Neuroscience

A previously undescribed scene-selective site is the key to encoding ego-motion in naturalistic environments

Bryan Kennedy, Sarala N. Malladi, Roger B. H. Tootell, Shahin Nasr

Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, MA, United States • Department of Radiology, Harvard Medical School, Boston, MA, United States

Reviewed Preprint

Revised by authors after peer review.

About eLife's process

Reviewed preprint version 2
March 7, 2024 (this version)

Reviewed preprint version 1
January 16, 2024

Posted to preprint server
September 25, 2023

Sent for peer review
August 28, 2023

Abstract

Current models of scene processing in the human brain include three scene-selective areas: the Parahippocampal Place Area (or the temporal place areas; PPA/TPA), the restrosplenial cortex (or the medial place area; RSC/MPA) and the transverse occipital sulcus (or the occipital place area; TOS/OPA). Here, we challenged this model by showing that at least one other scene-selective site can also be detected within the human posterior intraparietal gyrus. Despite the smaller size of this site compared to the other scene-selective areas, the posterior intraparietal gyrus scene-selective (PIGS) site was detected consistently in a large pool of subjects (n=59; 33 females). The reproducibility of this finding was tested based on multiple criteria, including comparing the results across sessions, utilizing different scanners (3T and 7T) and stimulus sets. Furthermore, we found that this site (but not the other three scene-selective areas) is significantly sensitive to ego-motion in scenes, thus distinguishing the role of PIGS in scene perception relative to other scene-selective areas. These results highlight the importance of including finer scale scene-selective sites in models of scene processing – a crucial step toward a more comprehensive understanding of how scenes are encoded under dynamic conditions.

eLife assessment

In this manuscript, the authors present a wealth of fMRI data at both 3T and 7T to identify a scene-selective region of the intraparietal gyrus ("PIGS") that appears to have some responsivity to characteristics of ego-motion. In a series of experiments, they delineate the anatomical location of PIGS and functionally differentiate it from nearby V6 and OPA. Evidence for these **important** findings is **solid**, but further investigations as to the role of this region in processing ego-motion will be needed to confirm this conclusion.

1. Introduction

In human and non-human primates (NHPs), fMRI has been used for many decades to localize the cortical regions that are preferentially involved in scene perception (Epstein and Kanwisher, 1998; Tsao et al., 2008; Rajimehr et al., 2009; Nasr et al., 2011). Early studies focused mainly on larger activity sites that were more easily reproducible across sessions and individuals, ignoring smaller sites that were not detectable in all subjects and/or were not reproducible across scan sessions, based on the techniques available at that time. This led to relatively simple models of neuronal processing solely based on larger visual areas.

These models suggested three scene-selective areas within the human visual cortex, with possible homologues in NHPs (Nasr et al., 2011; Kornblith et al., 2013; Li et al., 2022). The human cortical areas were originally named parahippocampal place area (PPA) (Epstein and Kanwisher, 1998), retrosplenial cortex (RSC) (Maguire, 2001) and transverse occipital sulcus (TOS) (Grill-Spector, 2003), based the local anatomical landmarks. However, subsequent studies noticed the discrepancy between the location of these functionally-defined areas and the anatomical landmarked, and instead named those regions temporal, medial and occipital place areas or TPA, MPA and OPA (Nasr et al., 2011; Dilks et al., 2013; Silson et al., 2016).

The idea that scene-selective areas are limited to these three regions is based largely on group-averaged activity maps, generated after applying large surface/volume-based smoothing to the data from individual subjects. In such group-averaged data, originally based on fixed-rather than random-effects, thresholds tended to be high to reduce the impact of nuisance artifacts (Nasr et al., 2011). Thus, though well founded, this approach conceivably may not have identified smaller scene-selective areas (**Figure 1A**).

However, at the single subject level, multiple smaller scene-selective sites can be detected outside these scene-selective areas, especially when drastic spatial smoothing is avoided (**Figure 1B**). This phenomenon is highlighted in a recent neuroimaging study in NHPs (Li et al., 2022) in which authors took advantage of high-resolution neuroimaging techniques using implanted head coils. Their findings suggested that scene-selective areas are likely not limited to the three expected sites, and that other, smaller, scene-selective areas may also be detected across the brain. Still, the reliability in detection of these smaller sites, their spatial consistency across large populations and their specific role in scene perception that distinguishes them from the other scene-selective areas, remain unclear.

Here, we used conventional (based on a 3T scanner) and high-resolution (based on a 7T scanner) fMRI to localize and study additional scene-selective site(s) that were detected outside PPA/TPA, RSC/MPA and TOS/OPA. We focused our efforts on the posterior portion of the intraparietal cortex mainly because multiple previous studies reported indirect evidence for scene and/or scene-related information processing within this region (Lescroart and Gallant, 2019; Pitzalis et al., 2020; Sulpizio et al., 2020; Park et al., 2022). Consistent with these studies, we found at least one additional scene-selective area within the posterior intraparietal gyrus, adjacent to the motion-selective area V6 (Pitzalis et al., 2010). This site was termed PIGS, reflecting its location (posterior intraparietal gyrus) and function (scene-selectivity). PIGS was detected consistently across individual subjects and populations and localized reliably across scan sessions. Besides its distinct location relative to two major anatomical landmarks (i.e. intraparietal sulcus (IPS) and parieto-occipital sulci (POS)) and the retinotopic visual areas (IPS0-4) that distinguishes it from other scene-selective areas (e.g. TOS/OPA and RSC/MPA), PIGS showed sensitivity to ego-motion within naturalistic visual scenes, a phenomenon not detectable in other scene-selective areas.

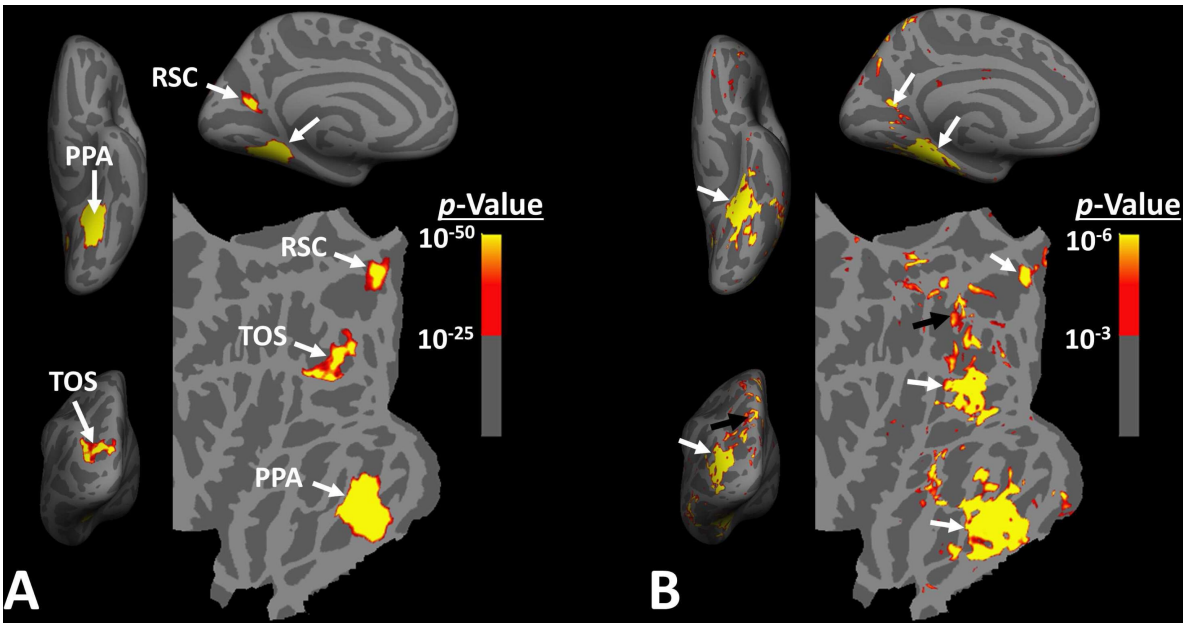

Figure 1

Distribution of scene-selective areas within the human visual cortex.

Panel **A** shows the group-averaged (n=14) response to 'scenes > faces' contrast (Experiment 1). Areas PPA/TPA, RSC/MPA and TOS/OPA are localized within the temporal, medial and posterior-lateral brain surfaces, respectively. To show consistency with our previous reports (Nasr et al., 2011), data from individual subjects was largely smoothed (FWHM=5mm) and the group-averaged maps were generated based on fixed- rather than random-effects (see also **Figure 3**). The resultant map was thresholded at $p<10^{-25}$ and overlaid on the common brain template (fsaverage). Panel **B** shows the activity map in one randomly-selected subject (see also **Figure 2**), evoked in response to the same stimulus contrast as in Panel **A**. Here, the activity map was only minimally smoothed (FWHM=2mm). Consequently, multiple smaller scene-selective sites could be detected across the cortex, including PIGS (black arrowhead), located within the posterior intraparietal gyrus. Traditionally, these smaller activity patches are treated as noise in measurement and discarded. For ease in comparing the two panels, the individual's data was also overlaid on the fsaverage.

2. Methods

2.1. Participants

Fifty-nine human subjects (33 females), aged 22-68 years, participated in this study. All subjects had normal or corrected-to-normal vision and radiologically normal brains, without any history of neuropsychological disorder. All experimental procedures conformed to NIH guidelines and were approved by Massachusetts General Hospital protocols. Written informed consent was obtained from all subjects before the experiments.

2.2. General procedure

This study consists of 7 experiments during which we used fMRI to localize and study the evoked scene-selective responses. During these experiments, stimuli were presented via a projector (1024 × 768 pixel resolution, 60 Hz refresh rate) onto a rear-projection screen. Subjects viewed the stimuli through a mirror mounted on the receive coil array. Details of these stimuli are described in the following sections.

During all experiments, to ensure that subjects were attending to the screen, they were instructed to report color changes (red to blue and vice versa) for a centrally presented fixation object (0.1° × 0.1°) by pressing a key on the keypad. Subject detection accuracy remained above 75%, and showed no significant difference across experimental conditions (p>0.10). MATLAB (MathWorks; Natick, MA, USA) and the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) were used to control stimulus presentation.

2.2.1. Experiment 1 – Localization of scene-selective areas

In fourteen subjects (6 females), we localized scene-selective areas PPA/TPA, RSC/MPA and TOS/OPA by measuring their evoked brain activity, using a 3T fMRI scanner, as they were presented with 8 colorful images of real-world (indoor) scenes vs. (group) faces. Scene and face images were retinotopically centered and subtended 20° × 26° of visual field without any significant differences between their root mean square (RMS) contrast (t(14) =1.10, p=0.29). Scene and face stimuli were presented in different blocks (16 s per block and 1 s per image). Each subject participated in 4 runs and each run consisted of 10 blocks plus 32 s of blank presentation at the beginning and at the end of each block. Within each run, the sequence of blocks and the sequence of images within them was randomized.

2.2.2. Experiment 2 – Reproducibility of PIGS across scan sessions (3T vs. 7T)

To localize PIGS with higher spatial resolution and to enhance the signal/contrast to noise ratio (relative to Experiment 1), four subjects were randomly selected from those who participated in Experiment 1 and were scanned in a 7T scanner. These individuals were presented with 300 grayscale images of scenes and 48 grayscale images of (single) faces other than those used in Experiment 1. Here, scene images included pictures of indoor (100 images), manmade outdoor (100 images) and natural outdoor (100 images) scenes, selected from the Southampton-York Natural Scenes (SYNS) dataset (Adams et al., 2016).

As in Experiment 1, all images were retinotopically centered, and subtended 20° × 26° of visual field and there was no significant difference between the RMS contrast across the two categories (t(346) =0.75, p=0.38). Scene and face images were presented across different blocks. Each block contained 24 stimuli (1 s per stimuli), with no blank presentation between the stimuli. The sequence of stimuli was randomized within the blocks. Each subject participated in 12 runs (11

blocks per run; 24 s per block; 1 s per stimulus), beginning and ending with an additional block (12 s) of uniform black presentation. In each run, the sequence of blocks and the sequence of images within them were randomized.

2.2.3. Experiment 3 – PIGS localization relative to area V6 and retinotopic visual areas

Experiment 3a was designed to clarify the relative localization of PIGS vs. area V6 (Pitzalis et al., 2010). All fourteen subjects who participated in Experiment 1 were examined again in a separate scan session using a 3T scanner. During this scan session, we localized area V6 by contrasting the response evoked by coherent radially moving (optic flow) vs. randomly moving white dots (20° × 26°), presented against a black background. The experiment was block-designed, and each block took 16 s. Each subject participated in 5 runs (14 blocks per run), beginning and ending with an additional block of 16 s uniform black presentation.

Experiment 3b was designed to compare the localization of PIGS relative to the border of retinotopic visual areas such as V3A/B and IPS0-4. Two subjects who had participated in Experiment 2 were randomly selected and scanned again in a 7T scanner, during which we defined the border of retinotopic visual areas using a phase encoding approach (Sereno et al., 1995; Engel et al., 1997). Specifically, subjects were presented with rotating (CW and CCW) wedge-shaped (45°) apertures that revolved over 28 seconds, followed by a 4 s blank presentation. Instead of using a flashing checkerboard, we used naturalistic stimuli consisting of color objects presented against a pink-noise background, updated at 15 Hz (Benson et al., 2018). Each subject participated in 10 runs (4 blocks per run).

2.2.4. Experiment 4 – Localization of PIGS in a larger population

Considering the small size of PIGS, it was important to show that this area could survive group-averaging over larger populations, compared to Experiment 1. Accordingly, Experiment 4 localized this area in a large pool of subjects, consisting of thirty-one individuals (19 females) other than those who participated in Experiment 1. The stimuli and procedure were identical to Experiment 1.

2.2.5. Experiment 5 – Response to two independent sets of scenes and non-scene objects

Experiments 1-4 used the response evoked by scenes vs. faces to localize PIGS. However, it remained unknown whether PIGS also showed a selective response to the 'scenes vs. objects' contrast. Accordingly, in two independent groups of subjects (no overlap), Experiment 5 tested the response evoked by scenes vs. non-scene objects in PIGS and the adjacent areas (i.e. V6, TOS/OPA and RSC/MPA).

Specifically, in Experiment 5a, thirteen subjects (7 females), other than those who participated in Experiment 1, were scanned in a 3T scanner. They were presented with 22 grayscale images of indoor/outdoor scenes, other than those presented in Experiments 1-4, and 88 grayscale images that included either a single or multiple everyday non-animate (non-face) objects. All stimuli were retinotopically centered and presented within a circular aperture (diameter=20°). The RMS contrast of the objects was significantly higher than the scenes ($t(108)=3.72$, $p<10^{-3}$). Scene and object images were presented in different blocks according to their category (22 s per block and 1 s per image). Each subject participated in 12 runs and each run consisted of 9 blocks, plus 16 s of blank presentation at the beginning and the end of each block. As in other experiments, the sequence of blocks and the sequence of images within them was randomized.

In Experiment 5b, fourteen subjects (8 females), other than those who participated in Experiment 1 and 5a, were scanned in a 3T scanner. Each subject was presented with 32 grayscales images of indoor/outdoor scenes, 32 images of everyday (non-face) objects plus also their scrambled versions, and 32 images of single faces. Scene and non-scene stimuli were different than those used in Experiment 1-4 and 5a. In contrast to Experiment 5a, all non-scene images included only one single object and there was no significant difference between the RMS contrasts of scenes and the three object categories ($F(3, 111)=0.42$, $p=0.74$). Other details were similar to those in Experiment 5a.

2.2.6. Experiment 6 – Coherently vs. incoherently changing scenes

This experiment was designed to differentiate the role of PIGS in scene perception from TOS/OPA, RSC/MPA and PPA/TPA. Twelve subjects, from the fourteen subjects who participated in Experiment 1, participated in this experiment. The excluded two subjects could not participate further in our tests for personal reasons. Subjects were scanned in a 3T scanner on a different day relative to Experiments 1-3. During this scan, they were presented with rapidly 'coherently vs. incoherently changing scenes' (100 ms per image), across different blocks (16 s per block).

Coherently changing scenes implied ego-motion (fast walking) along 3 different outdoor natural trails. Stimuli (20° × 26°) were generated as one of the experimenters walked through the trails while carrying a camera mounted on his forehead, taking pictures every 2 meters. Incoherently changing scenes consisted of the same images as the coherently changing blocks, but with randomized order. In other words, the only difference between the coherently vs. incoherently changing scenes was the sequence of stimuli within the block. For both coherently and incoherently changing scenes, images from different trails were presented across different blocks.

In separate blocks, subjects were also presented with 80 images that included multiple faces (20° × 26°) with the same timing as the scene images (i.e., 100 ms per image; 16 s per block). All stimuli were grayscaled. Each subject participated in 6 runs and each run consisted of 9 blocks, plus 8 s of blank presentation at the beginning and the end of each block and 4 s of blank presentation between blocks.

On different runs (within the same session), subjects were also presented with concentric rings, extending 20° × 26° (height × width) in the visual field, presented against a light gray background (40 cd/m^2). In half of the blocks (16 s per block), rings moved radially (centrifugally vs. centripetally; 4°/s) and the direction of motion changed every 4 s to reduce the impact of motion after-effects. In the remaining half of the blocks, rings remained stationary throughout the whole block. Each subject participated in 2 runs and each run consisted of 8 blocks, plus 16 s of uniform gray presentation at the beginning and the end of each run. The sequence of moving and stationary blocks was pseudo-randomized across runs.

2.2.7. Experiment 7 – Response to biological motion

To test whether PIGS also responds selectively to biological motion, twelve individuals were selected randomly and were scanned in a 3T scanner while they were presented with the moving point-lights that represented complex biological movements such as crawling, cycling, jumping, paddling, walking, etc. (Jastorff and Orban, 2009). Each action was presented for 2 s and the sequence of actions was randomized across the blocks (20 s per block). As a control, in different blocks, the subjects were shown the same stimuli when all of the point-lights moved in the same direction (i.e., translation motion). Each subject participated in 11 runs and each run consisted of 12 blocks, plus 10 s of blank presentation at the beginning and the end of each run.

2.3. Imaging

2.3.1. 3T scans

In Experiments 1, 3a and 4-6, subjects were scanned in a horizontal 3T scanner (Tim Trio, Siemens Healthcare, Erlangen, Germany). Gradient echo EPI sequences were used for functional imaging. Functional data were acquired using single-shot gradient echo EPI with nominally 3.0 mm isotropic voxels (TR=2000 ms; TE=30 ms; flip angle=90°; band width (BW)=2298 Hz/pix; echo-spacing= 0.5 ms; no partial Fourier; 33 axial slices covering the entire brain; and no acceleration). During the first 3T scan (see the General Procedure), structural (anatomical) data were acquired for each subject using a 3D T1-weighted MPRAGE sequence (TR=2530 ms; TE=3.39 ms; TI=1100 ms; flip angle=7°; BW=200 Hz/pix; echo-spacing=8.2 ms; voxel size=1.0×1.0×1.33 mm).

2.3.2. 7T scans

In Experiments 2 and 3.2, subjects were scanned in a 7T Siemens whole-body scanner (Siemens Healthcare, Erlangen, Germany) equipped with SC72 body gradients (maximum gradient strength, 70 mT/m; maximum slew rate, 200 T/m/s) using a custom-built 32-channel helmet receive coil array and a birdcage volume transmit coil. Voxel dimensions were nominally 1.0 mm, isotropic. Single-shot gradient-echo EPI was used to acquire functional images with the following protocol parameter values: TR=3000 ms; TE=28 ms; flip angle=78°; BW=1184 Hz/pix; echo-spacing=1 ms; 7/8 phase partial Fourier; 44 oblique-coronal slices; and acceleration factor r=4 with GRAPPA reconstruction and FLEET-ACS data (Polimeni et al., 2015) with 10° flip angle. The field of view included the occipital-parietal brain areas to cover PIGS, RSC/MPA and TOS/OPA (but not PPA/TPA).

2.4. Data Analysis

2.4.1. Structural data analysis

For each subject, inflated and flattened cortical surfaces were reconstructed based on the high-resolution anatomical data (Dale et al., 1999; Fischl et al., 1999; Fischl et al., 2002), during which the standard pial surface was generated as the gray matter border with the surrounding cerebrospinal fluid or CSF (i.e., the GM-CSF interface). The white matter surface was also generated as the interface between white and gray matter (i.e. WM-GM interface). In addition, an extra surface was generated at 50% of the depth of the local gray matter (Dale et al., 1999).

2.4.2. Individual-level functional data analysis

All functional data were rigidly aligned (6 df) relative to subject's own structural scan, using rigid Boundary-Based Registration (Greve and Fischl, 2009), and then were motion corrected. Data collected in the 3T (but not 7T) scanner was spatially smoothed using a 3D Gaussian kernel (2 mm FWHM). To preserve the spatial resolution, data collected within the 7T scanner was not spatially smoothed.

Subsequently, a standard hemodynamic model based on a gamma function was fit to the fMRI signal to estimate the amplitude of the BOLD response. For each individual subject, the average BOLD response maps were calculated for each condition (Friston et al., 1999). Finally, voxel-wise statistical tests were conducted by computing contrasts based on a univariate general linear model.

The resultant significance maps based on 3T scans were sampled from the middle of cortical gray matter (defined for each subject based on their structural scan (see section 2.4.1)). For 7T scans, the resultant significance maps were sampled from deep cortical layers at the gray-white matter interface. This procedure reduced the spatial blurring caused by superficial veins (Koopmans et

al., 2010; Polimeni et al., 2010; De Martino et al., 2013; Nasr et al., 2016). For presentation, the resultant maps were projected either onto the subject's reconstructed cortical surfaces or onto a common template (fsaverage; Freesurfer (Fischl, 2012)).

2.4.3. Group-level functional data analysis

To generate group-averaged maps, functional maps were spatially normalized across subjects, then averaged using weighted least square (WLS) random-effects models (using the contrast effect size and the variance of contrast effect size as the input parameters) and corrected for multiple comparisons (Friston et al., 1999). For **Figure 1A** and to replicate our original finding (Nasr et al., 2011), the group-average maps were generated using fixed-effects. The resultant significance maps were projected onto a common human brain template (fsaverage).

2.4.4. Region of interest (ROI) analysis

The main ROIs included area PIGS, the two neighboring scene-selective areas (RSC/MPA, TOS/OPA), and area V6. In Experiment 6, we also included area PPA/TPA in our analysis. These ROIs were localized in two different ways: (1) functionally, for each subject based on their own evoked activity (section 2.4.4.1), and (2) probabilistically, based on activity measured in a different group of subjects (section 2.4.4.2).

2.4.4.1. Functionally localized ROIs

For those subjects who participated in Experiments 6 and 7, we localized scene-selective areas PIGS, TOS/OPA, RSC/MPA, and PPA/TPA based on their stronger response to scenes compared to faces at a threshold level of $p<10^{-2}$, using the method described in Experiment 1. For subjects in Experiment 6, we also localized area V6 based on the expected selective response in this region to coherent radially vs. incoherently moving random dots (see Section 2.2.3). In those subjects in which PIGS and V6 showed partial overlap, the overlapping parts were excluded for the analysis.

2.4.4.2. Probabilistically localized ROIs

For those subjects who participated in Experiments 4 and 5, we tested the consistency of PIGS locations across populations, using probabilistic labels for areas PIGS, TOS/OPA, RSC/MPA and V6. These labels were generated based on the results of Experiment 1 (for PIGS, TOS/OPA and RSC/MPA) and Experiment 3a (for V6). Specifically, we localized the ROIs separately for the individual subjects who participated in Experiments 1 and 3a. Then the labels were overlaid on a common brain template (fsaverage). We computed the probability that each vortex within the cortical surface belonged to one of the ROIs. The labels for PIGS, TOS/OPA, RSC/MPA and V6 were generated based on those vertices that showed a probability higher than 20%. This method assured us that our measurements were not biased by those subjects who showed stronger scene-selective responses. Moreover, by selecting a relatively low threshold (i.e., 20%), we avoided confining our ROIs to the center of activity sites.

2.4.5. Statistical tests

To test the effect of independent parameters, we applied paired t-tests and/or a repeated-measures ANOVA, with Greenhouse-Geisser correction whenever the sphericity assumption was violated.

2.5. Data sharing statement

All data, codes and stimuli are ready to be shared upon request.

MATLAB (RRID: SCR_001622; *https://www.mathworks.com*).

FreeSurfer (RRID:SCR_001847; *https://surfer.nmr.mgh.harvard.edu/fswiki/FsFast*).

Psychophysics Toolbox (RRID:SCR_002881; *http://psychtoolbox.org/docs/Psychtoolbox*).

3. Results

This study consists of seven experiments. Experiment 1 focused on localizing the scene-selective site (PIGS) within the posterior intraparietal region. Experiment 2 showed consistency in the spatial location of PIGS across sessions. Experiment 3 examined PIGS location relative to V6, an area involved in motion coherency and optic flow encoding, and also relative to the retinotopic visual areas IPS0-4. Experiment 4 showed that, despite its small size, PIGS is detectable in group-averaged maps in large populations. Experiment 5 showed that scenes and non-scene objects are differentiable from each other based on the evoked response evoked within PIGS. Experiment 6 tested the response in PIGS to ego-motion in scenes, yielding a result that differentiated PIGS from the other scene-selective regions. Finally, Experiment 7 showed that PIGS does not respond selectively to biological motion.

3.1. Experiment 1 – Small scene-selective sites are detectable within the posterior intraparietal gyrus

When the level of spatial smoothing is relatively low, scene-selective sites (other than PPA/TPA, TOS/OPA and RSC/MPA) are detectable across the brain, especially within the posterior intraparietal gyrus (**Figure 1B**). To test the consistency in location of these scene-selective sites across individuals, fourteen subjects were presented with scene and face stimuli while we collected their fMRI activity. Considering the expected small size of the scene-selective sites within the intraparietal region, we used limited signal smoothing in our analysis (FWHM = 2 mm; see **Methods**) to increase the chance of detecting these sites.

Figure 2 shows the activity maps evoked by the 'scenes > faces' contrast in seven exemplar subjects. All activity maps were overlaid on a common brain template to clarify the consistency in location of scene-selective sites across individuals. In all tested individuals, besides areas RSC/MPA and TOS/OPA, we detected at least one scene-selective site within the posterior portion of the intraparietal gyrus, close to (but outside) the parieto-occipital sulcus (POS). Accordingly, we named this site the posterior interparietal gyrus scene-selective site or PIGS.

When measured at the same threshold levels ($p<10^{-2}$), the relative size of PIGS was 73.86% ± 49.01% (mean ± S. D.) of RSC/MPA, 28.26% ± 15.67% of TOS/OPA, and 19.45% ± 8.43% of PPA/TPA. Considering the proximity of PIGS to the skull and head coil surface (**Figure 1**), the relatively small size of PIGS could not be ascribed to the lower signal/contrast to noise ratio in that region.

To better clarify the consistency of PIGS localization across subjects, we also generated group-averaged activity maps based on random-effects, and after correction for multiple comparisons. As demonstrated in **Figure 3A**, PIGS was also detectable in the group-averaged activity maps, in almost the same location as in the individual subject maps. Overall, these results suggest that, despite the relatively small size of this scene-selective site, PIGS is consistently detectable across subjects in the same cortical location.

3.2. Experiment 2 – PIGS reproducibility across scan sessions

To test the reproducibility of our results, four subjects were selected randomly from those who participated in Experiment 1. These subjects were scanned again (on a different day), using a 7T (rather than a 3T) scanner, and a different set of scenes and faces (**Figure 4A**).

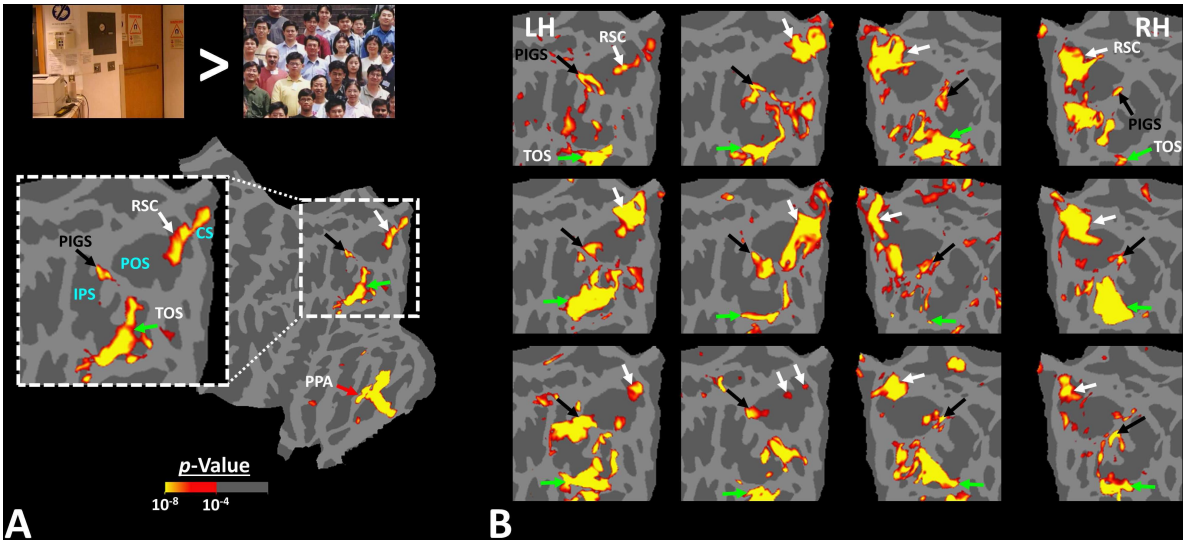

Figure 2

Activity evoked by 'scene > face' contrast in seven individual subjects, other than the one shown in Figure 1.

Panel **A** shows the significance of evoked activity in the right hemisphere of one individual subject. The inset shows the enlarged activity map within the intraparietal region. The three scene-selective areas, along with area PIGS, are indicated in the map with arrowheads. The location of adjacent sulci the parieto-occipital sulcus (POS), the intraparietal sulcus (IPS) and the calcarine sulcus (CS)) are also indicated in the inset. Panel **B** shows the result from six other individuals. In this panel, the first two columns show the activity within the left hemisphere, while the next two columns show the activity within the right hemisphere of the same subjects. In all subjects, PIGS is detectable bilaterally within the posterior portion of the intraparietal gyrus, near (but outside) the POS. For all of the subjects, threshold level was set at $p<10^{-4}$. All activity maps were overlaid on the fsaverage to highlight the consistency in PIGS location across the subjects.

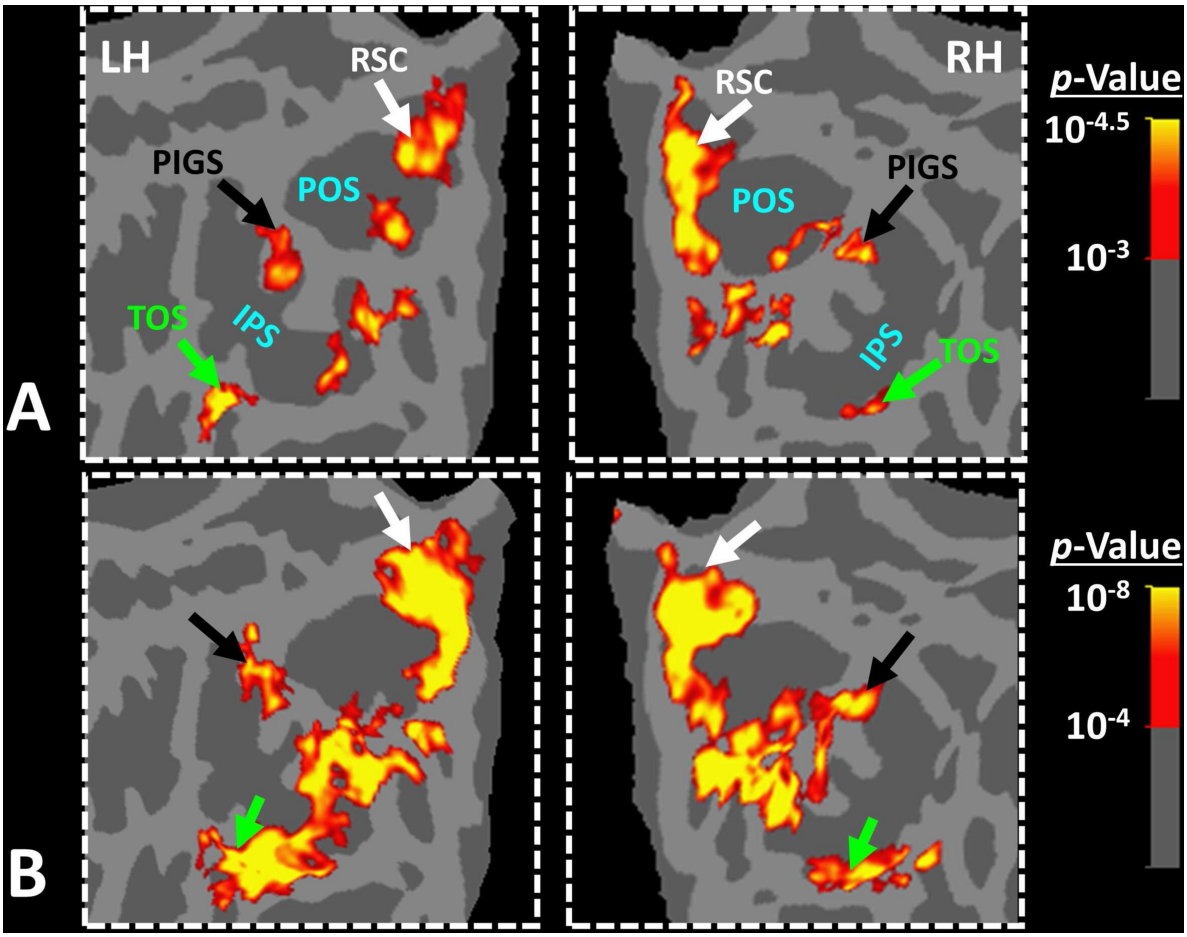

Figure 3

PIGS was detected in group-averaged activity maps across two non-overlapping populations.

Panel **A** shows the group-averaged activity, evoked within the intraparietal region of fourteen subjects who participated in Experiment 1. Panel **B** shows the group-averaged activity, evoked within the intraparietal region of thirty-one subjects who participated in Experiment 4. Importantly, PIGS was evident in both groups bilaterally in the corresponding location (black arrows). Thus, despite its small size, this area was detectable even in the group-averaged activity maps based on large populations. Notably, in both panels, maps were generated based on random-effects, after correction for multiple comparisons. In both maps, the location of RSC/MPA and TOS/OPA are respectively indicated with white and green arrowheads.

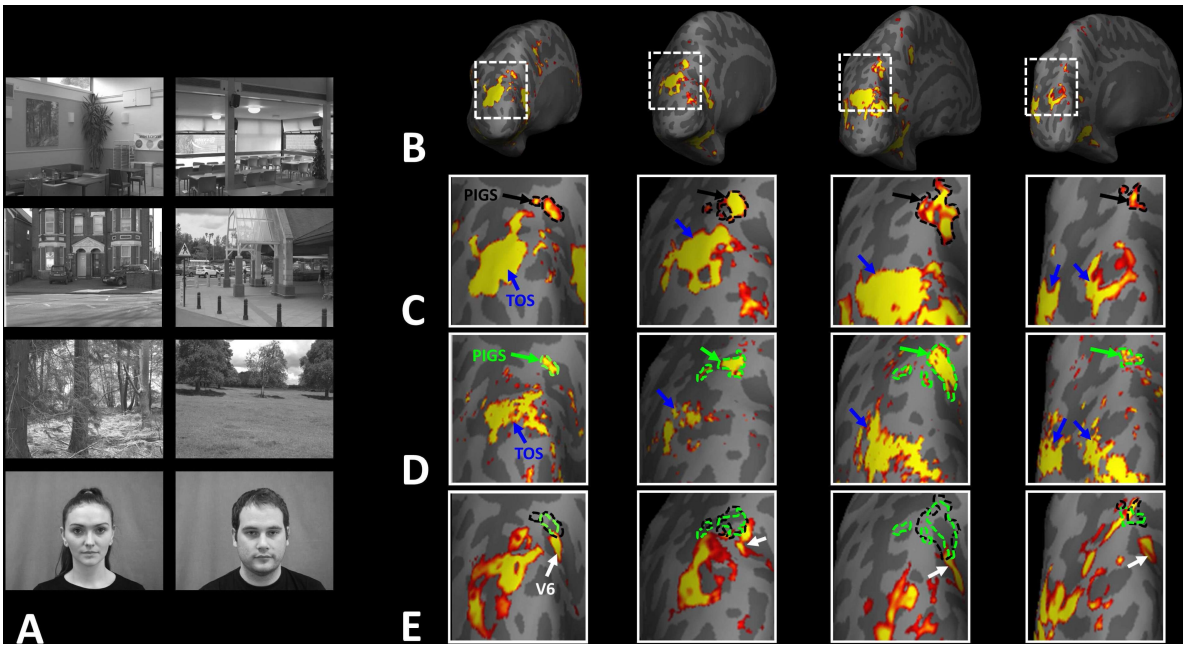

Figure 4

PIGS was detected consistently across sessions.

Panel **A** shows the stimuli used for localizing PIGS during 7T scans. Stimuli including indoor, manmade outdoor and natural outdoor scenes and faces other than those used in Experiment 1. Panels **B and C** show the significance ($p<10^{-2}$) of activity evoked by 'scene > face' contrast in the 3T scans (Experiment 1), overlaid on subjects own reconstructed brain (left hemisphere). Panel **D** shows the significance ($p<0.05$) of activity evoked by 'scene > face' contrast during 7T scans (Experiment 2). Despite the difference in scanners (3T vs. 7T) and stimuli, the location of PIGS remained mostly unchanged. Panel **E** shows the location of PIGS, measured in 3T (black dashed lines) and 7T (green dashed lines) relative to the location of area V6 (white arrowhead), localized functionally based on the response to 'optic-flow > random motion' (Experiment 3a). In all subjects, the center of scene- and optic-flow-selective responses was adjacent, but not overlapping.

As demonstrated in **Figure 4**, despite utilizing a different scanner and a different set of stimuli, PIGS was still detectable in the same location (**Figure 4B-D**). Here again, PIGS was localized within the posterior portion of the intraparietal gyrus and close to the posterior lip of parieto-occipital sulcus. Considering the higher contrast/signal to noise ratio of 7T (compared to 3T) scans, this result strongly suggested that the PIGS evidence was not simply a nuisance artifact in fMRI measurements.

3.3. Experiment 3 – Localization of areas PIGS vs. V6 and retinotopic visual areas

Posterior intraparietal cortex accommodates area V6, which is involved in motion coherency (optic-flow) encoding (Pitzalis et al., 2010). Recent studies have suggested that scene stimuli evoke a strong response within V6 (Sulpizio et al., 2020). Moreover, the intraparietal cortex accommodates multiple retinotopically organized visual areas (Swisher et al., 2007), including IPS0-4 that are believed to be involved in spatial attention control and higher-level object information processing (Silver et al., 2005; Konen and Kastner, 2008). Previous studies have suggested that the area TOS/OPA overlaps with the retinotopic visual areas V3A/B and IPS0 (V7) (Nasr et al., 2011; Silson et al., 2016). In Experiment 3, we clarified the location of PIGS relative to these regions.

In Experiment 3a we localized V6 in all subjects who participated in Experiment 1, based on visual presentation of random vs. radially moving dots (see Methods). **Figure 4D** shows the co-localization of V6 and PIGS in four individual subjects. Consistent with previous studies (Pitzalis et al., 2010; Pitzalis et al., 2015), V6 was localized *within* the posterior portion of the POS without any overlap between its center and PIGS. To test the relative localization of these two regions at the group level, we generated probabilistic labels for PIGS and V6 (see Methods). As demonstrated in **Figure 5**, the probabilistic label for PIGS was localized within the intraparietal gyrus and outside the POS (**Figure 5A**), while V6 was located within the POS (**Figure 5B**). We also did not find any overlap between area V6 and areas RSC/MPA and TOS/OPA (**Figure 5C**). Thus, despite the low threshold level used to generate these labels (probability > 20%), the areas PIGS and V6 were located side-by-side (**Figure 5D**), without any overlapping between their centers.

In Experiment 3b we scanned two subjects, randomly selected from those who had participated in Experiment 2, using a 7T scanner to map the borders of retinotopic visual areas (see **Methods**). As demonstrated in **Figure 6**, in both subjects, PIGS was located adjacent to IPS3 and IPS4. In comparison, TOS/OPA was located more ventrally relative to PIGS, overlapping with areas V3A/B and IPS0 (V7). Considering these differences in the localization of PIGS vs. TOS/OPA, relative to the anatomical and functionally defined landmarks, our results further suggest that PIGS and TOS/OPA are two distinct visual areas.

3.4. Experiment 4 – PIGS localization in a larger population

The results of Experiments 1-3 suggest that PIGS can be localized consistently across individual subjects, and this area appears to be distinguishable from the adjacent area V6. However, considering the small size of this area, it appears necessary to test whether this area was detectable based on group averaging in a larger population. Accordingly, in Experiment 4 we scanned thirty-one individuals (other than those who participated in Experiments 1-3) while they were presented with the same stimuli as in Experiment 1 (**Figure 2**).

As demonstrated in **Figure 3B**, PIGS was also detectable in this new population in almost the same location as in Experiment 1. Specifically, PIGS was detected bilaterally within the posterior portion of the intraparietal gyrus, adjacent to the POS. We did not find a significant difference

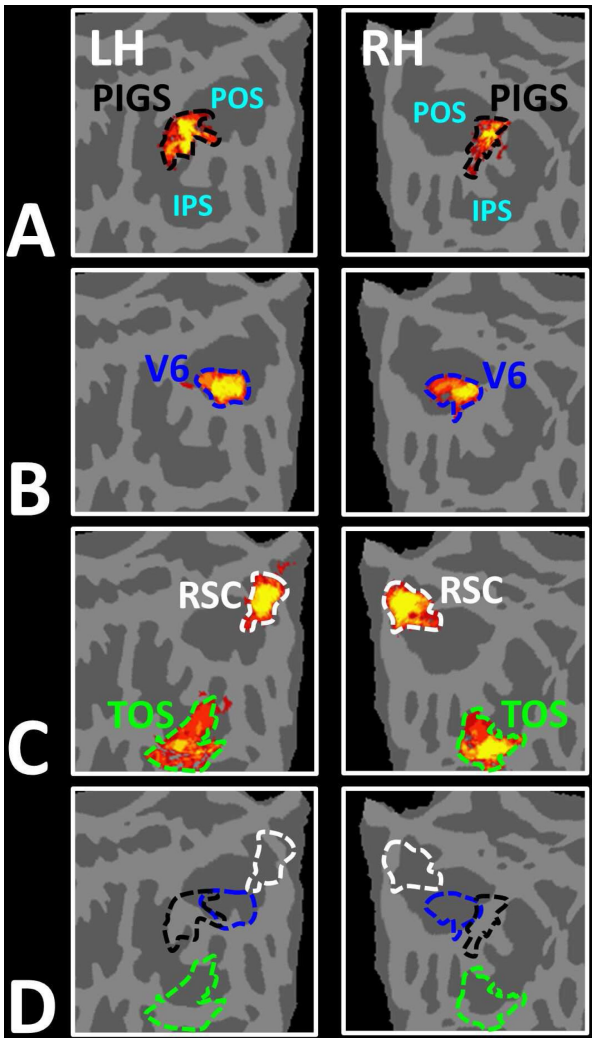

Figure 5

Area PIGS is located outside the POS and adjacent to the functionally-localized area V6.

Panels **A and B** show the probabilistic localization of areas PIGS and V6, respectively (see **Methods**). Panel **C** shows the probabilistic localization of areas RSC/MPA and TOS/OPA. All probability maps are thresholded at 20%-50% (red-to-yellow) and overlaid on the fsaverage. Panel **D** shows the relative location of these sites. Consistent with the results from the individual maps (**Figure 4E**), PIGS and V6 were located adjacent to each other, such that V6 was located within the POS and PIGS located outside the POS (within the intraparietal gyrus) with minimal overlap between the two regions.

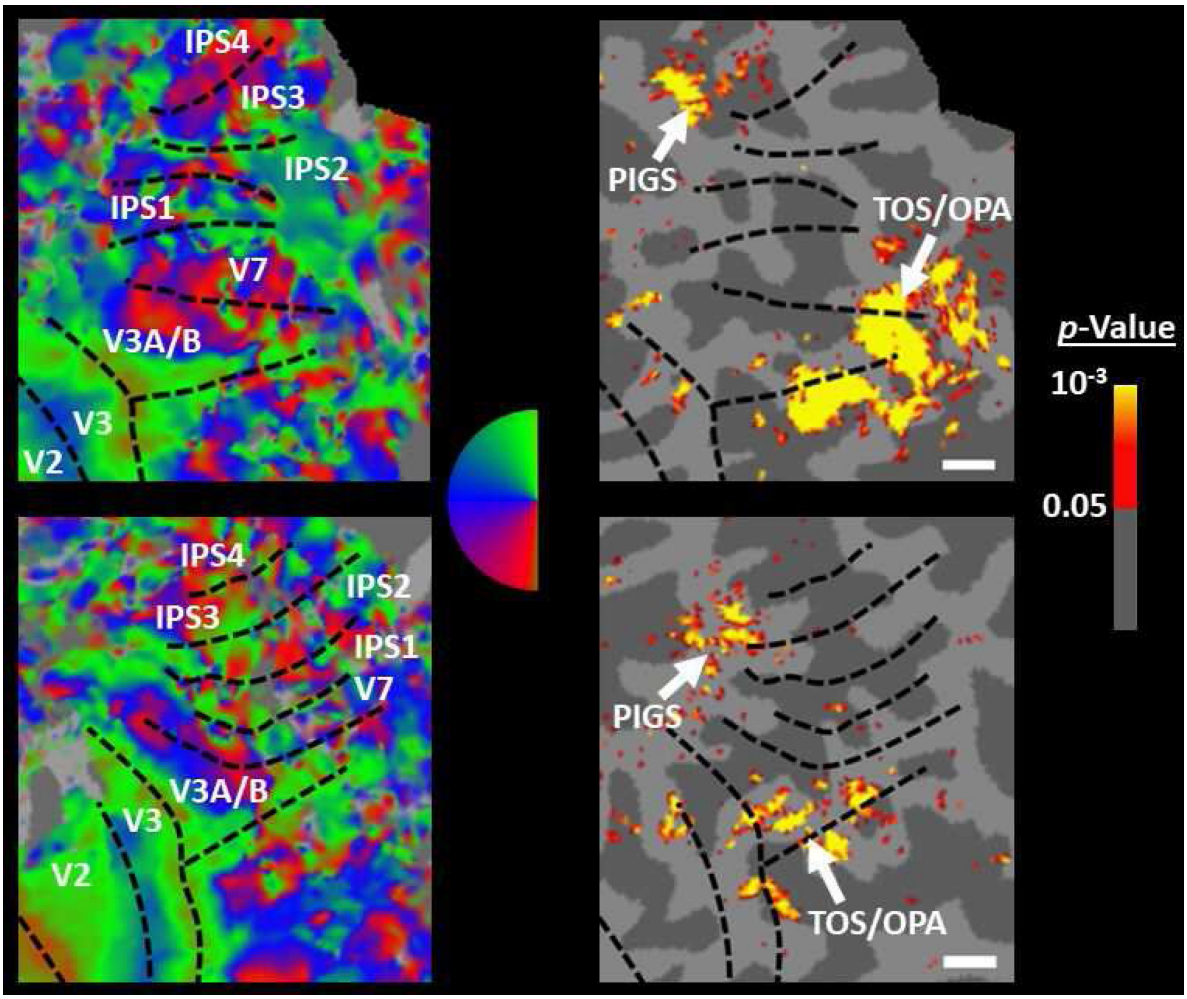

Figure 6

Localization of PIGS and TOS/OPA relative to the retinotopic visual areas in the right hemisphere of two subjects.

The right and left columns show respectively the polar angle and scene>face response mapping, collected in a 7T scanner on two different days. In both columns, the borders of visual areas (defined based on the polar angle mapping) are indicated by dashed black lines. For both subjects, maps were overlaid on their own reconstructed flattened cortex. No activity smoothing was applied to the collected data (i.e., FWHM = 0; see **Methods**). Similar results were also found in the opposite hemispheres (not shown here). On the right column, the scale bars indicate 1 cm.

between the two populations in the size of PIGS when normalized either relative to the size of RSC/MPA (t(43)=0.98, p=0.33), or TOS/OPA (t(43)=0.26, p=0.80) or PPA/TPA (t(43)=0.52, p=0.61). Thus, the location and relative size of PIGS appeared to remain unchanged across populations.

These results suggest that one may rely on the probabilistically generated labels to examine the evoked activity within PIGS. To test this hypothesis, we measured the level of scene-selective activity in PIGS, along with the areas TOS/OPA, RSC/MPA and V6, using the probabilistic labels generated based on the results of Experiments 1 and 3a (see **Methods** and **Figure 5**). As demonstrated in **Figure 7A-B**, results of this ROI analysis showed a significant scene-selective activity in PIGS (t(31)=8.11, $p<10^{-8}$), TOS/OPA (t(31)=7.91, $p<10^{-7}$) and RSC/MPA (t(31)=9.11, $p<10^{-8}$). Importantly, despite the proximity of PIGS and V6, the level of scene-selective activity in PIGS was significantly higher than that in V6 (t(11)=5.03, $p<10^{-4}$). Thus, it appears that the probabilistically generated ROIs can be used to examine PIGS response, and to differentiate it from adjacent areas such as V6 (see also Experiment 5).

3.5. Experiment 5 – Selective response to scenes compared to non-scene objects in PIGS

Thus far, we localized PIGs in multiple experiments by contrasting the response evoked by scenes vs. faces. In Experiments 5a and 5b, we examined whether PIGS also showed a selective response to scenes compared to objects (not just faces). In Experiment 5a, twelve individuals, other than those who participated in Experiments 1-3, were scanned while viewing pictures of scenes (other than those used to localize PIGS) and everyday objects (**Figure 8A**) (see **Methods**).

As demonstrated in **Figures 8B** and **8C** for one individual subject, 'scenes vs. objects' and 'scenes vs. faces' (Experiment 4) contrasts generated similar activity maps. Importantly, in both maps, PIGS was detectable in a consistent location adjacent to (but outside) the parieto-occipital sulcus. Moreover, results of an ROI analysis, using the probabilistically generated labels based on the results of Experiments 1 and 3a, yielded significant scene-selective activity within PIGS (t(11)=6.57, $p<10^{-4}$), RSC/MPA (t(12)=11.00, $p<10^{-6}$) and TOS/OPA (t(12)=6.26, $p<10^{-3}$) (**Figures 9A** and **9B**). We also found that the level of scene-selective activity within PIGS is significantly higher than that in the adjacent area V6 (t(11)=2.42, p=0.03). Thus, scenes and (non-face) objects are differentiable from each other, based on the activity evoked within PIGS.

In Experiment 5b, fifteen individuals (other than those who participated in Experiments 1 and 5a), were scanned while viewing a new set of stimuli that included pictures of scenes, faces, everyday objects and scrambled objects (**Figure 8D**). In contrast to Experiment 5a in which the number of objects within each image could vary, here, each image contained only one object (see **Methods**). Despite this change, contrasting the response to scene vs. non-scene images (averaged over objects, scrambled objects and faces) evoked a similar activity pattern, as Scene vs. Faces (**Figures 8E** and **8F**). Moreover, the ROI analysis yielded a significant scene-selective activity within PIGS (t(14)=2.37, p=0.03), RSC/MPA (t(14)=10.33, $p<10^{-7}$) and TOS/OPA (t(14)=4.79, $p<10^{-3}$) (**Figures 9**). Here again, the level of scene-selective activity within PIGS was higher than V6 (t(14)=2.27, p=0.04). Together, results of Experiments 1-5 suggest that PIGS responds selectively to a wide range of scenes compared to non-scene objects, and that the level of this activity is higher than in the adjacent area V6.

3.6. Experiment 6 – PIGS response to ego-motion

Experiments 1-5 clarified the location of PIGS, and its general functional selectivity for scenes. However, a more specific role of this area in scene perception remains undefined. Experiment 6 tested the hypothesis that area PIGS is involved in encoding ego-motion within scenes. This hypothesis was motivated by the fact that PIGS is located adjacent to V6 (**Figure 5D**), an area

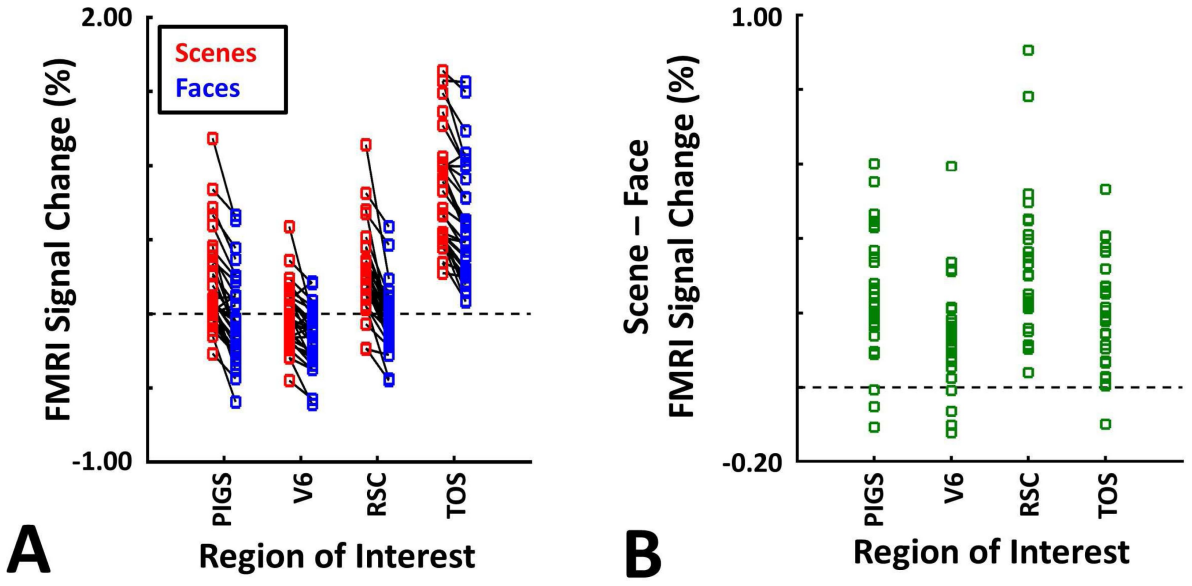

Figure 7

Probabilistically generated labels can be used to detect PIGS.

Panel **A** shows the activity evoked by 'scenes vs. faces' stimuli, across PIGS, V6, RSC/MPA and TOS/OPA. Panel **B** shows the level of scene-selective activity, measured as 'scene – face', within these regions. Despite the small size of PIGS, the probabilistic label could detect the scene-selective activity within this area and the level of this activity was significantly higher than the adjacent area V6. In all panels, each dot represents the activity measured in one subject.

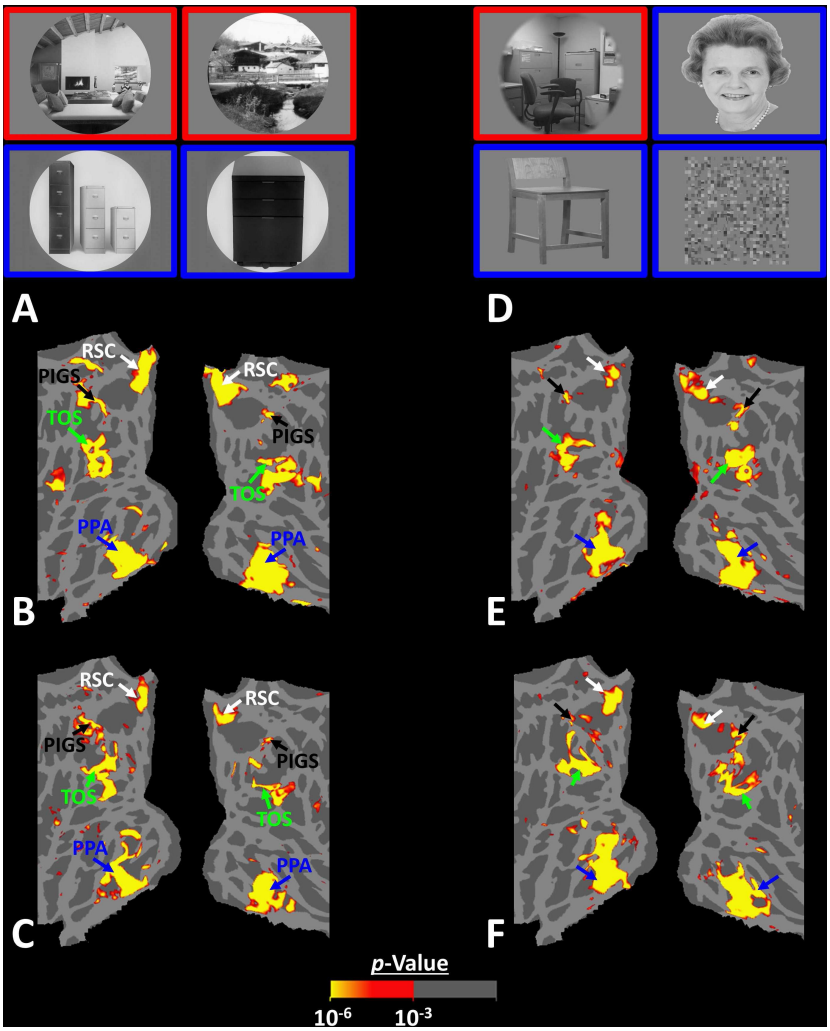

Figure 8

PIGS could also be detected based on the 'scene > object' contrast.

Panels **A and D** show the stimuli used in Experiments 5a and 5b respectively. Panels **B and E** show the activity maps evoked by 'scene > object' contrast in two different individuals who participated in Experiment 5a and 5b. Panels **C and F** show the activity maps evoked by and a different set of scenes and faces (used in Experiments 1 and 4) in the same individuals. The location of PIGS remained unchanged between the two maps.

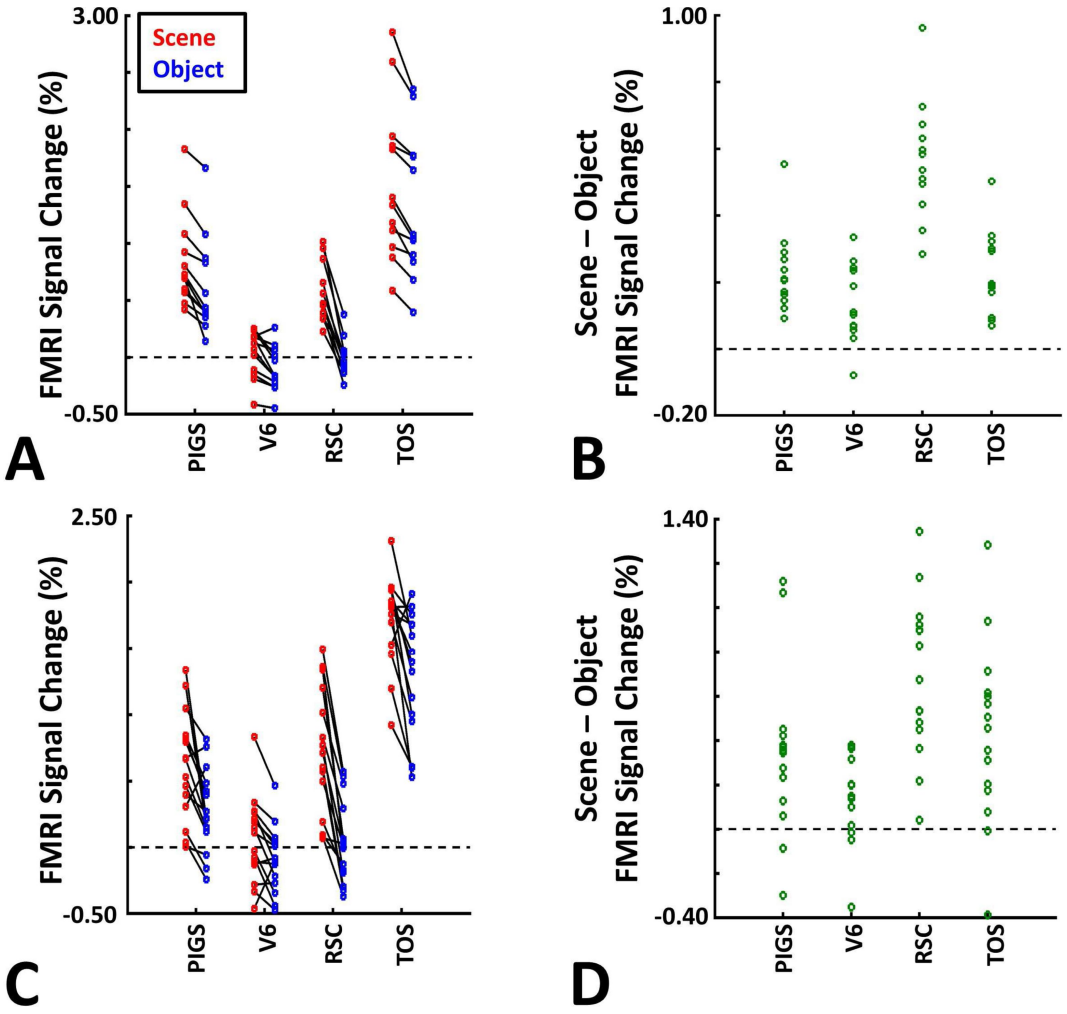

Figure 9

The application of probabilistically generated labels to measure the PIGS response to 'scene vs. object' stimuli.

Panels **A and C** show the activity evoked by 'scenes vs. object' stimuli in Experiments 5a and 5b, respectively. Panels **B and D** show the level of scene-selective activity within the regions of interest. As in Experiment 4, the probabilistic label detected the scene-selective activity within PIGS and the level of this activity was significantly higher than the adjacent area V6. Other details are similar to **Figure 7**.

involved in encoding optic flow. Other studies have also suggested that ego-motion may influence the scene-selective activity within this region, without clarifying whether this activity was centered either within or outside V6 (Pitzalis et al., 2020; Sulpizio et al., 2020).

Twelve individuals, from those who participated in Experiment 1, took part in this experiment (see Methods). These subjects were presented with coherently changing scene stimuli that implied ego-motion across different outdoor trails (**Figure 10**). In separate blocks, they were also presented with incoherently changing scenes and faces. **Figure 9** shows the group-averaged scene-selective activity, evoked by coherently (**Figure 11A**) and incoherently changing scene stimuli (**Figure 11B**). Consistent with our hypothesis, PIGS showed a significantly stronger response (bilaterally) to coherently (compared to incoherently) changing scenes that implied ego-motion (**Figure 11C**). However, the level of activity within RSC/MPA and TOS/OPA did not change significantly between these two conditions.

Consistent with the group-averaged activity maps, results of an ROI analysis (**Figure 12**) yielded a significantly stronger response to coherently (vs. incoherently) changing scenes in PIGS ($t(11)=5.97$, $p<10^{-4}$) but not in RSC/MPA ($t(11)=0.12$, $p=0.90$) and TOS/OPA ($t(11)=0.48$, $p=0.64$). Interestingly, area PPA/TPA showed a stronger response to incoherently (compared to coherently) changing scenes ($t(11)=3.48$, $p<0.01$). To better clarify the difference between scene-selective areas, we repeated this test by applying a one-way repeated measures ANOVA to the differential response to 'coherently vs. incoherently changing scenes', measured across these four scene-selective areas. This test yielded a significant effect of area on the evoked differential activity ($F(3, 11)=53.89$, $p<10^{-10}$). Post hoc analysis, with Bonferroni correction, showed that the level of differential activity evoked by 'coherently vs. incoherently changing scenes' was significantly higher within PIGS than all other scene-selective areas ($p<10^{-6}$). These results suggest a distinctive role for area PIGS in ego-motion encoding, that differentiates it from the other scene-selective areas. The absence of activity modulation in the other scene-selective areas also ruled out the possibility that the activity increase in PIGS was simply due to attentional modulation during coherently vs. incoherently changing scenes (see **Discussion**).

In addition to PIGS, we also found a significantly stronger response to coherently (rather than incoherently) changing scenes in area V6 ($t(11)=3.57$, $p<0.01$). However, the level of this selectivity was significantly weaker in V6 compared to that in PIGS ($t(11)=2.63$, $p=0.02$). Moreover, in the group-averaged activity maps, the contrast between coherently vs. incoherently changing scenes yielded a stronger response outside (rather than inside) the POS and also in area MT, located at the tip of medial temporal sulcus (**Figure 11C**). Together, these results suggest that the impact of ego-motion on scene processing is stronger in PIGS than that in V6.

In the same session (but different runs), we also tested the selectivity of the PIGS response for simpler forms of motion. In different blocks, subjects were presented with radially moving vs. stationary concentric rings (see **Methods**). Consistent with the previous studies of motion perception (Pitzalis et al., 2010; Hacialihafiz and Bartels, 2015), the results of an ROI analysis here, did not yield any strong (significant) motion-selective activity within PIGS ($t(11)=1.84$, $p=0.10$), RSC/MPA ($t(11)=1.97$, $p=0.08$), PPA/TPA ($t(11)=1.93$, $p=0.08$) and V6 ($t(11)=2.03$, $p=0.07$). In contrast, we found strong motion selectivity within area TOS/OPA ($t(11)=4.57$, $p<10^{-3}$), likely due to its overlap with the motion-selective area V3A/B (Nasr et al., 2011). Thus, in contrast to optic flow and ego-motion, simpler forms of motion only evoke weak-to-no selective activity within PIGS and V6.

3.7. Experiment 7 – PIGS response to biological motion

The results of Experiment 6 showed that PIGS responds selectively to ego-motion in scenes, but not strongly to radially moving rings. However, it could be argued that PIGS may also respond to the other types of complex motion, e.g., biological motion. To test this hypothesis, we measured the PIGS response to biological vs. translational motion in twelve subjects (see Methods). As

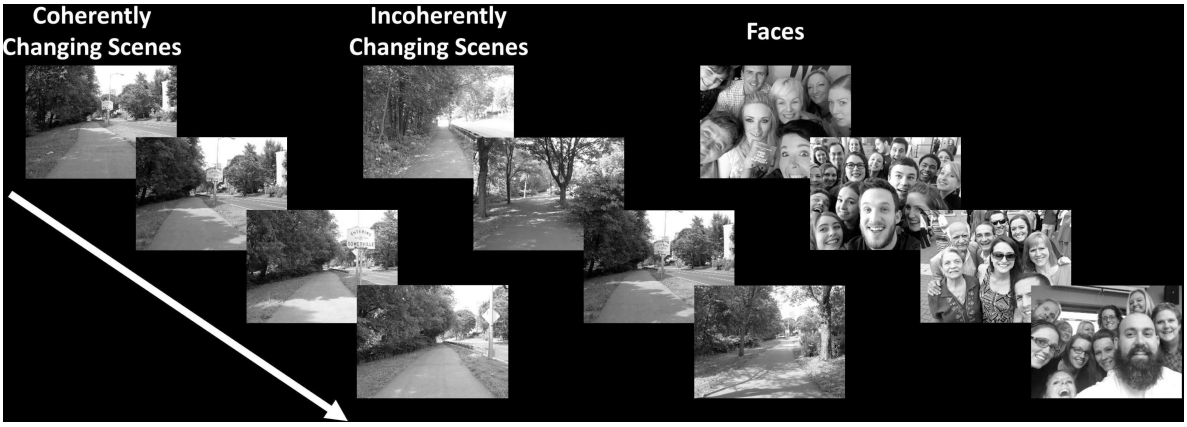

Figure 10

Example of stimuli used in Experiment 6.

Coherently changing scenes implied ego-motion, as if the observer was jogging through a trail. Incoherently changing scenes consisted of the same scene images as the coherently changing scenes, but presented in a pseudo-random order. Face stimuli consisted of a mosaic of faces. These stimuli were different than those used in the previous experiments.

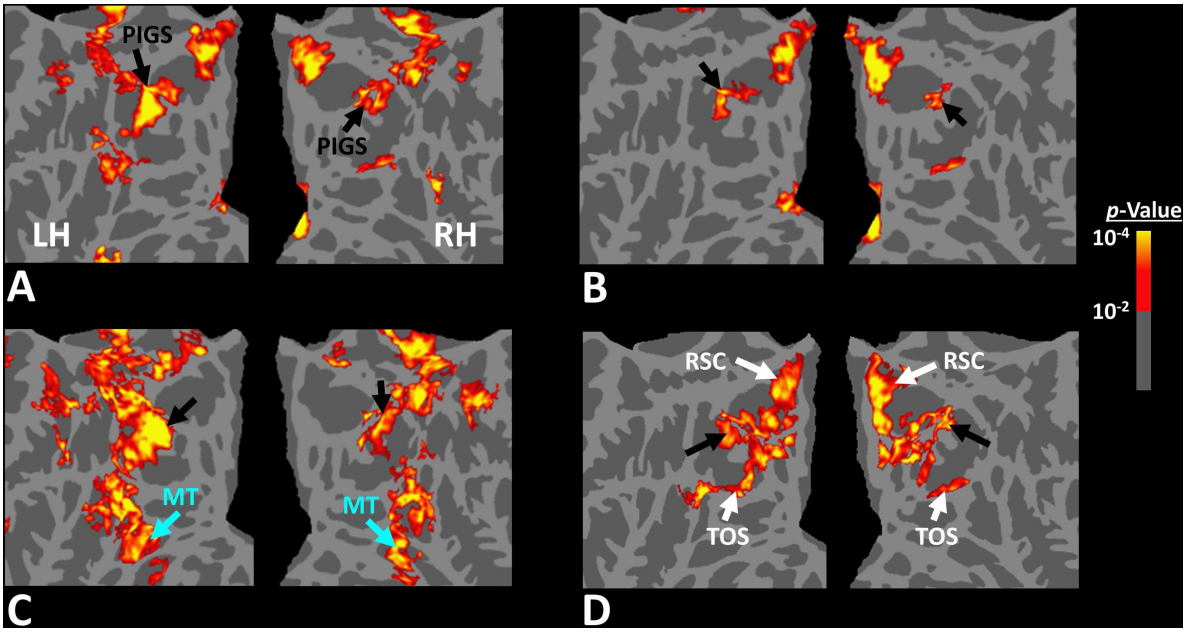

Figure 11

Scene-selective response to coherently vs. incoherently changing scenes within the intraparietal region (Experiment 6).

Panels **A and B** show respectively the group-averaged activity evoked by coherently and incoherently changing scenes relative to faces. Panel **C** shows the group-averaged response evoked by the 'coherently > incoherently changing scenes' contrast. Among scene-selective areas, only PIGS showed significant sensitivity to the observer ego-motion. Besides PIGS, this contrast also evoked activity within area MT (cyan arrowhead), also more dorsal portions of the parietal cortex. Panel **D** shows the location of scene-selective areas in the same group of subjects based on an independent set of scene and face stimuli (Experiment 1) and generated based on random-effects. The location of PIGS (outside the POS) and RSC/MPA (within the POS) are indicated by black and white arrowheads, respectively. All maps were generated based on random-effects, after correction for multiple comparisons.

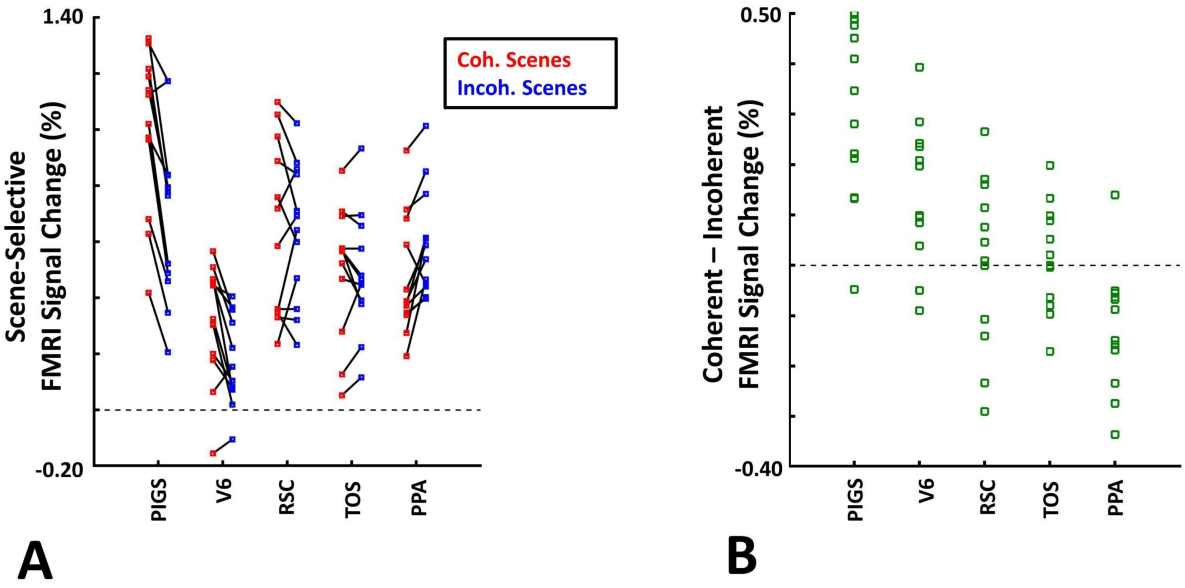

Figure 12

The scene-selective activity evoked within PIGS is influenced by the observer ego-motion.

Panel **A** shows the scene-selective activity evoked by the coherently (red) and incoherently changing scenes (blue), measured relative to the response to the faces, across areas PIGS, V6, RSC/MPA, TOS/OPA and PPA/TPA. Panel **B** shows the level of difference between the response evoked by 'coherently – incoherently' changing scenes across the regions of interest. While all regions showed a significantly stronger response to scenes compared to faces, PIGS showed the strongest impact of ego-motion on the scene-selective response. Other details are similar to **Figure 7**.

illustrated in **Figure 13**, and consistent with the previous studies of biological motion (Puce et al., 1998; Beauchamp et al., 2003; Puce and Perrett, 2003; Pelphrey et al., 2005; Jastorff and Orban, 2009; Kamps et al., 2016), biological motion evoked a stronger response bilaterally within area MT and superior temporal sulcus but not within the posterior intraparietal gyrus. Consistent with the maps, an ROI analysis (based on the functionally-defined labels) showed no significant difference between the response to biological vs. translational motion within PIGS ($t(11)=1.27$, $p=0.23$), TOS/OPA ($t(11)=1.63$, $p=0.13$), RSC/MPA ($t(11)=1.40$, $p=0.18$), and PPA/TPA ($t(11)=0.41$, $p=0.69$). These results indicated that PIGS does not respond to all types of complex motion.

4. Discussion

These data suggest that selective scene processing is not limited to areas PPA/TPA, RSC/MPA and TOS/OPA, and that additional smaller scene-selective sites can also be found across the visual system. By focusing on one small scene-selective site, we showed that this site (PIGS) was consistently identifiable across individuals and groups. We also showed that inclusion of this site in the models of scene processing may clarify how ego-motion influences scene perception.

4.1. FMRI and all that "noise, noise, noise"!

The early fMRI studies dealt with a considerable amount of noise in measurements, partly due to using lower magnetic field scanners and imperfect hardware and software. This noise in measurements affected the reliability of the findings. Consequently, those early studies focused on larger activity sites that were more reliably detectable across subjects/sessions. The smaller sites were either ignored or eliminated by excessive signal smoothing, applied to enhance the level of contrast to noise ratio.

However, advances in neuroimaging techniques have now made it possible to detect and distinguish fMRI activity at the spatial scale of cortical columns (Yacoub et al., 2007; Zimmermann et al., 2011; Nasr et al., 2016). Although the reliability of the fMRI signal still depends on the number of trial repetitions, a spatially confined, but extensively repeated, evoked response can be detected reliably across different sessions (Nasr et al., 2016; Kennedy et al., 2023).

The present data shows that PIGS could be localized consistently across multiple subjects and across different sessions and scanners. Furthermore, our results indicated that the probabilistic labels, generated based on one population, can be used to localize PIGS, and to distinguish its function from the adjacent regions (e.g., V6) in a second population. Together, these results highlight the reliability of current fMRI techniques in detecting smaller cortical regions, in the level of individual subjects.

4.2. PIGS responds selectively to a variety of scene stimuli

To establish a true category-selective response, the stimulus set should sample enough variety to reflect the range and variability among the category members. Consistent with this are the many (and continuing) studies seeking to define the range and fundamental aspects of 'place selective' (Epstein and Kanwisher, 1998; Troiani et al., 2014) and 'face selective' (Kanwisher et al., 1997; Yue et al., 2011) stimuli in extrastriate visual cortex, decades after their first discovery.

Accordingly, here we tested five different scene stimulus sets across our experiments, including a wide variety of indoor/outdoor and natural/manmade scenes. In all cases, we were able to evoke a selective response within PIGS, and the level of this response was comparable to that in the adjacent scene-selective areas RSC/MPA and TOS/OPA. Thus, the scene-selective response in PIGS

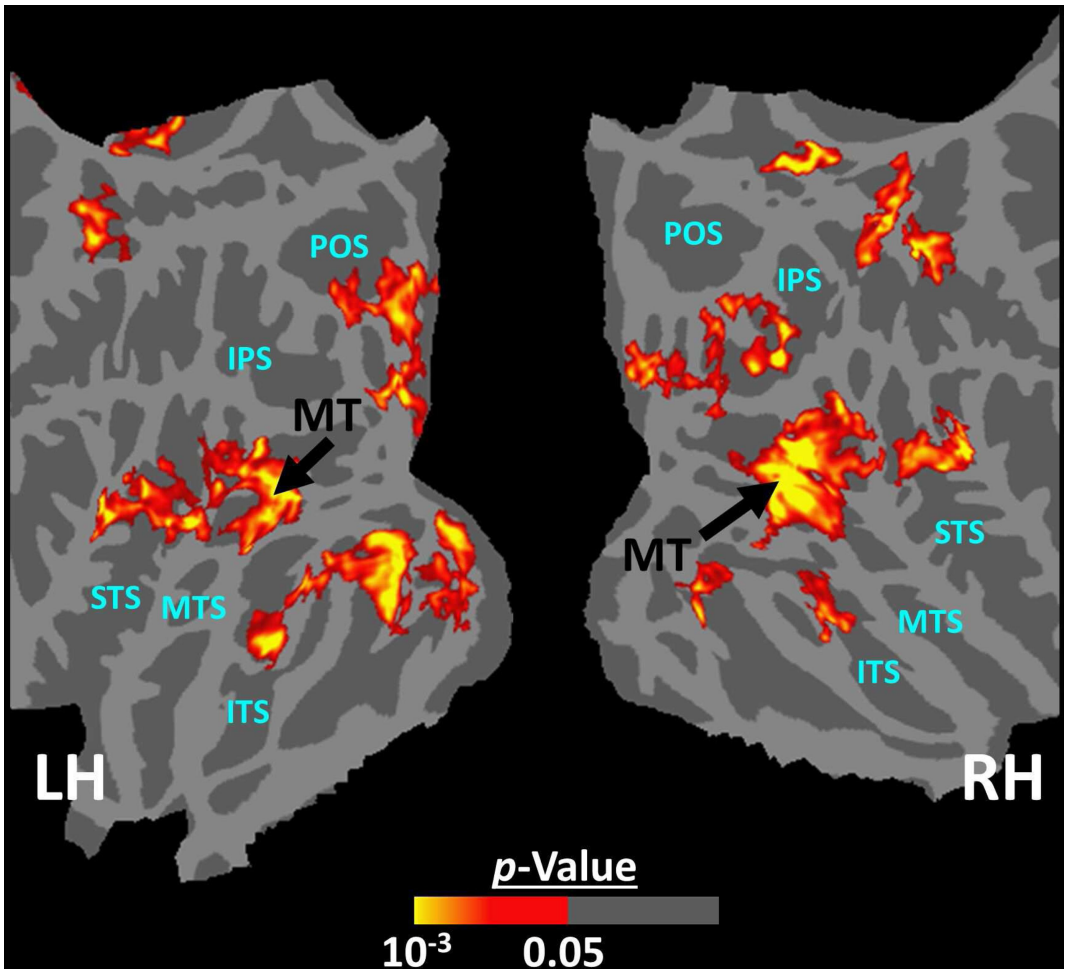

Figure 13

The group-averaged activity map evoked by the 'biological > translational motion' contrast.

Despite the low threshold used to generate these maps, we did not detect any significant activity evoked by the 'biological > translational motion' contrast within the PIGS and/or the other scene-selective areas. Rather, this contrast evoked a significant activity mainly within the inferior temporal sulcus (ITS), medial temporal sulcus (MTS) and superior temporal sulcus (STS).

appeared not to be limited to a single subset of scenes. However, it remains unclear whether scene stimuli are differentiable from each other based on the pattern of evoked response in this region. More experiments are necessary to test this hypothesis (see also the Limitations).

4.3. PIGS and TOS/OPA are two different areas

Our results clearly showed that PIGS and TOS/OPA are two distinct scene-selective areas based on multiple criteria: First, anatomically, TOS/OPA is located mostly anterior to the IPS, whereas PIGS is located more dorsally and posterior to the IPS. Second, TOS/OPA overlaps with areas V3A/B and IPS0 whereas PIGS was located adjacent to IPS3-4. Third, these two areas respond distinctly to moving stimuli. Specifically, while TOS/OPA responds selectively to moving concentric rings and less selectively to ego-motion, PIGS shows the opposite pattern and responds selectively to ego-motion within the naturalistic scenes but not to moving rings (see below). Considering these anatomical and functional differences, these two areas appear to be two distinct hubs within the scene processing networks.

Also notably, PIGS is located relatively far from the lateral place memory area (LPMA), which is located anterior to the IPS and close to the tip of the superior temporal sulcus (Steel et al., 2021; Steel et al., 2023). Considering this, and the fact that there was no memory demand in our paradigms, PIGS and LPMA also appear to be two distinct visual areas.

4.4. PIGS is not just another scene selective area

Our results (Experiment 6) suggest that ego-motion can significantly influence the activity evoked within PIGS. This phenomenon distinguishes the role of PIGS in scene perception, relative to other scene-selective regions. Specifically, previous studies have shown that PPA/TPA and RSC/MPA show weak-to-no sensitivity to motion *per se* (Hacialihafiz and Bartels, 2015). In comparison, area TOS/OPA shows a stronger motion-selective response, presumably related to its (partial) overlap with area V3A/B (Tootell et al., 1997; Nasr et al., 2011). Instead, the current data show that the ego-motion related activity within PIGS is stronger than in TOS/OPA.

This finding is consistent with the fact that PIGS is located adjacent to area V6 (**Figures 4** and **5**), an area that contributes to encoding optic flow (Pitzalis et al., 2010). Considering PIGS and V6 proximity, hypothetical inputs from V6 may contribute to the strong ego-motion selective response in PIGS. This said, the current data also suggests that the role of PIGS differs from that in V6, in terms of ego motion encoding. Compared to V6, PIGS showed a stronger impact of ego-motion on scene processing, while V6 shows a stronger response to optic flow induced by random dot arrays. Thus, PIGS contributes to scene encoding and ego motion within scenes, while V6 is likely involved in detecting optic flow caused by ego-motion.

4.5. Ego-motion Encoding in PIGS vs. TOS/OPA

We showed that PIGS and TOS/OPA are located on two different sides of the IPS with TOS/OPA located more ventrally compared to PIGS. We also showed a stronger impact of ego-motion on activity within PIGS compared to TOS/OPA. In contrast, TOS/OPA (but not PIGS) responded selectively to simpler forms of motion. These results suggest that PIGS and TOS/OPA are likely two different visual areas, with PIGS being involved in encoding higher-level ego-motion cues.

However, at least two previous studies suggested that area TOS/OPA may also contribute to ego-motion encoding in scenes. Specifically, Kamps and colleagues have shown increased response in TOS/OPA during ego-motion vs. static scene presentation (Kamps et al., 2016). Jones et al. have also shown that ego-motion (and not other types of movements) enhances TOS/OPA activity when compared to scrambled scenes (Jones et al., 2023). In contrast to these findings, our tests showed weak-to-no ego-motion related activity enhancement in area TOS/OPA.

This difference may well reflect methodological discrepancies. Specifically, in the study by Kamps et al., the static and ego-motion stimuli were presented with two different refresh rates. While in our study, the coherently and incoherently changing stimuli were refreshed with the same temporal frequency (see Methods). In the study by Jones et al., the response to scrambled scenes was used as a control condition, whereas our stimuli were more equivalent, differing only in the sequence of image presentation. Moreover, these studies used higher levels of spatial smoothing (FWHM = 5 mm), compared to the values we used here during pre-processing. Also, for understandable reasons, they limited their analysis to previously known scene-selective areas. These technical differences make it difficult to directly compare the two sets of results.

4.6. Ego-motion but not attention

Experiment 6 showed stronger scene-selective activity within PIGS when subjects were presented with coherently (compared to incoherently) changing scenes. It could be argued that coherently changing scenes attract more attention compared to incoherently changing scenes. On the face of it, this hypothesis appears to be consistent with the expected contribution of the intraparietal cortex in controlling spatial attention (Behrmann et al., 2004; Szczepanski et al., 2010). But if true, attention to scenes should also increase the level of activity within the scene-selective areas (O'craven et al., 1999; Nasr and Tootell, 2012a; Baldauf and Desimone, 2014). While here, we did not find any significant activity increases in response to coherently (vs. incoherently) changing scenes in PPA/TPA, RSC/MPA and TOS/OPA. Thus, modulation of attention, per se, could not be responsible for the enhanced activity within PIGS in response to coherently (compared to incoherently) changing scenes.

4.7. Direction-selective response within the intraparietal cortex

Motion-selective sites are expected to show at least some level of sensitivity to motion direction (Albright et al., 1984; Zimmermann et al., 2011). We did not test the sensitivity of PIGS to the direction of ego motion. However, Pitzalis et al. have shown evidence for motion direction encoding within the V6+ region (Pitzalis et al., 2020). Furthermore, Tootell et al. reported evidence for motion direction (approaching vs. withdrawing) encoding within posterior intraparietal cortex (Tootell et al., 2022). Although none of these studies showed any evidence for a new scene-selective area, they raised the possibility that PIGS may also contribute towards encoding ego-motion direction, and even higher-level cognitive concepts such as detecting an intrusion to personal space (Holt et al., 2014).

4.8. Limitations

In the past, many studies have scrutinized the response function of scene-selective areas to numerous stimulus contrasts. According to these studies, scene-selective areas can differentiate many object categories based on their low-, mid-, and/or higher-level visual features such as their natural size (Konkle and Oliva, 2012), (non-)animacy (Yue et al., 2020; Coggan and Tong, 2023), rectilinearity (Nasr et al., 2014), spatial layout (Harel et al., 2013), orientation (Nasr and Tootell, 2012b), spikiness (Coggan and Tong, 2023), location within the visual field (Levy et al., 2001), and spatial content (Bar et al., 2008). Our findings are only a first step toward characterizing PIGS in greater detail. More tests are required to reach the current (yet incomplete) knowledge about the response function of PIGS.

5. Conclusion

Neuroimaging studies of scene perception have typically focused on linking scene perception to the evoked activity within PPA/TPA, TOS/OPA and RSC/MPA. Although other scene-selective sites are detectable across the visual cortex, they are largely ignored because of their relatively small

size. Our data suggests that the future inclusion of these small sites in models of scene perception may help clarify current models of scene processing in dynamic environments.

Acknowledgements

This work was supported by NIH NEI (grants R01 EY017081 and R01 EY030434), and by the MGH/HST Athinoula A. Martinos Center for Biomedical Imaging. Crucial resources were made available by a NIH Shared Instrumentation Grant S10-RR019371. We thank Ms. Azma Mareyam for help with hardware maintenance during this study. We also thank Dr. Claudio Galletti for his helpful comments.

References

Adams WJ, Elder JH, Graf EW, Leyland J, Lugtigheid AJ, Muryy A (2016) **The southampton-york natural scenes (syns) dataset: Statistics of surface attitude** *Sci Rep* **6**:1–17

Albright TD, Desimone R, Gross CG (1984) **Columnar organization of directionally selective cells in visual area MT of the macaque** *J Neurophysiol* **51**:16–31

Baldauf D, Desimone R (2014) **Neural mechanisms of object-based attention** *Science* **344**:424–427

Bar M, Aminoff E, Schacter DL (2008) **Scenes unseen: the parahippocampal cortex intrinsically subserves contextual associations, not scenes or places per se** *J Neurosci* **28**:8539–8544

Beauchamp MS, Lee KE, Haxby JV, Martin A (2003) **FMRI responses to video and point-light displays of moving humans and manipulable objects** *J Cogn Neurosci* **15**:991–1001

Behrmann M, Geng JJ, Shomstein S (2004) **Parietal cortex and attention** *Curr Opin Neurobiol* **14**:212–217

Benson NC, Jamison KW, Arcaro MJ, Vu AT, Glasser MF, Coalson TS, Van Essen DC, Yacoub E, Ugurbil K, Winawer J (2018) **The Human Connectome Project 7 Tesla retinotopy dataset: Description and population receptive field analysis** *Journal of vision* **18**:23–23

Brainard DH (1997) **The Psychophysics Toolbox** *Spat Vis* **10**:433–436

Coggan DD, Tong F (2023) **Spikiness and animacy as potential organizing principles of human ventral visual cortex** *Cereb Cortex:bhad108*

Dale AM, Fischl B, Sereno MI (1999) **Cortical surface-based analysis. I. Segmentation and surface reconstruction** *Neuroimage* **9**:179–194

De Martino F, Zimmermann J, Muckli L, Ugurbil K, Yacoub E, Goebel R (2013) **Cortical depth dependent functional responses in humans at 7T: improved specificity with 3D GRASE** *PLoS One* **8**

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

Engel SA, Glover GH, Wandell BA (1997) **Retinotopic organization in human visual cortex and the spatial precision of functional MRI** *Cereb Cortex* **7**:181–192

Epstein R, Kanwisher N (1998) **A cortical representation of the local visual environment** *Nature* **392**:598–601

Fischl B (2012) **FreeSurfer** *Neuroimage* **62**:774–781

Fischl B, Sereno MI, Dale AM (1999) **Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system** *Neuroimage* **9**:195–207

Fischl B *et al.* (2002) **Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain** *Neuron* **33**:341–355

Friston KJ, Holmes AP, Price CJ, Buchel C, Worsley KJ (1999) **Multisubject fMRI studies and conjunction analyses** *Neuroimage* **10**:385–396

Greve DN, Fischl B (2009) **Accurate and robust brain image alignment using boundary-based registration** *Neuroimage* **48**:63–72

Grill-Spector K (2003) **The neural basis of object perception** *Curr Opin Neurobiol* **13**:159–166

Hacialihafiz DK, Bartels A (2015) **Motion responses in scene-selective regions** *Neuroimage* **118**:438–444

Harel A, Kravitz DJ, Baker CI (2013) **Deconstructing visual scenes in cortex: gradients of object and spatial layout information** *Cereb Cortex* **23**:947–957

Holt DJ, Cassidy BS, Yue X, Rauch SL, Boeke EA, Nasr S, Tootell RB, Coombs G (2014) **Neural correlates of personal space intrusion** *J Neurosci* **34**:4123–4134

Jastorff J, Orban GA (2009) **Human functional magnetic resonance imaging reveals separation and integration of shape and motion cues in biological motion processing** *J Neurosci* **29**:7315–7329

Jones CM, Byland J, Dilks DD (2023) **The occipital place area represents visual information about walking, not crawling** *Cereb Cortex:bhad055*

Kamps FS, Lall V, Dilks DD (2016) **The occipital place area represents first-person perspective motion information through scenes** *Cortex* **83**:17–26

Kanwisher N, McDermott J, Chun MM (1997) **The fusiform face area: a module in human extrastriate cortex specialized for face perception** *J Neurosci* **17**:4302–4311

Kennedy B, Bex P, Hunter D, Nasr S (2023) **Two fine-scale channels for encoding motion and stereopsis within the human magnocellular stream** *Prog Neurobiol:102374*

Konen CS, Kastner S (2008) **Two hierarchically organized neural systems for object information in human visual cortex** *Nat Neurosci* **11**:224–231

Koopmans PJ, Barth M, Norris DG (2010) **Layer-specific BOLD activation in human V1** *Hum Brain Mapp* **31**:1297–1304

Kornblith S, Cheng X, Ohayon S, Tsao DY (2013) **A network for scene processing in the macaque temporal lobe** *Neuron* **79**:766–781

Lescroart MD, Gallant JL (2019) **Human scene-selective areas represent 3D configurations of surfaces** *Neuron* **101**:178–192

Levy I, Hasson U, Avidan G, Hendler T, Malach R (2001) **Center–periphery organization of human object areas** *Nat Neurosci* **4**

Li X, Zhu Q, Vanduffel W (2022) **Submillimeter fMRI reveals an extensive, fine-grained and functionally-relevant scene-processing network in monkeys** *Prog Neurobiol* **211**

Maguire E (2001) **The retrosplenial contribution to human navigation: a review of lesion and neuroimaging findings** *Scand J Psychol* **42**:225–238

Nasr S, Tootell RB (2012) **Role of fusiform and anterior temporal cortical areas in facial recognition** *Neuroimage* **63**:1743–1753

Nasr S, Tootell RB (2012) **A cardinal orientation bias in scene-selective visual cortex** *J Neurosci* **32**:14921–14926

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

Nasr S, Polimeni JR, Tootell RB (2016) **Interdigitated Color- and Disparity-Selective Columns within Human Visual Cortical Areas V2 and V3** *J Neurosci* **36**:1841–1857

Nasr S, Liu N, Devaney KJ, Yue X, Rajimehr R, Ungerleider LG, Tootell RB (2011) **Scene-selective cortical regions in human and nonhuman primates** *J Neurosci* **31**:13771–13785

O'craven KM, Downing PE, Kanwisher N (1999) **fMRI evidence for objects as the units of attentional selection** *Nature* **401**:584–587

Park J, Josephs E, Konkle T (2022) **Ramp-shaped neural tuning supports graded population-level representation of the object-to-scene continuum** *Sci Rep* **12**

Pelli DG (1997) **The VideoToolbox software for visual psychophysics: transforming numbers into movies** *Spat Vis* **10**:437–442

Pelphrey KA, Morris JP, Michelich CR, Allison T, McCarthy G (2005) **Functional anatomy of biological motion perception in posterior temporal cortex: an fMRI study of eye, mouth and hand movements** *Cereb Cortex* **15**:1866–1876

Pitzalis S, Fattori P, Galletti C (2015) **The human cortical areas V6 and V6A** *Vis Neurosci*

Pitzalis S, Sereno MI, Committeri G, Fattori P, Galati G, Patria F, Galletti C (2010) **Human V6: the medial motion area** *Cereb Cortex* **20**:411–424

Pitzalis S, Serra C, Sulpizio V, Committeri G, de Pasquale F, Fattori P, Galletti C, Sepe R, Galati G (2020) **Neural bases of self-and object-motion in a naturalistic vision** *Hum Brain Mapp* **41**:1084–1111

Polimeni JR, Fischl B, Greve DN, Wald LL (2010) **Laminar analysis of 7T BOLD using an imposed spatial activation pattern in human V1** *Neuroimage* **52**:1334–1346

Polimeni JR, Bhat H, Witzel T, Benner T, Feiweier T, Inati SJ, Renvall V, Heberlein K, Wald LL (2015) **Reducing sensitivity losses due to respiration and motion in accelerated echo planar imaging by reordering the autocalibration data acquisition** *Magn Reson Med*

Puce A, Perrett D (2003) **Electrophysiology and brain imaging of biological motion** *Philos Trans R Soc Lond B Biol Sci* **358**:435–445

Puce A, Allison T, Bentin S, Gore JC, McCarthy G (1998) **Temporal cortex activation in humans viewing eye and mouth movements** *J Neurosci* **18**:2188–2199

Rajimehr R, Young JC, Tootell RB (2009) **An anterior temporal face patch in human cortex, predicted by macaque maps** *Proc Natl Acad Sci U S A* **106**:1995–2000

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

Silson EH, Groen II, Kravitz DJ, Baker CI (2016) **Evaluating the correspondence between face-, scene-, and object-selectivity and retinotopic organization within lateral occipitotemporal cortex** *Journal of vision* **16**:14–14

Silver MA, Ress D, Heeger DJ (2005) **Topographic maps of visual spatial attention in human parietal cortex** *J Neurophysiol* **94**:1358–1371

Steel A, Billings MM, Silson EH, Robertson CE (2021) **A network linking scene perception and spatial memory systems in posterior cerebral cortex** *Nature communications* **12**

Steel A, Garcia BD, Goyal K, Mynick A, Robertson CE (2023) **Scene perception and visuospatial memory converge at the anterior edge of visually responsive cortex** *J Neurosci* **43**:5723–5737

Sulpizio V, Galati G, Fattori P, Galletti C, Pitzalis S (2020) **A common neural substrate for processing scenes and egomotion-compatible visual motion** *Brain structure and Function* **225**:2091–2110

Swisher JD, Halko MA, Merabet LB, McMains SA, Somers DC (2007) **Visual topography of human intraparietal sulcus** *J Neurosci* **27**:5326–5337

Szczepanski SM, Konen CS, Kastner S (2010) **Mechanisms of spatial attention control in frontal and parietal cortex** *J Neurosci* **30**:148–160

Tootell RB, Nasiriavanaki Z, Babadi B, Greve DN, Nasr S, Holt DJ (2022) **Interdigitated Columnar Representation of Personal Space and Visual Space in Human Parietal Cortex** *J Neurosci* **42**:9011–9029

Tootell RB, Mendola JD, Hadjikhani NK, Ledden PJ, Liu AK, Reppas JB, Sereno MI, Dale AM (1997) **Functional analysis of V3A and related areas in human visual cortex** *J Neurosci* **17**:7060–7078

Troiani V, Stigliani A, Smith ME, Epstein RA (2014) **Multiple object properties drive scene-selective regions** *Cereb Cortex* **24**:883–897

Tsao DY, Moeller S, Freiwald WA (2008) **Comparing face patch systems in macaques and humans** *Proc Natl Acad Sci U S A* **105**:19514–19519

Yacoub E, Shmuel A, Logothetis N, Ugurbil K (2007) **Robust detection of ocular dominance columns in humans using Hahn Spin Echo BOLD functional MRI at 7 Tesla** *Neuroimage* **37**:1161–1177

Yue X, Robert S, Ungerleider LG (2020) **Curvature processing in human visual cortical areas** *Neuroimage* **222**

Yue X, Cassidy BS, Devaney KJ, Holt DJ, Tootell RB (2011) **Lower-level stimulus features strongly influence responses in the fusiform face area** *Cereb Cortex* **21**:35–47

Zimmermann J, Goebel R, De Martino F, van de Moortele PF, Feinberg D, Adriany G, Chaimow D, Shmuel A, Ugurbil K, Yacoub E (2011) **Mapping the organization of axis of motion selective features in human area MT using high-field fMRI** *PLoS One* **6**

Article and author information

Bryan Kennedy
Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, MA, United States

Sarala N. Malladi
Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, MA, United States

Roger B. H. Tootell
Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, MA, United States, Department of Radiology, Harvard Medical School, Boston, MA, United States

Shahin Nasr
Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, MA, United States, Department of Radiology, Harvard Medical School, Boston, MA, United States
For correspondence: shahin.nasr@mgh.harvard.edu

Copyright

Editors

Reviewing Editor
Margaret Schlichting
University of Toronto, Toronto, Canada

Senior Editor
Tirin Moore
Howard Hughes Medical Institute, Stanford University, Stanford, United States of America

Reviewer #2 (Public Review):

Summary

The authors report an extensive series of neuroimaging experiments (at both 3T and 7T) to provide evidence for a scene-selective visual area in human posterior parietal cortex (PIGS) that is distinct from the main three (parahippocampal place area, PPA; occipital place area, OPA; medial place area, MPA) typically reported in the literature. Further, they argue that in

comparison with the other three, this region may specifically be involved in representing ego-motion in natural contexts. The characterization of this scene-selective region provides a useful reference point for studies of scene processing in humans.

Strengths

One of the major strengths of the work is the extensive series of experiments reported, showing clear reproducibility of the main finding and providing functional insight into the region studied. The results are clearly presented and convincing with careful comparison to retinotopic and scene-selective regions described in prior studies.

Weaknesses

While the results are strong and clear, the claim in the title ("A previously undescribed scene-selective site is the key to encoding ego-motion in naturalistic environments") is not fully supported. The results show that this scene-selective region is sensitive to visual cues that reflect ego-motion but not that it is "key" to encoding ego-motion. Further, there are many differences between the two types of stimuli used to test ego-motion and greater characterization of this scene-selective region will be needed to confirm this conclusion.

https://doi.org/10.7554/eLife.91601.2.sa1

Reviewer #3 (Public Review):

Summary:
The authors report a scene-selective areas in the posterior intraparietal gyrus (PIGS). This area lies outside the classical three scene-selective regions (PPA/TPA, RSC/MPA, TOS/OPA), and is selective for ego motion.

Strengths:
The authors firmly establish the location and selectivity of the new area through a series of well-crafted controlled experiments. They show that the area can be missed with too much smoothing, thus providing a case for why it has not been previously described. They show that it appears in much the same location in different subjects, with different magnetic field strengths, and with different stimulus sets. Finally, they show that it is selective for ego motion - defined as series of sequential photographs of an egocentric trajectory along a path. They further clarify that the area is not generically motion selective by showing that it does not respond to biological motion without an egomotion component to it. All statistics are standard and sound; the evidence presented is strong.

Weaknesses:
There are a few weaknesses in this work. If pressed, I might say that the stimuli depicting ego motion do not, strictly speaking, depict motion, but only apparent motion between 2s apart photographs. However, this choice was made to equate frame rates and motion contrast between the 'ego motion' and a control condition, which is a useful and valid approach to the problem.

This is a very strong paper.

https://doi.org/10.7554/eLife.91601.2.sa0

Author Response

The following is the authors' response to the original reviews.

> ***Reviewer 1:***
>
> *Comment 1.1: The distinction of PIGS from nearby OPA, which has also been implied in navigation and ego-motion, is not as clear as it could be.*

Response1.1: The main "functional" distinction between TOS/OPA and PIGS is that TOS/OPA responds preferentially to moving vs. stationary stimuli (even concentric rings), likely due to its overlap with the retinotopic motion-selective visual area V3A, for which this is a defining functional property (e.g. Tootell et al., 1997, J Neurosci). In comparison, PIGS does not show such a motion-selectivity. Instead, PIGS responds preferentially to more complex forms of motion within scenes.

Moreover, PIGS and TOS/OPA are located in differently relative to the retinotopic visual areas. Briefly, PIGS is located adjacent to areas IPS3-4 while TOS/OPA overlaps with areas V3A/B and IPS0 (V7). This point is now highlighted in the new experiment 3b and the new Figure 6. In this revision, we also tried to better highlight these point in sections 4.3, 4.4 and 4.5. (see also the response to the first comment from Reviewer #2).

> ***Reviewer 2:***
>
> *Comment 2.1: First, the scene-selective region identified appears to overlap with regions that have previously been identified in terms of their retinotopic properties. In particular, it is unclear whether this region overlaps with V7/IPS0 and/or IPS1. This is particularly important since prior work has shown that OPA often overlaps with v7/IPS0 (Silson et al, 2016, Journal of Vision). The findings would be much stronger if the authors could show how the location of PIGS relates to retinotopic areas (other than V6, which they do currently consider). I wonder if the authors have retinotopic mapping data for any of the participants included in this study. If not, the authors could always show atlas-based definitions of these areas (e.g. Wang et al, 2015, Cerebral Cortex).*

Response 2.1: We thank the reviewers for reminding us to more clearly delineate this issue of possible overlap, including the information provided by Silson et al, 2016. The issue of possible overlap between area TOS/OPA and the retinotopic visual areas, both in humans and non-human primates, was also clarified by our team in 2011 (Nasr et al., 2011). As you can see in Figure 6 (newly generated), and consistent with those previous studies, TOS/OPA overlaps with visual areas V3A/B and V7. Whereas PIGS is located more dorsally close to IPS3-4. As shown here, there is no overlap between PIGS and TOS/OPA and there is no overlap between PIGS and areas V3A/B and V7.

To more directly address the reviewer's concern, in this revision, we have added a new experiment (Experiment 3b) in which we have shown the relative position of PIGS and the retinotopic areas in two individual subjects (Figure 6). All the relevant points are also discussed in section 4.3.

> *Comment 2.2: Second, recent studies have reported a region anterior to OPA that seems to be involved in scene memory (Steel et al, 2021, Nature Communications; Steel et al, 2023, The Journal of Neuroscience; Steel et al, 2023, biorXiv). Is this region distinct from PIGS? Based on the figures in those papers, the scene memory-related region is inferior to V7/IPS0, so characterizing the location of PIGS to V7/IPS0 as suggested above would be very helpful here as well. If PIGS overlaps with either of V7/IPS0 or the scene memory-related area described by Steel and colleagues, then arguably it is not a newly defined region (although the characterization provided here still provides new information).*

Response 2.2: The lateral-place memory area (LPMA) is located on the lateral brain surface, anterior relative to the IPS (see Figure 1 from Steel et al., 2021 and Figure 3 from Steel et al.,

2023). In contrast, PIGS is located on the posterior brain surface, also posterior relative to the IPS. In other words, they are located on two different sides of a major brain sulcus. In this revision we have clarified this point, including the citations by Steel and colleagues in section 4.3.

> *Comments 2.3: Another reason that it would be helpful to relate PIGS to this scene memory area is that this scene memory area has been shown to have activity related to the amount of visuospatial context (Steel et al, 2023, The Journal of Neuroscience). The conditions used to show the sensitivity of PIGS to ego-motion also differ in the visuospatial context that can be accessed from the stimuli. Even if PIGS appears distinct from the scene memory area, the degree of visuospatial context is an alternative account of what might be represented in PIGS.*

Response 2.3: The reviewer raises an interesting point. One minor confusion is that we may be inadvertently referring to two slightly different types of "visuospatial context". Specifically, the stimuli used in the ego-motion experiment here (i.e. coherently vs. incoherently changing scenes) represent the same scenes, and the only difference between the two conditions is the sequence of images across the experimental blocks. In that sense, the two experimental conditions may be considered to have the same visuospatial "context". However, it could be also argued that the coherently changing scenes provide more information about the environmental layout. In that case, considering the previous reports that PPA/TPA and RSC/MPA may also be involved in layout encoding (Epstein and Kanwisher 1998; Wolbers et al. 2011), we expected to see more activity within those regions in response to coherently compared incoherently changing scenes. These issues are now more explicitly discussed in the revised article (section 4.6).

> ***Reviewer 3:***
>
> *Comment 3.1: There are few weaknesses in this work. If pressed, I might say that the stimuli depicting ego-motion do not, strictly speaking, depict motion, but only apparent motion between 2s apart photographs. However, this choice was made to equate frame rates and motion contrast between the 'ego-motion' and a control condition, which is a useful and valid approach to the problem. Some choices for visualization of the results might be made differently; for example, outlines of the regions might be shown in more plots for easier comparison of activation locations, but this is a minor issue.*

Response 3.1: We thank the reviewer for these constructive suggestions, and we agree with their comment that the ego-motion stimuli are not smooth, even though they were refreshed every 100 ms. However, the stimuli were nevertheless coherent enough to activate areas V6 and MT, two major areas known to respond preferentially to coherent compared to incoherent motion.

> ***Reviewer #1 (Recommendations For The Authors):***
>
> *I enjoyed reading this article. I have a few suggestions for improvement:*
>
> *(1) Delineation from OPA: The OPA has been described in quite similar terms as PIGS, with its involvement in ego-motion (e.g., crawling, walking) and navigation in general (e.g., Dilks' recent work; Bonner and Epstein). The authors address the distinction in section 4.4. Unlike Kamps et al. (2016) and Jones et al. (2023), the authors found weak or no evidence for ego-motion in OPA. They explain this discrepancy with differences in refresh rates and different levels of spatial smoothing of the fMRI data. It is not clear why these fairly small methodological differences would lead to different findings of ego-motion in the OPA. Arguably, the OPA is the closest of the "established" scene areas to*

> *PIGS, both in anatomical location and in function. I would therefore appreciate a more detailed discussion of the differences between these two areas.*

Response: Jones et al. have also shown that ego-motion TOS/OPA activity when compared to scrambled scenes. This is fundamentally different than what we have shown here, which coherently vs. incoherently changing scenes (i.e. not a small difference). Also, Kamps et al. used static scenes as a control which, considering TOS/OPA motion-selectivity, have a large impact on TOS/OPA response.

> *(2) Random effects analysis: The authors mention using a "random effects analysis" for several of their experiments. I would ask them to provide more details on what statistical models were used here. Were they purely random-effects models or actually mixed-effects models? What were the factors that entered into the analysis? Providing more detail would make the analysis techniques more transparent.*

Response: This point is now clarified in the Methods section.

> *(3) Data and code availability: The authors write that data and code "are ready to be shared upon request." (section 2.5) In the spirit of transparency and openness, I strongly encourage the authors to make the data publicly available, e.g., on OSF or OpenNeuro. In particular, having probabilistic maps of PIGS available will allow other researchers to include PIGS in their analysis pipelines, making the current work more impactful.*

Response: We have made the probabilistic labels available to the public. This point is now highlighted in section 2.5.

> *(4) Minor comments on the writing that caught my eye while reading the article:*
>
> - *Line 27: "in the human brain".*

Response: Done.

> *-Line 30: I don't agree with the characterization of the previous model of scene perception as "simplistic." Adding one additional ROI makes it no less simplistic. Perhaps the authors can rephrase to make this slightly less antagonistic?*

Response: Done.

> - *Line 71: it is not clear why NHPs are relevant here.*

Response: We decided to keep the text intact.

> - *Line 138" "were randomized".*

Response: Done.

> - *Line 152: "consisting".*

Response: Done.

> - *Line 155: "sets" (plural).*

Response: Done.

- *Lines 253-255: Why were the 3T spatially smoothed but not the 7T data? This seems odd.*

Response: We kept the text intact.

- *Line 481: "we found strong motion selectivity" (remove "a").*

Response: Done.

- *Line 564: a word is missing, probably: "a stronger effect of ego-motion".*

Response: Done.

- *Line 591: "controlling spatial attention" (remove "the").*

Response: Done.

- *Line 591 and 594: Both sentences start with "However". I think the first of these should not because it is setting up the contrast for the second sentence.*

Response: Done.

- *Line 607: "higher-level" (hyphen).*

Response: Done.

- *Throughout the manuscript: adverbial phrases such as "(in)coherently changing" or "probabilistically localized" do not get a hyphen.*

Response: Done.

Reviewer #2 (Recommendations For The Authors):

The authors state that "All data, codes and stimuli are ready to be shared upon request". Ideally, these materials should be deposited in appropriate repositories (e.g. OpenMRI, GitHub) and not require readers to contact the authors to obtain such materials.

Other Comments:

(a) The title ("A previously undescribed scene-selective site is the key to encoding ego-motion in natural environments") is potentially misleading - the work was not conducted in a natural environment. At best, you could say they are 'naturalistic stimuli'. Also, in what sense is PIGS "key" to encoding ego-motion - the study just shows sensitivity to this factor.

Response: We changed the title to “naturalistic environments”.

(b) Figure 1 - I'm not sure what point the authors are trying to make with Figure 1. The comparison is between a highly smoothed, group fixed-effects analysis and a less-smoothed individual subject analysis. The differences between the two could reflect

> *group vs. individual, highly-smoothed (5 mm) versus less-smoothed (2 mm), or differences in thresholding. If the thresholding were lower for the group analysis, it would probably start to look more similar to the individual subject. As it stands, this figure isn't particularly informative, it seems redundant with Figure 2, and Figure 1A is not even referenced in the main text. Further, fixed effects analyses are relatively uncommon in the recent literature, so their inclusion is unusual.*

Response: Figure 1A is a replication of the data/method used in Nasr et al., 2011 and it will help the readers see the difference between the "traditional" scene-selectivity maps generated based on group-averaging" vs. data from individual subjects. In this case, we decided not to change the Figure.

> *(c) Figure 3 - why are the two sets of maps shown at different thresholds? For 3B given the larger sample size, it is expected that the extent of the significant activations will increase. Currently the higher threshold for 3B and the smaller range for 3A is making the sets of maps look more comparable.*

Response: As the reviewer noticed, the number of subjects is larger in Figure 3B compared to 3A. The main point of this figure is to show that the PIGS activity center does not vary across populations. Considering this point, we decided not to change this figure.

> *(d) Figure 10 - why is the threshold lower than used for other figures? It would be helpful if there was consistent thresholding across figures.*

Response: Experiment 6 and Experiment 1 are based on different stimuli (see Methods). Also, among those subjects who participated in Experiment 1, two subjects did not participate in Experiment 6. These points are already highlighted in the text.

> *(e) Figures - how about the AFNI approach of thresholding and showing sub-threshold data at the same time? (Taylor et al, 2023, Neuroimage).*

Response: We highly appreciate the methodology suggested by Taylor and colleagues. However, our main point here is to show the center of PIGS activity. In this condition, showing an unthresholded activity map doesn't have any advantage over the current maps. Considering these points, we decided not to change the figures.

> *(f) Coherent versus incoherent scenes - there are many differences between the coherent and incoherent scenes. Arguing that it must be ego-motion seems a little premature without further investigation. Activity anterior to OPA has been associated with the construction of an internal representation of a spatial environment (Steel et al., 2023, The Journal of Neuroscience). Could it be that this is the key effect, not really the ego-motion?*

Response: In this revision, we discussed the study by Steel et al., 2021 and 2023 in section 4.3.

> ***Reviewer #3 (Recommendations For The Authors):***
>
> *Overall, I think this is already an excellent contribution. The suggestions I have are minor and may help with the clarity of the results.*
>
> *(1) My main request of the authors would be to provide more points of reference in some of the figures with cortical maps. In many cases, the authors use arrows to point to the locations of activations of interest. However, the arrows in adjacent figures are often not placed in exactly the same places on maps that are meant to be compared. It would very much help the viewer to compare activations if the arrows pointing to activations or*

> *regions of interest were placed in identical locations for the same brains appearing in different sub-panels (e.g. in panels A and B of Figure 1). The underlying folds of the cortical surface provide some points of reference, but these are often occluded to different extents by data in figures that are meant to be compared.*

Response: To address the reviewer's concern, we regenerated Figure 8 (Figure 7 in the previous submission) and we tried to put arrowheads in identical locations, as much as possible. Especially for PIGS, this point was also considered in Figures 2 and 3.

> *(2) Outlines (such as those in Figure 5) are also very useful, and I would encourage broader use of them in other figures (e.g. Figures 7, 10, and 12). Figures 10 and 12 are on the fsaverage surface, so the same outlines could be used for them as for Figure 5.*
>
> *To be clear, it's possible to apprehend the results with the figures as they are, but I think a few small changes could help a lot.*

Response: In this revision, we added outlines to Figures 11 and 13 (Figure 10 and 12 in the previous submission). We did not add the outline to Figure 8 because it made it hard to see PIGS. Rather we used arrows (see the previous comment).

> *Other minor points:*
>
> *In the method for Experiment 4, the authors write: "Other details of the experiment were similar to those in Experiment 1.". Similar or the same? The authors should clarify this statement, e.g. "the number of images per block, the number of blocks, the number of runs were the same as Experiment 1" - with any differences noted.*

Response: This point is now addressed in the Methods section.

> *In Figure 8, it would be better to have the panel labels (A, B, C, D) in the upper left of each panel rather than the lower left.*

Response: We tried to keep the panels arrangement consistent across the figures. That is why letters are positioned like this.

> *A final gentle suggestion: pycortex (http://github.com/gallantlab/pycortex) provides a means to visualize the flattened fsaveage surface with outlines for localized regions of interest and overlaid lines for major sulci. Though it is by no means necessary for publication, It would be lovely to see these results on that surface, which is freely available and downloadable via a pycortex command (surface here: https://figshare.com/articles/dataset/fsaverage_subject_for_pycortex/9916166)*

Response: We thank the reviewer for bringing pycortex to our attention. We will consider using it in our future studies.