked around while interacting with off-screen toys or adults, and the Dynamic Object stimuli depicted moving toys against a black background (Pitcher et al. 2011). The Static Face and Static Object stimuli were created and presented using the exact same procedure and parameters described for the Static Scene condition above.

Each participant completed 6 runs, except for one participant who only completed 5 runs due to time constraints. Each run was 330 s long and consisted of 2 blocks per stimulus category. For each run, the order of the first six blocks was pseudorandomized (e.g. Static Objects, Dynamic Faces, Dynamic Scenes, Static Faces, Dynamic Objects, Static Scenes), and the order of the remaining blocks was the palindrome of the first six (e.g. Static Scenes, Dynamic Objects, Static Faces, Dynamic Scenes, Dynamic Faces, Static Objects). Dynamic blocks (i.e. Dynamic Scene, Dynamic Face, Dynamic Object) contained six 3-s video clips from the same category. In the static blocks, each image was presented for 1 s, and there were 6 sets of 3 images of Static Scenes, Static Faces, or Static Objects. For all blocks, there was a 1.5-s blank screen at the beginning of each block and an ISI of 500 ms (total block duration of 22 s). For all runs, we also included three 22-s fixation blocks: one at the beginning, one after the first six blocks, and one at the end. Participants were instructed to passively view the stimuli and were not required to perform any task throughout the scan session.

fMRI scanning

All scan sessions were carried out using a 3T Siemens MAGNETOM Prisma scanner in 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 Center for Magnetic Resonance Research (CMRR), Department of Radiology, University of Minnesota multiband accelerated EPI sequence (60 slices, TR = 2 s, TE = 30 ms, voxel size = 2.0 × 2.0 × 2.0 mm, multiband acceleration factor = 2) (Xu et al. 2013). For all scans, slices were acquired parallel to the anterior commissure–posterior commissure line, covering all the occipital and parietal lobes, as well as most of the temporal lobe. Whole-brain, high-resolution anatomical images were also acquired for each participant for purposes of registration and anatomical localization.

Data analysis

fMRI data analysis was conducted using the Analysis of Functional NeuroImages (AFNI) software (Cox 1996; Cox and Jesmanowicz 1999) (version 20.3.02). Prior to statistical analysis, images were motion-corrected using 3dVolreg and then spatially smoothed with a 6-mm kernel with 3dMerge. After preprocessing, the SPL, OPA, PPA, and MPA were bilaterally defined in all 16 participants by identifying regions that responded significantly more to Dynamic Scenes than Dynamic Objects ($P < 10^{-4}$, uncorrected) (Fig. 1). Anatomically, the SPL is located adjacent and anterior to the parieto-occipital sulcus (see Fig. 1). Finally, the SPL could also be defined using the contrast of Static Scenes > Static Objects in all 16 participants (see Supplementary Fig. 3).

For each participant, we used 3 (either odd or even runs; counterbalanced across participants) of the 6 total runs to define each ROI and then used the remaining set of independent runs to investigate the responses in each ROI to the six conditions. We also defined three control regions: primary visual cortex (V1) and the middle temporal (MT) area (Tootell et al. 1995) using the probabilistic atlas from Wang et al. (2015) and V6 using the parcel from Glasser et al. (2016). Critically, V6 voxels were excluded from each participant's SPL, and SPL voxels were excluded from each participant's V6, to ensure that the two regions do not share any voxels, though the overlap between the regions was minimal.

Next, for each ROI for each participant, the average response across voxels for each condition was extracted and converted to percent signal change relative to the fixation baseline. Statistical analyses on the percent signal changes in the ROIs were conducted using R (R Core Team, 2013) (version 2022.07.2) and JASP statistical software (JASP Team, 2023) (version 0.17.1).

Finally, a 2 (Hemisphere: Left, Right) × 4 (ROI: SPL, PPA, OPA, MPA) × 6 (Condition: Dynamic Scenes, Dynamic Faces, Dynamic Objects, Static Scenes, Static Faces, Static Objects) repeated-measures ANOVA did not reveal a significant hemisphere by ROI by condition interaction [$F_{(15, 225)} = 2.42$, $P = 0.065$, Greenhouse–Geisser-corrected]; thus, data from each hemisphere were collapsed for all further analysis.

Experiment 2

In Experiment 2, using an fMRI adaptation paradigm, we asked whether the SPL represents sense information in scenes—another kind of information necessary for visually guided navigation.

Participants

Sixteen additional healthy participants (ages 21 to 41; $M = 27.94$, 11 females) were recruited for this experiment. This sample

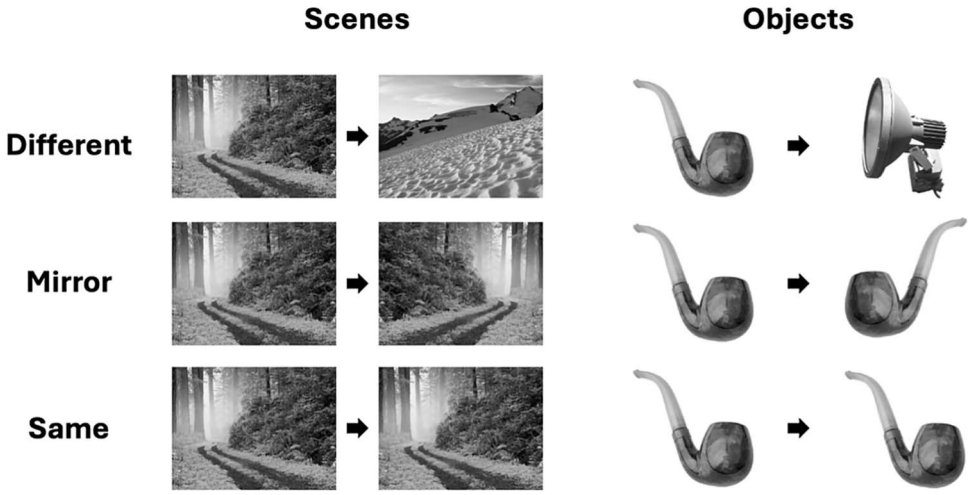

Fig. 3. Example stimuli, adapted from Dilks et al. 2011. The conditions included Different Scenes, Mirror Scenes, and Same Scenes (left column, from top to bottom) as well as Different Objects, Mirror Objects, and Same Objects (right column, from top to bottom). The "Different" conditions consisted of repeated pairs of different scene or object images; the "Mirror" conditions consisted of repeated pairs of different scene or object images followed by their mirror-reversed images (the original image's reflection about the vertical axis); and the "same" conditions consisted of repeated pairs of the same scene or object image. Adapted from Dilks et al. 2011. Available under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License: https://creativecommons.org/licenses/by-nc-sa/3.0/

size was determined based on previous work using a similar paradigm (Dilks et al. 2011; Persichetti and Dilks 2016). Of the 16 participants, 6 had previously participated in Experiment 1. All participants gave informed consent. All participants had normal or corrected-to-normal vision and had no history of neurological or psychiatric conditions. Participants were compensated for their participation upon completing the study. All procedures were approved by the Emory University Institutional Review Board.

Design and stimuli

We used an ROI approach in which we used a set of localizer runs to identify category-selective regions, SPL, PPA, OPA, and MPA, and then used another independent set of experimental runs to investigate the responses in each of these ROIs, as detailed in the Data Analysis section.

Each participant completed 2 runs of the localizer scan. Each run was 312 s long and consisted of 3 blocks per stimulus category. For each run, the order of the stimulus category blocks in each run was palindromic (e.g. Dynamic Faces, Dynamic Objects, Dynamic Scenes, Dynamic Scrambled Objects and Dynamic Scrambled Objects, Dynamic Scenes, Dynamic Objects, Dynamic Faces). Each block contained six 3-s video clips from the same category for a total of 18 s. For all runs, we also included four 12-s fixation blocks: one at the beginning, one in the middle, and one at the end of the run.

For the experimental runs, participants completed 6 runs, each of which was 296 s long and consisted of 2 blocks per condition (Different Scenes, Mirror Scenes, Same Scenes, Different Objects, Mirror Objects, Same Objects). Each block consisted of 6 repetitions of a pair of images that were either entirely different, mirror-reversed images, or exactly the same (Fig. 3). All stimuli were grayscale images, 15° × 10° in size, and were presented for 800 ms. Participants performed an orthogonal task (not related to whether an image was a scene or object or whether it was mirror-reversed), responding via button box whether a presented image differed in luminosity compared to the rest of the images in the block. The trial with the different level of luminosity was presented after at least two trials of images so that the participants could establish a point of comparison.

fMRI scanning

Scanning was identical to Experiment 1.

Data analysis

fMRI data analysis was conducted using the AFNI software (Cox 1996; Cox and Jesmanowicz 1999) (version 20.3.02). Prior to statistical analysis, images were motion-corrected using 3dVolreg and then spatially smoothed with a 6-mm kernel with 3dMerge. After preprocessing, the SPL, OPA, PPA, and MPA were bilaterally defined in all 16 participants by identifying regions that responded significantly more to Dynamic Scenes than Dynamic Objects ($P < 10^{-4}$, uncorrected). For each participant, we used 2 localizer runs to define each ROI and then used the remaining set of independent runs to investigate the responses in each ROI to the six conditions. Next, for each ROI for each participant, the average response across voxels for each condition was extracted and converted to percent signal change relative to the fixation baseline. Statistical analyses on the percent signal changes in the ROIs were conducted using R (R Core Team, 2013) (version 2022.07.2) and JASP statistical software (JASP Team, 2023) (version 0.17.1).

Finally, a 2 (Hemisphere: Left, Right) × 4 (ROI: SPL, PPA, OPA, MPA) × 6 (Condition: Different Scenes, Mirror Scenes, Same Scenes, Different Objects, Mirror Objects, Same Objects) repeated-measures ANOVA did not reveal a significant hemisphere by ROI by condition interaction [$F_{(15, 225)} = 1.84$, $P = 0.113$, Greenhouse–Geisser-corrected, $\eta_p^2 = 0.11$); thus, data from each hemisphere were collapsed for all further analyses.

Experiment 3

In Experiment 3, using resting-state fMRI, we investigated the functional connectivity between SPL and OPA, compared to both PPA and MPA, asking whether SPL and OPA are preferentially connected and hence both part of the visually guided navigation system.

Participants

Twenty healthy participants were recruited for the resting-state scan. All participants gave informed consent. All participants had normal or corrected-to-normal vision and had no history of neurological or psychiatric conditions. Two participants were excluded because the SPL could not be defined in these individuals. Thus, we report results from 18 participants (ages 19 to 39; $M = 24.7$, 9 females) for the resting-state connectivity analysis. This sample size was determined based on previous resting-state studies that examined high-level visual areas (Baldassano et al. 2013; Steel et al. 2021). Of the 18 participants, five had previously participated in Experiment 1, and two had previously participated in Experiment 2. Participants were compensated for their participation upon completing the study. All procedures were approved by the Emory University Institutional Review Board.

Design and stimuli

We used an ROI approach in which we used a localizer scan to define SPL, OPA, PPA, and MPA and then used a resting-state scan to investigate how SPL is functionally connected to each of the other scene-selective regions. For the localizer scan, we used a blocked design in which participants viewed stimuli from four conditions: Dynamic Faces, Dynamic Objects, Dynamic Scenes, and Dynamic Scrambled Objects, as previously described (Pitcher et al. 2011).

Each participant completed 2 runs of the localizer scan. Each run was 312 s long and consisted of 3 blocks per stimulus category. For each run, the order of the stimulus category blocks in each run was palindromic (e.g. Dynamic Faces, Dynamic Objects, Dynamic Scenes, Dynamic Scrambled Objects; and Dynamic Scrambled Objects, Dynamic Scenes, Dynamic Objects, Dynamic Faces). Each block contained six 3-s video clips from the same category for a total of 18 s. For all runs, we also included four 12-s fixation blocks: one at the beginning, one in the middle, and one at the end of the run. For the resting-state scan, each participant completed a 6-min scan where they were instructed to close their eyes and to not think about anything in particular while refraining from falling asleep.

fMRI scanning

All scan sessions were carried out using a 3T Siemens MAGNETOM Prisma scanner in the FERN at Emory University (Atlanta, Georgia). 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 × 1.5 × 2.5 mm, and a 0.5 interslice gap) for the localizer scans, and (60 slices, TR = 2 s, TE = 30 ms, voxel size = 2 × 2 × 2 mm) for the resting-state scan. For all scans, slices were acquired parallel to the anterior commissure–posterior commissure line, covering all the occipital and parietal lobes, as well as most of the temporal lobe. Whole-brain, high-resolution anatomical images were also acquired for each participant for purposes of registration and anatomical localization.

Data analysis

fMRI data analysis was conducted using the AFNI software (Cox 1996; Cox and Jesmanowicz 1999) (version 20.3.02) for both the localizer scan and the resting-state scan. Prior to statistical analysis, fMRI data from the localizer runs were registered to a T1w reference (that is already aligned to the resting-state scan) using align_epi_anat.py (AFNI) and corrected for head-motion using 3dDespike and 3dvolreg (AFNI). For the localizer scans, data were spatially smoothed using a 6-mm kernel with 3dmerge. After preprocessing, the SPL, OPA, PPA, and MPA were bilaterally defined in all 18 participants by identifying regions that responded significantly more to Dynamic Scenes than Dynamic Objects ($P < 10^{-4}$, uncorrected).

For the resting-state scan, fMRI data were first corrected for head motion using 3dvolreg (AFNI). Before motion correction, volumes with movement >2 mm were corrected via interpolation between the nearest nonaffected volumes to reduce abrupt signal changes caused by head motion (3dDespike, AFNI). Temporal smoothing was performed to retain only low-frequency signals between 0.01 and 0.08 Hz. No spatial smoothing was applied. Nuisance regression was performed to remove motion artifacts (six motion parameter estimates) and the mean signal from a ventricular and white matter region. T1w image was registered to the middle image of the resting-state scan using align_epi_anat.py (AFNI). After preprocessing, we next extracted the continuous time series data from all the voxels from each ROI, averaged them, and calculated an average time series for each ROI. We then assessed the connectivity (or the cofluctuation) between the SPL and each of the other scene-selective ROIs by calculating the correlation between the time series for each pair (i.e. SPL and OPA, SPL and PPA, and SPL and MPA) while partialing out the time series data from the remaining ROIs. For example, when correlating the time series data from SPL and OPA, we partialed out the time series of PPA and MPA to observe the connectivity between those areas above and beyond the shared responses across the other ROIs. Correlation coefficients (r) were then transformed to Gaussian-distributed z-scores via Fisher's transformation for further analyses.

Experiment 4

In Experiment 4, using a preexisting dataset from Persichetti and Dilks (2018), we asked whether the SPL would respond more while participants performed a visually guided navigation task compared to both a scene categorization task and a baseline task.

Participants

The preexisting dataset included 11 healthy participants (ages 20 to 36; 4 females), who completed all three task conditions (discussed in detail below). None of the participants had participated in any of the other experiments described earlier in this paper. All participants gave informed consent and had normal or corrected-to-normal vision and no history of neurological or psychiatric conditions. Participants were compensated for their participation upon completing the study. All procedures were approved by the Emory University Institutional Review Board.

Design and stimuli

As detailed in Persichetti and Dilks (2018), we used an ROI approach in which we used a set of localizer runs to identify PPA, OPA, and MPA and then used another independent set of experimental runs to investigate the responses in each of these ROIs. For the current study, we further defined the SPL in each participant using a probabilistic map, derived from localizer data from several previous studies in our lab, including data from Experiments 1 to 3. This approach was necessary because we could not functionally define SPL (when using the localizer data from the Persichetti and Dilks study) in 6 of 11 participants, most likely due to an insufficient amount of localizer data. Unlike in Experiment 1 here, in which three localizer runs were available to define SPL using the static stimuli, the Persichetti and Dilks study had only two localizer runs.

For the Localizer runs to define PPA, OPA, and MPA—again, as detailed in Persichetti and Dilks (2018) but reiterated here for convenience—participants viewed images of faces, objects, scenes, and scrambled objects in separate blocks, as previously described (Epstein and Kanwisher 1998). Each participant completed two runs, each lasting 336 s and consisting of four blocks per stimulus category presented in random order. Each block included 20 images from the same category, shown for 300 ms each with a 500 ms ISI, for a total of 16 s per block. Five 16-s fixation blocks were interspersed: one at the beginning, three in the middle, and one at the end of each run. During the localizer task, participants performed a one-back task, responding whenever the same image appeared twice in a row. Again, however, to define SPL in each participant, we used a probabilistic map, derived from localizer data from several previous studies in our lab, including data from Experiments 1 to 3.

For the Experimental runs, participants completed 10 runs consisting of three task types: Navigation, Categorization, and Baseline. Each run included three blocks of each task presented in random order. Blocks were 24 s long and included 16 images displayed for 500 ms each, followed by a 1-s ISI. Fixation blocks (6 s each) were interleaved between task blocks, with additional fixation blocks at the start and end of each run. During fixation, a central white cross (0.5° visual angle) appeared on a neutral gray background. Instructions for the upcoming task were displayed before each block for 4 s.

In the Navigation task, participants imagined walking along a continuous path through a room and pressed a button to indicate whether the exit door was on the left, center, or right wall. In the Categorization task, participants imagined standing in the room and pressed a button to identify the room type (bedroom, kitchen, or living room). During the Baseline task, participants performed a one-back task, indicating whether the current image matched the previous one.

Stimuli were computer-generated images (12° × 9° visual angle) created using Sims 3 software (Electronic Arts), depicting rooms with furniture, doors, and a continuous path to one of the doors. Before training, participants watched first-person perspective videos of motion through rooms to familiarize themselves with the Navigation task. They then completed training sessions to ensure 80% accuracy on both Navigation and Categorization tasks before scanning. Following training, all participants reported successfully imagining navigating or categorizing on each trial.

fMRI scanning and data analysis

All scan sessions were carried out using a 3T Siemens Trio scanner in the FERN at Emory University (Atlanta, GA). 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, flip angle = 90°, voxel size = 1.5 × 1.5 × 2.5-mm, and a 0.2-mm interslice gap). Slices were oriented approximately between perpendicular and parallel to the calcarine sulcus, covering the occipital, temporal, parietal, and most of the frontal lobes. Whole-brain, high-resolution T1-weighted anatomical images (TR = 1900-msec, TE = 2.27 ms, inversion time = 900 ms, voxel size = 1 × 1 × 1 mm) were also acquired for each participant for purposes of registration and anatomical localization.

Data analysis

fMRI data analysis was conducted using FSL software (Smith et al. 2004) and custom MATLAB code for both the localizer and experimental runs. Prior to statistical analysis, images were skull-stripped (Smith 2002) and registered to T1-weighted anatomical images using FSL's registration tools (Jenkinson et al. 2002). Data were spatially smoothed with a 5-mm kernel, de-trended, and fit with a general linear model containing covariates convolved with a double-gamma function to approximate the hemodynamic response function.

After preprocessing, scene-selective ROIs (PPA, OPA, MPA) were bilaterally defined for each participant using the independent localizer runs. ROIs were identified as regions responding more strongly to scenes than objects ($P = 10^{-4}$, uncorrected), following the method of Epstein and Kanwisher (1998). PPA, MPA, and OPA were identified in at least one hemisphere in all 11 participants. The SPL, using our probabilistic map, was defined in all 11 participants.

For each ROI, average responses across voxels for each task were extracted and converted to percentage signal change relative to a fixation baseline. Repeated-measures ANOVAs were conducted to analyze task-specific responses for each ROI.

Results

Experiment 1

SPL is scene selective

If SPL is scene selective as previously shown (Persichetti and Dilks 2018; Kennedy et al. 2024), like the previously identified scene-selective regions (i.e. PPA, OPA, and MPA), then it will respond significantly more to scenes than face or object stimuli. We tested this hypothesis using two different sets of stimuli: "dynamic" and "static" (Fig. 2; see Materials and Methods for more details). Indeed, a 3-level (Condition: Dynamic Scenes, Dynamic Faces, Dynamic Objects) repeated-measures ANOVA revealed a significant effect of condition ($F_{(2, 30)} = 83.24$, $P < 0.001$, Greenhouse–Geisser-corrected; $\eta_p^2 = 0.85$), with the SPL responding significantly more to Dynamic Scenes than to both Dynamic Faces and Dynamic Objects (main effect contrasts, both $Ps < 0.001$) (Fig. 4A). Likewise, a 3-level (Condition: Static Scenes, Static Faces, Static Objects) repeated-measures ANOVA revealed a significant effect of condition [$F_{(2, 30)} = 49.94$, $P < 0.001$, $\eta_p^2 = 0.77$], with SPL responding significantly more to Static Scenes than to both Static Faces and Static Objects (main effect contrasts, both $Ps < 0.001$) (Fig. 4A). Note that SPL is also scene selective even when it is defined using static stimuli, ruling out the possibility that the scene selectivity in SPL is only exhibited when it is defined using dynamic stimuli (Supplementary Fig. 4). For completeness, we also confirmed the scene selectivity of the well-established OPA, PPA, and MPA, to ensure our paradigm was working. In all three regions, for both dynamic and static sets of stimuli, a 3-level ANOVA revealed a significant effect of condition (all $Ps < 0.001$), with each region responding significantly more to scenes than face and object stimuli (main effect contrasts, all $Ps < 0.001$) (Fig. 4B–D). Taken together, these findings suggest—across two different sets of stimuli—that SPL is scene selective.

SPL responds to first-person perspective motion information through scenes

If SPL supports visually guided navigation, then it should respond to first-person perspective motion information through scenes (i.e. information mimicking the actual visual experience of walking through a place), responding significantly more to Dynamic Scenes depicting first-person perspective motion information through scenes than to Static Scenes without first-person perspective motion information through scenes, relative to both Dynamic Faces versus Static Faces and Dynamic versus Static Objects. To test this hypothesis, we compared the difference

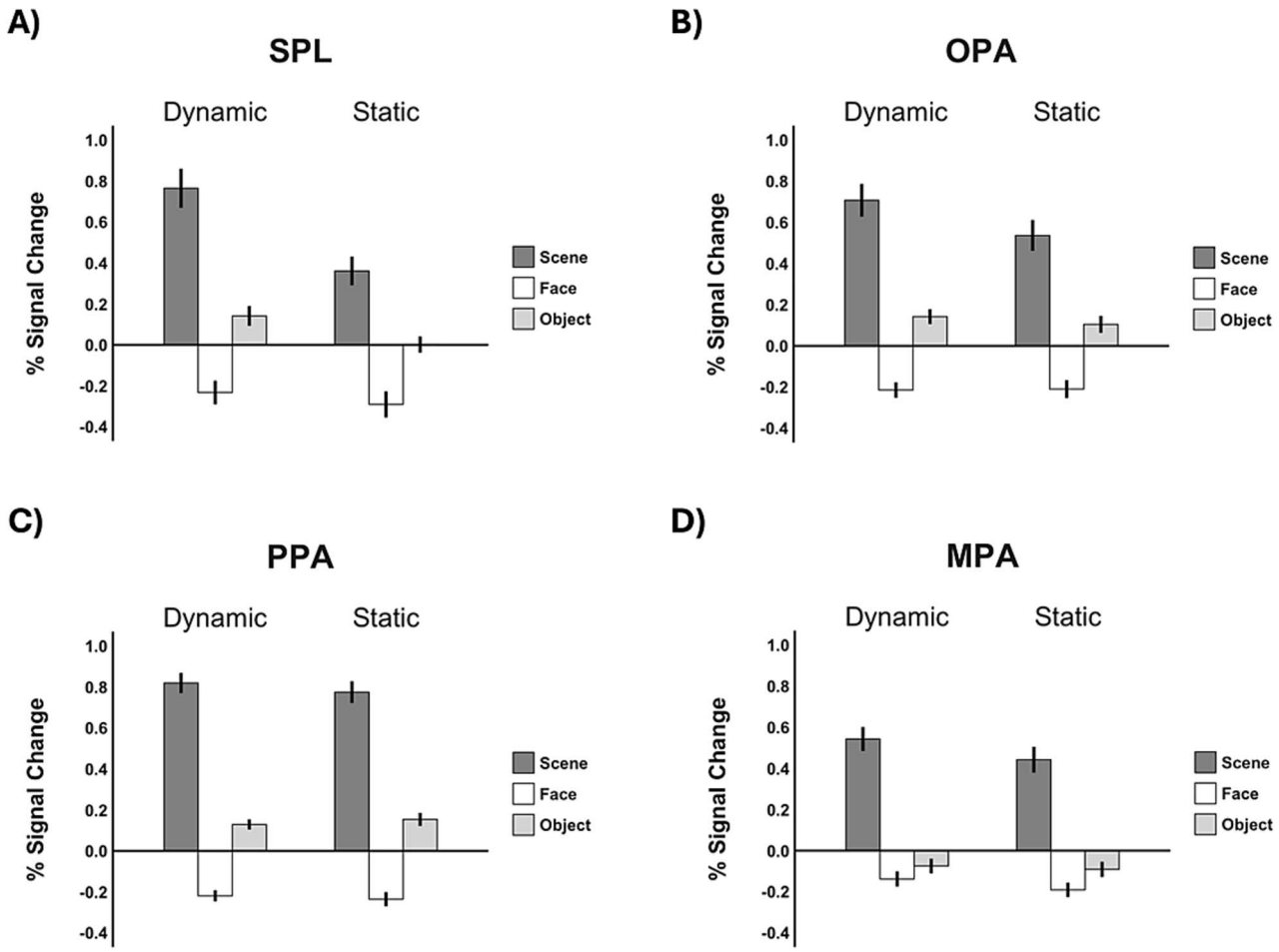

Fig. 4. Percent signal change for Dynamic stimuli (Dynamic Scenes, Dynamic Faces, Dynamic Objects) and for Static stimuli (Static Scenes, Static Faces, Static Objects), labeled accordingly, in SPL A), OPA B), PPA C), and MPA D). SPL, like OPA, PPA, and MPA, responded significantly more to both Dynamic and Static Scenes than to Dynamic and Static Faces and Dynamic and Static Objects—revealing that the SPL is scene selective. Error bars reflect the standard error of the mean. Created by the authors.

in response in SPL to Dynamic Scenes versus Static Scenes ("Scene Motion") with the difference in response to Dynamic Faces versus Static Faces ("Face Motion") and Dynamic Objects versus Static Objects ("Object Motion"). Supporting our hypothesis, a 3-level (Condition: Scene Motion, Face Motion, Object Motion) repeated-measures ANOVA revealed a significant effect of condition [$F_{(2, 30)} = 17.04$, $P < 0.001$, $\eta_p^2 = 0.5$], with the SPL responding significantly more to Scene Motion than to Face Motion and Object Motion (main effect contrasts, both *P*s < 0.001) (Fig. 5A). Note that SPL, when defined using static stimuli (versus dynamic stimuli, as just above) remains selective to scene motion, ruling out the possibility that the selectivity to scene motion is only exhibited when SPL is defined using dynamic stimuli (Supplementary Fig. 5).

Such first-person perspective motion selectivity was also found in OPA [$F_{(2, 30)} = 6.51$, $P = 0.004$, $\eta_p^2 = 0.30$], with OPA responding significantly more to Scene Motion than to Face Motion and Object Motion (main effect contrasts, $P = 0.002$, $P = 0.01$, respectively) (Fig. 5B), replicating a previous finding (Kamps et al. 2016b) and further supporting the hypothesis that OPA is involved in visually guided navigation. By contrast, no such first-person perspective motion selectivity was found in either PPA [$F_{(2, 30)} = 0.90$, $P = 0.42$, $\eta_p^2 = 0.06$] or MPA [$F_{(2, 30)} = 1.08$, $P = 0.35$, $\eta_p^2 = 0.07$] (Fig. 5C and D), as would be expected given neither PPA nor MPA are hypothesized to be involved in visually guided navigation. Taken together, these results suggest that SPL is involved in visually guided navigation, insofar as it responds to at least one kind of information necessary for visually guided navigation.

Next, we compared the scene motion selectivity in SPL to the other scene-selective regions, investigating whether (i) SPL is responding preferentially to scene motion over the non-scene motions—consistent with its hypothesized role in visually guided navigation—relative to the two other scene-selective regions known not to be involved in visually guided navigation (i.e. PPA and MPA) and (ii) the extent of scene motion selectivity is similar between SPL and OPA, both hypothesized to be involved in visually guided navigation. To address these two questions, we conducted a 4 (ROI: SPL, OPA, PPA, MPA) × 3 (Condition: Scene Motion, Face Motion, Object Motion) repeated-measures ANOVA and found a significant interaction [$F_{(6, 90)} = 5.72$, $P = 0.002$, $\eta_p^2 = 0.28$, Greenhouse–Geisser corrected]. Interaction contrasts then revealed that SPL, perhaps not surprisingly, responded significantly more to Scene Motion than to both Face and Object Motion, compared to both PPA and MPA (all *P*s < 0.005), consistent with the distinct role of SPL, but not PPA and MPA, in visually guided navigation. Next, interaction contrasts also revealed that SPL responded significantly more to Scene Motion than to both Face and Object Motion, compared to OPA (both *P*s < 0.05). Interestingly, this finding of differential selectivity between SPL and OPA suggests that each of these regions may play a distinct role even within visually guided navigation. Finally, replicating

a previous finding (Kamps et al. 2016b), OPA also responded, compared to PPA and MPA significantly more to Scene Motion than to Face or Object Motion (all *P*s < 0.04), consistent with the distinct role of OPA, but not PPA and MPA, in visually guided navigation.

But given that we did not precisely match the amount of motion information across the scene, face, and object stimuli, might SPL be responding more to Scene Motion than to Face and Object Motion because Dynamic Scenes simply have more motion information than Dynamic Faces or Dynamic Objects, rather than responding specifically to Scene Motion? To address this possibility, we compared the responses to Scene Motion, Face Motion, and Object Motion in SPL with those in the middle temporal (MT) area—a domain-general motion-selective region. A 2 (ROI: SPL, MT) × 3 (Condition: Scene Motion, Face Motion, Object Motion) repeated-measures ANOVA revealed a significant interaction [$F_{(2, 30)} = 35.76$, $P < 0.001$, $\eta_p^2 = 0.70$], with SPL responding significantly more to Scene Motion than to both Face Motion and Object Motion, compared to MT (interaction contrasts, both *P*s < 0.001) (Fig. 5E). In fact, MT shows a qualitatively different pattern of response than SPL, responding more to Object Motion, followed by Scene Motion followed by Face Motion. Moreover, and perhaps not surprisingly, this response profile of MT mirrors the amount of motion energy in each type of stimuli (i.e. Object Motion > Scene Motion > Face Motion), estimated by the Farnebäck algorithm (Farnebäck 2003; see Supplementary Fig. 6). Thus, taken together these results suggest that SPL's preferential response to first-person perspective motion information through scenes cannot be explained by general motion sensitivity. Future research, however, is needed to more definitively rule out this possibility by independently varying the amount and category of motion information.

While the above findings demonstrate that SPL responds to first-person perspective motion information through scenes and does not respond to just any motion information, might it still be the case that the response of SPL simply reflects visual information inherited from low-level visual regions? To address this possibility, we compared the responses to Scene Motion, Face Motion, and Object Motion in SPL with those in primary visual cortex (V1). A 2 (ROI: SPL, V1) × 3 (Condition: Scene Motion, Face Motion, Object Motion) repeated-measures ANOVA revealed a significant interaction [$F_{(2, 30)} = 7.87$, $P = 0.002$, $\eta_p^2 = 0.34$], with SPL responding significantly more to Scene Motion than to both Face (interaction contrast, $P = 0.004$) and Object Motion, compared to V1 (interaction contrast, $P < 0.001$) (Fig. 5F). Note that V1 also exhibits a qualitatively different pattern of response than SPL, responding most to both Scene and Object Motion, and least to Face Motion. Not surprisingly, this pattern of response in V1 reflects the amount of high spatial frequency information in each type of the stimuli, which was estimated using the Natural Image Statistical Toolbox (Bainbridge and Oliva 2015) (see Supplementary Fig. 7). Taken together, these findings suggest that SPL's preferential response to first-person perspective motion through scenes is not simply inherited from V1 but instead reflects distinct, specialized processing. Again, however, future research is needed to more definitively rule out this possibility by varying the amount of low-level visual information in the stimulus.

Finally, although we removed any voxels that overlapped between the SPL and nearby V6 (see Materials and Methods), we nevertheless asked whether the SPL might merely be an extension of its neighboring V6 and respond to multiple kinds of ego-motion, not simply first-person perspective motion in scenes. To rule out this possibility, we compared the responses to Scene Motion, Face Motion, and Object Motion in SPL with those in V6. A 2 (ROI: SPL, V6) × 3 (Condition: Scene Motion, Face Motion, Object Motion) repeated-measures ANOVA revealed a significant interaction [$F_{(2, 30)} = 6.555$, $P = 0.004$, $\eta_p^2 = 0.304$], with SPL responding significantly more to Scene Motion than to both Face Motion and Object Motion, compared to V6 (interaction contrasts, $P = 0.001$, $P = 0.017$, respectively) (Fig. 5G), revealing that the SPL is not a mere extension of V6. However, given the quantitative, not qualitative, difference between SPL and V6, it could still be the case that SPL and V6 are not entirely functionally distinct regions, with V6 either sharing scene selectivity (see Fig. 4A) or scene motion selectivity (see Fig. 5A) with SPL. To address this possibility further, we investigated both scene selectivity and scene motion selectivity in V6. For scene selectivity, a 3-level (Condition: Dynamic Faces, Dynamic Objects, Dynamic Scenes) repeated-measures ANOVA revealed a significant main effect of condition [$F_{(2, 30)} = 11.31$, $P < 0.001$, $\eta_p^2 = 0.43$; see Supplementary Fig. 8], with V6 responding significantly more to Dynamic Scenes than to both Dynamic Faces (main effect contrast, $P < 0.001$) and Dynamic Objects (main effect contrast, $P = 0.010$), most likely reflecting its general ego-motion sensitivity (Sulpizio et al. 2023). By contrast, no such scene selectivity was found for the static stimuli. Indeed, a 3-level (Condition: Static Faces, Static Objects, Static Scenes) repeated-measures ANOVA did not reveal a significant main effect of condition [$F_{(2, 30)} = 1.63$, $P = 0.212$, $\eta_p^2 = 0.10$], which is contradictory to the dynamic stimuli findings. Recall that all the scene-selective regions, including SPL, showed scene selectivity for both the dynamic and static stimuli, unlike V6. For scene motion selectivity, a 3-level (Condition: Scene Motion, Face Motion, Object Motion) repeated-measures ANOVA revealed no significant effect of condition, [$F_{(2, 30)} = 2.65$, $P = 0.087$, $\eta_p^2 = 0.150$], unlike SPL (and confirmed by the above direct comparison). Thus, taken together, these findings demonstrate that V6, unlike SPL, does not show either scene selectivity or scene motion selectivity, suggesting that this region is neither part of the scene processing network nor the visually guided navigation system (Dilks et al. 2022).

Experiment 2

SPL represents sense information in scenes

If the SPL supports visually guided navigation, then it should represent sense (left/right) information in scenes—another kind of information necessary for visually guided navigation—but not objects. To test this hypothesis, using fMRI adaptation, we first compared the responses in SPL to three scene conditions: Different, Same, and Mirror (Fig. 3; see Materials and Methods for more details). We predicted that SPL would discriminate an image of a scene from its mirror-reversed image (i.e. represent sense information in scenes), responding more to the Different than Same condition (i.e. show adaptation) and also responding more to the Mirror than Same condition. Consistent with our prediction, a 3-level (Condition: Different, Mirror, Same) repeated-measures ANOVA revealed a significant effect of condition [$F_{(2, 30)} = 5.98$, $P = 0.023$, Greenhouse–Geisser corrected; $\eta_p^2 = 0.285$], with a significantly greater response to the Different than Same condition (main effect contrast, $P = 0.003$)—demonstrating the expected fMRI adaptation effect in SPL (Fig. 6A), and a significantly greater response to the Mirror than Same condition (main effect contrast, $P = 0.025$). There was no significant difference between the responses for the Different and Mirror conditions (main effect contrast, $P = 0.376$). Taken together, these findings indicate that SPL indeed treats mirror images of scenes as two different images,

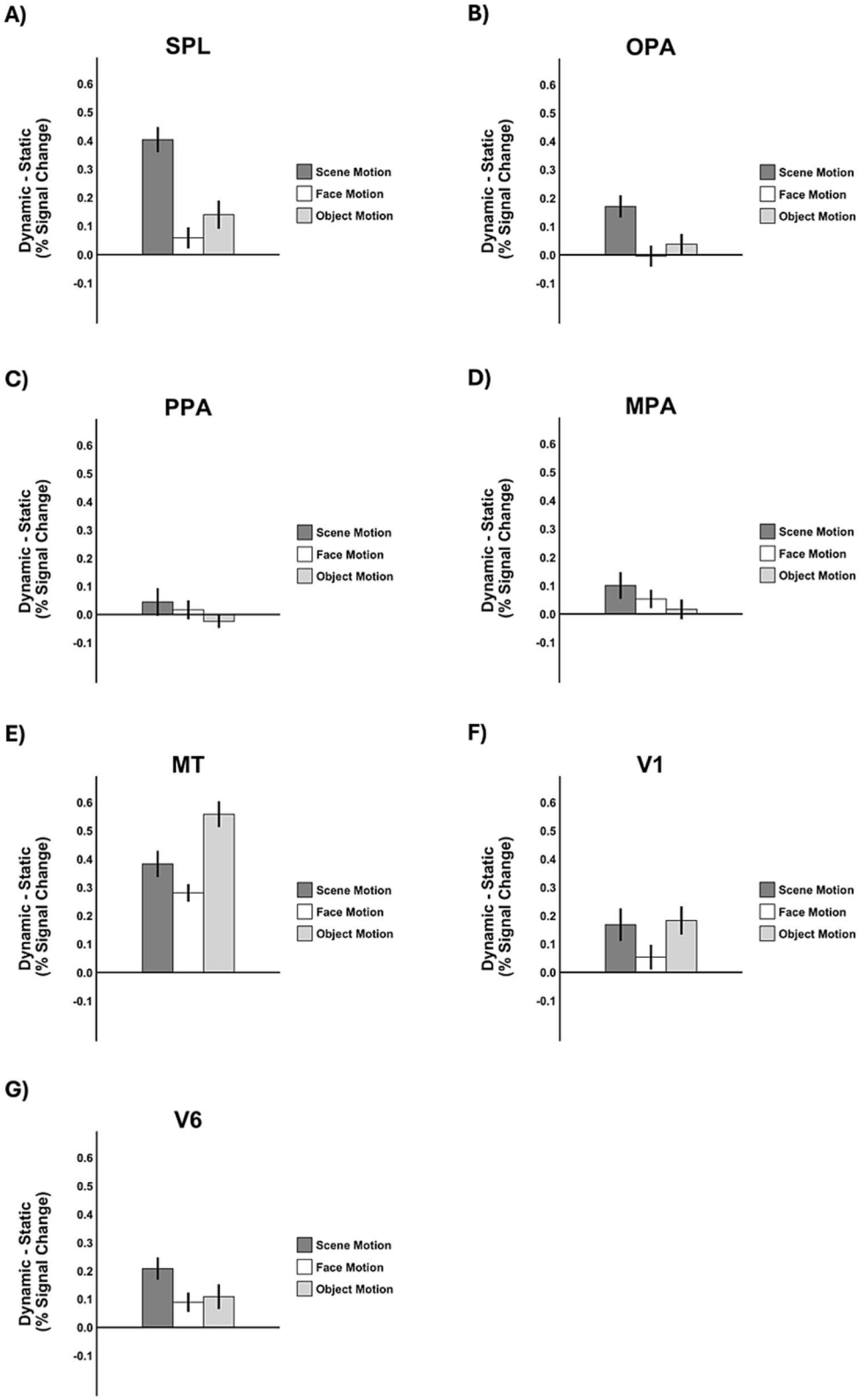

Fig. 5. Percent signal change for Dynamic Scenes minus Static Scenes ("Scene Motion"), Dynamic Faces minus Static Faces ("Face Motion"), and Dynamic Objects minus Static Objects ("Object Motion") in SPL A), OPA B), PPA C), MPA D), MT E), V1 F), and V6 (G). SPL responded significantly more to Scene Motion than to Face and Object Motion. Moreover, replicating the findings in previous research, OPA also responded significantly more to Scene Motion than to Face and Object motion. Unlike both SPL and OPA, neither PPA, MPA, MT, V1 nor V6 showed such Scene Motion selectivity. Created by the authors.

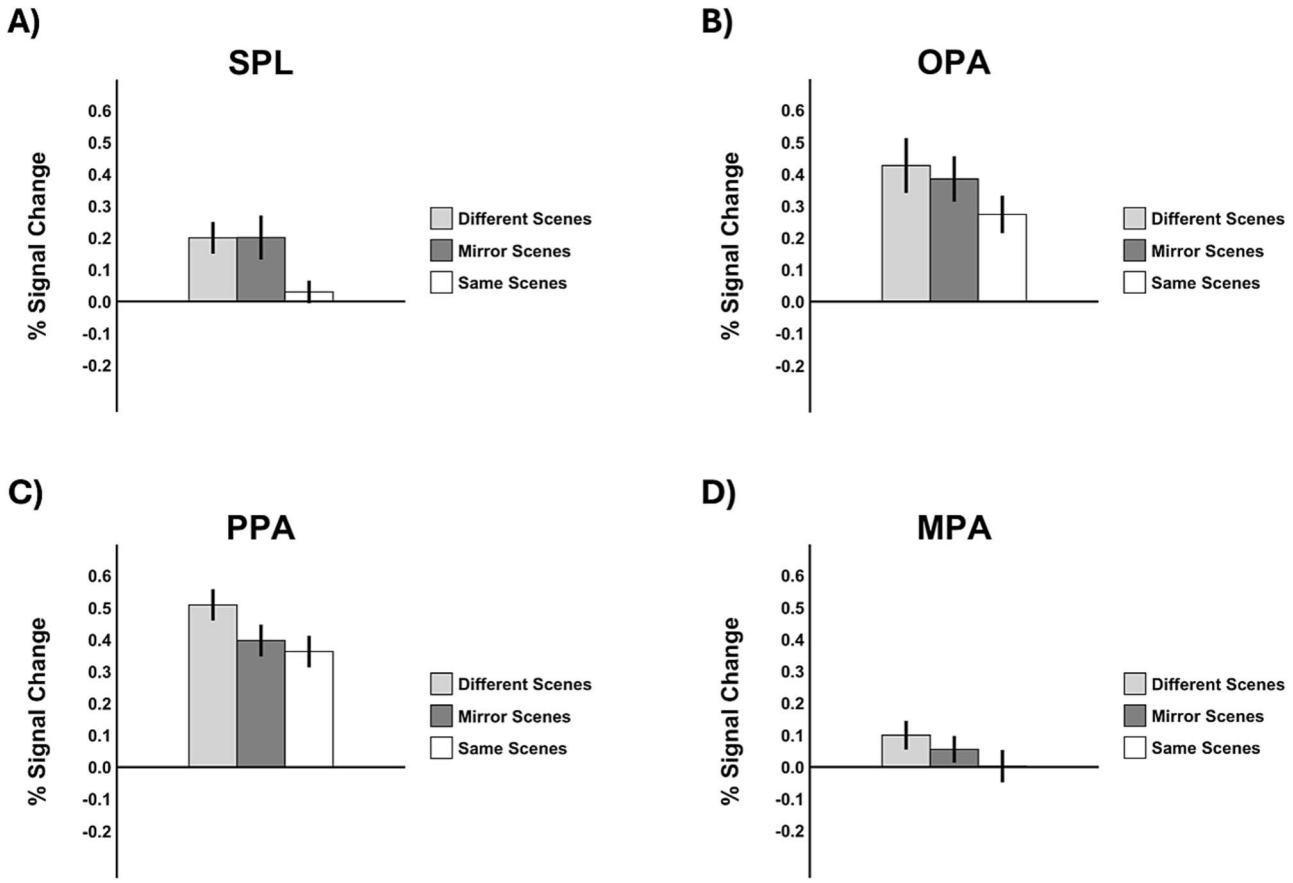

Fig. 6. Percent signal change for Different Scenes, Mirror Scenes, and Same Scenes in SPL A), OPA B), PPA C), and MPA D). SPL responded significantly more to both the Different Scenes and Mirror Scenes conditions than to the Same Scenes condition, but similarly between the Different Scenes and Mirror Scenes conditions—revealing that SPL has mirror-image sensitivity for scene images and therefore represents sense information in scenes. Error bars reflect the standard error of the mean. Created by the authors.

thus representing sense information in scenes and providing another piece of evidence supporting our hypothesis that SPL may be involved in visually guided navigation.

Replicating previous findings (Dilks et al. 2011), such mirror image sensitivity in scenes was also found for OPA, but not PPA. For OPA, a 3-level (Condition: Different, Mirror, Same) repeated-measures ANOVA revealed a significant main effect of condition [$F_{(2, 30)} = 5.67$, $P = 0.008$, $\eta_p^2 = 0.27$], with a significantly greater response to the Different than Same condition (main effect contrasts, $P = 0.003$), a significantly greater response to the Mirror than Same condition (main effect contrast, $P = 0.025$), but no significant difference between the responses for the Different and Mirror conditions (main effect contrast, $P = 0.376$) (Fig. 6B). By contrast, PPA showed tolerance to mirror images of scenes. A 3-level (Condition: Different, Mirror, Same) repeated-measures ANOVA revealed a significant main effect of condition [$F_{(2, 30)} = 15.55$, $P < 0.001$, $\eta_p^2 = 0.51$], with a significantly greater response to the Different condition than Same condition (main effect contrast, $P < 0.001$) but no significant difference between the Mirror and Same condition (main effect contrast, $P = 0.216$) (Fig. 6C). For MPA, while mirror-image sensitivity was previously reported (Dilks et al. 2011), our findings were inconclusive. A 3-level (Condition: Mirror, Different, Same) repeated-measures ANOVA revealed a marginal effect of condition [$F_{(2, 30)} = 5.67$, $P = 0.057$, $\eta_p^2 = 0.17$], with a significantly greater response to the Different than Same condition (main effect contrast, $P = 0.018$)—again demonstrating the expected fMRI adaptation effect—but no significant difference between the responses for Mirror condition compared to either the Different or Same conditions (main effect contrasts, $P = 0.263$ and 0.182, respectively) (Fig. 6D). However, are SPL and OPA (the two scene-selective regions sensitive to sense information in scenes) truly sensitive to sense information in scenes—as expected from a scene-selective region involved in visually guided navigation—or might they simply represent sense information more generally (e.g. even in their nonpreferred category, objects)? Given that our stimuli set included objects as well as scenes, we were then able to investigate how SPL and OPA (as well as PPA and MPA) respond to mirror reversals of objects. None of the scene-selective regions revealed mirror-image sensitivity for objects, with no significant difference between the responses for the Mirror condition compared to the Same condition (main effect contrasts, all Ps > 0.07).

Finally, we compared the mirror-image sensitivity for scenes in SPL to the other scene-selective regions, investigating whether (i) the SPL is indeed responding preferentially to sense information in scenes—consistent with its hypothesized role in visually guided navigation—relative to PPA, known not to be involved in visually guided navigation and (ii) the extent of mirror-image sensitivity for scenes is similar between SPL and the two other scene-selective regions involved in navigation (OPA and MPA), albeit different kinds, and hence requiring sense information. To address these two questions, we conducted a 4 (ROI: SPL, OPA, PPA, MPA) × 3 (Condition: Different, Mirror, Same) repeated-measures ANOVA and found a significant interaction [$F_{(6, 90)} = 3.68$, $P = 0.003$, $\eta_p^2 = 0.20$]. Interaction contrasts then revealed that the SPL, perhaps not surprisingly, responded significantly more to Mirror Scenes than to Same Scenes (i.e. represents left/right information), compared to PPA ($P < 0.001$), consistent with the

distinct role of SPL, but not PPA, in visually guided navigation. By contrast, interaction contrasts revealed that SPL did not respond significantly different to Mirror and Same Scenes, compared to OPA ($P > 0.09$), consistent with their hypothesized roles in visually guided navigation. Finally, interaction contrasts revealed that SPL responded significantly more to the Mirror Scenes than to Same Scenes, compared to MPA ($P = 0.001$), perhaps reflecting their involvement in two different kinds of navigation (i.e. visually guided navigation and map-based navigation, respectively).

Experiment 3

SPL is functionally connected to OPA, not PPA or MPA

If SPL and OPA are both involved in visually guided navigation, then we should see stronger functional connectivity between SPL and OPA (since OPA is involved in visually guided navigation) than between SPL and either PPA or MPA (since neither of these regions are hypothesized to be involved in visually guided navigation). To test this hypothesis, we calculated the correlation between the activity time series for SPL and each of the other scene-selective ROIs while partialing out the activation from the remaining ROIs, which is critical since it is known that OPA, PPA, and MPA are functionally connected to one another (Baldassano et al. 2016; Kamps et al. 2020a). To maximize the statistical power of our analysis, within-hemisphere (e.g. left SPL–left OPA) and between-hemisphere (e.g. left SPL–right OPA) correlations were averaged together for each pair of regions in each subject. Note that between-hemisphere correlations did not include homotopic connections (e.g. left SPL–right SPL).

Consistent with our hypothesis, we found that the SPL is preferentially connected to OPA, compared to PPA or MPA (Fig. 7). Specifically, a 3-level (Functional Connectivity: SPL-OPA, SPL-PPA, SPL-MPA) repeated-measures ANOVA revealed a significant main effect [$F_{(2, 34)} = 18.53$, $P < 0.001$, $\eta_p^2 = 0.52$], with the functional connectivity between SPL and OPA significantly greater than the functional connectivity between SPL and both PPA and MPA, respectively (main effect contrasts, both $Ps < 0.001$). We did not find a significant difference in functional connectivity between SPL and either PPA or MPA (main effect contrast, $P = 0.79$). Note that the preferential connectivity between SPL and OPA holds even when we separately investigate the between-hemisphere functional connectivity rather than the average of within- and between-hemisphere functional connectivity, ruling out the possibility of local distance confounds. Again, a 3-level (Functional connectivity: SPL-OPA, SPL-PPA, SPL-MPA) repeated-measures ANOVA revealed a significant main effect [$F_{(2, 34)} = 14.35$, $P < 0.001$, $\eta_p^2 = 0.45$], with the functional connectivity between SPL and OPA significantly greater than the functional connectivity between SPL and both PPA and MPA, respectively (main effect contrasts, both $Ps < 0.001$). We did not find a significant difference in functional connectivity between SPL and either PPA or MPA (main effect contrast, $P = 0.98$). Taken together then, these results suggest that SPL and OPA share information more with each other than with the two other scene-selective regions and are thus may be part of the same visually guided navigation system.

Experiment 4

SPL responds selectively during a visually guided navigation task

Thus far, we have shown that SPL processes at least two kinds of information necessary for visually guided navigation and is preferentially connected to OPA—all supporting our hypothesis of a second scene-selective region that, like OPA, may be involved in visually guided navigation. To provide further evidence for our hypothesis, we used data from a previously published paper (Persichetti and Dilks 2018) that demonstrated the precise roles of PPA and OPA in scene categorization and visually guided navigation, respectively. In this preexisting dataset, we predicted that SPL would respond significantly more while participants performed a visually guided navigation task compared to both a scene categorization task and a baseline task (see Materials and Methods for more details). By contrast, PPA would show the complete opposite, as previously described (Persichetti and Dilks 2018).

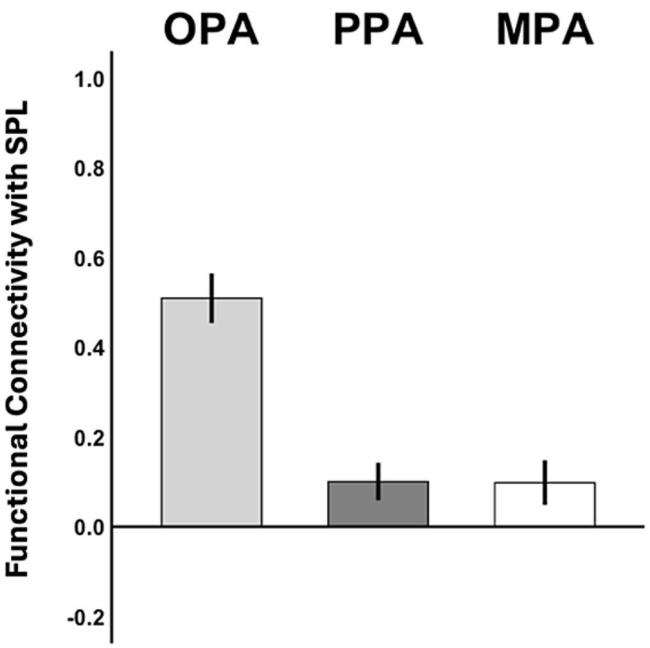

Fig. 7. Functional connectivity between SPL and each of the three scene-selective regions (i.e. OPA, PPA, MPA). Error bars indicate the standard error of the mean. Created by the authors.

Confirming our prediction, for SPL, a 3-level (Task: Navigation, Categorization, Baseline) repeated-measures ANOVA revealed a significant main effect of task [$F_{(2, 20)} = 14.90$, $P < 0.001$, $\eta_p^2 = 0.60$], with the SPL responding significantly more during the visually guided navigation task than during both the scene categorization and baseline tasks (main effect contrasts, both $Ps < 0.001$) (Fig. 8). For OPA, a 3-level (Task: Navigation, Categorization, Baseline) repeated-measures ANOVA also revealed a significant main effect of task [$F_{(2, 20)} = 5.57$, $P = 0.030$, $\eta_p^2 = 0.36$], with a significantly greater response during the visually guided navigation task during both the scene categorization and baseline tasks (main effect contrasts, $P = 0.026$, $P = 0.004$, respectively)—replicating the previous finding that OPA is involved in visually guided navigation (Persichetti and Dilks 2018). Taken together, these results provide strong evidence that the SPL may specifically be involved in visually guided navigation.

By contrast, PPA showed the exact opposite pattern than SPL and OPA. A 3-level (Task: Navigation, Categorization, Baseline) repeated-measures ANOVA revealed a significant main effect of task [$F_{(2, 20)} = 8.97$, $P = 0.002$, $\eta_p^2 = 0.47$], with a significantly greater response during the scene categorization task than during both visually guided navigation and baseline tasks (main effect contrasts, $P < 0.001$, $P = 0.019$, respectively)—replicating a previous report that PPA is involved in scene categorization (Persichetti and Dilks 2018). For completeness, for MPA, a 3-level (Task: Navigation, Categorization, Baseline) repeated-measures ANOVA did not

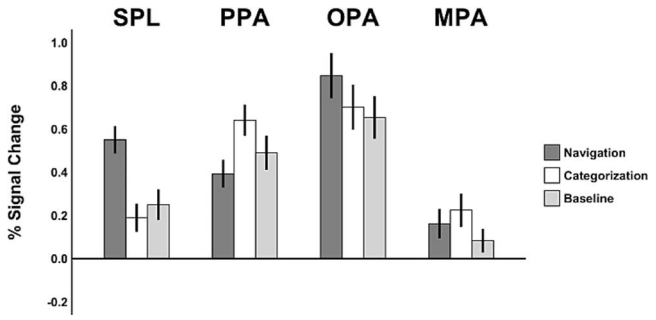

Fig. 8. Percent signal change for the visually guided navigation, scene categorization, and baseline tasks in SPL, PPA, OPA, and MPA. SPL and OPA responded significantly more during the visually guided navigation task than during both the scene categorization and baseline tasks, whereas PPA responded significantly more during the scene categorization task than during the visually guided navigation and baseline tasks. Responses in MPA did not differ between the three tasks. Created by the authors.

reveal a significant main effect of task ($F_{(2, 16)} = 2.60$, $P = 0.106$, $\eta_p^2 = 0.25$), thus finding no significant differences across the three tasks in MPA, which is expected given that MPA is hypothesized to not be involved in either visually guided navigation or scene categorization.

Discussion

In this study, we explored whether a newly uncovered scene-selective region in the SPL is involved in visually guided navigation. First, we confirmed its scene selectivity across two different sets of stimuli (i.e. Dynamic and Static Scenes, Faces, and Objects). In both sets of stimuli, we found that the SPL responds significantly more to Scenes than Faces and Objects, revealing that SPL is indeed a fourth scene-selective cortical region, joining PPA, MPA, and OPA. Second, we then tested the hypothesis that SPL is involved in visually guided navigation by asking if it represents two kinds of information critical for visually guided navigation. Indeed, we found that SPL represents (i) first-person perspective motion information through scenes but not faces or objects in motion and (ii) sense information in scenes but not objects. Such selectivity to first-person motion and sense information could not be explained by scene selectivity alone (not all scene-selective regions showed such selectivity), domain-general motion sensitivity, or low-level visual information.

To further understand the functional role of SPL, we next employed resting-state fMRI to assess its connectivity with the three other known scene-selective cortical regions. Our analysis showed that SPL is preferentially connected to OPA, compared to both PPA and MPA. Based on a growing body of evidence suggesting that OPA is involved in visually guided navigation (Dilks et al. 2022), our functional connectivity data suggest that the SPL, like OPA, may be part of the same system.

Finally, we leveraged a preexisting fMRI dataset (Persichetti and Dilks 2018) to evaluate the response of SPL while participants performed a visually guided navigation task, a scene categorization task, and a baseline task. We hypothesized that if the SPL is involved in visually guided navigation, then it should be modulated by task, responding more during a visually guided navigation task than any other. Indeed, our results indicated that the SPL responded significantly more during the visually guided navigation task than during both the scene categorization and baseline tasks, offering further evidence for our hypothesis that the SPL may be involved in visually guided navigation in a separate dataset.

Recent studies provide complementary insights into the role of SPL in visually guided navigation. For example, Sulpizio et al. (2023) suggested that areas within the dorsomedial parietal cortex, which include the SPL, are involved in visually guided action. Specifically, they proposed that subregions within the dorsomedial parietal cortex—V6, V6A, and PEc—each support distinct visually guided actions: V6 processes optic flow and motion information related to navigation. V6A (including both V6Av and V6Ad) is involved in reaching and grasping and perhaps navigation, since it responds preferably to peripheral visual field stimulation (Pitzalis et al. 2013), which has been shown to be critical for navigation (Stoffregen et al. 1987; Warren et al. 2001; Franchak and Adolph 2010; Pitzalis et al. 2013; Cao and Händel 2019). Finally, PEc is associated with leg movement and locomotion. Our findings extend this framework by demonstrating that the SPL—a scene-selective region within the dorsomedial parietal cortex—is specifically involved in visually guided navigation. Interestingly, the MNI coordinates of SPL in the current study (see Supplementary Table 1) show overlap with those of V6A, raising the intriguing possibility then that these two regions may be the same region.

Similarly, as briefly discussed in the Introduction, Kennedy et al. (2024) recently identified a scene-selective region in the posterior intraparietal gyrus (PIGS) that was sensitive to ego-motion in scenes. However, while Kennedy and colleagues demonstrated that there is a scene-selective region within the SPL, their findings raise several questions. First, the lack of controls for nonscene motion (e.g. face in motion and objects in motion) leaves it unclear whether PIGS is specifically involved in processing egocentric motion through scenes or processing motion information across domains. Second, their use of "unnatural" motion stimuli (i.e. moving through a scene at 44.7 miles per hour, which is incompatible with "ego-motion," see Gibson 1950) limits the ecological validity of their findings. And third, Kennedy and colleagues found that V6, like PIGS, was also both scene selective and sensitive to ego-motion in scenes, making it unclear whether PIGS and V6 actually support distinct functions. Our study addresses these gaps by including appropriate controls, using ecologically valid motion stimuli, and demonstrating that SPL is both scene selective and selective to scene motion, unlike V6. Finally, the different anatomical positions of each of these regions—that is, SPL, adjacent and anterior to the parieto-occipital sulcus; PIGS in the posterior intraparietal gyrus; and V6 in the posterior bank of the parieto-occipital sulcus—raise the interesting possibility that each of these regions supports distinct functions. Thus, future research directly testing the function of each of these regions, within subject and on the same tasks, is needed to investigate this possibility.

In conclusion, across four fMRI experiments, we have shown that a region in SPL (i) is scene selective, demonstrated across two datasets, (ii) responds to first-person perspective motion information through scenes, (iii) encodes sense information in scenes, (iv) is preferentially connected to OPA over PPA and MPA, and (v) is selectively engaged during a visually guided navigation task. Together, these results provide converging evidence that SPL is a scene-selective region in human cortex that is involved in visually guided navigation. Future research then needs to elucidate the specific roles that SPL and OPA play within the visually guided navigation system.

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. We would also like to thank Rebecca Rennert, Stuart Gakio, Lily Su, Quyen Cao, Boris Botzanowski, and Chris Jones for their insight and/or their assistance in data collection.

Author contributions

Hee Kyung Yoon (Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Visualization, Writing—original draft, Writing—review & editing), Yaelan Jung (Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Visualization, Writing—original draft, Writing—review & editing), Andrew S. Persichetti (Data curation, Formal analysis, Writing—review & editing), and Daniel D. Dilks (Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Writing— review & editing).

Supplementary material

Supplementary material is available at *Cerebral Cortex* online.

Funding

This work was supported by grants from the National Eye Institute (R01 EY29724 to D.D.D.)

Conflict of interest statement. None declared.

Data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Ethics approval disclosure statement

This study was approved by the Institutional Review Board at Emory University under approval number IRB00069592. All participants provided informed consent prior to participation.

References

Bainbridge WA, Oliva A. 2015. A toolbox and sample object perception data for equalization of natural images. *Data in Brief*. 5: 846–851. https://doi.org/10.1016/j.dib.2015.10.030.

Baldassano C, Beck DM, Fei-Fei L. 2013. Differential connectivity within the Parahippocampal place area. *NeuroImage*. 75:228–237. https://doi.org/10.1016/j.neuroimage.2013.02.073.

Baldassano C, Esteva A, Fei-Fei L, Beck DM. 2016. Two distinct scene-processing networks connecting vision and memory. *eNeuro*. 3:ENEURO.0178–ENEU16.2016. https://doi.org/10.1523/ENEURO.0178-16.2016.

Bonner MF, Epstein RA. 2017. Coding of navigational affordances in the human visual system. *Proc Natl Acad Sci USA*. 114:4793–4798. https://doi.org/10.1073/pnas.1618228114.

Cao L, Händel B. 2019. Walking enhances peripheral visual processing in humans. *PLoS Biol*. 17:e3000511. https://doi.org/10.1371/journal.pbio.3000511.

Cox RW. 1996. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res*. 29:162–173. https://doi.org/10.1006/cbmr.1996.0014.

Cox RW, Jesmanowicz A. 1999. Real-time 3D image registration for functional MRI. *Magn Reson Med*. 42:1014–1018. https://doi.org/10.1002/(sici)1522-2594(199912)42:6<1014::aid-mrm4>3.0.co;2-f.

Dilks DD, Julian JB, Kubilius J, Spelke ES, Kanwisher N. 2011. Mirror-image sensitivity and invariance in object and scene processing pathways. *J Neurosci*. 31:11305–11312. https://doi.org/10.1523/JNEUROSCI.1935-11.2011.

Dilks DD, Julian JB, Paunov AM, Kanwisher N. 2013. The occipital place area is causally and selectively involved in scene perception. *J Neurosci*. 33:1331–1336. https://doi.org/10.1523/JNEUROSCI.4081-12.2013.

Dilks DD, Kamps FS, Persichetti AS. 2022. Three cortical scene systems and their development. *Trends Cogn Sci*. 26:117–127. https://doi.org/10.1016/j.tics.2021.11.002.

Dillon MR, Persichetti AS, Spelke ES, Dilks DD. 2018. Places in the brain: bridging layout and object geometry in scene-selective cortex. *Cereb Cortex*. 28:2365–2374. https://doi.org/10.1093/cercor/bhx139.

Epstein R, Kanwisher N. 1998. A cortical representation of the local visual environment. *Nature*. 392:598–601. https://doi.org/10.1038/33402.

Farnebäck G. 2003. Two-frame motion estimation based on polynomial expansion. In: *Image analysis* Bigun J, Gustavsson T (eds). Springer, Berlin, Heidelberg, pp 363–370.

Franchak JM, Adolph KE. 2010. Visually guided navigation: head-mounted eye-tracking of natural locomotion in children and adults. *Vis Res*. 50:2766–2774. https://doi.org/10.1016/j.visres.2010.09.024.

Gibson JJ. 1950. *The perception of the visual world*. Houghton Mifflin, Oxford, England, (The perception of the visual world).

Glasser MF et al. 2016. A multi-modal parcellation of human cerebral cortex. *Nature*. 536:171–178. https://doi.org/10.1038/nature18933.

Jenkinson M, Bannister P, Brady M, Smith S. 2002. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*. 17:825–841. https://doi.org/10.1006/nimg.2002.1132.

Jones CM, Byland J, Dilks DD. 2023. The occipital place area represents visual information about walking, not crawling. *Cereb Cortex*. 33:7500–7505. https://doi.org/10.1093/cercor/bhad055.

Jung Y, Larsen B, Walther DB. 2018. Modality-independent coding of scene categories in prefrontal cortex. *J Neurosci*. 38:5969–5981. https://doi.org/10.1523/JNEUROSCI.0272-18.2018.

Jung Y, Hsu D, Dilks DD. 2024. "Walking selectivity" in the occipital place area in 8-year-olds, not 5-year-olds. *Cereb Cortex*. 34:bhae101. https://doi.org/10.1093/cercor/bhae101.

Kamps FS, Julian JB, Kubilius J, Kanwisher N, Dilks DD. 2016a. The occipital place area represents the local elements of scenes. *NeuroImage*. 132:417–424. https://doi.org/10.1016/j.neuroimage.2016.02.062.

Kamps FS, Lall V, Dilks DD. 2016b. The occipital place area represents first-person perspective motion information through scenes. *Cortex*. 83:17–26. https://doi.org/10.1016/j.cortex.2016.06.022.

Kamps FS, Hendrix CL, Brennan PA, Dilks DD. 2020a. Connectivity at the origins of domain specificity in the cortical face and place networks. *Proc Natl Acad Sci USA*. 117:6163–6169. https://doi.org/10.1073/pnas.1911359117.

Kamps FS, Pincus JE, Radwan SF, Wahab S, Dilks DD. 2020b. Late development of navigationally relevant motion processing in

the occipital place area. *Curr Biol*. 30:544–550.e3. https://doi.org/10.1016/j.cub.2019.12.008.

Kennedy B, Malladi SN, Tootell RB, Nasr S. 2024. A previously undescribed scene-selective site is the key to encoding ego-motion in naturalistic environments. *eLife*. 13:e91601. https://doi.org/10.7554/eLife.91601.

Maguire EA. 2001. The retrosplenial contribution to human navigation: a review of lesion and neuroimaging findings. *Scand J Psychol*. 42:225–238. https://doi.org/10.1111/1467-9450.00233.

Persichetti AS, Dilks DD. 2016. Perceived egocentric distance sensitivity and invariance across scene-selective cortex. *Cortex*. 77: 155–163. https://doi.org/10.1016/j.cortex.2016.02.006.

Persichetti AS, Dilks DD. 2018. Dissociable neural Systems for Recognizing Places and Navigating through them. *J Neurosci*. 38: 10295–10304. https://doi.org/10.1523/JNEUROSCI.1200-18.2018.

Pitcher D, Dilks DD, Saxe RR, Triantafyllou C, Kanwisher N. 2011. Differential selectivity for dynamic versus static information in face-selective cortical regions. *NeuroImage*. 56:2356–2363. https://doi.org/10.1016/j.neuroimage.2011.03.067.

Pitzalis S et al. 2013. The human homologue of macaque area V6A. *NeuroImage*. 82:517–530. https://doi.org/10.1016/j.neuroimage.2013.06.026.

Silson EH, Steel AD, Baker CI. 2016. Scene-selectivity and Retinotopy in medial parietal cortex. *Front Hum Neurosci*. 10:412. https://doi.org/10.3389/fnhum.2016.00412.

Smith SM. 2002. Fast robust automated brain extraction. *Hum Brain Mapp*. 17:143–155. https://doi.org/10.1002/hbm.10062.

Smith SM et al. 2004. Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*. 23: S208–S219. https://doi.org/10.1016/j.neuroimage.2004.07.051.

Steel A, Billings MM, Silson EH, Robertson CE. 2021. A network linking scene perception and spatial memory systems in posterior cerebral cortex. *Nat Commun*. 12:2632. https://doi.org/10.1038/s41467-021-22848-z.

Stoffregen TA, Schmuckler MA, Gibson EJ. 1987. Use of central and peripheral optical flow in stance and locomotion in young walkers. *Perception*. 16:113–119. https://doi.org/10.1068/p160113.

Sulpizio V, Fattori P, Pitzalis S, Galletti C. 2023. Functional organization of the caudal part of the human superior parietal lobule. *Neurosci Biobehav Rev*. 153:105357. https://doi.org/10.1016/j.neubiorev.2023.105357.

Suzuki S, Kamps FS, Dilks DD, Treadway MT. 2021. Two scene navigation systems dissociated by deliberate versus automatic processing. *Cortex*. 140:199–209. https://doi.org/10.1016/j.cortex.2021.03.027.

Tootell RB et al. 1995. Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *J Neurosci*. 15:3215–3230. https://doi.org/10.1523/JNEUROSCI.15-04-03215.1995.

van Assche M, Kebets V, Vuilleumier P, Assal F. 2016. Functional dissociations within posterior parietal cortex during scene integration and viewpoint changes. *Cereb Cortex*. 26:bhu215–bhu598. https://doi.org/10.1093/cercor/bhu215.

Walther DB, Caddigan E, Fei-Fei L, Beck DM. 2009. Natural scene categories revealed in distributed patterns of activity in the human brain. *J Neurosci*. 29:10573–10581. https://doi.org/10.1523/JNEUROSCI.0559-09.2009.

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*. 108:9661–9666. https://doi.org/10.1073/pnas.1015666108.

Wang L, Mruczek REB, Arcaro MJ, Kastner S. 2015. Probabilistic maps of visual topography in human cortex. *Cereb Cortex*. 25:3911–3931. https://doi.org/10.1093/cercor/bhu277.

Warren WH, Kay BA, Zosh WD, Duchon AP, Sahuc S. 2001. Optic flow is used to control human walking. *Nat Neurosci*. 4:213–216. https://doi.org/10.1038/84054.

Xu J et al. 2013. Evaluation of slice accelerations using multiband echo planar imaging at 3 T. *NeuroImage*. 83:991–1001. https://doi.org/10.1016/j.neuroimage.2013.07.055.

Cerebral Cortex, 2025, **35**, bhaf082
https://doi.org/10.1093/cercor/bhaf082
Original Article