Jordan Pincus March 26, 2019

Differential Development of Scene Processing Regions in the Human Brain

By

Jordan Elyse Pincus

Daniel Dilks
Adviser

Neuroscience and Behavioral Biology

Daniel Dilks

Adviser

Leah Roesch

Committee Member

Philippe Rochat

Committee Member

2019

Differential Development of Scene Processing Regions in the Human Brain

By

Jordan Elyse Pincus

Daniel Dilks

Adviser

An abstract of
a thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Science with Honors

Neuroscience and Behavioral Biology

2019

Abstract

Differential Development of Scene Processing Regions in the Human Brain

By Jordan Pincus

Humans can effortlessly recognize the kind of place (or “scene”) they are in, as well as navigate through that scene. Losing these abilities – as a result of stroke, disease, or developmental disorder – has devastating effects on quality of life. Evidence suggests that two specialized brain regions support these remarkable cognitive processes: the parahippocampal place area (PPA), supporting “scene categorization” (i.e., recognizing a scene as a kitchen or a beach), and the occipital place area (OPA), supporting “visually-guided navigation” (i.e., navigating through the immediately visible scene). While the functions of these two systems are increasingly clear, almost nothing is known about how these systems develop over childhood. Some behavioral work suggests that the navigation ability develops slower than the categorization ability. In the current project, we look for neural evidence of such differential development using functional magnetic resonance imaging (fMRI) in children ages 5-8. We predict that OPA will develop later than PPA. We investigate the development of these regions by measuring their responses to i) video clips of first-person perspective motion through scenes (“Dynamic Scenes”) and ii) still frames taken from these same movies (“Static Scenes”). We found that OPA’s sensitivity to scene motion information, a critical cue for its role in navigation, increases from ages 5 to 8, while no developmental changes were observed in PPA across this same age range. These findings support our prediction that the navigation system (including OPA) develops independently and differentially from the categorization system (including PPA).

Differential Development of Scene Processing Regions in the Human Brain

By

Jordan Elyse Pincus

Daniel Dilks

Adviser

A thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Science with Honors

Neuroscience and Behavioral Biology

2019

Acknowledgements

I would like to thank the Facility for Education and Research in Neuroscience (FERN) Imaging Center in the Department of Psychology, Emory University, Atlanta, GA. I would also like to thank Dr. Daniel Dilks, Dr. Leah Roesch, and Dr. Philippe Rochat for being on my thesis committee this year and for their helpful feedback. Lastly, I would like to thank Dr. Daniel Dilks and Frederik Kamps for their endless support, guidance, and mentorship.

Table of Contents

1. Introduction

Humans can effortlessly recognize not only the kind of place ("or scene") they are in, but also how to navigate through that scene. Being able to understand where we are and to walk around places without disorientation and confusion is critical to our survival. Losing these abilities – as a result of stroke, disease, or developmental disorder – has devastating effects on quality of life. Therefore, it is not surprising that there are neural regions in humans dedicated to scene processing, including the parahippocampal place area (PPA) (Epstein & Kanwisher, 1998), and the occipital place area (OPA) (Dilks et al., 2013), otherwise known as the transverse occipital sulcus (Grill-Spector, 2003). Extensive work has established the selective involvement of these regions in scene processing, and recent evidence has elucidated the more precise functions of these regions within scene processing. However, almost nothing is known about how these regions develop. Thus, the present project addresses that question. Here, I review the evidence that OPA and PPA are selectively involved in scene processing, and hypotheses about the precise functions of these regions, setting the stage for the current project investigating the development of these systems.

1.1 Scene-selective Cortex

In 1998, Epstein and Kanwisher published the landmark discovery of a cortical region responding approximately two to four times stronger to images of scenes than to images of objects or faces. The discovery of this region, which they named the parahippocampal place area (PPA), launched the study of scene perception. The initial question that this work sought to address was perhaps the most fundamental: what exactly is a "scene"? Beginning to address this question, Epstein and Kanwisher (1998) found that a critical feature for PPA activation is not the number of objects in the space, but rather the layout of the local space itself; PPA responded equally to images of furnished and unfurnished rooms, but more to images of rooms in which the walls, floors, and ceilings formed a coherent space than to images in which these same walls, floors, and ceilings were rearranged, disrupting the coherent spatial arrangement of the place. Further work in computer vision and robotics supported this “spatial layout” of a scene (Oliva and Torralba, 2001, 2002; Thrun, 2002). However, later studies revealed that while PPA does not respond to objects per se, PPA is nevertheless also sensitive to scene “content” information (i.e., the objects, textures, and colors within the spatial layout) (Bar and Aminoff, 2003; Mullaly and Maguire, 2011; Konkle and Oliva, 2012; Troiani et al., 2014; Kamps et al., 2016a). Thus, the current view is that a “scene” can be thought of in terms of these two descriptors, that is, a particular spatial layout with particular scene content (Oliva and Torralba, 2001, 2002).

Importantly, PPA was discovered using fMRI, and fMRI data are only correlational. Later studies of patients with damage to the area in and around PPA revealed that these patients present with deficits in identifying scenes or landmarks (Aguirre and D’Esposito, 1999; Mendez and Cherrier, 2003) and in knowing where they are (Habib and Sirigu, 1987; Epstein et al., 2001). This work established critical, causal evidence for the selective involvement of PPA in scene processing.

It was later discovered that PPA is not the only scene-selective region. In particular, many studies also observed activation to scenes in the transverse occipital sulcus. Initially this region was thought to reflect low-level visual differences between scene and object stimuli, rather than a proper scene-selective region, due to its proximity to early visual cortex.

However, two studies using transcranial magnetic stimulation (TMS), a technique for safely and temporarily impairing activation in human cortex, found that TMS over the transverse occipital sulcus impaired participants' ability to discriminate scenes, but not faces or objects (Dilks et al., 2013; Ganaden et al., 2013). This work therefore established that the transverse occipital sulcus was causally involved in scene perception and processing, and the region was consequently renamed the occipital place area (OPA) (Dilks et al., 2013).

1.2 Two Systems for Scene Processing

It is not only important to understand that these scene-selective regions exist, but also to discern the more precise functions of these regions within scene processing, and thus how these regions allow us to recognize and navigate through our surroundings. To that end, considerable evidence suggests that each of these scene processing regions has a distinct role in scene processing, and further, that the roles of these regions can be divided into two scene processing systems: a navigation system (involving OPA) and a categorization system (involving PPA).

Many studies support the role of OPA in navigation (Epstein, 2008; Dilks et al., 2011; Persichetti and Dilks, 2016). First, two fMRI adaptation studies were conducted to determine whether this region encodes navigationally relevant information, particularly sense (left-right) and egocentric distance (proximal-distal) information. Sense and egocentric distance information are essential for navigation (Loomis, 2008; Wang, 2000; Fajen & Warren, 2003). For example, manipulating sense information, by flipping an image, changes the navigability of a scene (i.e., changing how one can walk to the left versus the right). Further, egocentric distance information is critical to navigation, because it reflects the changes in distance to obstacles and boundaries, as one navigates through a scene. Evidence from these experiments reveals that OPA is sensitive to navigationally-relevant information, while PPA is not (Dilks et al., 2011; Persichetti and Dilks, 2016). That is, OPA "saw" two mirrored images as different from each other, while PPA "saw" them as the same (Dilks et al., 2011). A similar adaptation study found that the OPA distinguished images of varying distances (i.e., the same beach, but from a close or a proximal perspective), while the PPA was insensitive to such distance information (Persichetti and Dilks, 2016).

Notably, this navigation system, involving OPA, can be distinguished from other systems involved in navigation. Particularly, OPA is responsible for navigating through the immediately visible scene (what we refer to as "visually-guided navigation"). Building on the evidence above, another study tested the hypothesis that OPA is involved in navigation, but PPA is not. The study assessed OPA's sensitivity to first-person perspective motion, that is, the exact visual information perceived during visually-guided navigation (Gibson, 1950; Kamps et al., 2016a). Indeed, this study found that OPA responded more to such first-person perspective motion, whereas PPA did not. This finding further confirms the role of OPA in visually-guided navigation. Further, this study supports the division of labor between these systems for categorization and visually-guided navigation, with OPA being involved in navigation and PPA not. Further supporting OPA's role in navigation, another study reveals that OPA is sensitive to the local elements of a scene, an important cue for avoiding obstacles as one navigates through the immediately visible scene, while PPA is not (Kamps et al., 2016a).

As mentioned above, while OPA is sensitive to navigationally-relevant information, PPA is not, challenging its role in navigation (Dilks et al., 2011; Persichetti and Dilks, 2016). Instead, PPA encodes information critical for scene categorization. For example, PPA discriminates between open and closed spatial layouts (i.e., a desert vs a forest) (Kravitz et al., 2011; Park et al., 2011), responds greater to intact spatial layouts than to ones where walls, floors, and ceiling have been fragmented and rearranged (Epstein and Kanwisher, 1998; Kamps et al., 2016a), and is sensitive to relative changes in length and angle of scenes (Dillon et al., 2018). Computer vision studies suggest that such spatial layout information is a critical feature for determining the category to which a scene belongs (Oliva and Torralba, 2001, 2002). Therefore, it is proposed that PPA plays a role in categorization (i.e. recognizing a scene as a kitchen or a beach).

The type of information encoded by PPA and OPA suggests their roles in categorization and visually-guided navigation, respectively. A recent fMRI study, however, looked directly at categorization and navigation tasks providing stronger evidence for these two dissociable neural systems. fMRI data were collected while participants performed a navigational task (deciding which door to exit through following a path on the ground) and a categorization task (deciding if a room is a bedroom, kitchen, or living room) (Persichetti & Dilks, 2018). Supporting their proposed roles, PPA responded significantly more during the categorization task than during the navigation task, while OPA showed the opposite pattern of results.

1.3 Development of Scene Processing Abilities

While abundant studies have explored the nature of the scene processing system (including OPA and PPA) in adulthood, a fundamental question about this system remains unexplored: How does this system develop across the lifespan? The idea of independent systems for navigation (including OPA) and categorization (including PPA) suggests a developmental hypothesis. In particular, if the two systems are functionally independent, then they may likewise develop differentially across the lifespan.

One clue for the idea of differential development of these two systems comes from the classic division of labor in the visual system between the dorsal stream (Action System), which supports visually-guided action (i.e., recognizing how one would reach out to grab a cup), and the ventral stream (Perception System), which supports visual perception (i.e., recognizing that an object is a cup) (Ungerleider and Mishkin, 1982; Milner & Goodale, 1995). In a behavioral study of the development of these two systems, Dilks et al. (2008) found that the Action System develops later than the Perception System. In particular, 4 year olds performed disproportionately worse on a task testing visually-guided reaching than a similar task testing visual perception, whereas 8 year olds were relatively matched on the two tasks. Given that OPA is found in the dorsal stream, whereas PPA is found in the ventral stream, we predict that in scene-selective cortex, OPA will be slower to develop than PPA.

While substantial developmental work has investigated object processing, almost nothing is known about the development of scene processing. Intriguingly, however, some evidence, regarding the development of locomotion (i.e., the ability to walk around obstacles), suggests that children under age 8 make significantly more errors when attempting to avoid obstacles than older children and adults (Pryde et al., 1997; Berard and Vallis, 2006). Another study looks at the use of optic flow in locomotion (i.e., the pattern of relative motion between an observer and the objects, surfaces, and edges in a scene) (Higgins et al., 1996), which reflects the first-perspective motion information used by OPA in navigating through the immediately visible scene (Kamps et al., 2016b). Higgins et al. (1996) suggest that the later development of navigation is due to locomotor experience. That is, with experience, walkers are more responsive to specific aspects of stimuli relevant to navigation, such as peripheral optic flow. Overall these data suggest that navigation (putatively in the dorsal stream) undergoes prolonged development (Pryde et al., 1998; Berard and Vallis, 2006; Higgins et al., 1996; Vasudevan et al., 2011). In contrast, categorization (putatively in the ventral stream) develops early (Weigelt et al., 2014; Carey and Diamond, 1977; Golarai et al., 2007). Weigelt et al. (2014) found that the ability to perceptually discriminate scenes from one another improved around age 5.

Recent behavioral work in our lab has shown that this prediction is true. That is, scene categorization abilities mature earlier in development than scene navigation abilities. However, complementary neural evidence is lacking. Such neural evidence would strengthen these claims by showing directly that the dorsal stream scene region (OPA) is later developing than the ventral stream scene region (PPA).

1.4 Present Study

Thus, here we looked for neural evidence of such development. We assessed development in two key ways. First, we studied how selective each region is for scenes (what we refer to as “scene domain selectivity”). Scene domain selectivity is a region’s response to scenes over faces or objects. Second, we studied how selective each region is for scene motion information, particularly first-person perspective motion (what we refer to as “scene motion selectivity”). In adulthood, OPA, but not PPA, is selective for scene motion (Kamps et al., 2016b). Scene motion selectivity is a region’s response to motion information in scenes (i.e., the response to dynamic videos of scenes minus still images of scenes), relative to motion information in faces or objects. Responses in OPA and PPA were measured using fMRI while participants viewed i) 3-sec video clips of dynamic scenes, faces, or objects and ii) 3, 1-sec still images taken from these same video clips. We predicted that OPA (supporting visually-guided navigation) develops later than PPA (supporting categorization). Specifically, we predicted that OPA will develop between ages 5 through 8, while PPA will show no evidence of development.

2. Methods

2.1 Participants

Two groups of participants were recruited for this study: 24 5-year-old children (ages 57-72 months; mean = 62.7 months; SD = 3.2; 14 females) and 15 8-year-old children (ages 93-112 months; mean = 101.2 months; SD = 5.0; 9 females). 7 5-year-olds were excluded from our analysis due to excessive motion and lack of attention during runs (see 2.4 Data Quality Control), giving us a group of 17 5-year-old children (ages 57-72 months, mean = 62.6 months; SD = 3.3; 11 females) with high-quality data. All participants were healthy, had no history of neurological or psychiatric conditions, and had no structural brain abnormalities. All participants were recruited through Emory's Child Study Center. Consent was given for all children by their parent or guardian. In addition, verbal assent was collected for children ages 6 and older. All procedures have been approved by the Emory University Institutional Review Board.

2.2 Experimental Design

2.21 Stimuli

Our experiment consisted of dynamic and static stimuli in three categories: faces, scenes, and objects (Figure 1). The stimuli for each condition were 3-sec video clips. The Static

Scene and Dynamic Scene stimuli were the same as those used in Kamps, Lall, and Dilks (2016). The Dynamic Scene stimuli consisted of 60, 3-sec video clips depicting first-person perspective motion, as would be experienced during locomotion through a scene. The video clips were filmed by walking at a typical pace through 8 different places (e.g., a parking garage, a hallway, etc.) with the camera (a Sony HDR XC260V HandyCam with a field of view of 90.3 x 58.9 degrees) held at eye level. The video clips subtended 23 x 15.33 degrees of visual angle. The Static Scene stimuli were created by taking stills from each Dynamic Scene video clips at 1-, 2- and 3-sec time points, resulting in 180 images. These still images were presented in groups of three images taken from the same place, and each image was presented for 1 sec with no ISI, thus equating the presentation time of the static images with the duration of the movie clips from which they were made. Importantly, the still images were presented in a random sequence such that first-person perspective motion could not be inferred. Like the video clips, the still images subtended 23 x 15.33 degrees of visual angle.

We also included static and dynamic faces and objects in order to test the specificity of any observed scene motion selectivity (i.e., response differences between Dynamic Scene and Static Scene conditions) and in order to calculate scene domain selectivity (i.e., greater response to scenes than faces or objects). The Dynamic Face stimuli and the Dynamic Object stimuli were the same as those used in Pitcher, Dilks, Saxe, Triantafyllou, and Kanwisher (2011). The Dynamic Face stimuli depicted only the faces of 7 children against a black background as they smiled, laughed, and looked around while interacting with off-screen toys or adults. The Dynamic Object stimuli depicted only toys against a black background, as they were played with, ensuring that no body part or face was in the image. The Static Faces and Static Objects were created and presented using the exact same procedure and parameters described for the Static Scene condition above.

2.22 Procedure

During each scanning session, we first took a high resolution anatomical scan while the children watched a movie or show of their choice. Then, we collected fMRI data while participants viewed 4 block-design runs consisting of 3-sec video clips of our dynamic and static stimuli (i.e., faces, scenes, and objects). After 2 runs, the children were offered a break where they could watch a movie or show of their choice for some time. Each run was 360-sec long and contained 2 blocks for each stimulus category. The order for the first set of blocks was pseudorandomized across runs (e.g., Dynamic Faces, Static Objects, Dynamic Scenes, Static Scenes, Static Faces, Dynamic Objects) and the order for the second set was the palindrome of the first (e.g., Dynamic Objects, Static Faces, Static Scenes, Dynamic Scenes, Static Objects, Dynamic Faces). Each block consisted of 6, 3-sec video clips in the particular stimulus category (i.e., Dynamic Faces, Static Scenes). Each run also included 3, 19.8-sec fixation blocks: one at the beginning, one in the middle, and one at the end.

During all runs, participants were instructed to actively view the stimuli. They were told to imagine themselves walking through the places in the scene video clips, playing with the children in the face video clips, and playing with the toys in the object video clips. These instructions were in place to enhance the children’s attention during the videos.

2.3 fMRI Scanning

All scanning was performed on a 3T Siemens Trio scanner in the Facility for Education and Research in Neuroscience at Emory University. The functional images were collected using a 32-channel head matrix coil and a gradient-echo single-shot echoplanar imaging sequence (28 slices, TR = 2 sec, TE = 30 msec, voxel size = 1.5 x 1.5 x 2.5mm, and a .25 interslice gap). For all scans, slices were oriented approximately between perpendicular and parallel to the calcarine sulcus, covering all of the occipital and temporal lobes, as well as the lower portion of the parietal lobe. Whole-brain, high-resolution anatomical images were also acquired for each participant for the purposes of registration and anatomical localization (see 2.5 Data Analysis).

2.4 Data Quality Control

Children move more and have lower attentional capacity than adults. To control for subject motion during fMRI scans, we excluded runs where the average frame-to-frame motion exceeded the size of one voxel (2mm: x, y, z). To verify that participants paid attention across the whole run, and ensure that data were of sufficient quality to detect fMRI activation, we investigated activation in the primary visual cortex (V1) to the contrast of all conditions minus fixation. If participants paid attention during the run and saw the stimuli, then we should see strong activity in both hemispheres of primary visual cortex ($Z>2.3$ and a corrected cluster significance threshold of $p=0.05$). Therefore, we excluded runs where participants did not show V1 activity.

For our analysis, we only included children that had two or more good runs after these exclusionary criteria. At least two runs were required in order to use two independent halves of data for localizing the regions of interest (Localizer runs, see 2.5 Data Analysis) and for measuring the response of these regions to the various experimental stimuli (Experimental runs, see 2.5 Data Analysis).

2.5 Data Analysis

Analysis of functional data was conducted using a combination of tools from the FSL software (FMRIB's Software Library; www.fmrib.ox.ac.uk/fsl) (Smith et al., 2004), the FreeSurfer Functional Analysis Stream (FS-FAST; http://surfer.nmr.mgh.harvard.edu/), and custom written MATLAB code. Before analyzing the data, functional images were motion corrected (Cox and Jesmanowicz, 1999), detrended and fit using a double gamma function, and spatially smoothed with a 5-mm kernel. After preprocessing, we used a region of interest (ROI) approach in which we used one half of the runs to define the OPA and PPA (Localizer Runs) and used the second, independent half of runs (Experimental Runs) to measure the response magnitudes of each region to the six experimental conditions.

For the Localizer runs, we defined the scene-selective regions, OPA and PPA, using a standard contrast of scenes greater than objects (Epstein and Kanwisher, 1998). Given that responses in OPA or PPA may not yet be fully developed in younger children, we employed a group-constrained subject-specific method (GSS) for defining ROIs (Julian et al, 2012). This method allows for the functional definition of each region even if responses in that region would not be detected using a standard threshold approach used in adults (cf. Kamps et al., 2018). Specifically, we identify a unique search space for each scene-selective ROI and its surrounding cortex using a previously published atlas of functional "parcels" (Julian et al., 2012) (Figure 2). These parcels describe the vicinity in which OPA and PPA are likely to fall in any individual, given the typical distribution found in a large, independent sample of adults. For each specific subject, the voxels that fall within each parcel were ranked based on the statistical significance of that voxel's response to scenes > objects. The top 100 most selective voxels were defined as the subject's specific ROI. Using this approach, we were able to bilaterally define regions in each child regardless of how strongly those regions respond to scenes greater than objects, as is standard in studies of adults. We also defined face-selective and object-selective regions for control analyses.

For the remaining, independent Experimental runs, we calculated the magnitude of response (percent signal change) within each ROI to the six experimental conditions: Dynamic Scenes, Static Scenes, Dynamic Faces, Static Faces, Dynamic Objects, Static Objects. This split half, "define and test" procedure was done for every possible permutation of the 2-4 usable runs for each subject (e.g., if a subject had three usable runs, then it was possible to conduct 3 different permutations using 2 runs for localization and one run for test). Results from each permutation were then averaged together, resulting in a final estimate of activation to each condition in each region. For all analyses, we averaged responses in each region across hemispheres. For our domain selectivity analysis, we further averaged response magnitudes across the dynamic and static stimuli for each condition. For our motion selectivity analysis, we calculated a motion difference score for each condition (i.e., the response magnitude to the dynamic stimuli minus response magnitude to the static stimuli).

3. Results

3.1 Scene Domain Selectivity

First, we assessed development by measuring whether there are changes in scene domain selectivity in OPA and PPA from ages 5 through 8. If OPA develops later than PPA, then OPA's domain selectivity for scenes will develop between ages 5 through 8, while the PPA will show no evidence of development. To test this prediction, we first compared responses to scenes over faces and objects in each region (i.e., scene domain selectivity). Then, we explored whether these selective responses increased from ages 5 through 8 in each region, using a mixed model ANOVA.

For OPA, a 2 (group: 5-year-olds, 8-year-olds) x 3 (domain: scenes, faces, objects) mixed model ANOVA revealed a significant main effect of domain ($F_{(2,60)} = 39.95$, $p < 0.001$), with greater responses to scenes than to faces or objects (pairwise comparisons, both p's < 0.05), indicating that OPA responds selectively to scenes in both the 5-year-olds and 8-year-olds (Figure 3A). Critically, however, this same ANOVA did not reveal a significant group x domain interaction ($F_{(2,60)}=0.38$, $p=0.68$), suggesting that scene domain selectivity in OPA does not change from ages 5 through 8 (Figure 3A). This finding is inconsistent with our hypothesis, and rather suggests that scene selectivity in OPA is already present by age 5, and does not develop between ages 5 and 8.

For PPA, a 2 (group: 5-year-olds, 8-year-olds) x 3 (domain: scenes, faces, objects) mixed model ANOVA also revealed a significant main effect of domain ($F_{(2,60)} = 150.97$, $p < 0.001$), with greater responses to scenes than faces or objects (pairwise comparisons, both p's < 0.05), confirming that PPA responds selectively to scenes in both 5 and 8-year-olds (Figure 3B). Further, this same ANOVA surprisingly revealed a significant interaction ($F_{(2,60)}=4.87$, $p=0.01$), suggesting that scene domain selectivity in the PPA might be increasing from ages 5 through 8 (Figure 3B).

The analyses above suggest that counter to our hypothesis, scene domain selectivity in PPA is increasing from ages 5 through 8, while the OPA shows no such development. To test this possibility, we next compared the regions directly, in order to explore whether the apparent developmental change in PPA is significantly greater than that in OPA. Critically, a 2 (region: OPA, PPA) x 3 (domain: scenes, faces, objects) x 2 (group: 5-year-olds, 8-year-olds) mixed model ANOVA failed to reveal a significant interaction ($F_{(2,60)}$=0.97, p=0.38), suggesting that scene domain selectivity in PPA does not develop significantly more than that in OPA. Taken together, our results therefore are most consistent with the idea that neither PPA nor OPA show developmental increases in scene selectivity from ages 5 to 8.

3.2 Scene Motion Selectivity

Next, we assessed development by measuring whether there are changes in scene motion selectivity in OPA and PPA from ages 5 through 8**.** If OPA develops later than PPA, then OPA's motion selectivity for scenes will develop between ages 5 through 8, while PPA will show no evidence of development. To test this prediction, we first compared responses to scene motion over face motion and object motion in each region (i.e., scene motion selectivity). Then, we explored whether these selective responses increased from ages 5 through 8 in each region, using a mixed model ANOVA. In order to perform these analyses, we calculated a scene motion difference score as the difference in responses to dynamic scenes and static scenes. We calculated a face motion difference score and an object motion difference score in the same way.

For OPA, a 2 (group: 5-year-olds, 8-year-olds) x 3 (motion: face motion, scene motion, object motion) mixed model ANOVA revealed a significant interaction ($F_{(2,60)}$=5.23, p=.01), suggesting that OPA's scene motion selectivity changes from ages 5 through 8 (Figure 4A). Critically, this interaction was driven by scene motion selectivity in the 8-year-olds (pairwise comparisons: scene motion > object motion, $p=0.002$; scene motion > face motion, $p=0.004$); no significant selectivity for scene motion was observed in the 5-year-olds (pairwise comparisons: scene motion > object motion, $p=0.39$; scene motion > face motion, $p=0.98$) (Figure 4A). These findings support our hypothesis that OPA is developing from ages 5 through 8. Further, the pattern of responses in the 8-year-olds indicate that OPA is not responding to motion in general, but rather to motion information specific to scenes and not to objects or faces.

For PPA, a 2 (group: 5-year-olds, 8-year-olds) x 3 (motion: face motion, scene motion, object motion) mixed model ANOVA did not reveal a significant interaction ($F_{(2,60)}=0.10$, $p=0.91$), indicating that PPA's scene motion selectivity does not change from ages 5 through 8 (Figure 4B). This finding is consistent with our hypothesis that PPA will not show evidence of development between ages 5 and 8. Interestingly, the same ANOVA revealed a marginally significant main effect of motion ($F_{(2,60)}= 3.01$, $p=0.06$), suggesting that PPA is weakly selective for scene motion in both the 5-year-olds and 8-year-olds (Figure 4B), consistent with previous findings in adults (Kamps et al., 2016b).

The analyses above suggest that in support of our hypothesis, scene motion selectivity in OPA is increasing from ages 5 through 8, while PPA shows no such development. However, to confirm that the observed developmental change in OPA is significantly greater than that in PPA, we next compared the regions directly. A 2 (region: OPA, PPA) x 2 (group: 5-year-olds, 8-year-olds) x 3 (motion: face motion, scene motion, object motion) mixed model ANOVA revealed a marginally significant region x group x motion interaction ($F_{(2,60)}=2.80$, $p=0.07$) with a medium effect size ($\eta_p^2=0.09$), suggesting that with a larger sample, the interaction would be significant. These findings suggest that scene motion selectivity in OPA develops significantly more than that in PPA. Taken together, these results strongly support our hypothesis that OPA is developing between ages 5 and 8, while the PPA is not.

However, could it be that the OPA's increase in scene motion selectivity is driven by better data quality in the 8-year-olds? Due to our exclusionary criteria for motion and attention (see 2.4 Data Quality Control), it is unlikely that the changes in scene motion selectivity are a result of data quality. Further, our findings above that scene domain selectivity was just as strong in 5-year-olds as it was in 8-year-olds, suggest that the 5-year-old data are reasonably well matched to the 8-year-old data. Nevertheless, to ensure that differences in data quality between the groups was not driving these results, we measured data quality directly by investigating the temporal signal to noise ratio (tSNR). To calculate tSNR, we took the mean response of each voxel in a region of interest during the fixation blocks and divided this mean by the standard deviation of each voxel's response during those blocks. This calculation gave us a measure of signal to noise in each voxel of a scene-selective region. Then we averaged the tSNR of each voxel in the region to get the overall tSNR for the ROI in each individual participant. We then analyzed our tSNR data in three ways. First, we compared tSNR in OPA between 5-year-olds and 8-year-olds, and found no significant difference ($t_{(30)}=1.60$, $p=0.12$) between the two groups. Second, we assessed whether there was a correlation between i) tSNR in OPA and ii) the strength of the response to scene motion in OPA, where a positive correlation would indicate that better signal quality predicts better detection of the scene motion effect in OPA. This analysis did not reveal a significant correlation ($r=0.10$, $p=0.59$), indicating that tSNR does not explain changes in OPA's scene motion selectivity across 5 and 8-

year-olds. Third, we again performed the critical 2 (region: OPA, PPA) x 2 (group: 5-year-olds, 8-year-olds) repeated measures ANOVA presented above, but now included tSNR in OPA as a covariate. We still observed a marginally significant interaction ($F_{(1,29)}=3.34$, $p=0.08$), with a medium effect size ($\eta_p^2 = .10$), further ensuring that this effect was not driven by data quality. Taken together, differences in data quality between the 5 year olds and 8 year olds are unlikely to explain the development of motion selectivity in OPA, but not PPA, across this age range.

4. Discussion

Here, we investigated how the well-established adult navigation system (including the OPA) and categorization system (including the PPA) develop across the lifespan. In particular, we compared responses in OPA and PPA to i) video clips of first-person perspective motion through scenes, a critical cue for navigating through the immediately visible scene (“Dynamic Scenes”) and ii) still frames taken from these same movies (“Static Scenes”), rearranged such that first-person perspective motion cannot be inferred. We also measured the responses of OPA and PPA to dynamic and static faces and objects. We assessed developmental changes at two levels: scene domain selectivity and scene motion selectivity. We found scene domain selectivity in both OPA and PPA by age 5, and found no developmental changes from ages 5 through 8 in either region. Critically, we found that OPA’s scene motion selectivity develops between ages 5 and 8, an indication of its role in visually-guided navigation (Kamps et al., 2016b), with no such evidence of development in PPA. These findings are consistent with the idea that the navigation and categorization systems are dissociable, and that they develop differentially, with the navigation system (including the OPA) developing later than the categorization system (including the PPA). Importantly, our findings are consistent with recent behavioral evidence in our lab, showing that 4-year-olds perform worse on a navigation task than a categorization task, relative to 8-year-olds.

In regards to scene motion selectivity, OPA did not respond greater to scene motion than face motion or object motion at age 5, but did show this effect by age 8. This increase in scene motion selectivity supports our hypothesis that OPA is developing from ages 5 through 8, particularly at the level of its sensitivity for scene motion. On the other hand, PPA showed a marginal response for scene motion by age 5, but this weak selectivity did not change from ages 5 through 8. Critically, this marginal selectivity for scene motion is found in adult PPA (Kamps et al., 2016b), and therefore suggests that PPA may already be adultlike by age 5. In our scene motion selectivity analysis, we critically found that OPA only responded to motion information in scenes, and not to motion on faces or objects. We also found that changes in scene motion selectivity could not be explained by better data and signal quality in the older group of children. OPA's changes in scene motion selectivity are exciting because they suggest that OPA 's response to motion, a critical cue for its role in navigation (Kamps et al., 2016), is developing across this age range. The independent and later developing navigation system is consistent with the general late development of the dorsal stream. In object processing, the dorsal stream (Action System) develops later than the ventral stream (Perception System) (Ungerleider and Mishkin, 1982; Milner & Goodale, 1995; Dilks et al., 2008). Given that the OPA is found in the dorsal stream, whereas the PPA is found in the ventral stream, the developmental trend found in scene motion selectivity mirrors that in object processing.

In regards to scene domain selectivity, OPA responded greater to stimuli of scenes than faces or objects by age 5, with no significant development from ages 5 through 8. PPA also responded greater to stimuli of scenes than faces or objects by age 5, but in contrast to OPA, we found a significant change in scene domain selectivity from ages 5 through 8. Critically, the change observed in PPA was not significantly different than that in OPA, suggesting that both OPA and PPA are domain selective for scenes by age 5 with no changes from ages 5 through 8. Therefore, if OPA develops between ages 5 through 8, it is not happening at the level of scene domain selectivity. This pattern of development is surprising, and suggests a novel developmental hypothesis where the scene processing system (and perhaps high-level visual systems in general) first develops selectivity for a category overall (i.e., scenes) and then later bifurcates into more specific functions (i.e., motion selectivity for navigation). In short, our results suggest that scene-selective regions are first established in the same locations they will exist in adults (i.e., by age 5) and then more specific functions develop in each region (i.e., OPA further develops into a navigation region from ages 5 through 8).

While this study elucidates the development of OPA from ages 5 through 8, future work is required to address the precise nature of PPA's development. A limitation of our work is that we do not test directly for scene categorization information in PPA. By contrast, we do test for scene navigation information in OPA using motion. Even though categorization (and PPA) cannot be measured directly, one clue from our work that PPA is in fact developed by age 5 is that scene domain selectivity is present by age 5 and does not change from 5 to 8. A second clue is that PPA showed a slight motion effect, like that seen in adults, by age 5, with no changes from 5 to 8. But, because scene domain and scene motion selectivity are not specific to PPA, PPA's development still needs to be rigorously tested, and stronger evidence that PPA encodes scene category information by age 5 is needed. One approach is to assess PPA's sensitivity to other scene category stimuli, such as relative length and angle changes in scenes (Dillon et al., 2018) at ages 5 and 8. Stronger evidence, however, would be to look in younger children (e.g, ages 3 -5) to assess whether PPA's selectivity for scenes and other scene category information develops at a younger age. This approach would be difficult given that collecting quality data in 5-year-olds is already a challenge. The strongest evidence would be having kids perform a navigation and a categorization task during fMRI data collection (Persichetti and Dilks, 2018). However, it would be challenging to get young children to perform this task while staying still in the scanner.

Other future directions for our study include comparing the responses in 8-year-olds to adults. While our data suggest that OPA develops between ages 5 through 8, are the responses adultlike by age 8? Numerous studies have assessed when PPA development becomes adultlike, but to our knowledge, no one has investigated when OPA becomes fully developed. Some of these studies suggest that the scene region is adultlike by age 8 (Scherf et al., 2007, 2011; Jiang et al., 2014) while others find that this scene region is later developing (Golarai et al., 2007; Chai et al., 2010). Recently, Meissner et al. (2019) suggested that both OPA and PPA are later developing. A critical issue in addressing the development from childhood to adulthood is carefully matching data quality between children and adults. The lack of careful data quality matching in this previous literature may explain the inconsistent findings, and future work can resolve this problem. Further, we would like to extend our current findings to William's syndrome. Behavioral work in our lab has shown that people with William's syndrome are impaired on a navigation task, but not a categorization task. We can use the methods in this experiment to assess whether these patients accordingly have damage to OPA and not PPA, providing both casual and neural evidence that these two systems are dissociated and have different developmental trajectories.

In conclusion, we found that the visually-guided navigation system (including OPA) develops differentially from the categorization system (including PPA), with the navigation system developing later. Specifically, we found that unlike in PPA, OPA's sensitivity to scene motion develops between ages 5 through 8. The increasing response of OPA to scene motion information, particularly to first-person perspective motion, suggests the development of OPA's role in visually-guided navigation across this age range. On the other hand, we did not find developmental changes in OPA or PPA for scene domain selectivity. This finding suggests the novel hypothesis that high-level visual systems develop selectivity for a category overall before developing the particular functions within that category (i.e., categorization vs navigation).

Figure 1

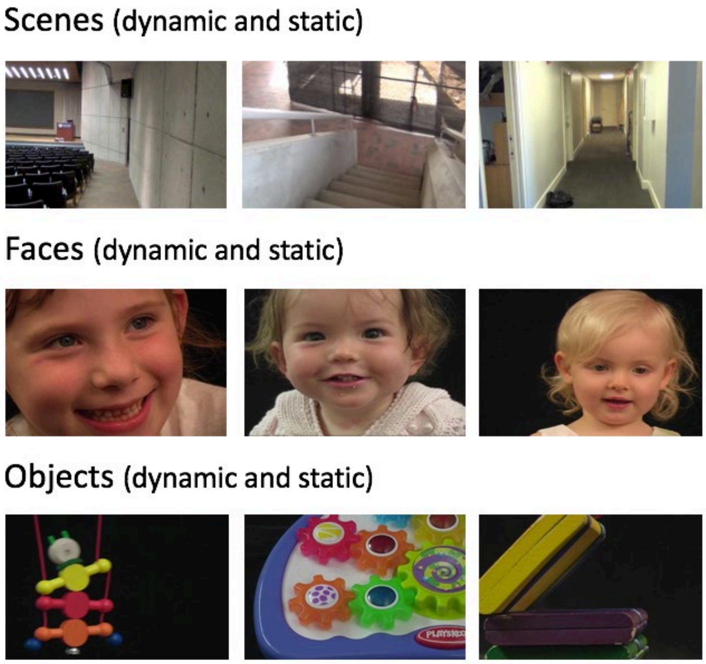

Figure 1. Example stimuli used in each blocked design run. From top to bottom: scenes (dynamic and static), faces (dynamic and static), objects (dynamic and static). Dynamic stimuli were 3-sec video clips of first-person perspective motion through a place (or “scene”), of a child’s face smiling and laughing, or of moving toys. Static stimuli were 3 1-sec still images taken from these video clips, rearranged such that no two frames came consecutively, and so that first-person perspective motion could not be inferred.

Figure 2

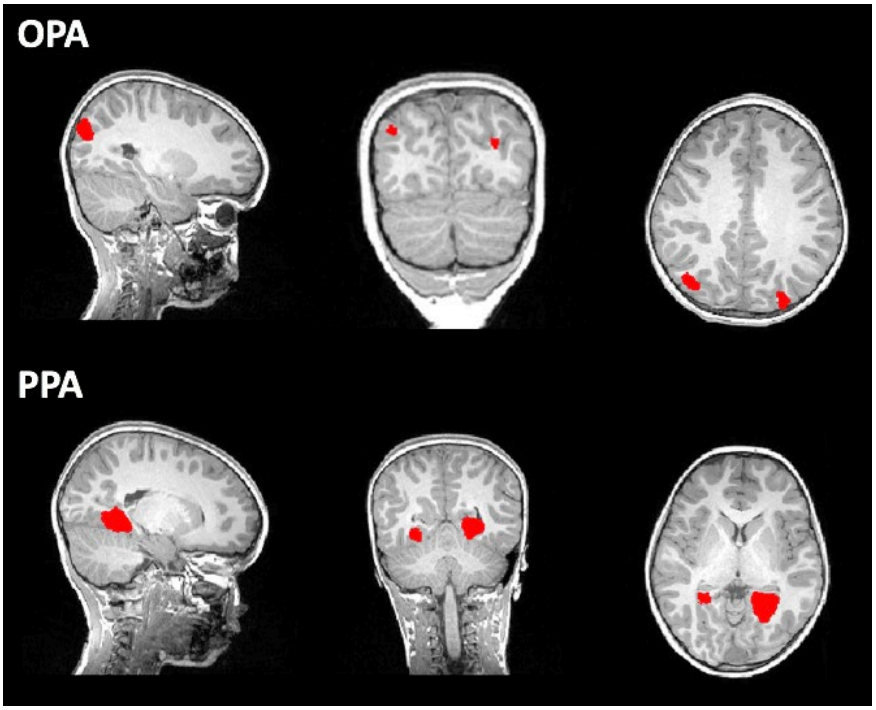

Figure 2. Parcels for scene-selective regions of interest (Julian et al., 2012) on a sample 5-year-old participant's anatomical scan. These parcels represent the unique search space for each scene-selective ROI, describing the vicinity in which OPA and PPA are likely to fall in any individual. Using these parcels, we define each individual's ROI by selecting the top 100 most scene-selective voxels within that parcel during half of the runs. Responses in that ROI were then tested using the remaining independent, set of runs.

Figure 3

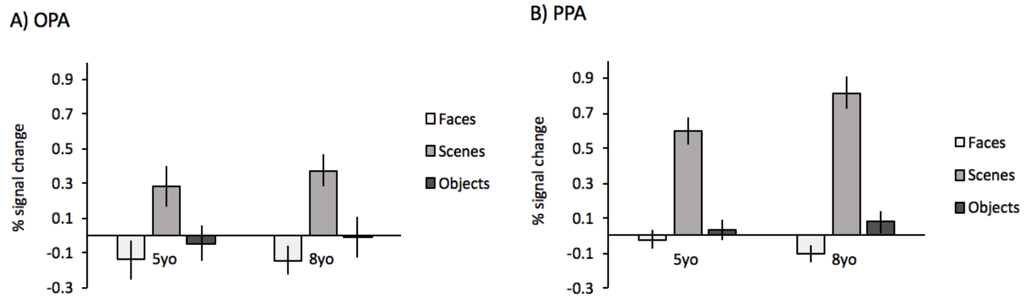

Figure 3 - Scene Domain Selectivity. Average percent signal change (PSC) in OPA and PPA to faces, scenes, and objects. Each bar is the average response to the static and dynamic stimuli in the given domain (faces, scenes or objects). A) OPA responded more to scenes than faces or objects in both the 5yo and 8yo groups ($F_{(2,60)} = 39.95$, $p < 0.001$). Selective responses did not change from age 5 through 8 ($F_{(2,60)}=0.38$, $p=0.68$), suggesting that scene domain selectivity in OPA is present by age 5, and does not develop between ages 5 and 8. B) PPA responded more to scenes than faces or objects in both the 5-year-old and 8-year-old groups ($F_{(2,60)} = 150.97$, $p < 0.001$), Although we did see significant developmental change in PPA from age 5 to 8 ($F_{(2.60)}=4.87$, $p=0.01$), such developmental change in PPA was not greater than that in OPA ($F_{(2,60)}=0.97$, $p=0.38$), suggesting that neither PPA nor OPA show developmental increases in scene-selectivity from ages 5 through 8.

Figure 4

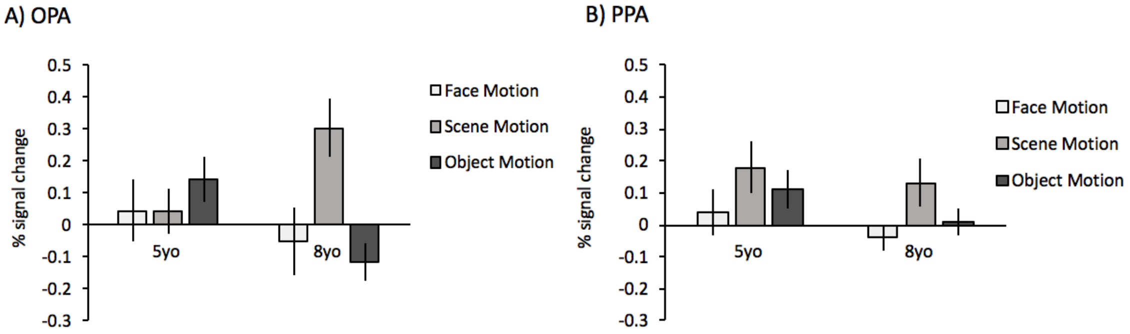

Figure 4 – Scene Motion Selectivity. Difference scores (percent signal change) for Dynamic Scenes minus Static Scenes (“Scene Motion”), Dynamic Faces minus Static Faces (“Face Motion”), and Dynamic Objects minus Static Objects (“Object Motion”). A) OPA responded greater to scene motion than object motion and face motion at age 8 (pairwise comparisons: scene motion > object motion, p=0.002; scene motion > face motion, p=0.004), but not at age 5 (pairwise comparisons: scene motion > object motion, p=0.39; scene motion > face motion, p=0.98), suggesting that OPA’s scene motion selectivity increases from ages 5 through 8 ($F_{(2,60)}$=5.23, p=.01). B) By contrast, PPA was marginally selective for scene motion in both the 5 and 8-year-olds ($F_{(2,60)}$= 3.01, p=.06), and no significant increase in scene motion selectivity was observed from ages 5 through 8 ($F_{(2,60)}$=.10, p=.91). Overall, this data suggests that scene motion selectivity develops in OPA, but not PPA ($F_{(2,60)}$=2.80, p=0.07, η_p^2=0.09).

References

Aguirre, G. K., & D'Esposito, M. (1999). Topographical disorientation: a synthesis and taxonomy. *Brain*, *122*(9), 1613-1628.

Bar, M., Aminoff, E., 2003. Cortical analysis of visual context. Neuron 38, 347–358.

Berard, J. R., & Vallis, L. A. (2006). Characteristics of single and double obstacle avoidance strategies: a comparison between adults and children. *Experimental brain research*, *175*(1), 21-31.

Carey, S., & Diamond, R. (1977). From piecemeal to configurational representation of faces. Science, 195 (4275), 312–314.

Chai, X. J., Ofen, N., Jacobs, L. F., & Gabrieli, J. D. (2010). Scene complexity: influence on perception, memory, and development in the medial temporal lobe. *Frontiers in human neuroscience*, *4*, 21.

Cox, R.W., Jesmanowicz, A., 1999. Real-time 3D image registration for functional MRI. Magn. Reson. Med. 42, 1014–1018.

Dilks, D.D., Hoffman, J.E., & Landau, B. (2008). Vision for perception and vision for action: Normal and unusual development. *Developmental Science*, *11*, 474-486.

Dilks, D.D., Julian, J.B., Kubilius, J., Spelke, E.S., Kanwisher, N., 2011. Mirror-image sensitivity and invariance in object and scene processing pathways. J. Neurosci. 31, 11305–11312.

Dilks, D.D., Julian, J.B., Paunov, A.M., Kanwisher, N., 2013. The occipital place area is causally and selectively involved in scene perception. J. Neurosci. 33, 1331–1336a.

Dillon, M.R.*, Persichetti, A.S.*, Spelke, E.S., & Dilks, D.D. (2018). Places in the brain: Bridging layout and object geometry in scene-selective cortex. *Cerebral Cortex, 28,* 2365-2374.

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

Epstein, R., DeYoe, E. A., Press, D. Z., Rosen, A. C., & Kanwisher, N. (2001). Neuropsychological evidence for a topographical learning mechanism in parahippocampal cortex. *Cognitive neuropsychology*, *18*(6), 481-508.

Fajen, B. R., & Warren, W. H. (2003). Behavioral dynamics of steering, obstable avoidance, and route selection. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(2), 343.

Ganaden, R. E., Mullin, C. R., & Steeves, J. K. (2013). Transcranial magnetic stimulation to the transverse occipital sulcus affects scene but not object processing. *Journal of cognitive neuroscience*, *25*(6), 961-968.

Gibson, J. J. (1950). Perception of the visual world. Boston: Houghton- Mifflin.

Golarai, G., Ghahremani, D.G., Whitfield-Gabrieli, S., Reiss, A., Eberhardt, J.L., Gabrieli, J.D., & Grill-Spector, K. (2007). Differential development of high-level visual cortex correlates with category-specific recognition memory. Nature Neuroscience, 10 (4), 512–522.

Grill-Spector, K., 2003. The neural basis of object perception. Curr. Opin. Neurobiol. 13, 159-166.

Habib, M., & Sirigu, A. (1987). Pure topographical disorientation: a definition and anatomical basis. *Cortex*, *23*(1), 73-85.

Higgins CI, Campos JJ, Kermoian R. Effect of self-produced locomotion on infant postural compensation to optic flow. Developmental Psychology. 1996;32:836–841.

Jiang, P., Tokariev, M., Aronen, E. T., Salonen, O., Ma, Y., Vuontela, V., & Carlson, S. (2014). Responsiveness and functional connectivity of the scene-sensitive retrosplenial complex in 7–11-year-old children. *Brain and cognition*, *92*, 61-72.

Julian JB, Fedorenko E, Webster J, Kanwisher N (2012) An algorithmic method for functionally defining regions of interest in the ventral visual pathway. Neuroimage 60:2357–2364. CrossRef Medline.

Julian, J.B., Kamps, F.S., Epstein, R.A., & Dilks, D.D. (2018). Dissociable spatial memory systems revealed by typical and atypical human development. *Developmental Science, e1273,* 1- 12.

Kamps, F.S., Julian, J.B., Kubilius, J., Kanwisher, N., & Dilks, D.D. (2016a). The occipital place area represents the local elements of scenes. *NeuroImage, 132,* 417-424.

Kamps, F.S., Lall, V., & Dilks, D.D. (2016b). The occipital place area represents first person perspective motion information through scenes. *Cortex, 83,* 17-26.

Konkle, T., Oliva, A., 2012. A real-world size organization of object responses in occipitotemporal cortex. Neuron 74, 1114–1124.

Kravitz, D.J., Peng, C.S., Baker, C.I., 2011. Real-world scene representations in high-level visual cortex: it's the spaces more than the places. J. Neurosci. 31, 7322–7333.

Loomis, J. M., & Philbeck, J. W. (2008). Measuring spatial perception with spatial updating and action. In *Carnegie Symposium on Cognition, 2006, Pittsburgh, PA, US*. Psychology Press.

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

Meissner, T. W., Nordt, M., & Weigelt, S. (2019). Prolonged functional development of the parahippocampal place area and occipital place area. *NeuroImage*.

Mendez, M. F., & Cherrier, M. M. (2003). Agnosia for scenes in topographasnosia. *Neuropsychologia*, *41*(10), 1387-1395.

Milner, A. D., & Goodale, M. A. (1995). Oxford psychology series, No. 27.

Mullally, S.L., Maguire, E.A., 2011. A new role for the parahippocampal cortex in representing space. J. Neurosci. 31, 7441–7449.

Oliva, A., Torralba, A., 2001. Modeling the shape of a scene: a holistic representation of the spatial envelope. Int. J. Comput. Vis. 42, 145–175.

Oliva, A., Torralba, A., 2002. Scene-centered description from spatial envelope properties. Lect. Notes Comput. Sci. 2525, 263–272.

Pitcher, D., Dilks, D.D., Saxe, R., Triantafyllou, C., & Kanwisher, N. (2011). Differential selectivity for dynamic versus static information in face selective cortical regions. *Neuroimage*, *56*, 2356-63.

Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E.J., Johansen-Berg, H., Bannister, P.R., De Luca, M., Drobnjak, I., Flitney, D.E., Niazy, R.K., Saunders, J.,Vickers, J., Zhang, Y.Y., De Stefano, N., Brady, J.M., Matthews, P.M., 2004. Advances in functional and structural MR image analysis and implementation as FSL. NeuroImage, 23, S208–S219.

Thrun, S., 2002. Robotic Mapping: A Survey. In: Lakemeyer, G., Nebel, B. (Eds.), Exploring Artificial Intelligence in the New Millennium. Morgan Kaufmann.

Troiani, V., Stigliani, A., Smith, M.E., Epstein, R.A., 2014. Multiple object properties drive scene-selective regions. Cereb. Cortex 24, 883–897.

Park, S., Brady, T.F., Greene, M.R., Oliva, A., 2011. Disentangling scene content from spatial boundary: complementary roles for the parahippocampal place area and lateral occipital complex in representing real-world scenes. J. Neurosci. 31, 1333–1340.

Persichetti, A.S. & Dilks, D.D. (2016). Perceived egocentric distance sensitivity and invariance across scene-selective cortex. *Cortex, 77,* 155-163.

Persichetti, A.S., & Dilks, D.D. (2018). Dissociable neural systems for recognizing places and navigating through them. *Journal of Neuroscience, 38,* 10295-10304.

Pryde, K. M., Roy, E. A., & Patla, A. E. (1997). Age-related trends in locomotor ability and obstacle avoidance. *Human movement science*, *16*(4), 507-516.

Scherf, K. S., Behrmann, M., Humphreys, K., & Luna, B. (2007). Visual category-selectivity for faces, places and objects emerges along different developmental trajectories. *Developmental science*, *10*(4), F15-F30.

Scherf, K. S., Luna, B., Avidan, G., & Behrmann, M. (2011). "What" precedes "which": developmental neural tuning in face-and place-related cortex. *Cerebral Cortex*, *21*(9), 1963-1980

Ungerleider LG, Mishkin M (1982) Two cortical visual systems. In Analysis of visual behavior (Ingle D, Goodale M A, Mansfield R, eds), pp 549 –587. Cambridge, MA: MIT.

Vasudevan, E. V., Torres-Oviedo, G., Morton, S. M., Yang, J. F., & Bastian, A. J. (2011). Younger is not always better: development of locomotor adaptation from childhood to adulthood. *Journal of neuroscience*, *31*(8), 3055-3065.

Wang, R. F., & Spelke, E. S. (2000). Updating egocentric representations in human navigation. *Cognition*, *77*(3), 215-250.

Weigelt, S., Koldewyn, K., Dilks, D. D., Balas, B., McKone, E., & Kanwisher, N. (2014). Domain-specific development of face memory but not face perception. *Developmental Science*, *17*(1), 47-58.