Distribution Agreement

Signature:

Frederik S. Kamps Date

The functional organization and development of the human visual scene processing system

By

Frederik S. Kamps
Doctor of Philosophy

Psychology

Daniel D. Dilks
Advisor

Stella F. Lourenco
Committee Member

Joseph R. Manns
Committee Member

Hillary R. Rodman
Committee Member

Michael T. Treadway
Committee Member

Accepted:

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

Date

The functional organization and development of the human visual scene processing system

By

Frederik S. Kamps
M.A., Emory University, 2015

Advisor: Daniel D. Dilks, Ph.D.

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Psychology
2019

Abstract

The functional organization and development of the human visual scene processing system

By Frederik S. Kamps

Recognizing the local visual environment (or “scene”) and navigating through it is essential for survival. Perhaps not surprisingly then, human cortex contains a set of cortical regions dedicated to visual scene processing, including the occipital place area (OPA), parahippocampal place area (PPA), and retrosplenial complex (RSC). However, beyond establishing the general involvement of these regions in scene processing (i.e., responding more to images of scenes than to images of everyday objects and faces in neuroimaging studies), two fundamental questions remain unanswered: What is the precise function of each region in adult human visual scene processing? And how does that function get wired up in development? Here I present converging neural, causal, and developmental evidence for two hypotheses. First, I test the hypothesis that scene processing depends on dissociable systems for navigating and categorizing scenes (Papers 1-2). Consistent with this hypothesis, I found that the adult OPA responds more strongly than PPA and RSC to first-person perspective motion information (i.e., videos mimicking visual experience of actually walking through a scene), suggesting that OPA is involved in what I call “visually-guided navigation”, while PPA and RSC support other aspects of scene recognition and navigation. Further, I found that adults with Williams syndrome (a developmental disorder involving cortical thinning in and around OPA) are impaired on a visually-guided navigation task, but not a scene categorization task (e.g., recognizing a kitchen versus a living room), providing causal evidence for dissociable visually-guided navigation and scene categorization systems. Second, I test the further hypothesis that the visually-guided navigation and scene categorization systems develop independently, with visually-guided navigation maturing later than scene categorization (Papers 2-3). Consistent with this hypothesis, I indeed found that visually-guided navigation ability emerges more slowly than scene categorization ability across childhood. Moreover, I found that first-person perspective motion processing in OPA is still developing across this same age range, further revealing that the visually-guided navigation system undergoes protracted development. Taken together, these findings indicate that human visual scene processing is supported by at least two dissociable systems: an earlier-developing scene categorization system in PPA, and a later-developing visually-guided navigation system in OPA.

The functional organization and development of the human visual scene processing system

By

Frederik S. Kamps
B.A. Macalester College 2013

Advisor: Daniel D. Dilks, Ph.D.

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Psychology

Acknowledgments

First, and foremost, I would like to thank everyone in the Dilks lab: I have been so lucky to work with such a smart, driven, and inspiring group of people, and could not have accomplished this work without their help and collaboration. I would like to give a special thanks to my advisor, Danny Dilks, who taught me what it means to be a mentor, and how to truly think like a scientist. Beyond the Dilks lab, I want to thank my dissertation committee for their constant support and helpful insight, as well as the Facility for Education and Research in Neuroscience (FERN) Imaging Center, the Emory Child Study Center, and the funding sources that supported my work as a graduate student: NICHD grant T32HD071845 and NEI grant T32EY7092. Last but not least, I would like to thank my family and friends for all of their love and support. I love you all so much.

Table of Contents

General Introduction

Our ability to perceive the visual environment is remarkable: We can recognize a place (e.g., a kitchen, a beach, city hall) within a fraction of a second, even if we have never seen that particular place before, and almost simultaneously be ready to navigate through it, finding our way through both the immediately visible surroundings and the broader spatial environment relative to the currently visible scene. Given the ecological importance of place or "scene" perception, it is perhaps not surprising then that particular regions of the human brain are specialized for processing visual scene information, including the occipital place area (OPA) (Dilks et al., 2013), the parahippocampal place area (PPA) (Epstein and Kanwisher, 1998a), and the retrosplenial complex (RSC) (Maguire, 2001) (Figure 1). However, while two decades of neuroimaging research has established the general involvement of these regions in scene processing (i.e., responding more to images of scenes than to images of everyday objects and faces in human neuroimaging studies), two fundamental and yet unanswered questions remain. First, what is the particular functional role of each region within adult human visual scene processing? And second, how does that functional organization get wired up in development? The goal of this thesis is to address these two foundational questions about the nature of the human visual scene processing system.

A central assumption since the discovery of scene-selective cortex has been that the adult human visual scene processing system is a monolithic system entirely devoted to navigation (Epstein and Kanwisher, 1998a; Epstein, 2008; Epstein and Vass, 2015). However, recent work from myself and others in my lab has challenged this widely held assumption. While two scene-selective regions – the OPA and RSC – represent information that is indeed essential for navigation, including egocentric distance (i.e., near versus far from me) and "sense" (i.e., left versus right) information, another scene-selective region – the PPA – does *not* represent such information (Dilks et al., 2011; Persichetti and Dilks, 2016). This finding indicates that PPA lacks the basic information necessary for orienting oneself to the current scene, and therefore cannot play a direct role in navigation. As such, my lab and I have proposed that the human visual scene processing system is not a monolithic system in support of navigation, but rather is composed of at least two dissociable systems: one for navigation, including OPA and RSC, and a second for scene categorization (e.g., recognizing a kitchen versus beach), including PPA (Dilks et al., 2011; Kamps et al., 2016a; Persichetti and Dilks, 2016). Moreover, I have suggested that a division of labor can be found even within the navigation system, with the OPA supporting "visually-guided navigation" (i.e., how to get to the other side of the room without running into walls or tripping over obstacles) and the RSC supporting "memory-guided navigation" (i.e., how to get from my current location to another place beyond the current view) (Kamps et al, 2016b).

Further, while the past two decades have begun to shed light on the functional organization of the adult scene processing system, very few studies to date have explored how human visual scene processing develops across the lifespan. Moreover, all developmental studies to date have focused exclusively on the development of scene selectivity (i.e., greater responses to scenes than objects or faces), and the majority have focused only on PPA. As a consequence, critical questions remain unexplored. How does the more precise function of each region emerge over development? Do the scene regions develop along similar timelines, or independent timelines? How does the emergence of the selective function in each region coincide with the development of scene processing behavior (e.g., navigation and scene categorization)? Indeed, answers to these questions are critical for beginning to understand the relative contributions of genetics and experience in shaping specialized cognitive and neural systems for visual scene

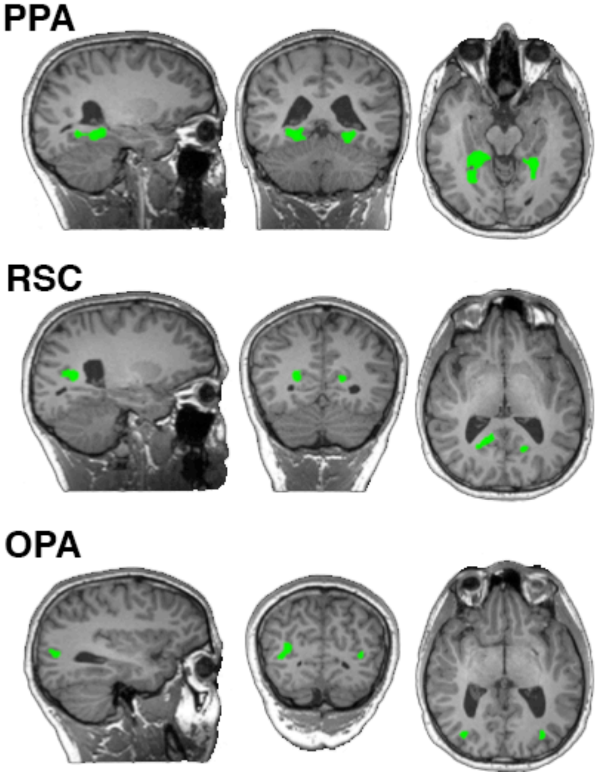

Figure 1. Scene-selective cortical regions in a sample participant. The Parahippocampal Place Area (PPA), Retrosplenial Complex (RSC), and Occipital Place Area (OPA) are defined as those regions responding significantly more to scenes than objects ($p < .0001$). Image adapted from Kamps et al. (2016).

processing. Moreover, understanding the development of the human visual scene processing system may provide a powerful test of the “two systems for scene processing hypothesis” above by showing that these systems develop independently. For example, one strong source of evidence for the classic division of labor in object processing between the dorsal “vision-for-action” system and the ventral “vision-for-perception” system comes from findings that the dorsal action system is later developing than the ventral perception system (Diamond and Goldman-Rakic, 1989; Diamond et al., 1989; Bertenthal, 1996; Gilmore and Johnson, 1997a, b; Csibra et al., 1998; Atkinson et al., 2003; Dilks et al., 2008). Given our analogous proposal that the human visual scene processing system can be divided into a more dorsal navigation system and a more ventral scene categorization system, we therefore predict that the navigation system (including OPA, RSC, or both, in the dorsal stream) will similarly emerge later in development than the scene categorization system (including PPA, in the ventral stream).

Finally, all of the work discussed above involves only correlational data. Thus, a critical open question is whether one system can be *causally* impaired (e.g., as a result of neurological insult), independent of the other. For example, if the navigation system is causally distinct from the scene categorization system, then it should be possible to find cases of selective impairment in navigation, but not scene categorization, and vice versa.

In the first three sections of this General Introduction, I review the literature on each scene-selective region individually, considering what is currently known about the selectivity of each region, the particular kinds of scene information each represents, the role each region might play in scene processing, and finally how each region develops. In the final section, and in view of the current literature, I propose the two hypotheses outlined above: i) that the scene processing system can be divided into at least two distinct systems: one for navigation, including OPA and RSC, and a second for scene categorization, including PPA; and ii) that these two scene processing systems develop independently, with the navigation system maturing later than the scene categorization system.

Parahippocampal place area

Scene selectivity

The parahippocampal place area (PPA) was the first scene-selective region of the brain to be discovered. PPA is functionally defined as the region near the parahippocampal-lingual boundary that responds significantly more to images of scenes than objects. In their original report, Epstein and Kanwisher (1998) discovered that PPA responds strongly to images of outdoor scenes (e.g., cityscapes and landscapes) and indoor scenes (e.g., an apartment), but only weakly to images of everyday objects (e.g., a blender), and not at all to images of faces. Subsequent studies found that PPA also responds to scenes made out of Lego blocks (versus objects made from those same Lego blocks) (Epstein et al., 1999), isolated buildings (Aguirre et al., 1998; Haxby et al., 1999), tabletop scenes (Epstein et al., 2003), line drawings of scenes (Walther et al., 2011), haptically explored scenes (both in sighted and in blind patients) (Wolbers et al., 2011), and imagined scenes (O'Craven and Kanwisher, 2000; Johnson and Johnson, 2014). PPA also responds more to "full" scenes (e.g., a kitchen) than to close up views of local portions of those scenes (e.g., a stove and the immediate, surrounding countertop) or to diagnostic objects from those scenes (e.g., a stove) presented in isolation (Henderson et al., 2008). While PPA has been shown to be sensitive to multiple sources of low- and mid-level visual information (found mostly in images of scenes compared to images of objects or faces), such as higher spatial frequencies (Rajimehr et al., 2011), cardinal orientations (Nasr and Tootell, 2012), rectilinearity (Nasr et al., 2014), texture (Cant and Goodale, 2011), and information presented in the upper visual field (Silson et al., 2015), sensitivity to this lower-level visual information does not completely explain responses in the region. For example, PPA responds reliably to line drawings of scenes where almost no low-level visual information is present (Walther et al., 2011), as well as to haptically explored scenes with which subjects have had no previous visual experience (Wolbers et al., 2011). As such, these findings do not challenge the claim that PPA is scene selective, but rather explicate the particular visual "building blocks" that PPA might extract in order to construct a representation of the scene. Finally, in addition to fMRI evidence for its selective role in scene processing, which is only correlational, studies of stroke patients with damage in and around the parahippocampal cortex have shown an impairment in scene, but not object processing (Aguirre and D'Esposito, 1999; Epstein et al., 2001), and an electroencephalographic study of a patient with electrodes over PPA showed that stimulation of the region lead to hallucinations of indoor and outdoor scenes (Megevand et al., 2014), suggesting that PPA is causally involved in scene processing.

Information processing

While these findings demonstrate that PPA is scene-selective, precisely what scene information does PPA represent? A central hypothesis is that PPA encodes the “spatial layout”, or geometry, of the local visual environment (Epstein and Kanwisher, 1998a; Kamps et al., 2016b). Supporting this hypothesis, these studies found that PPA responds significantly more to images of empty apartment rooms than to images of these same rooms in which the coherent spatial layout of the room has been disrupted by fracturing and rearranging the major surfaces (i.e., walls, floors and ceilings) that define the space (Figure 2). Importantly, these studies also found that PPA responds similarly to images of empty rooms and these same rooms fully furnished, and comparatively weakly to images of multiple objects presented on a white background; thus, they argued that PPA encodes the spatial layout of the local visual environment, but not the individual objects that make up the spatial layout themselves. A variety of other findings support spatial layout representation in PPA (Epstein et al., 1999; Wolbers et al., 2011; Harel et al., 2012; Kamps et al., 2016b). For example, two fMRI studies using multivoxel pattern analysis found that patterns of activity in PPA could discriminate between scenes with “closed” (e.g., a cave) versus “open” (e.g., a grassy plain) layouts, but not between scenes with manmade (e.g., a city street) versus natural (e.g., a forest) content (Kravitz and Baker, 2011; Park et al., 2011). An important question for future research concerns what scene information in particular is used to extract the spatial layout of a scene. In pursuit of this question, a recent fMRI adaptation study found that PPA was sensitive to changes in the length and angle of walls in minimal scenes, suggesting that PPA may use length and angle information in particular to interpret the overall shape of the space (Dillon et al., 2018).

One prominent alternative to the spatial layout hypothesis is the "contextual association hypothesis," which states that rather than processing purely spatial information in scenes, PPA instead is part of a larger system in parahippocampal cortex that more generally processes contextual associations between objects and contexts, including both spatial (e.g., a steering wheel—an object—and the inside of car—a context, where the steering wheel is typically found in a certain location relative to the inside of a car) and non-spatial (e.g., a box of chocolates and Valentine's Day, where no spatial information is involved) associations (Bar, 2004). Evidence for this hypothesis comes from findings that a parahippocampal region closely corresponding to the PPA responds significantly more to objects strongly associated with particular contexts (e.g., a refrigerator, which is always found in a kitchen) than to weakly contextual objects (e.g., a camera, which may be found in a variety of environments) (Bar and Aminoff, 2003), as well as findings that PPA responds more to scenes with strong contextual associations (e.g., a barber chair in a barber shop) versus weak contextual associations (e.g., a laptop computer on a tabletop) (Bar et al., 2008). Beyond these studies, however, empirical support for the contextual association hypothesis is relatively weak. For example, contextual object effects in PPA fail to replicate at shorter presentation rates, which presumably limit the incidence of imagery, suggesting that these effects may be explained instead by participants' *imagining* the full scene in which these objects belong (Epstein and Ward, 2010). Further, when analyzed on a subject-by-subject basis, contextual object effects are only significant in a minority of individuals, unlike place effects (i.e., scene > object), which are significant in almost all individuals, suggesting that context effects are less reliable (Epstein and Ward, 2010). Another study found that responses in PPA to two objects from the same context (e.g., a refrigerator and a stove) are no more similar than its responses to two objects from different contexts (e.g., a refrigerator and a traffic light), providing further evidence against the claim that PPA processes contextual information about objects (MacEvoy and Epstein, 2011). Finally, studies of contextual associations rarely define PPA using the typical functional localizer (i.e., scenes > objects). Thus, in these cases, it is unknown whether the voxels reported (e.g., in parahippocampal cortex) are in fact the same as those that would be functionally defined as PPA, preventing strong inferences from being drawn across these regions.

Beyond the contextual association hypothesis, however, a variety of studies have at least made it clear that spatial layout information is not the only kind of scene information processed in PPA. Although PPA does not process objects per se, mounting evidence indicates that certain types of objects and object features modulate responses in PPA. For example, PPA is sensitive to furniture (Kamps et al., 2016a; Harel et al., 2012; MacEvoy and Epstein, 2011); larger, less portable objects (Mullally and Maguire, 2011; Konkle and Oliva, 2012); buildings (Aguirre et al., 1998); and large, fixed objects that are well suited to be “landmarks” (Troiani et al., 2014). Importantly, however, PPA does not simply process scenes as a collection of objects; for example, PPA responds as much to apartment rooms devoid of objects as it does to fully furnished rooms (Epstein and Kanwisher, 1998). Beyond these initial clues, however, the reason for such object sensitivity in PPA is unknown, and clearly warrants further investigation. One possibility is that PPA constitutes a specialized system for landmark recognition, and as such is sensitive to larger, fixed objects better suited to be landmarks (Epstein and Vass, 2015). Another possibility is that object information in PPA reflects a larger role for the region in scene

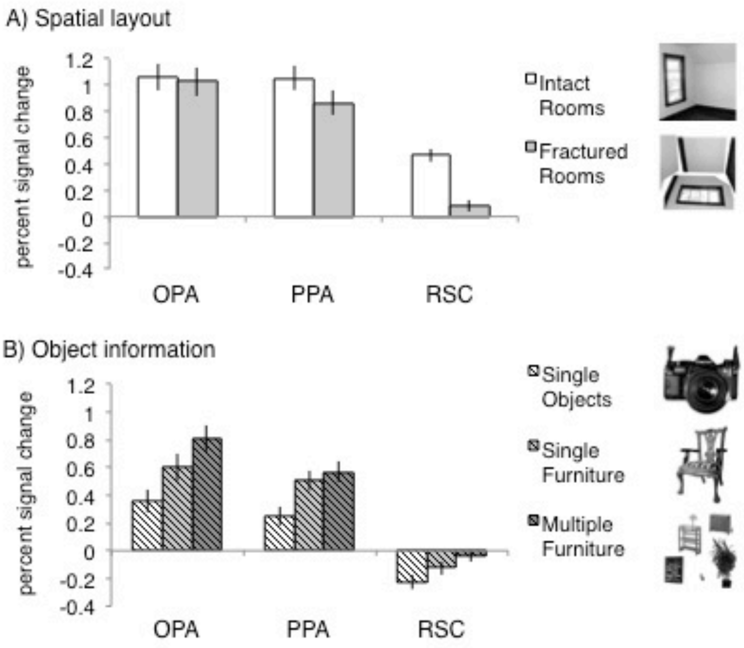

Figure 2. Dissociations of scene and object information processing across OPA, PPA, and RSC. (A) Spatial layout representation: While PPA and RSC are sensitive to spatial layout, responding more to intact than fractured rooms, OPA is relatively insensitive to such spatial layout information, responding to the walls, floors, and ceilings themselves, independent of their spatial arrangement. (B) Object representation: OPA and PPA both respond more to furniture than objects, while RSC does not respond above baseline to any object information. Further, OPA is more sensitive than PPA to the number of objects, responding more to multiple furniture than single pieces of furniture. Figure taken from Kamps et al. (2016).

categorization (e.g., recognizing a beach versus a kitchen). Indeed, larger, less portable, more contextual objects are more likely diagnostic of the category of scene. These possibilities are considered in more detail in the following section.

The role of PPA in scene processing

Ultimately the aim of cognitive neuroscience is to understand not just the kinds of information processing that are carried out in a cortical system, but the precise *role* that a system plays in cognition. In other words, we need to go beyond a simple list of scene properties (e.g., spatial layout, furniture) to which a region is sensitive, in order to understand more broadly *why* a region is sensitive to those kinds of scene information and not others. Given the information processing in PPA explored above, what particular scene behaviors and abilities might the PPA support? A central assumption has been that PPA plays a direct role in navigation (Ghaem et al., 1997; Ino et al., 2002; Janzen and van Turennout, 2004; Rosenbaum et al., 2004; Rauchs et al., 2008) and reorientation (Epstein and Kanwisher, 1998a; Cheng and Newcombe, 2005; Spelke et al., 2010). For example, it has been proposed that PPA is the neural locus of a specialized system for landmark recognition (Epstein and Vass, 2015). In this framework, landmarks are fixed elements in space that may take the form of either single, discrete objects or extended topographical features (e.g., a valley, ridge, or the arrangement of buildings at an intersection) – thus fitting well with the literature reviewed above, which finds that PPA represents both objects (especially large, space-defining objects, which may be more likely to remain fixed in space, and thus be useful as a landmark) and spatial layouts (e.g., the arrangement of extended surfaces that are unique to a particular place) in scenes. Perhaps the best evidence for this landmark hypothesis comes from a study showing a region in parahippocampal cortex, partially overlapping with the approximate anatomical location of the PPA, responds significantly more to objects that had previously been associated with navigational decision points (i.e., to turn left or right) than to objects that were not associated with navigational decisions (Janzen and van Turennout, 2004). However, this study did not use an independent ROI definition (i.e., the same data were used to localize the ROI as were used to explore responses in that ROI) (Kriegeskorte et al., 2009; Vul and Kanwisher, 2010), and no correction for multiple comparisons was conducted, rendering these findings invalid. Further, since the study did not functionally define PPA using the standard contrast (i.e., scenes > objects), it is unknown whether these voxels are indeed within PPA or in adjacent cortex. A subsequent study using a standard, independent definition of PPA found a similar effect for building stimuli previously associated with navigational decisions; however, behavioral performance was not matched between stimuli associated with decision points versus non-decision points, and this effect was also found in a host of other regions. As such, the greater response in PPA to buildings associated with navigational decisions than to buildings with no such association may not reflect landmark encoding in particular, but may rather be the result of general differences in task demands between the two conditions.

Recent findings from my lab have challenged the hypothesis that PPA plays a direct role in navigation. If a region plays a direct role in navigation, then it critically must be sensitive to the spatial relationship between the scene and the viewer. Initial reports of viewpoint-specific scene representations in PPA (Epstein et al., 2003; Epstein et al., 2005; Park and Chun, 2009) suggested that PPA might encode scenes in such an egocentric reference frame. However, these effects disappear after subjects gain familiarity with the scene across the experimental session (Epstein et al., 2005), or when scenes are presented discontinuously rather continuously (Park and Chun, 2009), leaving such findings difficult to interpret. More recently, two studies have shown that PPA is *not* sensitive to two essential kinds of egocentric spatial information – distance (i.e., near versus far) (Persichetti and Dilks, 2016) and "sense" (i.e., left versus right) (Dilks et al., 2011) information – indicating that PPA more likely encodes an allocentric representation of the scene (i.e., information about spatial relationships between elements of the scene, independent of the observer). Supporting this latter claim, PPA is sensitive to other sources of spatial information that are critical for analyzing the overall shape of local space, but not for determining one's position relative to that space, including length and angle information (Dillon et al., 2018).

These studies suggest that PPA is poorly suited to support navigation, and rather suggest that PPA may support the recognition of scene categories (e.g., kitchen versus beach). To more directly test the role of PPA in scene categorization versus navigation, my lab (Persichetti and Dilks, 2018a) recently conducted an fMRI experiment in which participants were asked to perform either a scene categorization task or navigation task while viewing images of scenes. In the scene categorization task, participants imagined standing in a room and indicated whether the scene was a kitchen, living room, or bedroom. In the navigation task, participants imagined walking along a path on the floor that lead to only one of three doors, and indicated the correct door. My colleagues hypothesized that if PPA is involved in scene categorization, but not navigation, then it should respond more during the scene categorization task than during the navigation task. Importantly, the exact same stimuli were used in the two tasks, behavioral performance was matched, and participants fixated on a central fixation cross in both tasks (as verified with eye tracking). Supporting the proposed role for PPA in scene categorization, activity in PPA increased selectively during the scene categorization, but not the navigation task. This finding fits well with other data showing that patterns of PPA activity, as measured using MVPA, can discriminate between scene categories, and are even correlated with behavioral performance on a scene categorization task (Walther et al., 2009).

Development

The growing literature on the function of PPA in adulthood begs the question: How does this region develop? To date, only a few studies have explored the developmental origins of PPA. A consistent finding has been that PPA is reliably detectable as early as age 5 (Golarai et al., 2007; Scherf et al., 2007; Chai et al., 2010; O'hearn et al., 2011; Scherf et al., 2011a; Ofen et al., 2012; Vuontela et al., 2013). Intriguingly, one recent study (Deen et al., 2017) found that scene preferences (i.e., greater responses to scenes than faces) can already be detected in parahippocampal cortex (consistent with the location of PPA) by just 3 to 8 months, but that scene selectivity (i.e., greater responses to scenes than objects) is not yet detectable in PPA at this age. Further, we recently found that connectivity underlying PPA is already intact in as little as 27 days of age – with PPA showing biased connectivity with RSC, as well as with peripheral early visual cortex – suggesting that the connectivity underlying this region develops prior to, and potentially scaffolds, the development of scene selectivity (Kamps et al., 2019b). Thus, while the foundations of the scene processing system may be present early in infancy, the more focal PPA may not emerge until later in childhood.

While PPA is *detectable* as early as 5, it is unclear whether it is already adultlike (i.e., functionally mature) in 5 year olds, or whether it undergoes significant developmental change from childhood to adulthood. To assess functional maturation, researchers have primarily focused on measures of size and selectivity. Studies of how PPA changes in *size* from childhood to adulthood have been remarkably inconsistent. One study found that PPA increases in size (Golarai et al., 2007), another found that PPA decreases in size (Vuontela et al., 2013), and several other studies report no changes in the size of PPA from childhood to adulthood (Scherf et al., 2007; Golarai et al., 2010; Scherf et al., 2011a). Studies of how PPA changes in *selectivity* across these groups have also been inconsistent. Several studies have found no change in selectivity between childhood and adulthood (Scherf et al., 2007; Golarai et al., 2010; Scherf et al., 2011a; Vuontela et al., 2013; Jiang et al., 2014), while two other studies *did* observe developmental increases in PPA selectivity between children and adults (Golarai et al., 2007).

The striking inconsistency of these results is likely due to methodological issues. A fundamental problem in pediatric neuroimaging is that younger children produce lower-quality data on average than older children or adults. In particular, younger children move their heads more and pay less attention than older children or adults (McKone et al., 2012). While challenging, these issues can be addressed by in a variety of ways. Head motion and data quality confounds can be addressed by matching groups on head motion and temporal signal-to-noise ratio (a direct measure of data quality). Attentional confounds can be addressed by careful use of control regions (e.g., if children paid attention to visual stimuli, then strong responses should be detectable in V1), eye-tracking (i.e., ensuring that participants were looking at the stimuli), and/or by matching performance with older groups on a concurrent task (e.g., a 1-back task). As a final note, these efforts can be enhanced through use of functionally defined ROIs, rather than traditional group analyses (which most previous studies of these regions employed). Functionally-defined ROIs offer many advantages over group analyses, most notably by assessing the split-half reliability of responses (i.e., since the effect is defined and tested on two independent halves of the data), and by overcoming issues of individual anatomical variability (i.e., since participants ROIs will not perfectly overlap in stereotaxic space, potentially washing out strong effects). Few studies to date have carefully controlled for these issues, and consequently, there is still much to learn about precisely when and how PPA develops.

Retrosplenial complex

Scene Selectivity

The retrosplenial complex (RSC) is the second best studied scene-selective region of the brain. Like PPA, RSC is a functionally defined region responding significantly more to scenes than objects. The term “retrosplenial complex” is used since the region is not always found directly in retrosplenial cortex, and may also extend into the adjacent posterior cingulate (Maguire, 2001;

Epstein, 2008). Although retrosplenial cortex has been implicated in a variety of cognitive functions beyond scene processing, including episodic memory, mind wandering, and future planning (for review, see Vann et al., 2009), the functionally-defined RSC shows clear scene selectivity, responding significantly more to indoor rooms than objects, furniture, or arrays of multiple pieces of furniture (Kamps et al., 2016b); indoor and outdoor scenes than faces (Henderson et al., 2007); minimal scenes than comparable non-scene gradients (Harel et al., 2012); haptically explored scenes than objects (Wolbers et al., 2011); and imagined scenes than imagined faces (Johnson and Johnson, 2014); reflecting its true involvement in scene processing. Indeed, while RSC shows biases for low-level visual information such as rectilinearity (Nasr et al., 2014) and high visual contrast (Kauffmann et al., 2015), such biases are typically only found in the context of scenes. Further, damage to retrosplenial cortex and posterior cingulate, which presumably encompasses the RSC, leads to specific scene processing deficits (especially related to navigation; see below) despite spared object recognition (for review, see Aguirre and D'Esposito, 1999).

Information Processing

While little is known about the specific information that RSC extracts from scenes, one early conclusion is that unlike PPA, which is sensitive to both spatial layout and object information, RSC is primarily sensitive to spatial layout, and not object information (Harel et al., 2012; Kamps et al., 2016). Evidence of spatial layout processing comes from the finding that RSC responds significantly more to images of intact apartment rooms than to images of these same rooms in which the walls, floors, and ceilings have been fractured and rearranged, thereby disrupting the coherent layout of space (Kamps et al., 2016) (Figure 2). This same study found that RSC did not respond above baseline (i.e., the response during fixation with no stimulus) to large, highly contextual objects such as furniture, or even to arrays of multiple pieces of furniture, indicating that RSC is relatively insensitive to object information. Similarly, another study found that patterns of activity in RSC reliably discriminate open versus closed layouts, but do not contain information about the presence versus absence of furniture in these same minimal scenes (Harel et al., 2012). In contrast to these findings, however, some studies have reported significant object information in RSC. For example, RSC may be sensitive to highly contextual objects (Bar et al., 2008), although these effects may alternately be explained by scene imagery (see discussion of similar findings in PPA above). Interestingly, two studies have found that objects better suited as landmarks (i.e., larger, fixed, more space-defining, and place-related objects) drive responses in RSC more than less "landmark suitable" objects (Auger et al., 2012; Troiani et al., 2014). While these findings are intriguing, more work is clearly necessary to clarify the nature of object representation in RSC, especially with regard to landmark recognition.

The role of RSC in scene processing

Most studies on RSC have explored the role of the region in scene processing, and while a number of proposals have been made, a precise role for RSC is far from established. In general, there is little contention surrounding the broad claim that RSC is involved in navigation (O'Craven and Kanwisher, 2000; Cooper et al., 2001; Maguire, 2001; Cain et al., 2006; Iaria et al., 2007; Dilks et al., 2011). For example, fMRI studies have shown that RSC is sensitive to egocentric spatial cues such as distance (i.e., near versus far) (Persichetti and Dilks, 2016) and sense (i.e., left versus right) (Dilks et al., 2011) information, which are vital for navigating. Likewise, patients with damage to retrosplenial cortex (which may include RSC) present with topographical disorientation, an inability to orient oneself relative to one's surroundings (Aguirre and D'Esposito, 1999; Maguire, 2001). Current work has therefore focused on uncovering the precise cognitive operations that RSC might support in the context of navigation.

One hypothesis is that RSC supports our ability to translate between egocentric representations of the current scene and allocentric representations of the broader spatial map, beyond the current view(i.e., what I call “memory-guided navigation”). This hypothesis was inspired by early neuroanatomical studies of the retrosplenial cortex and posterior cingulate in monkeys (putative homologues of the human RSC), which show strong connectivity with parietal areas thought to encode egocentric spatial information (including area 7a and the lateral intraparietal area), and medial temporal areas thought to encode allocentric spatial information (including entorhinal cortex, pre- and post-subiculum, and parahippocampal regions TF and TH), suggesting that RSC might constitute a natural bridge between these spatial codes (Suzuki and Amaral, 1994; Maguire, 2001; Kobayashi and Amaral, 2003, 2007; Epstein, 2008). Likewise, patients with pure topographical disorientation resulting from damage to retrosplenial cortex are able to recognize familiar landmarks, but are unable to use these landmarks to orient themselves, a function which presumably requires integrating an egocentric representation of one’s viewpoint relative to a landmark with an allocentric representation of the broader spatial map associated with that landmark (Aguirre and D'Esposito, 1999). Such a mechanism for translating between egocentric and allocentric spatial codes has even been made explicit in a computational model of the neural mechanisms underlying spatial cognition, with the authors positing a key role for retrosplenial cortex (Burgess et al., 2001; Byrne et al., 2007). Thus, retrosplenial cortex and surrounding areas are well suited for and have been strongly implicated in navigational operations that require translation between egocentric and allocentric spatial codes.

Other work has focused in particular on the functionally defined RSC, and has generally supported the claim that RSC integrates information about the current scene with representations of the broader spatial environment. For example, RSC shows strong familiarity effects, responding two times more strongly to familiar than unfamiliar scenes (Sugiura et al., 2005; Epstein et al., 2007b), suggesting that it encodes long-term knowledge about familiar scenes. Likewise, RSC is selectively recruited during tasks that require long-term knowledge about the broader spatial environment related to a familiar scene (e.g., determining whether the University of Pennsylvania bookstore is on the north or south side of campus) (Epstein et al., 2007a). Less clear, however, is the precise navigational function that RSC serves in integrating these sources of spatial information. One possibility is localization, where a scene is recognized as a particular place situated in the broader spatial map (e.g., "I am at the psychology building, which is located south of the chemistry building and to the west of the bookstore"). Another possibility is orientation, where one's heading direction is obtained relative to the broader environment (e.g., "I am facing west" or "The grocery store is in that direction [pointing] from here"). Overall, the evidence for location and orientation encoding in RSC is mixed. While some studies find location information in RSC (Epstein et al., 2007a; Vass and Epstein, 2013), others do not (Morgan et al., 2011). Similarly, while one study found orientation information in RSC (Epstein et al., 2007a), another did not (Vass and Epstein, 2013). Yet another study found evidence of both location and orientation information in RSC, but this information was grounded in the geometry of the current scene, independent of the scene's location in the broader spatial environment (Marchette et al., 2015a). Thus, while in general RSC is well suited to integrate information about the current scene with information about the broader spatial environment, the precise function that this translation serves is currently unclear.

Development

How does RSC develop? Relatively few pediatric neuroimaging studies have targeted RSC in particular, and consequently little is known about how scene selectivity develops in this region. Similar to PPA, these studies have found that a preference for scene information (i.e., scenes > faces) is detectable as early as 3 to 8 months old (although RSC was not investigated in as much detail as OPA or PPA in that study) (Deen et al., 2017), while scene selectivity is not detected until sometime later in childhood, with the earliest current reports finding scene selectivity by 7-8 years old (Jiang et al., 2014; Meissner et al., 2019). Likewise, I recently found that adultlike patterns of functional connectivity underlying RSC are present as early as 27 days, with RSC showing biased connectivity with PPA and with peripheral early visual cortex (Kamps et al., 2019b). Thus, connectivity underlying RSC may provide a scaffold for subsequent functional development of RSC extending into childhood.

Intriguingly, unlike PPA, fMRI studies in children have consistently found that RSC is already adultlike in size and selectivity by age 7-8. Given this finding, it has even been hypothesized that RSC develops before PPA (counter to the hypothesis of this thesis that the navigation system develops later than the scene categorization system) (Meissner et al., 2019). However, this study did not find the critical region x age group interaction necessary to support this claim, and it is not clear that this same pattern would hold when data quality is closely matched between children and adults. For example, the lack of a group difference in RSC may have resulted simply from the fact that there was more anatomical variability across subjects in the location of this region, making it harder to detect a group difference (Saxe et al., 2006).

Occipital Place Area

Scene Selectivity

Occipital place area (OPA), formerly known as transverse occipital sulcus, is the most recently discovered scene selective region in cortex (Dilks et al., 2013), and consequently the least is known about it. While its precise anatomical location may vary, OPA is generally found in and around the transverse occipital sulcus. Despite its proximity to early visual cortex – indeed, OPA is consistently located immediately anterior and ventral to the retinotopically defined area V3A (Grill-Spector, 2003; Nasr et al., 2011) – OPA clearly responds more to scenes than objects, both in perception and imagery, findings that are now widely replicated (Nakamura et al., 2000; Grill-Spector, 2003; Hasson et al., 2003; Epstein et al., 2005; Epstein et al., 2007b; MacEvoy and Epstein, 2007; Schwarzlose et al., 2008; Dilks et al., 2011; Bettencourt and Xu, 2013; Johnson and Johnson, 2014). Further, while OPA shows biases to certain low-level visual properties like high spatial frequencies (Kauffmann et al., 2015) and rectilinear features (Nasr et al., 2014), these sources of visual information naturally occur in scenes more than other categories, such as objects. Similarly, while OPA is biased toward information presented in the lower versus the upper visual field (Silson et al., 2015), these effects are stronger when scenes are used to map the retinotopic field, rather than simple checkerboards. Finally, the causal and selective role of OPA in scene processing has now been confirmed in two studies showing that transcranial magnetic stimulation (TMS) applied over OPA selectively impairs scene, but not object or face recognition (Dilks et al., 2013; Ganaden et al., 2013).

Information Processing

Researchers have only recently begun to explore scene information processing in OPA. Nevertheless, a central hypothesis is that unlike PPA and RSC, which represent global scene properties like spatial layout, OPA represents the local elements (e.g., major surfaces, large objects like furniture) of scenes (Kamps et al., 2016). Evidence for this local elements hypothesis comes from two key findings. First, OPA responds similarly to images of empty, intact apartment rooms and images of these same rooms where the major surfaces (i.e., walls, floors, ceilings) are fractured and rearranged, suggesting that OPA does not represent the spatial layout of the major surfaces per se, but rather represents the major surfaces *themselves* (Figure 2) (Kamps et al., 2016). By contrast, PPA and RSC are sensitive to such spatial layout information. Second, OPA, like PPA, is sensitive to particular object information, responding more to images of single pieces of furniture than images of single, non-furniture objects (Kamps et al., 2016), and more to objects that are larger, fixed, and “space-defining” (Konkle and Oliva, 2012; Troiani et al., 2014). Unlike PPA, however, OPA is also sensitive to the *number* of these objects, responding more to multiple pieces of furniture than single pieces of furniture, and more to multiple pieces of furniture than multiple everyday objects (Bettencourt and Xu, 2013; Kamps et al., 2016). These findings suggest that OPA encodes each piece of furniture individually, while PPA encodes global representations of furniture that are independent of the number of pieces presented.

The role of OPA in scene processing

Although the study of OPA is still in its early days, the existing evidence strongly points to a role for OPA in navigation through the local visual environment – what I call “visually-guided navigation”. The finding that OPA analyzes scenes at the level of local elements (i.e., the major surfaces that comprise a space, as well as large objects such as furniture), suggests that OPA encodes boundaries and obstacles in the local visual environment. Next, OPA is sensitive to mirror reversals (Dilks et al., 2011) and distal versus proximal views (Persichetti and Dilks, 2016) of the same scene, indicating that OPA further represents “sense” (i.e., left versus right) and egocentric distance (i.e., near versus far) information in scenes—information that is essential for determining one’s position relative to boundaries and obstacles in the environment.

Supporting this claim, one recent study found that TMS to OPA impaired participants' ability to localize objects relative to boundaries, but not landmarks (Julian et al., 2015). Another recent study found that OPA codes possible routes through a scene, or "navigational affordances" (Bonner and Epstein, 2017). Finally, a recent study found that OPA was activated during a task in which participants were asked to imagine walking through a scene by following a path along the floor, but not when participants were asked to categorize (e.g., bedroom, kitchen, or living room) the exact same scene (Persichetti and Dilks, 2018). Intriguingly, our hypothesized role for OPA in guiding navigation through the local visual environment is in line with neuropsychological data from patients with damage to posterior parietal cortex (which may encompass OPA), who present with egocentric disorientation, which is characterized by a profound inability to localize objects with respect to the self, despite intact scene and landmark recognition (Stark et al., 1996; Wilson et al., 2005). While this possibility is tantalizing, these patients have not been tested rigorously for scene processing deficits in particular, and may show deficits in nearly all egocentric-based visuospatial tasks, including those involving action on objects (For review, see Aguirre et al., 1999). Further, it is possible that damage to other regions beyond OPA could account for such deficits. Indeed, there is growing evidence for a scene-selective region in the superior parietal lobe/intraparietal sulcus (Baldassano et al., 2013; Marchette et al., 2015a), well positioned to play a role in navigation (Kamps et al., 2016), which may also be damaged in this patient group. Future work will therefore be necessary to confirm the precise relationship between OPA damage and egocentric disorientation.

Development

Like RSC, the development of OPA is poorly understood, with only two studies to date investigating its development. These studies, like those for PPA and RSC, have focused on the development of scene selectivity, with a similar pattern found for OPA as that in PPA and RSC: specifically, a basic preference for scenes (i.e., greater responses to scenes than faces) is found by 3 to 8 months, while a full-fledged scene-selective response does not emerge until later in childhood, with the earliest reports to date finding scene selectivity by 7 to 8 years old.

Distinct systems for scene navigation and scene categorization

As outlined above, a central assumption since the discovery of scene-selective regions in human cortex has been that the scene-processing system is a monolithic system in the service of navigation (for reviews and theoretical perspectives, see Epstein, 2008 and Epstein and Vass, 2014). However, recent work from our lab has directly challenged this widely held assumption. While two scene selective regions – OPA and RSC – are sensitive to critical information for navigation, such as sense and distance information, a third region – PPA – is *not* sensitive to such navigationally-relevant information (Dilks et al., 2011; Persichetti and Dilks, 2016). As such, PPA lacks essential information for orienting oneself to the current scene, and therefore cannot play a direct role in

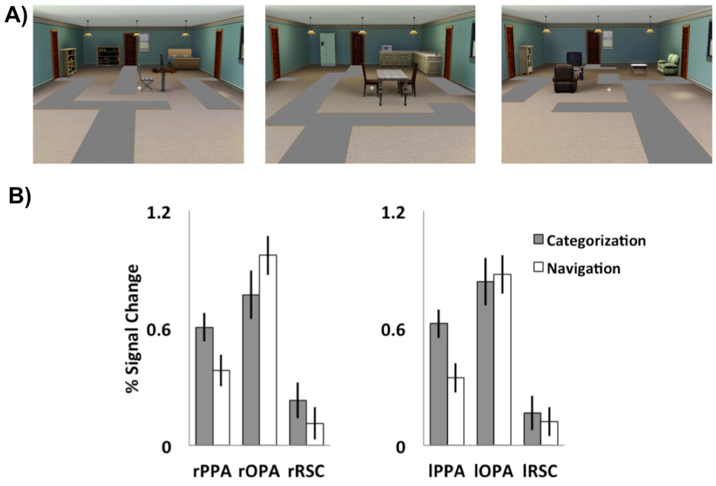

Figure 3. A double dissociation between recognizing places and navigating through them. A) Stimuli from the experiment (Persichetti et al., 2018). On these same stimuli, participants performed both a *categorization task*, in which they identified the room as either a bedroom, kitchen or living room, and a *navigation task*, in which they imagined walking along the path on the floor that could only be completed for one of the three doors (left, middle, right). B) rOPA responded significantly more during the navigation than categorization task, while rPPA responded significantly more during the categorization task. rRSC responded similarly to both tasks. Similar findings were obtained in the left hemisphere, with the exception of lOPA, which showed no difference in response between the two tasks. Figures adapted from Persichetti et al. (2018).

navigating scenes. We have therefore proposed that the human visual scene processing system contains at least two systems: one for navigation (including OPA and RSC), and a second for scene categorization (including PPA). In a stronger test of this “two systems for scene processing hypothesis”, my lab recently found a double dissociation between the two systems, with OPA responding more during a navigation task than a scene categorization task, and PPA responding more during a scene categorization task than a navigation task (Persichetti et al., 2018) (Figure 3). Interestingly, RSC responded similarly during both tasks, suggesting that OPA and RSC play different roles in scene navigation, a topic I explore in the following section.

While the involvement of OPA and RSC in navigation is relatively uncontroversial, our hypothesis that PPA plays a role in scene categorization, but not navigation, runs counter to numerous studies highlighting the role of this region in navigation (Ghaem et al., 1997; Epstein and Kanwisher, 1998a; Janzen and van Turennout, 2004; Rosenbaum et al., 2004; Cheng and Newcombe, 2005; Epstein, 2008; Rauchs et al., 2008; Spelke et al., 2010; Epstein and Vass, 2015). However, I argue that few studies provide direct evidence for this claim, and that many of these findings are in fact open to alternative interpretations. For example, Epstein and Kanwisher (1998) argued that spatial layout sensitivity in PPA signaled its role in reorientation. While spatial layout is indeed critical for reorientation (e.g., Hermer and Spelke, 1996), such information is also critical for distinguishing categories of scenes, as revealed by numerous behavioral and computational studies of scene processing (Oliva and Schyns, 1997; Oliva and Torralba, 2001; Greene and Oliva, 2009; Walther et al., 2011). Two other studies found that a region in posterior parahippocampal gyrus (only one study functionally defined the PPA) responded more strongly to objects that were previously associated with navigational decisions (i.e., landmarks) than objects that were not (Janzen and van Turennout, 2004; Schinazi and

Epstein, 2010). While these findings would indeed be interesting, these studies have significant methodological limitations as previously discussed. Finally, patients with damage to parahippocampal and lingual gyri (areas encompassing and surrounding the PPA) commonly present with landmark agnosia, an inability to recognize landmarks (Aguirre and D'Esposito, 1999). However, it is unclear whether such patients are deficient in landmark recognition in particular, or in scene recognition more generally; while some patients appear to show a deficit in landmark recognition despite preserved scene recognition (Pallis, 1955), others appear to be unable to represent a scene as a coherent whole (Whiteley and Warrington, 1978; Hecaen et al., 1980; Takahashi et al., 1989), which would likely affect scene categorization abilities as well, making it impossible to draw any clear conclusions from this literature. Future work is therefore required to test such patients systematically against the predictions outlined in this review, especially with regard to landmark recognition versus scene categorization.

We have therefore proposed a “multiple systems” framework for scene processing, with distinct systems for navigation and categorization. While navigation is no doubt crucial to our successful functioning, the ability to recognize the category of scenes (e.g., kitchen versus beach) is also crucial in everyday life, since this ability makes it possible to know what to expect from, and how to behave in, different environments (e.g., you wouldn’t want to behave in a classroom the way you behave in a bathroom) (Bar, 2004). This multiple systems for scene processing hypothesis is analogous to Milner and Goodale’s (1992) classic distinction between systems for recognition and action in object processing. Just as in object processing, where one need not recognize the category of an object (e.g., “cup”) in order to successfully reach out for it (and vice versa), in scene processing, one need not recognize the category of a scene (e.g., “beach”) in order to successfully navigate it (and vice versa). Further, while scene category information is potentially useful for navigation, as in landmark-based navigation (e.g., turn left at the coffee shop on 10th street), recognizing scene categories is not integral to navigation, since scene categories by definition do not describe unique places in the world, must abstracted from details about the location or orientation of the scene in space (i.e., understanding that a room is a "kitchen" tells you nothing about where you are relative to the living room), and are therefore no different from any other visual cue (e.g., no one would suggest that visual word form recognition is essential to navigation, despite common use of words as landmarks, as when directing someone to turn left on "Peachtree Street"). Note that our distinction between the navigation and scene categorization systems does not preclude interaction between these systems (or any other system); however, this interaction need not be essential to the operations performed in either system.

Two scene navigation systems

Beyond evidence that the human visual scene processing system contains separate processing systems for navigation and scene categorization, recent findings suggest a division of labor even within the navigation system, with two distinct systems supporting altogether different kinds of navigation. In particular, one system in OPA may support navigation through the currently visible scene, such as how to get from one side of the room to another, while a second system in RSC may support navigation from the current scene to locations in the broader spatial environment, such as how to get home from the grocery store.

Several lines of evidence support this "two scene navigation systems" hypothesis. First, OPA and RSC represent distinct kinds of scene information. While OPA is sensitive to local elements of scenes, such as major surfaces (i.e., boundaries) and large objects like furniture (i.e., obstacles), RSC does not represent any object information (Kamps et al., 2016a; Harel et al., 2012; although see Trolioni et al., 2012 and Bar, 2003). Further, while OPA does not represent the arrangement of walls, floors and ceilings that comprise a space, RSC shows strong sensitivity to such spatial layout information (Kamps et al., 2016). This effective double dissociation in object versus spatial layout information processing suggests that OPA and RSC support different navigational abilities. Indeed, in order to guide locomotion through the local visual environment, it is critical that a system represents boundaries and obstacles in that environment. By contrast, spatial layout information (i.e., local geometry) is critical to recovering heading during reorientation (e.g., Hermer and Spelke, 1996), and therefore is useful for a system that supports navigation from the current scene to locations in the broader spatial map.

A final, tentative piece of evidence for the "two scene navigation systems" hypothesis comes the neuropsychological literature on patients with topographical disorientation. Aguirre et al. (1999) proposed a taxonomy of spatial navigation deficits in which they argued for a distinction between cases of egocentric disorientation, where patients show a profound deficit in localizing objects in space relative to the body (e.g., Stark et al., 1996), and cases of heading disorientation, where participants are unable to derive directional information from the currently visible scene (e.g., case 2 in Takahashi et al., 1997). Egocentric disorientation is related to posterior parietal damage, whereas heading disorientation is related to retrosplenial and/or posterior cingulate damage, sites that may encompass OPA and RSC, respectively. However, while the similarity between this neurological dissociation and the proposed division of labor between OPA and RSC is striking, future experiments are need to rigorously test such patients in terms of our "two scene navigation systems" hypothesis, and clarify the precise contributions of OPA and RSC to these deficits.

Development of the multiple scene processing systems

Given previous work in object processing suggesting that the dorsal “action” system develops later than the ventral “perception” system (Dilks et al., 2008), I hypothesized that a similar developmental dissociation will be found in the case of scene processing, with the more dorsal navigation systems developing later than the more ventral categorization system. Critically, no pediatric neuroimaging studies to date have tested this hypothesis, for several reasons. First, most studies have only investigated a single scene-selective region (i.e., PPA), rather than multiple scene-selective regions, precluding comparisons between those regions. Second, even among those studies that did investigate multiple scene regions, none to date have directly compared these regions to one another (i.e., by testing for a statistical interaction of age group and region, which is required to support the claim that one region develops differently from another region across the age range tested). Third, most studies have only investigated the development of scene selectivity (i.e., greater responses to scenes than objects), and have not explored the development of the more specific information processing in each region. Indeed, it could be the case that scene selectivity develops early in all three regions, effectively marking each region as a future sight of scene processing more generally, while the more precise function of each region in categorization or navigation only emerges later in development. In this case, scene selectivity may simply reflect the developmental scaffold upon which each functional region is ultimately built, and as a consequence, understanding when scene selectivity emerges in each region will tell us little about how categorization and navigation systems develop.

Likewise, despite some behavioral work on the development of scene perception, and a large literature on the development of navigation more generally, no study has directly tested the hypothesis that the navigation system develops later than the scene categorization system. This lack of evidence likely stems from the fact that scene categorization and navigation abilities have traditionally been studied in separate literatures. For example, almost all behavioral studies of the development of scene perception have compared perception or memory for scenes to that for faces or objects, not navigation (Golarai et al., 2007; Weigelt et al., 2014). Furthermore, although there is a substantial literature on navigation more generally (Newcombe, 2019), this work has largely focused on the development of spatial memory systems that are thought to be critical for navigation through the large scale environment; far less work has focused on the development of the visually-guided navigation system for finding one's way through the immediately visible environment, and none of these studies have directly compared this visually-guided navigation to scene categorization.

Despite these limitations, the current literature provides some initial support for the hypothesis that the navigation system develops later than the categorization system. For example, in the navigation literature, while many remarkable abilities are at least operational very early in life, such as the ability to recover ones orientation relative to local boundaries (Hermer and Spelke, 1994; Hermer and Spelke, 1996), or to understand that a cliff is not safe to walk over (Gibson and Walk, 1960), other navigational abilities undergo protracted development. For example, the ability to remember precise spatial locations relative to environmental boundaries undergoes protracted development across childhood (Julian et al., 2019) – an ability that depends on OPA in adulthood (Julian et al., 2016). Likewise, obstacle avoidance behaviors are still being refined well into childhood (Pryde et al., 1997; Berard and Vallis, 2006), as is the ability to use peripheral vision to navigate obstacles during locomotion (Franchak and Adolph, 2010a). These findings are consistent with the hypothesis that the navigation system develops late. Other work suggests that basic scene perception abilities will be relative early developing. For example, scene recognition abilities (i.e., distinguishing two houses from one another based on the arrangement of walls, floors, and ceilings – information potentially represented by the PPA) mature more quickly than face perception abilities (Golarai et al., 2007; Weigelt et al., 2014).

Interim Summary

Building on the work above, in the following three Papers, I systematically test the hypotheses that i) scene processing depends on dissociable navigation and categorization systems, and ii) that the navigation system develops later than the categorization system. Given the hypothesis that the navigation system can be further divided into systems for visually-guided navigation (including OPA) and memory-guided navigation (including RSC), and the fact that the visually-guided navigation system in OPA is far less well understood, I will primarily test these hypotheses by drawing comparisons between the scene categorization system and the visually-guided navigation systems. Future work will compare scene categorization and memory-guided navigation for a complete test of the overall hypotheses. In Paper 1, I directly test the idea that OPA supports visually-guided navigation, while PPA does not, by measuring responses in these regions to first-person perspective motion information – a proxy for the visual experience of actually walking through a scene. Next, in Paper 2, I seek causal evidence for the proposed divison of labor between these systems by studying individuals with Williams syndrome (WS), a developmental disorder involving cortical thinning in and around OPA, predicting that WS adults will consequently be impaired in visually-guided navigation, but not scene categorization. Further, in Paper 2, I directly test the hypothesis that visually-guided navigation develops later than scene categorization by testing and comparing these abilities in children ages 4 to 7. Finally, in Paper 3, I seek converging neural evidence for hypothesis that the navigation system is late developing by studying the emergence of first-person perspective motion processing in children ages 5 to 8.

Paper 1

The occipital place area represents first-person perspective motion information through scenes

Frederik S. Kamps, Vishal Lall, Daniel D. Dilks

Abstract

Neuroimaging studies have identified multiple scene-selective regions in human cortex, but the precise role each region plays in scene processing is not yet clear. It was recently hypothesized that two regions, occipital place area (OPA) and retrosplenial complex (RSC), play a direct role in navigation, while a third region, parahippocampal place area (PPA), does not. Some evidence suggests a further division of labor even among regions involved in navigation: While RSC is thought to support navigational processes through the broader environment, OPA may be involved in navigation through the immediately visible environment, although this role for OPA has never been tested. Here we predict that OPA represents first-person perspective motion information through scenes, a critical cue for "visually-guided navigation", consistent with the hypothesized role for OPA. Response magnitudes were measured in OPA (as well as RSC and PPA) to i) video clips of first-person perspective motion through scenes ("Dynamic Scenes"), and ii) static images taken from these same movies, rearranged such that first-person perspective motion could not be inferred ("Static Scenes"). As predicted, OPA responded significantly more to the Dynamic than Static Scenes, relative to both RSC and PPA. The selective response in OPA to Dynamic Scenes was not due to domain-general motion sensitivity or to low-level information inherited from early visual regions. Taken together, these findings suggest the novel hypothesis that OPA may support visually-guided navigation, insofar as first-person perspective motion information is useful for such navigation, while RSC and PPA support other aspects of navigation and scene recognition.

Introduction

Recognizing the visual environment, or “scene”, and using that information to navigate is critical in our everyday lives. Given the ecological importance of scene recognition and navigation, it is perhaps not surprising then that we have dedicated neural machinery for scene processing: the occipital place area (OPA) (Dilks et al., 2013), the retrosplenial complex (RSC) (Maguire, 2001), and the parahippocampal place area (PPA) (Epstein and Kanwisher, 1998a). Beyond establishing the general involvement of these regions in scene processing, however, a fundamental and yet unanswered question remains: What is the precise function of each region in scene processing, and how do these regions support our crucial ability to recognize and navigate our environment?

Growing evidence indicates that OPA, RSC, and PPA play distinct roles in scene processing. For example, OPA and RSC are sensitive to two essential kinds of information for navigation: sense (i.e., left versus right) and egocentric distance (i.e., near versus far from me) information (Dilks et al., 2011; Persichetti and Dilks, 2016). By contrast, PPA is not sensitive to either sense or egocentric distance information. The discovery of such differential sensitivity to navigationally-relevant information across scene-selective cortex has lead to the hypothesis that OPA and RSC directly support navigation, while PPA does not (Dilks et al., 2011; Persichetti and Dilks, 2016). Further studies suggest that there may be a division of labor even among those regions involved in navigation, although this hypothesis has never been tested directly. In particular, RSC is thought to represent information about both the immediately visible scene and the broader spatial environment related to that scene (Maguire, 2001; Epstein, 2008), in order to support navigational processes such as landmark-based navigation (Auger et al., 2012; Epstein and Vass, 2015), location and heading retrieval (Epstein et al., 2007a; Vass and Epstein, 2013; Marchette et al., 2014), and the formation of environmental survey knowledge (Wolbers and

Buchel, 2005; Auger et al., 2015). By contrast, although little is known about OPA, it was recently proposed that OPA supports visually-guided navigation and obstacle avoidance in the immediately visible scene itself (Kamps et al., 2016b).

One critical source of information for such visually-guided navigation is the first-person perspective motion information experienced during locomotion (Gibson, 1950). Thus, here we investigated how OPA represents first-person perspective motion information through scenes. Responses in the OPA (as well as RSC and PPA) were measured using fMRI while participants viewed i) 3-s video clips of first-person perspective motion through a scene ("Dynamic Scenes"), mimicking the actual visual experience of locomotion, and ii) 3, 1-s still images taken from these same video clips, rearranged such that first-person perspective motion could not be inferred ("Static Scenes"). We predicted that OPA would respond more to the Dynamic Scenes than the Static Scenes, relative to both RSC and PPA, consistent with the hypothesis that OPA supports visually-guided navigation, since first-person perspective motion information is undoubtedly useful for such navigation, while RSC and PPA are involved in other aspects of navigation and scene recognition.

Methods

Participants

Sixteen healthy university students (ages 20- 35; mean age = 25.9; sd = 4.3; 7 females) were recruited for this experiment. All participants gave informed consent. All had normal or corrected to normal vision; were right handed (one reported being ambidextrous), as measured by the Edinburgh Handedness Inventory (mean = 0.74; SD = 0.31, where +1 is considered a neurological or psychiatric conditions. All procedures were approved by the Emory University Institutional Review Board.

Design

For our primary analysis, we used a region of interest (ROI) approach in which we used one set of runs (Localizer runs, described below) to define the three scene-selective regions (as described previously; Epstein and Kanwisher, 1998), and then used a second, independent set of runs (Experimental runs, described below) to investigate the responses of these regions to Dynamic Scenes and Static Scenes, as well as two control conditions: Dynamic Faces and Static Faces. As a secondary analysis, we performed a group-level analysis exploring responses to the Experimental runs across the entire slice prescription (for a detailed description of this analysis see Data analysis section below).

For the Localizer runs, we used a standard method used in many previous studies to identify ROIs (Epstein and Kanwisher, 1998a; Walther et al., 2009; Park et al., 2011; Kanwisher and Dilks, 2013; Kamps et al., 2016b). Specifically, a blocked design was used in which participants viewed static images of scenes, faces, objects, and scrambled objects. We defined scene-selective ROIs using static images, rather than dynamic movies for two reasons. First, using the standard method of defining scene-selective ROIs with static images allowed us to ensure that we were investigating the same ROIs investigated in many previous studies of cortical scene processing, facilitating the comparison of our results with previous work. Second, the use of dynamic movies to define scene-selective ROIs could potentially bias responses in regions that are selective to dynamic information in scenes, inflating the size of the "dynamic" effect. The same argument, of course, could be used for the static images (i.e., potentially biasing responses in regions that are selective to static information in scenes, again inflating the size of the "dynamic" effect). However, note that in either case, the choice of dynamic or static stimuli to define scene-selective ROIs would result in a main effect of motion (i.e., a greater response to

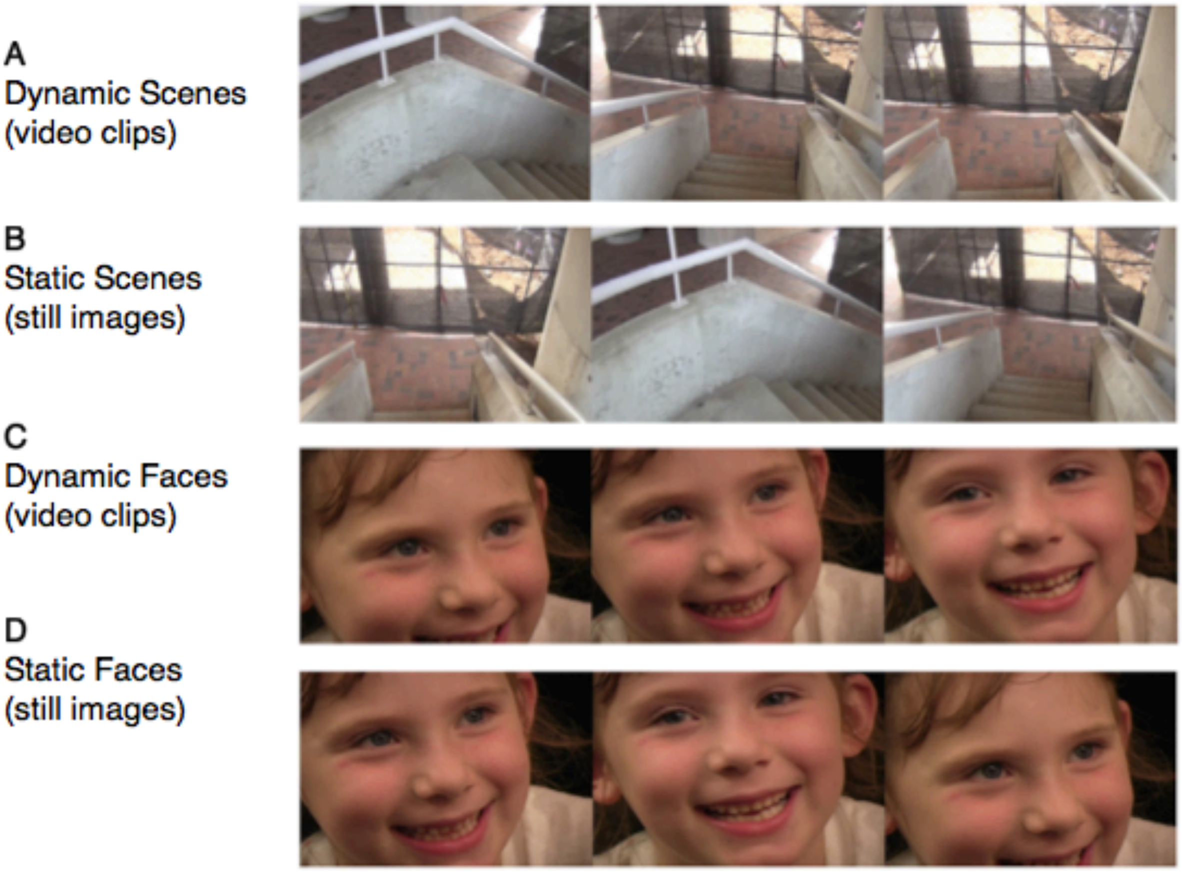

Figure 1. Example stimuli used in the experimental scans. The conditions included A) Dynamic Scenes, which consisted of 3-s video clips of first-person perspective motion through a scene; B) Static Scenes, which consisted of 3 1-s stills taken from the Dynamic Scenes condition and presented in a random order, such that first-person perspective motion could not be inferred; C) Dynamic Faces, which consisted of 3-s video clips of only the faces of children against a black background as they interacted with off-screen adults or toys; and D) Static Faces, which consisted of 3 1-s stills taken from the Dynamic Faces and presented in a random order.

Dynamic Scenes than Static Scenes in all three scene-selective regions, or vice versa), not an interaction of motion by ROI (i.e., a greater response in OPA to Dynamic Scenes than Static Scenes, relative to PPA and RSC), as predicted. Each participant completed 3 runs, with the exception of two participants who only completed 2 runs due to time constraints. Each run was 336 s long and consisted of 4 blocks per stimulus category. For each run, the order of the first eight blocks was pseudorandomized (e.g., faces, faces, objects, scenes, objects, scrambled objects, scenes, scrambled objects), and the order of the remaining eight blocks was the palindrome of the first eight (e.g., scrambled objects, scenes, scrambled objects, objects, scenes, objects, faces, faces). Each block contained 20 images from the same category for a total of 16 s blocks. Each image was presented for 300 ms, followed by a 500 ms interstimulus interval (ISI), and subtended 8 x 8 degrees of visual angle. We also included five 16 s fixation blocks: one at the beginning, and one every four blocks thereafter. Participants performed a one-back task, responding every time the same image was presented twice in a row.

For the Experimental runs, the Dynamic Scene stimuli consisted of 60, 3-s video clips depicting first-person perspective motion, as would be experienced during locomotion through a scene. Specifically, 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 clip at 1-, 2- and 3-s 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 one second 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. Next, to test the specificity of any observed differences between Dynamic Scene and Static Scene conditions, we also included Dynamic Face and Static Face conditions. The Dynamic Face stimuli were the same as those used in Pitcher et al. (2011), and 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 Static Face stimuli were created and presented using the exact same procedure and parameters described for the Static Scene condition above.

Participants completed 3 “dynamic” runs (i.e., blocks of Dynamic Scene and Dynamic Face conditions) and 3 “static” runs (i.e., blocks of Static Scene and Static Face conditions). The dynamic and static runs were interspersed within participant, and the order of runs was counterbalanced across participants. Separate runs of dynamic and static stimuli were used for two reasons. First, the exact same design had been used previously to investigate dynamic face information representation across face-selective regions (Pitcher et al., 2011), which allowed us to compare our findings in the face conditions to those of Pitcher and colleagues, validating our paradigm. Second, we wanted to prevent the possibility of “contamination” of motion information from the Dynamic Scenes to the Static Scenes, as could be the case if they were presented in the same run, reducing any differences we might observe between the two conditions. Each run was 315 s long and consisted of 8 blocks of each condition. For each run, the order of the first eight blocks was pseudorandomized (e.g., faces, scenes, scenes, faces, scenes, scenes, faces, faces), and the order of the remaining eight blocks was the palindrome of the first eight (e.g., faces, faces, scenes, scenes, faces, scenes, scenes, faces). In the dynamic runs, each block consisted of 5, 3-s movies of Dynamic Scenes or Dynamic Faces, totaling 15 s per block. In the static runs, each block consisted of 5 sets of 3-1s images of Static Scenes or Static Faces, totaling 15s per block. We also included five 15 s fixation blocks: one at the beginning, and one every four blocks thereafter. During the Experimental runs, participants were instructed to passively view the stimuli.

fMRI scanning.

All scanning was performed on a 3T Siemens Trio scanner in the Facility for Education and Research in Neuroscience at Emory University. Functional images were acquired using a 32-channel head matrix coil and a gradient-echo single-shot echoplanar imaging sequence (28 slices, TR = 2 s, TE = 30 ms, voxel size = 1.5 x 1.5 x 2.5 mm, and a 0.25 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 purposes of registration and anatomical localization (see Data analysis).

Data analysis.

fMRI data analysis was conducted using the FSL software (FMRIB's Software Library; www.fmrib.ox.ac.uk/fsl) (Smith et al., 2004) and the FreeSurfer Functional Analysis Stream (FS-FAST; http://surfer.nmr.mgh.harvard.edu/). ROI analysis was conducted using the FS-FAST ROI toolbox. Before statistical analysis, images were motion corrected (Cox and Jesmanowicz, 1999). Data were then detrended and fit using a double gamma function. All data were spatially smoothed with a 5-mm kernel. After preprocessing, scene-selective regions OPA, RSC, and PPA were bilaterally defined in each participant (using data from the independent localizer scans) as those regions that responded more strongly to scenes than objects ($p < 10^{-4}$, uncorrected), as described previously (Epstein and Kanwisher, 1998), and further constrained using a published atlas of "parcels" that identify the anatomical regions within which most subjects show activation for the contrast of scenes minus objects (Julian et al., 2012). We also defined several control regions. First, we functionally defined foveal cortex (FC) using the contrast of scrambled objects > objects, as previously described (Kamps et al., 2016b), using data from the localizer scans. Second, using an independent dataset from another experiment that included the same four experimental conditions used here (Dynamic Scenes, Static Scenes, Dynamic Faces, Static Faces), but that tested different hypotheses, we were able to functionally define middle temporal area (MT) (Tootell et al., 1995) as the region responding more to both Dynamic Scenes and Dynamic Faces than to both Static Scenes and Static Faces in 8 of our 16 participants. The number of participants exhibiting each region in each hemisphere was as follows: rOPA: 16/16; rRSC: 16/16; rPPA: 16/16; rFC: 14/16; rMT: 8/8; lOPA: 16/16; lRSC: 16/16; lPPA: 16/16; lFC: 15/16; lMT: 8/8. Within each ROI, we then calculated the magnitude of response (percent signal change) to the Dynamic Scenes and Static Scenes, using the data from the Experimental runs. A 2 (hemisphere: Left, Right) x 2 (condition: Dynamic Scenes, Static Scenes) repeated-measures ANOVA was conducted for each scene ROI. We found no significant hemisphere by condition interaction in OPA ($F_{(1,15)} = 0.04$, $p = 0.85$), RSC ($F_{(1,15)} = 0.38$, $p = 0.55$), PPA ($F_{(1,15)} = 2.28$, $p = 0.15$), FC ($F_{(1,13)} = 0.44$, $p = 0.52$), or MT ($F_{(1,7)} = 1.72$, $p = 0.23$). Thus, both hemispheres for each ROI were collapsed for further analyses. After collapsing across hemispheres, the number of participants exhibiting each ROI in at least one hemisphere was as follows: OPA: 16/16; RSC: 16/16; PPA: 16/16; FC: 15/16; MT: 8/8 (Supplemental Figure 1).

Finally, in addition to the ROI analysis described above, we also performed a group-level analysis to explore responses to the experimental conditions across the entire slice prescription. This analysis was conducted using the same parameters as were used in the ROI analysis, with the exception that the experimental data were registered to standard stereotaxic (MNI) space. We performed two contrasts: Dynamic Scenes vs. Static Scenes and Dynamic Faces vs. Static Faces. For each contrast, we conducted a nonparametric one-sample t-test using the FSL randomize program (Winkler et al., 2014) with default variance smoothing of 5 mm, which tests the t value at each voxel against a null distribution generated from 5,000 random permutations of group membership. The resultant statistical maps were then corrected for multiple comparisons (*p* < 0.01, FWE) using threshold-free cluster enhancement (TFCE), a method that retains the power of cluster-wise inference without the dependence on an arbitrary cluster-forming threshold (Smith and Nichols, 2009).

Results

OPA represents first-person perspective motion information through scenes, while RSC and PPA do not

Here we predicted that OPA represents first-person perspective motion information through scenes, consistent with the hypothesis that OPA plays a role in visually-guided navigation. Confirming our prediction, a 3 (ROI: OPA, RSC, PPA) x 2 (condition: Dynamic Scenes, Static Scenes) repeated-measures ANOVA revealed a significant interaction ($F_{(2,30)} = 8.09$, $p = 0.002$, $\eta_p^2 = 0.35$), with OPA responding significantly more to the Dynamic Scene condition than the Static Scene condition, relative to both RSC and PPA (interaction contrasts, both p's < 0.05). By contrast, RSC and PPA responded similarly to the Dynamic Scene and Static Scene conditions (interaction contrast, $p = 0.44$) (Figure 2). Importantly, the finding of a significant interaction across these regions rules out the possibility that differences in attention between the conditions drove these effects, since such a difference would cause a main effect of condition, not an interaction of region by condition. Thus, taken together, these findings demonstrate that OPA represents first-person perspective motion information through scenes—a critical cue for visually-guided navigation—while PPA and RSC do not, supporting the hypothesized role of OPA in visually-guided navigation.

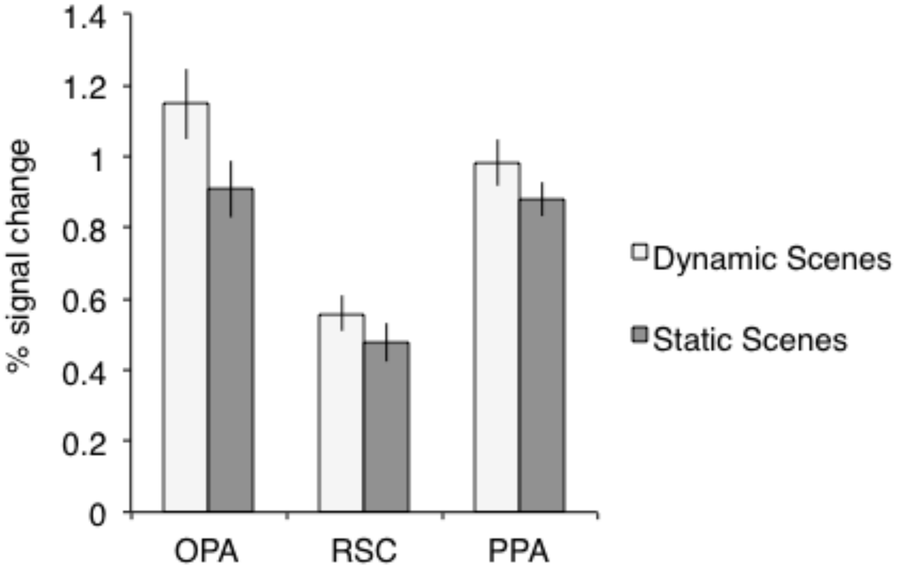

Figure 2. Average percent signal change in OPA, RSC, and PPA to the Dynamic Scenes condition, depicting first-person perspective motion information through scenes, and the Static Scenes condition, in which first-person perspective motion was disrupted. OPA responded more to the Dynamic Scenes than Static Scenes, relative to both RSC and PPA ($F_{(2,30)} = 8.091$, $p = 0.002$), suggesting that OPA selectively represents first-person perspective motion information, a critical cue for visually-guided navigation.

But might OPA be responding to motion information more generally, rather than motion information in scenes, in particular? To test this possibility, we compared the difference in response in OPA to the Dynamic Scene (with motion) and Static Scene (without motion) conditions ("Scene difference score"), with the difference in response to Dynamic Faces (with motion) and Static Faces (without motion) ("Face difference score") (Figure 4). A paired samples t-test revealed that the Scene difference score was significantly greater than the Face difference score in OPA ($t_{(15)} = 6.96$, $p < 0.001$, $d = 2.07$), indicating that OPA does not represent motion information in general, but rather selectively responds to first-person perspective motion in scenes. Of course, it is possible that OPA may represent other kinds of motion information in scenes beyond the first-person perspective motion information tested here, a question we explore in detail in the Discussion.

However, given that we did not precisely match the amount of motion information between scene and face stimuli, might OPA be responding more to Dynamic Scenes than Dynamic Faces because Dynamic Scenes have more motion information than Dynamic Faces, rather than responding specifically to scene-selective motion? To test this possibility, we compared the Scene difference score and the Face difference score in OPA with those in MT – a domain-general motion-selective region. A 2 (region: OPA, MT) x 2 (condition: Scene difference score, Face difference score) repeated-measures ANOVA revealed a significant interaction ($F_{(1,7)} = 95.41$, $p < 0.001$, $\eta_p^2 = 0.93$), with OPA responding significantly more to motion information in scenes than faces, and MT showing the opposite pattern (Bonferroni corrected post-hoc comparisons, both p's < 0.05). The greater response to face motion information than scene motion information in MT suggests that in fact there was more motion information present in the Dynamic Faces than Dynamic Scenes, ruling out the possibility that differences in the amount of motion information in the scene stimuli compared to the face stimuli can explain the selective response in OPA to motion information in scenes.

Responses in OPA do not reflect information inherited from low-level visual regions

While the above findings suggest that OPA represents first-person perspective motion information in scenes—unlike RSC and PPA—and further that OPA does not represent motion information in general, might it still be the case that the response of OPA simply reflects visual information inherited from low-level visual regions? To rule out this possibility, we compared the Scene difference score (Dynamic Scenes minus Static Scenes) and Face difference score (Dynamic Faces minus Static Faces) in OPA with those in FC (i.e., a low-level visual region)

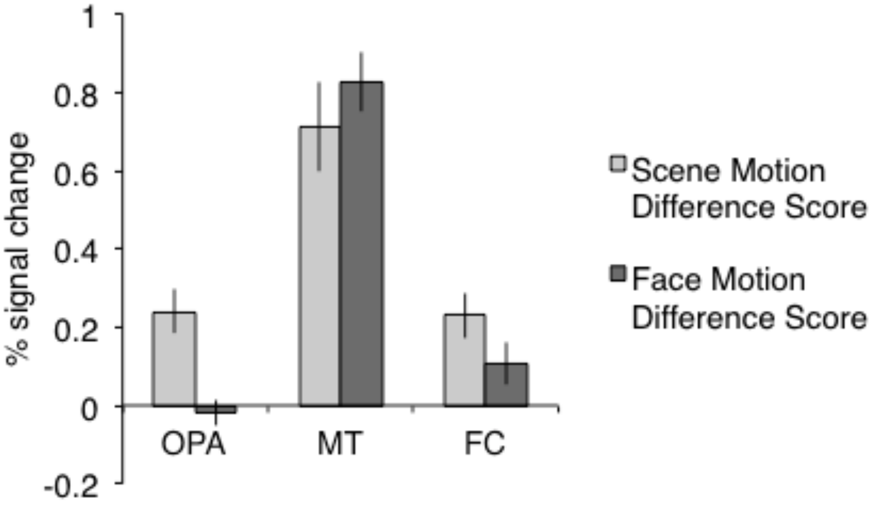

Figure 3. Difference scores (percent signal change) for Dynamic Scenes minus Static Scenes ("Scene Motion Difference Score") and Dynamic Faces minus Static Faces ("Face Motion Difference Score") in OPA, MT (a motion-selective region), and FC (a low-level visual region). OPA responded significantly more to scene motion than face motion, relative to both MT ($F_{(1,7)} = 95.41$, $p < 0.001$) and FC ($F_{(1,14)} = 9.96$, $p < 0.01$), indicating that the response to scene-selective motion in OPA does not reflect differences in the amount of motion information in the scene stimuli compared to the face stimuli, or information inherited from low-level visual regions.

(Figure 3). A 2 (ROI: OPA, FC) x 2 (condition: Scene difference score, Face difference score) repeated-measures ANOVA revealed a significant interaction ($F_{(1,14)} = 9.96$, $p < 0.01$, $\eta_p^2 = 0.42$), with the Scene difference score significantly greater than the Face difference score in OPA, relative to FC. This finding reveals that OPA is not simply inheriting information from a low-level visual region, but rather is responding selectively to first-person perspective motion information through scenes.

Do regions beyond OPA represent first-person perspective motion through scenes?

To explore whether regions beyond OPA might also be involved in representing first-person perspective motion through scenes, we performed a group-level analysis examining responses across the entire slice prescription (Figure 4, Table 1). If a region represents first-person perspective motion through scenes, then it should respond significantly more to the Dynamic Scene condition than the Static Scene condition ($p = 0.01$, FWE corrected). We found several regions showing this pattern of results: i) the left lateral superior occipital lobe (which overlapped with OPA as defined in a comparable group-level contrast of scenes vs. objects using data from the Localizer scans), consistent with the above ROI analysis; ii) a contiguous swath of cortex in both hemispheres extending from the lateral superior occipital lobe into the parietal lobe, including the intraparietal sulcus and superior parietal lobule, consistent with other studies implicating these regions in navigation (Spiers and Maguire, 2007; Burgess, 2008; Kravitz et al., 2011b; Marchette et al., 2014; Persichetti and Dilks, 2016; van Assche et al., 2016); and iii) the right and left precentral gyrus, perhaps reflecting motor imagery related to the task (Malouin et al., 2003). Crucially, none of these regions showed overlapping activation in the contrast of Dynamic Faces vs. Static Faces ($p = 0.01$, FWE corrected), suggesting that this activation is specific to motion information in scenes. Next, we observed two additional regions in right lateral middle occipital cortex, and one other region in left lateral middle occipital cortex, which responded more to Dynamic Scenes vs. Static Scenes. Importantly, however, these same regions also responded more to Dynamic Faces vs. Static Faces, revealing that they are sensitive to motion information in general. Indeed, consistent with the ROI analysis above, these regions corresponded to MT, as confirmed by overlaying functional parcels for MT that were created using a group-constrained method in an independent set of subjects (see Methods and Julian et al., 2012). Finally, we observed several regions responding more to Dynamic Faces vs. Static Faces, including bilateral posterior superior temporal sulcus, consistent with previous studies of dynamic face information processing (Pitcher et al., 2011) and thus validating our paradigm, as

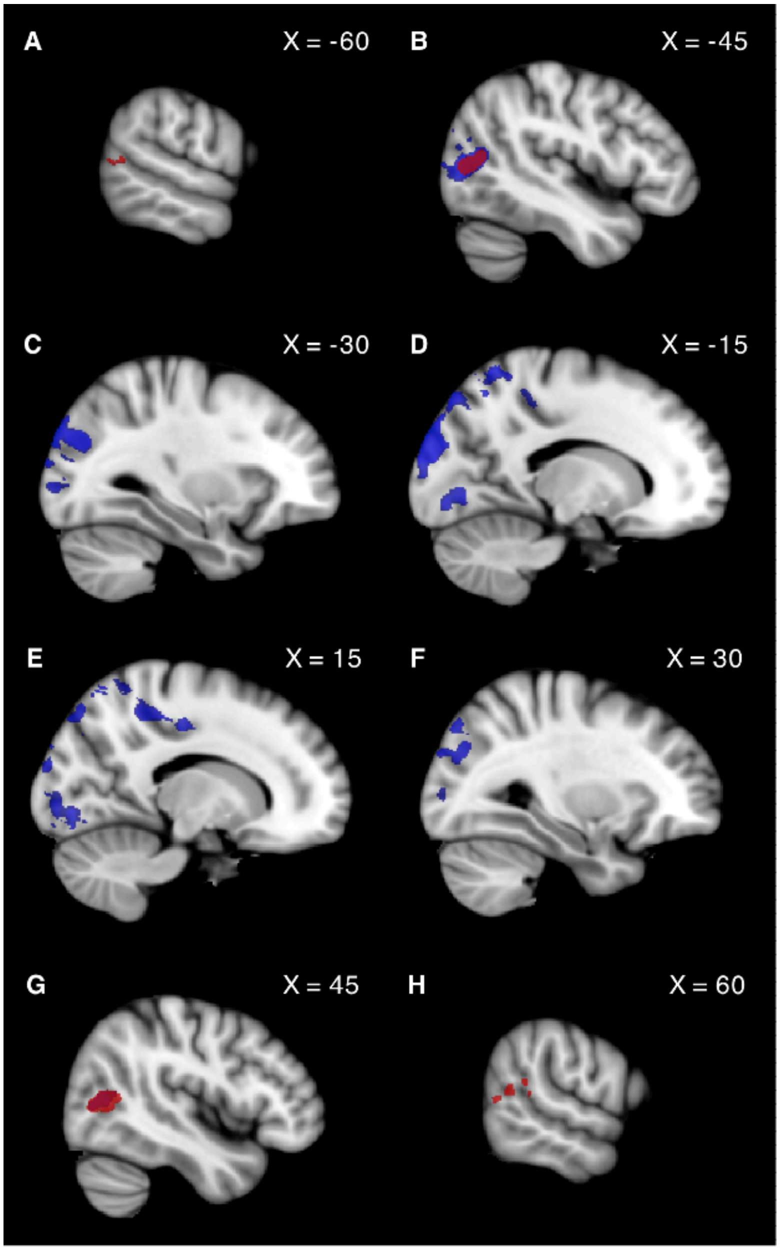

Figure 4. Group analysis exploring representation first-person perspective motion information beyond OPA. The contrast of "Dynamic Scenes > Static Scenes" is shown in blue ($p < 0.01$, FWE corrected), while the contrast of "Dynamic Faces > Static Faces" is shown in yellow ($p < 0.01$, FWE corrected). The right hemisphere is depicted in panels A-C, while the left hemisphere is depicted in panels D-H. X coordinates in MNI space are provided for each slice. A network of regions including lateral superior occipital cortex (corresponding to OPA; see F), superior parietal lobe, and precentral gyrus (see C-E) responded significantly more to "Dynamic Scenes > Static Scenes" (blue), but similarly to "Dynamic Faces vs. Static Faces" (yellow). One bilateral region in lateral occipital cortex (corresponding to motion-selective MT) showed overlapping activation across both contrasts (see B, G). Finally, regions in bilateral posterior superior temporal sulcus and anterior temporal pole responded more to "Dynamic Faces > Static Faces," but similarly to "Dynamic Scenes vs. Static Scenes", consistent with Pitcher et al. (2011) (see A and H).

well as a region in the right anterior temporal pole—a known face selective region (Sergent et al., 1992; Kriegeskorte et al., 2007)—suggesting that this region may also be sensitive to dynamic face information. Crucially, these same regions did not show overlapping activation with the contrast of Dynamic Scenes vs. Static Scenes, indicating that this activation is specific to stimuli depicting dynamic face information.

To further explore the data, we also examined activation to the contrast of Dynamic Scenes minus Static Scenes at lower thresholds (p = 0.05, uncorrected). Here we found the same network of regions responding more to Dynamic Scenes than Static Scenes, as well as additional regions in the right and left calcarine sulcus (consistent with the ROI analysis, insofar as FC also responded more to scene motion than face motion, albeit less so than OPA), right insula, right temporal pole, and right and left precentral gyrus.

Discussion

Here we explored how the three known scene-selective regions in the human brain represent first-person perspective motion information through scenes – information critical for visually-guided navigation. In particular, we compared responses in OPA, PPA, and RSC to i) video clips depicting first-person perspective motion through scenes, and ii) static images taken from these very same movies, rearranged such that first-person perspective motion could not be inferred. We found that OPA represents first-person perspective motion information, while RSC and PPA do not. Importantly, the pattern of responses in OPA was not driven by domain-general motion sensitivity or low-level visual information. These findings are consistent with a recent hypothesis that the scene processing system may be composed of two distinct systems: one system supporting navigation (including OPA, RSC, or both), and a second system supporting other aspects of scene processing, such as scene categorization (e.g., recognizing a kitchen versus a beach) (including PPA) (Dilks et al., 2011; Persichetti and Dilks, 2016). This functional division of labor mirrors the well-established division of labor in object processing between the dorsal ("how") stream, implicated in visually-guided action, and the ventral ("what") stream, implicated in object recognition (Goodale and Milner, 1992). Further, these data suggest a novel division of labor even among regions involved in navigation, with OPA particularly involved in guiding navigation through the immediately visible environment, and RSC supporting other aspects of navigation, such as navigation through the broader environment.

The hypothesized role of OPA in guiding navigation through the immediately visible environment is consistent with a number of recent findings. First, OPA represents two kinds of information necessary for such visually-guided navigation: sense (left versus right) and egocentric distance (near versus far from me) information (Dilks et al., 2011; Persichetti and Dilks, 2016). Second, OPA represents local elements of scenes, such as boundaries (e.g., walls) and obstacles (e.g., furniture)—which critically constrain how one can move through the immediately visible environment (Kamps et al., 2016b). Third, the anatomical position of OPA within the dorsal stream, which broadly supports visually-guided action (Goodale and Milner, 1992), suggests that OPA may support a visually-guided action in scene processing, namely visually-guided navigation. Thus, given the above findings, along with the present report that OPA represents the first-person perspective motion information through scenes, we hypothesize that OPA plays a role in visually-guided navigation, perhaps by tracking the changing sense and egocentric distance of boundaries and obstacles as one moves through a scene.

Critically, we found that OPA only responded to motion information in scenes, not faces. This finding rules out the possibility that OPA is sensitive to motion information in general, and suggests that OPA may selectively represent motion information in scenes. However, our study did not test other kinds of motion information within the domain of scene processing, and thus it may be the case that OPA represents other kinds of scene motion information in addition to the first-person perspective motion tested here. One candidate is horizontal linear motion (e.g., motion experienced when looking out the side of a car). However, one recent study (Hacialihafiz and Bartels, 2015) found that while OPA is sensitive to horizontal linear motion, OPA does not selectively represent such motion information in scenes, but also responds to horizontal linear motion in phase-scrambled non-scene images. This lack of specificity suggests that the horizontal linear motion sensitivity in OPA may simply be inherited from low-level visual regions (indeed, while many low-level features were matched between the stimuli, the study did not compare responses in OPA to those in a low-level visual region, such as FC), and thus may not be useful for scene processing in particular. Another candidate is motion parallax information, a 2D motion cue allowing inference of the surrounding 3D layout. Interestingly, however, a second recent study (Schindler and Bartels, 2016) found that OPA was not sensitive to such motion parallax information, at least in the minimal line drawing scenes tested there. Yet another candidate is optic flow information, which was likely abundant in our Dynamic Scene stimuli. Optic flow information provides critical cues for understanding movement through space (Britten and van Wezel, 1998), and thus may be a primitive source of information for a visually-guided navigation system. Indeed, while optic flow information is typically studied outside the context of scenes (e.g., using moving dot arrays), OPA has been shown to be sensitive to other scene “primitives” such as high spatial frequencies (Kauffmann et al., 2015) and rectilinear features (Nasr et al., 2014), supporting this possibility. Taken together, these findings on motion processing in OPA are thus far consistent with the hypothesis that OPA selectively represents motion information relevant to visually-guided navigation. However, future work will be required to address the precise types of motion information (e.g., optic flow information) that drive OPA activity in scenes.

As predicted, RSC did not respond selectively to first-person perspective motion through scenes, consistent with current hypotheses that RSC supports other aspects of navigation involving the integration of information about the current scene with representations of the broader environment (Burgess et al., 2001; Byrne et al., 2007; Marchette et al., 2014; Epstein and Vass, 2015). For example, RSC has been suggested to play a role in landmark-based navigation (Auger et al., 2012; Epstein and Vass, 2015), location and heading retrieval (Epstein et al., 2007a; Vass and Epstein, 2013; Marchette et al., 2014), and the formation of environmental survey knowledge (Wolbers and Buchel, 2005; Auger et al., 2015). Importantly, our stimuli depicted navigation through limited portions (each clip lasted only three seconds) of unfamiliar scenes. As such, it was not possible for participants to develop survey knowledge of the broader environment related to each scene, or what is more, to integrate cues about self-motion through the scene with such survey knowledge. The present single dissociation, with OPA, but not RSC, responding selectively to dynamic scenes, therefore suggests a critical, and previously unreported division of labor amongst brain regions involved in scene processing and navigation more generally. In particular, we hypothesize that while RSC represents the broader environment associated with the current scene, in order to support navigation to destinations outside the current view (e.g., to get from the cafeteria to the psychology building), OPA rather represents the immediately visible environment, in order to support navigation to destinations within the current view (e.g., to get from one side of the cafeteria to the other). Of course, since here we did not test how these regions support navigation through the broader environment, it might still be the case that OPA supports both navigation through the immediately visible scene and the broader environment. Future work will be required to test this possibility.

Finally, our group analysis revealed a network of regions extending from lateral superior occipital cortex (corresponding to OPA) to superior parietal lobe that were sensitive to first-person perspective motion information through scenes. This activation is consistent with a number of studies showing parietal activation during navigation tasks (Spiers and Maguire, 2007; Burgess, 2008; Kravitz et al., 2011b; Marchette et al., 2014; Persichetti and Dilks, 2016; van Assche et al., 2016). Interestingly, this activation is also consistent with neuropsychological data from patients with damage to posterior parietal cortex who show a profound inability to localize objects with respect to the self (a condition known as egocentric disorientation) (Stark et al., 1996; Aguirre and D'Esposito, 1999; Wilson et al., 2005; Ciaramelli et al., 2010).

In sum, here we found that OPA, PPA, and RSC differentially represent the first-person perspective motion information experienced while moving through a scene, with OPA responding more selectively to such motion information than RSC and PPA. This enhanced response in OPA to first-person perspective motion information, a critical cue for navigating the immediately visible scene, suggests the novel hypothesis that OPA is distinctly involved in visually-guided navigation, while RSC and PPA support other aspects of navigation and scene recognition.

Paper 2

Dissociable systems for recognizing places and navigating through them: causal and developmental evidence

Frederik Kamps, Samaher F. Radwan, Stephanie Wahab, Jordan E. Pincus, Daniel D. Dilks

Abstract

Recent neuroimaging evidence suggests that scene processing depends on dissociable systems for visually-guided navigation (including the occipital place area, OPA) and scene categorization (including the parahippocampal place area). Critically, however, a stronger test of this hypothesis would ask whether these systems are causally dissociable, and further, whether these systems develop independently across the lifespan. Thus, here we tested visually-guided navigation and scene categorization abilities in 36 adults with Williams syndrome (WS) – a developmental disorder involving cortical thinning in and around the OPA – as well as in 82 typically developing 4-8 year old children. We found that i) WS adults are impaired in visually-guided navigation, but not scene categorization, relative to mental-age matched children; and ii) visually-guided navigation matures later in typical development than scene categorization. These findings provide the first causal and developmental evidence for dissociable scene processing systems for recognizing places and navigating through them.

Introduction

A fundamental divide in the visual system lies between systems for perception versus action. For example, classic work from Milner and Goodale showed that recognizing objects and manipulating them involve distinct neural processes, with visual perception (e.g., recognizing that the object is a cup) supported by the ventral visual system, and visually-guided action (e.g., reaching out and grasping a cup) supported by the dorsal visual system (Ungerleider and Mishkin, 1982; Goodale and Milner, 1992). Although substantial fMRI evidence now supports this two visual systems hypothesis (Culham et al., 2003; Valyear et al., 2006; Cavina-Pratesi et al., 2007; Gallivan and Goodale, 2018), the strongest evidence came from studies of patients with damage to one or the other system, revealing that each system can be impaired independent of the other, and thus a double dissociation (Goodale et al., 1991; Goodale and Milner, 1992; Dilks et al., 2008). Subsequent work found evidence that the dorsal and ventral streams even develop independently, with the dorsal stream slower to develop than the ventral stream (Diamond and Goldman-Rakic, 1989; Diamond et al., 1989; Bertenthal, 1996; Gilmore and Johnson, 1997a, b; Csibra et al., 1998; Atkinson et al., 2003; Dilks et al., 2008). Thus, the two visual systems hypothesis is supported by converging evidence across neural, causal, and developmental studies.

Recent neural evidence suggests that this division of labor between systems for perception and action is not only found for object processing, but also extends to scene processing as well (Dilks et al., 2011; Kamps et al., 2016b; Kamps et al., 2016a; Persichetti and Dilks, 2016; Persichetti and Dilks, 2018a), contradicting the widely-held assumption that the scene processing system is a monolithic system in the service of navigation (Ghaem et al., 1997; Epstein and Kanwisher, 1998a; Rosenbaum et al., 2004; Cheng and Newcombe, 2005; Spelke et al., 2010; Epstein and Vass, 2015; Marchette et al., 2015b). In particular, it has been proposed that the more dorsal occipital place area (OPA) supports visually-guided navigation through scenes (e.g., walking through a place avoiding boundaries and obstacles), while the more ventral parahippocampal place area (PPA) supports scene categorization (e.g., recognizing that the place is a kitchen, and not a beach). For example, the more dorsal OPA responds more strongly when participants are asked to imagine navigating through a scene than when asked to judge the category of that same scene (i.e., a bedroom, kitchen, or living room), while the more ventral PPA shows the opposite pattern (Persichetti and Dilks, 2018a).

Critically, while the fMRI studies above provide strong initial evidence for the two scene processing systems hypothesis, fMRI evidence is only correlational, and the strongest test of this hypothesis would ask whether these systems are causally dissociable. That is, if visually-guided navigation and scene categorization are supported by distinct systems, then it should be possible to find cases of impairment to one system, coupled with relative sparing of the other system. To test this prediction, we measured visually-guided navigation and scene categorization abilities in adults with Williams syndrome (WS). WS is a genetic, developmental disorder involving impairment on a variety of dorsal stream functions (Atkinson et al., 1997; Atkinson et al., 2003; Dilks et al., 2008; O'Hearn et al., 2009; Kamps et al., 2017), and critically for the present experiment, cortical thinning in regions of the parietal cortex in and around the OPA (Meyer-Lindenberg et al., 2004). If visually-guided navigation and scene categorization are causally dissociable, then WS adults will show greater impairment on a visually-guided navigation task than a scene categorization task.

To further test the idea of dissociable scene processing systems, we also explored how these systems emerge in typical development. Indeed, findings that visually-guided navigation and scene categorization develop differentially across childhood would suggest that these systems must be functionally independent. The development of scene processing is relatively unexplored, but initial clues come from studies of the development of object processing, which suggest that visually-guided action develops later than visual perception, as discussed above. Thus, if the scene processing system is also divided between systems for action and perception, then the same developmental dissociation will be found, with visually-guided navigation maturing later than scene categorization. Some existing evidence provides initial support for this hypothesis. For example, we recently found that OPA undergoes protracted development across childhood, with key signatures of navigational function (i.e., first-person perspective motion processing) still emerging from 5 to 8 years old (Kamps et al., Submitted). Behavioral studies likewise find protracted development of locomotion, obstacle avoidance, and boundary-based spatial memory abilities well into childhood (Pryde et al., 1997; Berard and Vallis, 2006; Julian et al., 2019). By contrast, scene categorization appears to be relatively fast to mature. For example, scene recognition memory matures faster than face or object recognition memory (Golarai et al., 2007). Critically, however, no study to date has compared the development of visually-guided navigation and scene categorization directly. Thus, to test whether these systems are developmentally dissociable, the present study measured visually-guided navigation and scene categorization abilities in typically developing children ages 4 to 8. We predicted that if visually-guided navigation and scene categorization are developmentally dissociable, then these abilities will develop independently across childhood, with the visually-guided navigation system maturing later than the scene categorization system.

Finally, following our primary analysis testing the two predictions above, we conducted additional analyses addressing previous arguments that WS cannot be used as a neuropsychological model of typical cognitive systems. These arguments have been leveraged on the basis that WS individuals might develop differently from typically developing children from birth, leading to qualitative differences in the cognitive processes underlying their behavior (Karmiloff-Smith, 1997). Thus, in WS, it might be possible that any observed decrement in performance on the visually-guided navigation task results from a *qualitatively* different pattern of performance from the MA controls (e.g., WS might show a distinct pattern of performance across the trials, reflecting a distinct underlying mechanism), rather than a *quantitatively* different pattern of performance (e.g., WS might show the same overall pattern of performance across the trials as MA controls, but at reduced accuracy, reflecting a similar underlying mechanism that is less developed in the case of WS) (Musolino and Landau, 2012). To test this possibility, we compared detailed patterns of performance between WS adults and MA controls (around 7 years old), as well as an even younger group of typically developing children (i.e., 4 year olds)—an age at which WS adults have been observed to perform comparably on other tasks on which they show deficits (Bellugi et al., 1992; Dilks et al., 2008; Kamps et al., 2017).

Methods

Participants

Thirty-six adults with WS (23 females), 36 mental-age matched (MA) controls (i.e., typically developing 7 year olds; 22 females), and 36 typically developing 4 year olds (26 females) participated in the primary experiments. Participant characteristics are presented in Table 1. An additional group of 10 4 year olds (5 females; mean age = 53 months, range = 46-60 months) was also recruited for a control experiment. The WS adults were recruited through the Williams Syndrome Association, and all had been positively diagnosed by a geneticist and the FISH test,

Furnished rooms	WS Adults (N = 19)		MA children (N = 19)		4 year olds (N = 19)	
	M (±1 SEM)	Range	M (±1 SEM)	Range	M (±1 SEM)	Range
Chronological age (years)	28.58(1.82)	19-46	6.82(0.22)	5.58-8.42	4.54(0.07)	4.00-5.08
Matricies KBIT (raw score)	25.32(1.72)	14-42	25.21(1.74)	14-43	16.05(1.06)	9-31
Verbal KBIT (raw score)	71.44(2.68)	50-92	47.53(2.92)	30-71	27.05(0.90)	20-35

Empty rooms	WS Adults (N = 17)		MA children (N = 17)		4 year olds (N = 17)	
	M (±1 SEM)	Range	M (±1 SEM)	Range	M (±1 SEM)	Range
Chronological age (years)	27.69(1.42)	19-44	7.07(0.12)	6.17-8.00	4.61(0.08)	4.00-5.00
Matricies KBIT (raw score)	24.00(2.29)	14-32	23.41(1.90)	14-33	14.59(0.70)	8-20
Verbal KBIT (raw score)	69.50(2.37)	59-78	49.41(2.41)	34-71	30.19(1.51)	19-42

Table 1. Williams syndrome (WS), mental-age matched (MA) control, and 4 year old participant characteristics.

confirming a deletion in the characteristic WS region of chromosome 7 (Ewart et al., 1993). All adult participants and legal guardians of child participants gave informed consent.

Participants were tested on a standardized intelligence test, the Kaufman Brief Intelligence Test (KBIT) (Kaufman, 1990). This test yields an overall IQ score, as well as scores for two components, Verbal and Non-verbal (Matrices). The Verbal subtest requires participants to match words or descriptions to pictures, and the Matrices subtest requires participants to judge which objects or patterns "go together." Each WS adult was individually matched to a typically developing control participant based on the non-verbal component specifically because nonverbal IQ is particularly susceptible to impairment in WS (Jarrold et al., 1998). Matching of the raw non-verbal scores between the WS and MA groups was done as closely as possible ($t_{(36)} = 0.04$, $p = 0.97$, $d = 0.01$), with a mode of 0 points difference (maximum difference = 1, $N = 6$). Given the known relative strength of language abilities in WS compared to non-verbal abilities, the WS group unsurprisingly had significantly higher raw verbal scores than the MA children ($t_{(33)} = 5.95$, $p < 0.001$, $d = 2.03$).

***Figure 1*. *Example stimuli used in the visually-guided navigation and scene categorization tasks*.** *For the visually-guided navigation task, participants were asked to judge which door they could leave out of along a complete path on the floor. The answers for the stimuli above, from left to right, are "left", "center", and "right". For the scene categorization task, participants were asked to judge whether the room is a bedroom, kitchen, or living room. The answers for the stimuli above, from left to right, are "bedroom", "kitchen", and "living room".*

The sample size we obtained (N = 19 per group) is appropriate for three reasons. First, this sample size is consistent with the sample sizes (N = 12-18 per group) used in previous studies of impaired navigational or dorsal stream function in WS (Dilks et al., 2008; Kamps et al., 2017; Julian et al., 2019), which typically found large effect sizes ($\eta_p^2 = 0.29\text{-}0.53$). Indeed, a power analysis (alpha level = 0.05) assuming an effect size of half this magnitude ($\eta_p^2 = 0.15$) suggested we could collect 18 participants per group and have 90% power to detect the critical 2 group x 2 task interaction. Second, as discussed in detail below, we replicated our primarily finding (i.e., that WS adults and 4 year olds perform disproportionately worse on a visually-guided navigation task than a scene categorization task, relative to MA controls) in independent samples of WS adults, MA controls, and 4 year olds (N = 17 per group; i.e., the "Empty Rooms" task). Third, the finding of greater impairment on the visually-guided navigation task than the scene categorization task in WS adults and 4 year olds, relative to MA controls, suggests that our paradigm is sensitive enough to both detect real differences and reject false positives.

Design, stimuli, and procedure

Participants performed two tasks: a visually-guided navigation task and a scene categorization task, as previously used in an adult fMRI study (Persichetti and Dilks, 2018a) (Figure 1). However, in order to make the tasks more understandable and

accessible to younger participants, the visually-guided navigation task was referred to as the "doors" game, while the scene categorization task was referred to as the "rooms" game. During the visually-guided navigation task, participants were asked to imagine that they were walking on a continuous path through the room, and had to indicate whether they could leave through the door on the left, center, or right wall. To aid understanding, younger children were told that the ground was "lava", and that only the paths on the ground were safe to walk on. The visually-guided navigation task was designed in this way to simulate real-world navigation through the environment, similar to walking on a sidewalk, and not the grass, or on a clear path through a cluttered space. Furthermore, although our visually-guided navigation task did not actually require participants to navigate through the environment, we were confident that having participants simply look at pictures of places and imagine navigating through the rooms would be sufficient to recruit systems involved in visually-guided navigation because the task has previously been shown to activate OPA in adults (Persichetti and Dilks, 2018a). Participants responded by pointing in the direction of the door out of which they would exit, ensuring that performance on the task was not contaminated by difficulty mapping egocentric directions to words for "left" and "right". During the scene categorization task, participants were again asked to imagine standing in the room, but now to judge what kind of room they were standing in (i.e., a bedroom, kitchen, or living room). Participants responded by verbal report.

The two tasks were matched in all other aspects of the design, stimuli, and procedure. Immediately prior to the testing session, all participants first completed a short training phase of nonspeeded practice trials for each task, during which they were given feedback about whether or not their response was correct (participants were not given feedback during the testing session). Participants advanced from the training phase after responding correctly to at least 10 practice trials and after the experimenter was satisfied that they understood the task instructions. For the testing phase, both tasks were performed on the same set of 36 images. Stimuli were identical to those used in Persichetti and Dilks (2018), and were presented at a size of approximately 24 degrees visual angle on 11.97” x 8.36” LCD screen using custom software written for the Matlab Psychophysics Toolbox (Brainard, 1997). Participants completed 6 “blocks” of each task (12 blocks in total), and the order of blocks was pseudorandomized and counterbalanced across participants. An instruction screen appeared at the start of each block indicating which game/task (i.e., “doors”/visually-guided navigation or “rooms”/scene categorization) would come next. Each block consisted of 12 experimental trials, and each trial consisted of a stimulus presented for 500 ms, followed by a fixation screen, during which the participant gave a nonspeeded 3AFC response. The next trial began only after the participant had responded and was ready to start again. For both tasks, the experimenter recorded participant responses via button press. Finally, in order to verify that participants understood and were paying attention to the task, the testing phase was evenly interspersed with 12 “catch” trials (6 for each task, 1 per block), in which images were presented for 2000 ms, rendering the task trivially easy, and therefore assessing basic task understanding and attentiveness.

Results

The visually-guided navigation system is disproportionately impaired in Williams syndrome

If visually-guided navigation and scene categorization systems are causally dissociable, then adults with WS will show greater impairment on a visually-guided navigation task than a scene categorization task, relative to MA controls. Consistent with this prediction, WS adults were significantly more accurate on the scene categorization task than the visually-guided navigation task ($t_{(18)} = 3.95$, $p = 0.001$, $d = 0.91$), while MA controls were only marginally more accurate on the scene categorization task than the visually-guided navigation task ($t_{(18)} = 2.03$, $p = 0.06$, $d = 0.47$) (Figure 2). Critically, comparing across groups, a 2 (group: WS adults, MA controls) x 2 (task: visually-guided navigation, scene categorization) mixed-model ANOVA revealed a significant interaction ($F_{(1,36)} = 5.013$, $p = 0.03$, $\eta_p^2 = 0.12$). These results reveal that WS adults show greater impairment on the visually-guided navigation task than the scene categorization task, relative to MA controls, providing initial evidence that these systems are causally dissociable. Importantly, performance in both groups on both tasks was significantly below ceiling (one sample t-tests; *WS adults:* both *t's* > 4.63, both *p's* < 0.001, both *d's* > 1.09; *MA controls:* both *t's* > 4.77, both *p's* < 0.001, both *d's* > 1.06) and above floor (one sample t-tests, *WS adults:* $t_{(18)} = 9.58$, $p < 0.001$, both *d's* > 2.19; *MA controls:* both *t's* > 16.24, both *p's* < 0.001, both *d's* > 3.72), ensuring that this interaction was not driven by a restriction of range effect, which can produce false interactions due to ceiling or floor effects (McKone et al., 2012).

Attention or task understanding?

Might the reduced accuracy of WS adults on the visually-guided navigation task reflect a failure of the WS adults to understand or pay attention during the task, rather than a selective impairment of the visually-guided navigation system? Two lines of evidence rule out this possibility. First, as described above, despite their weaker performance on the visually-guided navigation task, the WS adults (like the MA controls) nevertheless

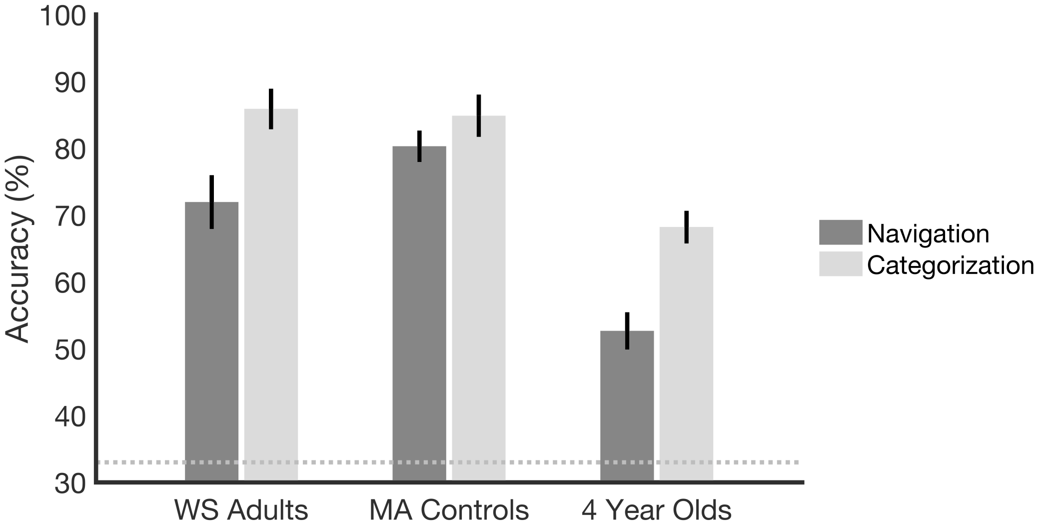

Figure 2. Average performance of WS adults, MA controls, and 4 year olds on the visually-guided navigation and scene categorization tasks. WS adults performed disproportionately worse on the visually-guided navigation task than the categorization task, relative to MA controls, providing causal evidence for the dissociation between these systems. Further, 4 year olds performed disproportionately worse on the visually-guided navigation task than the scene categorization task, relative to MA controls (who were 7 years old on average), providing developmental evidence for independent visually-guided navigation and scene categorization systems, with the visually-guided navigation system arising later in development than the scene categorization system. The grey dotted line indicates chance performance. Error bars indicate the standard error of the mean.

performed well above chance, ensuring that the WS adults understood, and were not simply guessing, on the task. Second, we included several additional "catch trials" throughout the course of the two tasks in which stimuli were presented for 2000 ms, rendering each task trivially easy, and therefore allowing us to probe task understanding and attention directly. Critically, WS adults' performance on the visually-guided navigation catch trials was similar to their performance on scene categorization catch trials ($t_{(18)} = 1.05$, $p = 0.31$, $d = 0.24$), and further, a 2 (group: WS adults, MA controls) x 2 (catch trials: visually-guided navigation, scene categorization) mixed-model ANOVA did not reveal a significant interaction ($F_{(1,36)} = 1.44$, $p =$

0.24, $\eta_p^2 = 0.04$). Caution should be taken in interpreting the lack of an effect (i.e., no significant group x catch trial interaction), and as such, we conducted an additional permutation F-test, which may be more powerful than ANOVA when sample size is relatively small, as might be the case here. To do so, we compared our observed F-statistic for the group x catch trial interaction against a distribution of F-statistics generated by randomly shuffling the accuracy data between groups and catch trials across 10,000 permutations. This analysis also failed to reveal a significant group x trial type interaction ($p = 0.23$). Taken together, these findings ensure that the WS adults understanding of or attentiveness during the visually-guided navigation task cannot explain the reduced performance on the visually-guided navigation task.

A general left versus right impairment?

Another possibility is that WS adults' impaired performance on the visually-guided navigation task was driven by impairment of a more general ability to distinguish left versus right. Indeed, the visually-guided navigation task required participants to distinguish left and right, whereas the scene categorization task did not, and WS individuals have been shown to make a disproportionate number of mirror image confusion errors (i.e., flipping left and right) on other "spatial" tasks, such as judgment of line orientation tasks (Palomares et al., 2009), potentially explaining the pattern of results we observed. To rule out this possibility, we conducted an additional analysis investigating performance on the center trials only, where left vs. right confusion (e.g., perceiving a "left" path as a "right" path, or pointing left when in reality one meant right) could not possibility affect performance. Even using data from the center trials only, a 2 (group: WS adults, MA controls) x 2 (task: visually-guided navigation center trials, categorization) mixed-model ANOVA revealed a significant interaction ($F_{(1,36)} = 4.87$, $p = 0.03$, $\eta_p^2 = 0.12$). Thus, our results do not likely reflect impairment of a more general ability to distinguish left versus right, and rather are consistent with the hypothesis that the visually-guided navigation system in particular is impaired in WS.

Executive function?

But is it still possible that WS adults' performance on the navigation task is explained by impairment in executive function, rather than impairment of visually-guided navigation system in particular, as we have suggested? Indeed, individuals with WS are known to have impaired executive functioning (Mobbs et al., 2007; Menghini et al., 2010). To address this concern, we considered three classic components of executive function ability – namely "shifting", "updating", and "inhibition" (Miyake et al., 2000) – and argue that impairment to any one of these components is unlikely to explain the pattern of results we observed. First, our results are not likely explained by *shifting* (i.e., the ability to shift back and forth between multiple tasks). Although our within-subject and interleaved design did require participants to switch between the visually-guided navigation and scene categorization tasks, this switching occurred slowly over the course of blocks (not individual trials), with an instruction screen interspersed between each block, alerting the participant to the next task. Furthermore, participants' responses were recorded by the experimenter, allowing the experimenter to immediately correct any failure to switch tasks (which happened rarely given the slow speed of switching). Second, our results cannot be explained by *updating* (i.e., the ability to monitor and code incoming information for relevance to the task at hand, and then appropriately revise the information held in working memory), since the visually-guided navigation task neither required participants to hold information in working memory (i.e., since participants responded to each stimulus as soon as they knew the answer), nor to activate manipulate information held in working memory (i.e., since participants only needed to report what they saw in a single stimulus at a single moment in time) (and in any case, these task demands did not differ between the visually-guided navigation task and the scene categorization task). Third, and finally, our results cannot be explained by *inhibition* (i.e., the ability to deliberately inhibit dominant, automatic, or prepotent responses when necessary). In theory, it is true that the visually-guided navigation task potentially depends on inhibition (since three possible options – the left, center, and right doors – are always in view), while the scene categorization task does not. However, the WS adults and MA controls were carefully matched on the nonverbal component of the KBIT, which requires participants to select a correct answer from 5 or 6 concurrently presented options. Despite such careful matching, however, we still saw differential performance between the two groups on the visually-guided navigation task, relative to the scene categorization task. To further rule out the possibility that inhibition explains these results, however, we also tested a second prediction: If the visually-guided navigation task tests inhibition ability rather than visually-guided navigation ability, then similar performance will be found across all trial types for the visually-guided navigation task (i.e., the left, center, and right trials), since all three trial types involve the same inhibition problem (i.e., choosing the correct answer among three options). Indeed, a one-way repeated measures ANOVA revealed a significant main effect of trail type ($F_{(2,36)} = 11.20$, $p < 0.001$, $\eta_p^2 = 0.38$), with WS adults performing better on center trials than left (pairwise comparison, $p < 0.001$) and right trials (pairwise comparison, $p = 0.02$). These reliable differences across trial types cannot be explained by inhibitory processing, and instead likely reflect differences in the difficulty of the paths to be followed in the navigation task. Taken together then, it is unlikely that the pattern of results we observed can be explained by the executive function deficit in WS.

Does WS reflect typical cognition?

Does the observed impairment of WS adults in visually-guided navigation, but not scene categorization, truly reflect dissociable systems in typical individuals? Indeed, given that WS is a genetic, developmental disorder, it has been argued that WS cannot be used as a neuropsychological model of the typical system, since genetic and developmental differences could lead to qualitatively different underlying systems between WS and typically developing individuals (Karmiloff-Smith, 1997). For example, it could be the case that WS adults' reduced ability on the visually-guided navigation task, relative to the scene categorization task, reflects operations of alternative, WS-specific cognitive mechanisms, rather the pattern of performance predicted by a model of typical scene processing systems in which one system is damaged, while the other is not. In the absence of specific proposals for these alternative mechanisms, our hypothesis is that despite the findings above that WS individuals perform *quantitatively* worse on the visually-guided navigation task than scene categorization task relative to MA controls, WS individuals will nevertheless perform the visually-guided navigation task *qualitatively* similar to typically developing individuals—suggesting that WS individuals employ a qualitatively similar mechanism. To test this hypothesis, we compared the pattern of performance of WS adults and MA controls across the three trial types of the visually-guided navigation task (i.e., left, center, and right) (Figure 3). A 2 (group: WS adults, MA controls) x 3 (trial type: left, center, right) mixed-model ANOVA revealed a significant main effect of trial type ($F_{(2,72)} = 17.09$, $p < 0.001$, $\eta_p^2 = 0.32$), with both groups not surprisingly showing greater accuracy on center trials than left or right trials (pairwise comparisons, both p's < 0.01). Critically, however, we failed to find a significant group x trial type interaction ($F_{(2,72)} = 0.03$, $p = 0.97$, $\eta_p^2 = 0.001$), indicating

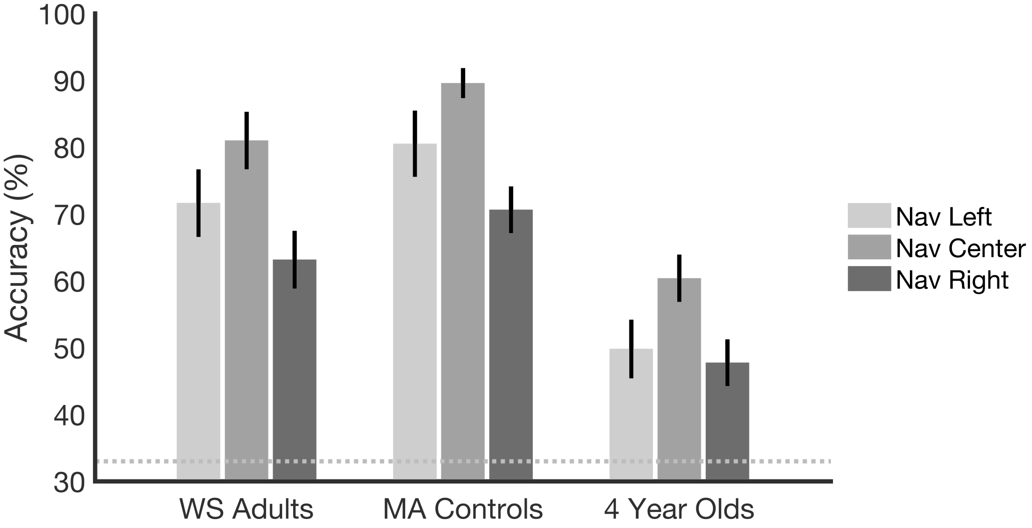

***Figure 3. Patterns of performance on each trial type of the visually-guided navigation task*.** All three groups showed a similar pattern of performance, with higher accuracy on center trials than left or right trials, supporting the idea that all three groups employed similar cognitive mechanisms to solve the task. The grey dotted line indicates chance performance. Error bars indicate the standard error of the mean.

that these patterns did not qualitatively differ between the two groups. A permutation F-test also failed to reveal a significant group x trail type interaction ($p = 0.97$). Taken together, then, we found no evidence that WS adults show qualitatively different patterns of performance compared with MA controls, despite their relative impairment. These findings suggest that this pattern of results in WS reflects a qualitatively similar underlying mechanism that is less developed in WS adults relative to MA controls, and begin to validate WS as a valid neuropsychological model of typical cognitive systems more generally.

Critically, an even stronger test of the hypothesis that WS involves quantitative impairment of typical cognitive systems, rather than qualitatively different underlying cognitive mechanisms, would evaluate whether the WS pattern of performance can be found at earlier stages of typical development. This finding would be consistent with the hypothesis that the WS cognitive profile arises from developmental arrest of systems that typically mature later in development, and thus qualitatively similar underlying mechanisms that did not fully mature (Landau and Ferrara, 2013). We therefore next compared the WS adults with a younger group of typically developing 4 year olds. Like the WS adults, 4 year olds were significantly more accurate on the scene categorization task than the visually-guided navigation task ($t(18) = 4.79$, $p < 0.001$, $d = 1.10$) (Figure 2). Interestingly, a 2 (group: WS adults, 4 year olds) x 2 (task: visually-guided navigation, scene categorization) mixed-model ANOVA did not reveal a significant interaction ($F_{(1,36)} = 0.11$, $p = 0.74$, $\eta_p^2 = 0.003$), and a permutation F-test corroborated this result ($p = 0.74$). These results therefore provide initial support for the idea that that the pattern of performance of WS adults (i.e., weaker performance on visually-guided navigation than scene categorization) is similar to that found in earlier stages of typical development (i.e., in typically developing 4 year olds).

We next compared the more detailed patterns of performance between WS adults and 4 year olds on the three trial types of the visually-guided navigation task. A 2 (group: WS adults, 4 year olds) x 3 (trial type: left, center, right) mixed-model ANOVA revealed a significant main effect of trial type ($F_{(2,72)} = 13.90$, $p < 0.001$, $\eta_p^2 = 0.28$), with both groups showing greater accuracy on center trials than left or right trials (pairwise comparisons, all p's < 0.01) (Figure 3). Critically, however, we failed to find a significant group x trial type interaction ($F_{(2,72)} = 0.67$, $p = 0.51$, $\eta_p^2 = 0.02$), indicating that these patterns did not qualitatively differ between the two groups. A permutation F-test also failed to reveal a significant group x trial type interaction ($p = 0.52$). Taken together then, across several analyses, we failed to find evidence that WS adults perform the task qualitatively differently from either MA controls or 4 year olds, suggesting that WS adults perform these tasks with a qualitatively similar underlying mechanism that is simply underdeveloped relative to their mental age. These findings therefore confirm the validity of WS as a neuropsychological model of the typical system, at least in the tasks used here.

Object recognition, rather than scene categorization?

A further alternative explanation for our findings is that the scene categorization task can potentially be solved entirely by recognizing the objects in the room (e.g., a "bed"), and then in turn inferring the scene category (e.g., only bedrooms have beds), rather than by directly recognizing the scene per se (e.g., recognizing the scene as a whole to be a bedroom). In this case, our findings might not reflect a dissociation *within* scene processing (i.e., between visually-guided navigation and scene categorization), but rather a more general dissociation between object and scene processing, with scene recognition operating entirely in the service of visually-guided navigation. Previous neuroimaging work in adults found that PPA responds strongly to the scene categorization task (Persichetti and Dilks, 2018), suggesting that this task indeed measures scene recognition ability in particular. Nevertheless, the possibility remains in the present behavioral study that participants solved the task using an alternative "object recognition" strategy. To address this alternative explanation, we created a second scene categorization task in which participants were asked to judge the category of empty rooms (i.e., containing no objects whatsoever) that differed based on their spatial layout (i.e., the arrangement of the walls, floor, and ceiling). For comparison, participants again performed a visually-guided navigation task on the same "empty room" stimuli. A 2 (experiment: furnished rooms, empty rooms) x 2 (group: WS adults, MA controls) x 2 (task: visually guided navigation, scene categorization) mixed-model ANOVA revealed a significant group x task interaction ($F_{(1,36)} = 5.96$, $p = 0.02$, $\eta_p^2 = 0.08$), with WS adults showing weaker performance on the visually-guided navigation task than the scene categorization task, relative to the MA controls –

but critically did not reveal a significant experiment x group x task interaction ($F_{(1,36)} = 0.40$, $p = 0.53$, $\eta_p^2 = 0.01$), suggesting that the strength of the group x task interaction effect (revealing greater impairment of the visually-guided navigation system than the scene categorization system) did not differ between furnished and empty rooms (Figure 4). A permutation F-test also failed to find a significant experiment x group x task interaction ($p = 0.53$). Note that these findings held even after excluding the "stairs" room trial type (*group x task interaction:* $F_{(1,68)} = 5.23$, $p = 0.03$, $\eta_p^2 = 0.07$; *group x task x experiment interaction:* $F_{(1,68)} = 0.71$, $p = 0.40$, $\eta_p^2 = 0.01$), ruling out the possibility that participants again completed the task by recognizing stairs as a kind of "object". Thus, our findings cannot be explained by an "object recognition" strategy, and instead reflect a dissociation within scene processing between systems for visually-guided navigation and scene categorization.

The visually-guided navigation system develops later than the scene categorization system in typical development

If visually-guided navigation and scene categorization are distinct cognitive systems, then these systems may arise along independent timelines in typical development. To test this hypothesis, we compared visually-guided navigation and scene categorization abilities in 4 year olds and the MA controls, who were 7 years old on average. As reported above, 4 year olds were significantly more accurate on the scene categorization task than the visually-guided navigation task ($t_{(18)} = 4.79$, $p < 0.001$, $d = 1.10$), while MA controls were marginally more accurate on the scene categorization task than the visually-guided navigation task ($t_{(18)} = 2.03$, $p = 0.06$, $d = 0.47$) (Figure 2). Critically, a 2 (group: 4 year olds, MA controls) x 2 (task: visually-guided navigation, scene categorization) mixed-model ANOVA revealed a significant interaction ($F_{(1,36)} = 7.71$, $p = 0.01$, $\eta_p^2 = 0.18$). Thus, the 4 year olds were disproportionately worse on the visually-guided navigation task than the scene categorization task, relative to the older MA controls, suggesting that visually-guided navigation is later developing than scene categorization. Four year olds performance on both tasks, like that for MA controls, was significantly below ceiling (one sample t-tests; both t's > 13.02, both p's < 0.001, both d's > 2.98) and above floor (one sample t-tests; both t's > 6.90, $p < 0.001$, both d's > 1.58), ensuring that this interaction was not driven by a restriction of range effect (McKone et al., 2012).

Attention or task understanding?

Two lines of evidence rule out the possibility that the 4 year olds simply did not understand or were not paying attention during the visually-guided navigation task. First, as described above, despite their weaker performance on the visually-guided navigation task relative to the scene categorization task, the 4 year olds (like the MA controls) performed well above chance, ensuring that the younger children understood, and were not simply guessing, on the task. Second, we included several additional "catch trials" throughout the course of the two tasks in which stimuli were presented for 2000 ms, rendering each task trivially easy, and therefore allowing us to probe task understanding and attention directly. Critically, 4 year olds' performance on the visually-guided navigation catch trials was similar to that on the scene categorization catch trials ($t_{(18)} = 0.66$, $p = 0.52$, $d = 0.15$), and further, a 2 (group: 4 year olds, MA controls) x 2 (catch trials: visually-guided navigation, scene categorization) mixed-model ANOVA did not reveal a significant interaction ($F_{(1,36)} = 0.07$, $p = 0.80$, $\eta_p^2 = 0.002$). A permutation F-test confirmed this result ($p = 0.80$), ensuring that the children's understanding or attentiveness during the visually-guided navigation task cannot explain the reduced performance on the visually-guided navigation task.

A general left versus right impairment?

Another possibility is that these findings are not explained by slower development of the visually-guided navigation system in particular, but rather by slower development of a more general ability to distinguish left versus right. Again, the visually-guided navigation task required participants to distinguish left and right, whereas the scene categorization task did not, potentially explaining the pattern of results we observed. To address this possibility, we conducted an additional analysis investigating performance on the center trials only, where left vs. right confusion (e.g., perceiving a "left" path as a "right" path, or pointing left when in reality one meant to point right) could not possibility lead to impaired performance. Using data from the center trials only, a 2 (group: 4 year olds, MA controls) x 2 (task: visually-guided navigation center trials, scene categorization) mixed-model ANOVA revealed a significant interaction ($F_{(1,36)} = 6.86$, $p = 0.01$, $\eta_p^2 = 0.16$). Thus, our results do not likely reflect slower development of a more general ability to distinguish left versus right, and rather support the hypothesis that visually-guided navigation in particular is slower developing than scene categorization.

Executive function?

Is it possible that the 4 year olds performance on the navigation task is explained by immature executive function ability, rather than immature visually-guided navigation ability in particular, as we have suggested? Indeed, executive functioning is thought to develop very slowly across childhood, with significant development occurring well into adolescence (Welsh, 1991; Zelazo et al., 2003; Diamond, 2006). To address this concern, we again considered three classic components of executive function ability – namely "shifting", "updating", and "inhibition" (Miyake et al., 2000) – and argue that impairment to any one of these components is unlikely to explain the pattern of results we observed. To begin, our results are not likely explained by slow development of *shifting* or *updating* ability for a variety of reasons related to the experimental

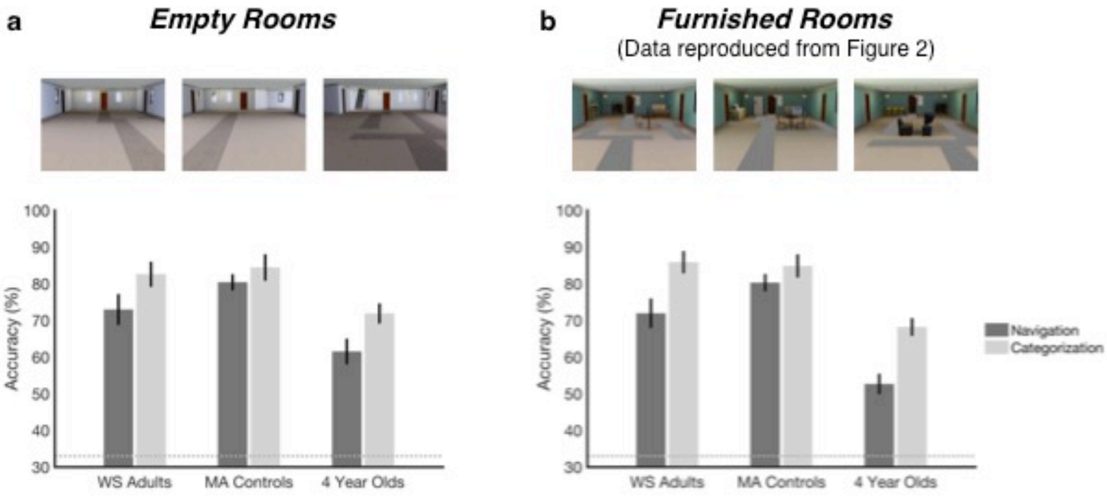

Figure 4. Average performance on furnished room versus empty room stimuli. Across both empty room stimuli (a) and furnished room stimuli (b; the same data as those presented in Figure 2), WS adults performed disproportionately worse on the visually-guided navigation task than the scene categorization task, relative to MA controls, and no differences were found between the Empty Rooms and Furnished Rooms tasks, indicating that WS adults' performance on the scene categorization task was not driven by object categorization ability. Likewise, 4 year olds performed disproportionately worse on the visually-guided navigation task than the scene categorization task, relative to MA controls, and no differences were found between the Empty Rooms and Furnished Rooms tasks, indicating that the 4 year olds' performance on the scene categorization task was not driven by object categorization ability. The grey dotted lines indicate chance performance. Error bars indicate standard error of the mean.

design, as discussed in detail above. Moreover, our results cannot be explained by slow development of *inhibition* for two reasons. First, if the visually-guided navigation task in fact tests inhibition ability, rather than visually-guided navigation ability, then one should expect similar performance across all trial types for the visually-guided navigation task (i.e., the left, center, and right trials), since all three trial types involve the same inhibition problem (i.e., choosing the correct answer among three options). Indeed, a one-way repeated measures ANOVA revealed a significant main effect of trail type ($F_{(2,36)} = 4.54$, $p = 0.02$, $\eta_p^2 = 0.20$), with 4 year olds performing better on center trials than left (pairwise comparisons, both p's < 0.05)

and right trials (pairwise comparisons, both *p*'s < 0.03). These reliable differences across trial types cannot be explained by inhibitory processing, and instead likely reflect differences in the difficulty of the paths to be followed in the navigation task. Second, given that we could not further rule out a role for inhibitory processing by matching 4 year olds and MA controls on nonverbal IQ, as was done for the WS adults, we next asked an additional group of 4 year olds (N = 10) to perform a follow-up "phone" task in which participants viewed an image of a phone, wires, and three outlets, and answered which of three outlets the phone was connected to via a complete wire. In this way, the phone task was closely matched to the visually-guided navigation task in virtually every respect (and inhibitory processing in particular), but did not involve visually-guided navigation. A paired-samples t-test revealed significantly stronger performance on the phone task than the visually-guided navigation task ($t_{(9)} = 2.90$, $p = 0.02$, $d = 0.92$), ruling out the possibility that performance on the visually-guided navigation task is explained by immature inhibition (since in that case 4 year olds would do equally poorly on both tasks), and consistent with the hypothesis of a slow-developing visually-guided navigation system. Taken together then, these results do not likely reflect late development of executive function across childhood, and rather support the hypothesis that the visually-guided navigation system in particular is later developing than the scene categorization system.

Different strategies?

Another possibility is that the 4 year olds attempted to solve the visually-guided navigation task using a different strategy than the MA controls, and that this alternative strategy lead to lower performance on the task. To assess this possibility, we compared performance of the two typically-developing groups across different trial types of the visually-guided navigation task (i.e., left, center, right) (Figure 3). We predicted that if these groups used distinct strategies, then such strategies would produce distinct patterns of performance across the trial types. Counter to this prediction, a 2 (group: 4 year olds, MA controls) x 3 (trial type: left, center, right) mixed-model ANOVA revealed a main effect of trial type ($F_{(2,72)} = 11.15$, $p < 0.001$, $\eta_p^2 = 0.24$), with greater performance on the center trials than the left or right trials (pairwise comparisons, both *p's* < 0.01), but critically failed to reveal a significant group x condition interaction ($F_{(2,74)} = 0.75$, $p = 0.48$, $\eta_p^2 = 0.02$). A permutation F-test also failed to reveal a significant group x trial type interaction ($p = 0.47$). Taken together then, we found no evidence that 4 year olds and MA controls used distinct strategies, consistent with the hypothesis that the visually-guided navigation system develops later than the scene categorization system.

Object recognition, not scene categorization?

Finally, might it be the case that children solved the scene categorization task entirely by recognizing the objects in the room (e.g., a "bed"), and then in turn inferring the scene category (e.g., only bedrooms have beds), rather than by directly recognizing the scene per se (e.g., recognizing the scene as a whole to be a bedroom)? If so, our findings might not reflect a developmental dissociation *within* scene processing (i.e., between visually-guided navigation and scene categorization), but rather a more general developmental dissociation between object and scene processing, with scene recognition operating entirely in the service of visually-guided navigation. To address this alternative explanation, we analyzed performance on a second scene categorization task in which participants were asked to judge the category of empty rooms (i.e., containing no objects whatsoever) that differed based on their spatial layout (i.e., the arrangement of the walls, floor, and ceiling). A 2 (experiment: furnished rooms, empty rooms) x 2 (group: 4 year olds, MA controls) x 2 (task: visually-guided navigation, scene categorization) mixed-model ANOVA revealed a significant group x task interaction ($F_{(1,68)} = 10.54$, $p = 0.002$,

$\eta_p^2 = 0.13$), with 4 year olds showing relatively weaker performance on the visually-guided navigation task than the scene categorization task, relative to the older MA controls – but critically did not reveal a significant experiment x group x task interaction ($F_{(1,68)} = 0.61$, $p = 0.44$, $\eta_p^2 = 0.01$), suggesting that the strength of the group x task interaction effect (revealing slower development of the visually-guided navigation system than the scene categorization system) did not differ between furnished and empty rooms (Figure 4). A permutation F-test also failed to find a significant experiment x group x task interaction ($p = 0.44$). These findings held even after excluding the "stairs" room trial type (*group x task interaction:* $F_{(1,68)} = 9.74$, $p = 0.003$, $\eta_p^2 = 0.13$; *group x task x experiment interaction:* $F_{(1,68)} = 0.83$, $p = 0.37$, $\eta_p^2 = 0.01$), ruling out the possibility that participants again completed the task by recognizing stairs as a kind of "object". Thus, our findings cannot be explained by an "object recognition" strategy, and instead reflect a developmental dissociation within scene processing between systems for visually-guided navigation and scene categorization.

Discussion

The present study tested the hypothesis that scene processing involves two functionally independent cognitive systems for visually-guided navigation and scene categorization. To do so, we studied the breakdown of these abilities in WS, as well as the development of these abilities in typically developing children. We found that WS adults are selectively impaired at visually-guided navigation, but not scene categorization, relative to MA controls. Further, we found that the visually-guided navigation system emerges later in typical development than the scene categorization system, with 4 year olds performing disproportionately worse on a visually-guided navigation task than a scene categorization task, relative to 7 year olds (i.e., the MA controls). Additional analyses ruled out the possibilities that these results are explained by impairment or slow development of attention, more general egocentric spatial abilities (i.e., left versus right), executive function, or qualitatively different underlying cognitive mechanisms or strategies. As such, our results provide the first causal and developmental evidence for the two-systems-for-scene-processing hypothesis.

The proposal that scene processing depends on separable systems for visually-guided navigation (including the more dorsal OPA) and scene categorization (including the more ventral PPA) dovetails with the classic division of labor found in object processing between dorsal systems for action and ventral systems for object perception (Ungerleider and Mishkin, 1982; Goodale and Milner, 1992). Indeed, just as the action system in object processing is selectively impaired in WS (Atkinson et al., 1997; Dilks et al., 2008), here we found that the visually-guided navigation system is selectively impaired in WS; likewise, just as the action system is slower to develop than the perception system in object processing (Dilks et al., 2008), here we found that the visually-guided navigation system is slower to develop than the scene categorization system. Our findings therefore support the broader hypotheses that i) the division between action and perception systems constitutes a general organizing principle for the visual system, extending to the case of scene processing as well as object processing; and further that ii) the action system in general is later developing than the perception system.

The finding that WS individuals are impaired in visually-guided navigation is consistent with studies showing a variety of navigational deficits in WS (Farran et al., 2010; Foti, 2011; Broadbent et al., 2014; Bostelmann et al., 2017). Most relevant to the current study, individuals with WS are impaired at using environmental boundaries to remember spatial locations (Julian et al., 2019) and to recover their orientation after becoming disoriented (Lakusta et al., 2010; Ferrara and Landau, 2015). Given that these studies, like the present study, used behavioral methods only, an open question concerns precisely which neural systems are damaged in WS, and whether the impairments observed on these various tasks result from damage to the same system or different systems. The present findings strongly suggest that WS adults have damage to the OPA. Indeed, the same visually-guided navigation task used to find impairment in WS here has been shown to selectively activate OPA in typically developing adults (Persichetti and Dilks, 2018b), and WS adults have reduced grey matter and reduced sulcal depth in and around the intraparietal sulcus near the typical location of OPA (Meyer-Lindenberg et al., 2004). Given proposals that OPA is critical for representing the egocentric distance and direction of environmental boundaries during navigation (Dilks et al., 2011; Persichetti and Dilks, 2016), it is possible that the impairment on the reorientation and boundary-based spatial memory tasks results from OPA damage as well. Indeed, disruption of OPA has been shown to selectively impair boundary-based spatial memory in adults (Julian et al., 2016). However, it is also possible that the impairments observed on these various tasks result from damage to neural systems beyond OPA. For example, reorientation and boundary-based spatial memory may also depend on spatial memory systems in the hippocampus (O'Keefe and Nadel, 1978; Doeller et al., 2008), and the hippocampus is thought to be dysfunctional in WS (Meyer-Lindenberg et al., 2006). Notably, given that our task tested the immediate perception of, rather than memory for the stimuli, it is unlikely hippocampal damage affected performance on the current visually-guided naviation task. In any case, future work will be required to fully understand which neural systems are damaged in WS, and precisely how those neural systems contribute to different kinds of navigational behavior.

In contrast to their impaired performance in visually-guided navigation, WS adults showed relatively spared ability in scene categorization. This sparing was found both when participants categorized rooms based on the spatial layout of the boundary walls that made up the rooms, and when participants categorized rooms based on the objects in the rooms, consistent with computer vision work suggesting that scene categories are represented by two independent, yet complementary descriptors: spatial boundary (i.e., the external shape, size, and scope of the space) and scene content (i.e., the internal features of the scene encompassing objects, textures, colors, and materials) (Oliva and Torralba, 2001, 2002). This finding also suggests that WS individuals do not have a deficit in boundary perception per se; rather, WS adults may be impaired in using boundaries for navigation (reflecting an impaired navigation system), despite a spared ability to use boundaries for scene categorization (reflecting an intact scene categorization system).

The finding that visually-guided navigation ability undergoes protracted development across childhood provides a striking complement to a recent finding that OPA is still developing sensitivity to navigationally-relevant information across this same age range (Kamps et al., Submitted). This finding is also consistent with evidence that boundary-based spatial memory (known to be supported by OPA in adulthood) and obstacle avoidance during locomotion are still developing late into childhood as well (Pryde et al., 1997; Berard and Vallis, 2006; Julian et al., 2019). Nevertheless, the idea of a late developing visually-guided navigation system may seem surprising, given that humans begin navigating early in life (e.g., crawling around 9 months, walking around 14 months) and show remarkably sophisticated navigational ability within the first few years. For example, infants as young as 18-24 months can use boundaries to recover their orientation after becoming disoriented (Hermer and Spelke, 1994; Hermer and Spelke, 1996), while infants as young as 6-14 months understand whether it is safe to locomote over a ‘visual cliff’ (Gibson and Walk, 1960). How then can we reconcile these findings and observations with the hypothesis that visually-guided navigation undergoes protracted development? One possibility is that these tasks rely on different systems, as suggested above in the case of WS. For example, the reorientation task may depend on a head direction system (e.g., for recovering orientation relative to the geometry of the environment), while the visual cliff task may depend simply on the ability to perceive depth. A second possibility is that these tasks *do* rely on the same visually-guided navigation system, and that despite the substantial development still occurring late in childhood (as detected here), the foundations of this system are nevertheless intact early, and sufficient to account for these early navigational behaviors. Indeed, future work will be required to fully understand the origins of the visually-guided navigation system and the more precise relationship between these various tasks used to measure navigation ability.

Finally, our study broadly supports a recent hypothesis about the nature of WS itself, which argues that the uneven WS cognitive profile emerges from developmental arrest of those cognitive abilities that are late emerging in typical development (Landau and Hoffman, 2012; Landau and Ferrara, 2013). Consistent with the predictions of this hypothesis, we found that the visually-guided navigation system, which is late developing in typical development, is likewise arrested in WS (resulting in weaker performance in WS, relative to MA controls), unlike the scene categorization system, which is earlier developing in typical development and relatively spared in WS. Although the precise age at which this developmental arrest occurs is not yet established, our results suggest that it may be around 4 years old, consistent with the results of many other studies across a variety of cognitive domains (Bellugi et al., 1992; Dilks et al., 2008; Landau and Hoffman, 2012; Landau and Ferrara, 2013; Kamps et al., 2017; Julian et al., 2019).

In conclusion, the present study provides the first causal and developmental evidence for dissociable visually-guided navigation and scene categorization systems. Future work will ask whether it is possible to find cases of impaired scene categorization ability, but spared visually guided navigation ability, and thus evidence of a double dissociation, complementing to the present findings in the case of WS. In any case, the current results present causal and developmental evidence that a key divide within the scene processing system lies between systems for visually guided navigation versus scene categorization.

Paper 3

Neural evidence for the protracted development of human visually-guided navigation

Frederik Kamps, Jordan E. Pincus, Samaher F. Radwan, Stephanie Wahab, Daniel D. Dilks

Abstract

Human adults flawlessly and effortlessly navigate through the immediately visible environment, a process we refer to as "visually-guided navigation". How does this astonishing ability develop? Here we explored the development of the occipital place area (OPA), a scene-selective region involved in visually-guided navigation in adulthood. We found that although OPA already responds significantly more to scenes than objects by 5 years old, responses to first-person perspective motion – a proxy for the visual experience of actually navigating the immediate environment – were not yet detectable at this same age, and rather only emerged by 8 years old. This protracted development was specific to first-person perspective motion through scenes, not motion on faces or objects, and was not found in other scene-selective regions (the parahippocampal place area or retrosplenial complex) or a motion-selective region (MT). Our findings therefore reveal novel neural evidence of surprisingly prolonged development of the visually-guided navigation system across childhood.

Introduction

Accurate visually-guided navigation is fundamental to everyday life, and the bedrock of virtually all independent behaviors. Indeed, by adulthood, we skillfully use vision to navigate our current surroundings, and do so apparently effortlessly, finding our way through even unfamiliar places (or "scenes"), without ever making errors like running into walls or tripping over furniture. Neuroimaging work in adults has shown that the occipital place area (OPA), a scene-selective region in the dorsal visual stream, is a critical region supporting visually-guided navigation. For example, OPA responds selectively to videos depicting the visual experience of navigation through scenes (i.e., first-person perspective motion information) (Kamps et al., 2016a); represents essential egocentric spatial information for visually-guided navigation, including sense (i.e., left vs. right) and egocentric distance information (i.e., near vs. far) (Dilks et al., 2011; Persichetti and Dilks, 2016); represents local scene elements that constrain navigation, such as boundaries and obstacles (Kamps et al., 2016b); codes "navigational affordances" (i.e., where one can and cannot walk through a scene) (Bonner and Epstein, 2017); and responds significantly more when participants imagine navigating through a scene than when participants categorize the kind of place they are in (Persichetti and Dilks, 2018a).

Despite this work on the neural basis of visually-guided navigation in adults, surprisingly little is known about how visually-guided navigation develops. Is visually-guided navigation present from the earliest days of life? Or rather does visually-guided navigation emerge slowly over the course of years of development? While three recent studies investigated the development of scene selectivity in OPA (i.e., significantly greater responses to scenes than objects) (Deen et al., 2017; Cohen et al., 2019; Meissner et al., 2019), nothing is known about how the more specialized visually-guided navigation function develops in this region. Given the ecological importance of visually-guided navigation, and the fact that infants begin crawling and walking within the first year of life, one might predict that visually-guided navigation emerges early, with little development across childhood. Consistent with this possibility, classic work using reorientation (Hermer and Spelke, 1994; Hermer and Spelke, 1996) and visual cliff tasks (Gibson and Walk, 1960) has shown that many navigational abilities are already present in infancy and early childhood; although, it is not clear whether or how these tasks relate to the visually-guided navigation system supported by OPA. By contrast, other work suggests that visually-guided navigation in particular may undergo a more protracted developmental trajectory across childhood. For example, the ability to locate oneself relative to environmental boundaries – an ability that depends on OPA in adulthood (Julian et al., 2016) – is still developing in children ages 6-10 (Julian et al., 2019). Likewise, behavioral studies of locomotion and obstacle avoidance find that even 7-11 year old children are still not adultlike in guiding navigation past obstacles (Pryde et al., 1997; Berard and Vallis, 2006). Finally, visually-guided reaching, another dorsal stream process, is still developing late in childhood, suggesting that visually-guided navigation may be late developing as well (Dilks et al., 2008). Thus, the question remains whether visually-guided navigation develops early or late.

To directly address this question, here we measured responses in OPA across childhood to first-person perspective motion – a proxy for the visual experience of actually navigating through a scene. Given previous behavioral work showing that childhood is a sensitive range for capturing the development of boundary-based navigation and other dorsal stream processes (Dilks et al., 2008; Julian et al., 2019), we scanned two groups of children: a group of 5-year-olds and a group of 8-year-olds. Responses in OPA were measured using functional magnetic resonance imaging (fMRI) while participants viewed i) 3-sec video clips of first-person perspective motion through scenes ("Dynamic Scenes"), mimicking the actual visual experience of visually-guided navigation, and ii) 3, 1-sec still images taken from these same video clips, rearranged such that first-person perspective motion could not be inferred ("Static Scenes"). As control stimuli, participants also viewed Dynamic and Static Objects and Dynamic and Static Faces (Figure 1). If OPA is early developing, then strong selectivity for first-person perspective motion information (i.e., significantly greater responses to dynamic than static scenes) will already be present by age 5, with little or no development from ages 5 to 8. By contrast, if OPA is late developing, then stronger selectivity for first-person perspective motion information will be observed in 8-year-olds than in 5-year-olds.

Critically, pediatric fMRI introduces a variety of methodological challenges, which could present confounds when comparing younger children to older children (e.g., 5-year-olds versus 8-year-olds). In particular, younger children i) are more likely to move and less likely to pay attention, reducing data quality; and ii) may not yet have developed strong, selective responses in a given region of interest, making it difficult or impossible to define regions of interest at standard thresholds used in adults, and increasing ambiguity in ROI selection for traditional, hand-picked ROI approaches. Here we addressed these concerns in two ways. First, to address data quality concerns, we only analyzed runs where participants moved less than a voxel (<1.5 mm average frame-to-frame displacement) – ensuring only low-motion data were included – and where activation could be detected in early visual cortex (all conditions > fixation, $Z = 2.3$, corrected) – ensuring that children were looking at the stimuli and that the data were of sufficient quality to detect fMRI responses in the visual system. As a result of these procedures, the 5- and 8-year-olds were matched on temporal signal to noise ratio (tSNR) in all regions of interest (all t's < 0.64, all p's > 0.52, all d's < 0.23), motion (average absolute frame-to-frame displacement for all usable runs) ($t_{(30)} = 0.75$, $p = 0.46$, $d = 0.26$), total usable runs (5-year-olds = 57, 8-year-olds = 61), and V1 activation (i.e., the average response in V1 across all conditions, $t_{(30)} = 0.90$, p = 0.38, $d = 0.32$) (Figure 2). Second, to address concerns related to ROI definition, we took advantage of a group-constrained, subject-specific (GSS) approach (see Methods) (Julian et al., 2012). This method allows all ROIs to be defined in all participants, even if they would not survive at standard thresholds used in adults, and is algorithmic, removing all ambiguity from ROI definition, providing the most accurate possible assessment of ROI responses at any stage of development.

Methods

Participants

Sixteen 5-year-olds (mean age: 63 months; range: 60-72 months; 6 females) and 16 8-year-olds (mean age: 101 months; range: 93-112 months; 8 females) participated in the study. Originally, 28 5-year-olds were scanned, but 12 were excluded due to excessive motion and/or lack of attention during runs (see Data Analysis). All participants were recruited through the Emory Child Study Center. Consent was given for all children by their parent or guardian, and verbal assent was additionally collected for the 8-year-olds. All participants had normal or corrected to normal vision, and no history of neurological or psychiatric conditions. All procedures were approved by the Emory University Institutional Review Board.

Design

We used a region of interest (ROI) approach in which we used one set of runs to localize scene-selective, motion-selective, and early visual regions, and a second set of runs to investigate the responses of these same voxels. This ROI approach was facilitated by a group-constrained, subject-specific (GSS) method (Fedorenko et al., 2010; Julian et al., 2012), as detailed in the

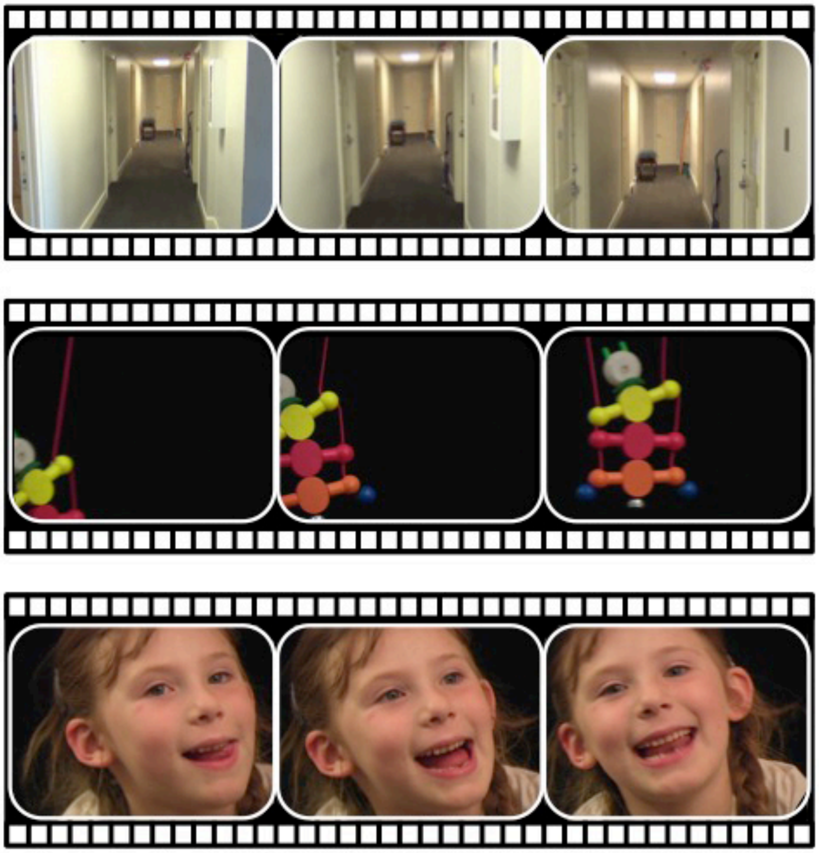

Figure 1. *Experimental stimuli.* The conditions included Dynamic and Static Scenes (top row), Dynamic and Static Objects (middle row), and Dynamic and Static Faces (bottom row). The Dynamic Scene stimuli consisted of 3-s video clips of first-person perspective motion through scenes. The Static Scene stimuli consisted of 3 still images taken from these same video clips, each presented for 1 s and in a random order such that first-person perspective motion could not be inferred. The Dynamic Face stimuli consisted of 3-s video clips of only the faces of children against a black background as they laughed, smiled, and looked around, while the Dynamic Object stimuli consisted of 3-s video clips of moving toys presented against a black background. The Static Object and Static Face stimuli were created following the same procedure described for the Static Scene stimuli.

Data Analysis section. In addition to the standard ROI analyses, we conducted a "volume-selectivity function" (VSF) analysis (Norman-Haignere et al., 2016; Saygin et al., 2016; Kamps et al., 2019a), described in the Data Analysis section.

For all runs, a blocked design was used in which participants viewed stimuli from 6 conditions: Dynamic Scenes, Static Scenes, Dynamic Objects, Static Objects, Dynamic Faces, Static Faces (Figure 1). The Dynamic Scene and Static Scene stimuli were taken from those used in Kamps, Lall, and Dilks (2016). The Dynamic Scene stimuli consisted of 16, 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 4 different places (e.g., a hallway, auditorium, 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 approximately 21 x 16 degrees of visual angle. The Static Scene stimuli were created by taking stills from Dynamic Scene video clips at 1-, 2- and 3-sec time points, resulting in 48 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 approximately 21 x 16 degrees of visual angle. Next, to test whether any observed differences between Dynamic Scenes and Static scenes were specific to first-person perspective motion, and further to allow us to measure scene selectivity more generally, we also included Dynamic Object, Static Object, Dynamic Face, and Static Face conditions. The Dynamic Object stimuli and the Dynamic Face stimuli were taken from those used in Pitcher, Dilks, Saxe, Triantafyllou, and Kanwisher (2011). The Dynamic Object stimuli depicted 7 different toys moving against a black background. The Dynamic Face stimuli depicted only the faces of 4 children against a black background as they smiled, laughed, and looked around. The Static Objects and Static Faces were created and presented using the exact same procedure and parameters described for the Static Scene condition above. Each run was 297 s 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-s video clips from a single condition (e.g., Dynamic Scenes), with an ISI of 0.3 s, resulting in 19.8 s blocks. Each run also included 3, 19.8-s fixation blocks: one at the beginning, one in the middle, and one at the end.

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 runs. To maintain children's interest in the study, children were given opportunities to "take a break" in between runs on an as-needed basis, during which time they could watch a few minutes of a movie or show of their choice. Further, to enhance the children's attention to the stimuli during the runs, participants were encouraged to view the stimuli "actively" by imagining themselves walking through 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.

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. Additionally, whole-brain, high-resolution anatomical images were acquired for each participant for the purposes of registration and anatomical localization (Data Analysis).

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) and custom written MATLAB code. Before analyzing the data, the following pre-statistics processing was applied: motion correction; slice-timing correction; non-brain removal; spatial smoothing using a Gaussian kernel of 5mm FWHM; intensity normalization of the volume at each timepoint; and highpass temporal filtering. To ensure that we included only the highest quality data in our sample, we only analyzed runs where the average absolute frame-to-frame displacement was less than 1.5mm (i.e., the approximate size of one voxel), and where activation could be detected in primary visual cortex (V1) ($Z>2.3$, corrected cluster significance threshold of $p=0.05$). Further, we only included children that had at least two runs that met these criteria, since at least two runs are required for the GSS method, which uses independent sets of runs to localize and test responses in each ROI (see Data Analysis). These criteria resulted in the exclusion of 12 5-year-olds (all 8-year-olds met these criteria).

Given that scene-selective cortex may not yet be fully developed in younger children, and that hand-defining ROIs can be ambiguous, particularly at earlier stages of development, ROIs were defined using a GSS method that circumvents these challenges (Fedorenko et al., 2010; Julian et al., 2012). The GSS analysis was conducted using the following procedure. First, we identified a unique search space for each ROI using previously published probabilistic atlases that predict the vicinity in which each ROI is likely to fall given the typical distribution found in a large, independent sample of adults. Search spaces for scene-selective regions were derived

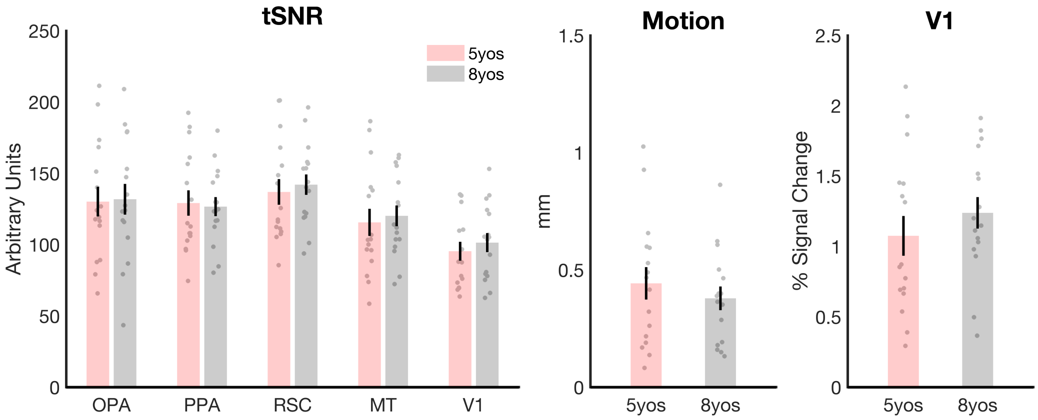

Figure 2. *Data quality did not differ between the 5- and 8-year-old children.* No significant differences were found between the 5- and 8-year-olds for any of the following measures: A) the temporal signal to noise ratio (tSNR) in any region of interest (all t's < 0.64, all p's > 0.52); B) participant head motion (average absolute frame-to-frame displacement for all usable runs) ($t_{(30)} = 0.75$, $p = 0.46$), or C) V1 activation (i.e., the average response in V1 across all conditions minus fixation; $t_{(30)} = 0.90$, p = 0.38). Error bars depict the standard error of the mean.

from Julian et al. (2012), while search spaces for MT and V1 were derived from Wang et al. (2015). Second, for each search space in each participant, voxels were ranked using half of the runs based on parameter estimates for the contrasts of either Dynamic Scenes > Dynamic Objects (for the scene-selective regions), all Dynamic conditions > all Static conditions (for MT), or all conditions > fixation (for V1). After ranking the voxels in this way, the top 10% were selected as the subject-specific ROI. Third, responses to each condition in each ROI and participant were measured using the remaining, independent half of the runs. Fourth, this same define-then-test procedure was repeated across every possible permutation of the runs. For participants with four good runs, two runs were used to define and two runs were used to test, resulting in 6 permutations; for participants with three good runs, two runs were used to define and one run was used to test, resulting in 3 permutations; and for participants with two good runs, one run was to define and one run used to test, resulting in 2 permutations. Fifth, the results of each permutation analysis were averaged together, resulting in the final estimate of responses to the 6 conditions for each ROI in each participant. For each ROI, GSS analysis was conducted separately in each hemisphere, and responses from the ROIs in each hemisphere were subsequently averaged together. Critically, because this analysis uses only a rank ordering of significance of the voxels in each participant, not an absolute threshold for voxel inclusion, all participants could be included in the analysis – not only those who show the ROI significantly. Likewise, because voxel selection is conducted algorithmically, not by hand, all ambiguity is removed from the ROI selection process. Finally, given that data quality is a key concern in developmental populations, this GSS analysis allowed us to assess the within-subject replicabilty of our findings, further ensuring that the data were reliable.

In addition to our primary GSS analysis, which estimates the average univariate response of an entire ROI, we also conducted VSF analyses (Norman-Haignere et al., 2016; Saygin et al., 2016; Kamps et al., 2019a), which allowed us to investigate responses in individual voxels extending from the peak response out into the surrounding cortex. VSF analyses were conducted using the same procedures as the GSS analysis above, with the exception that responses were never averaged across voxels, but rather were calculated for each voxel individually, and were not limited to the top 10% of voxels, but rather were explored across the top 152 voxels in OPA (as limited by participant with the smallest OPA search space).

Results

OPA is scene-selective by 5 years old

Prior to testing first-person perspective motion processing in OPA, we first asked whether scene selectivity (i.e., significantly greater responses to scenes than objects) could be detected in OPA at age 5. One previous study found that OPA is scene selective by around ages 7-8

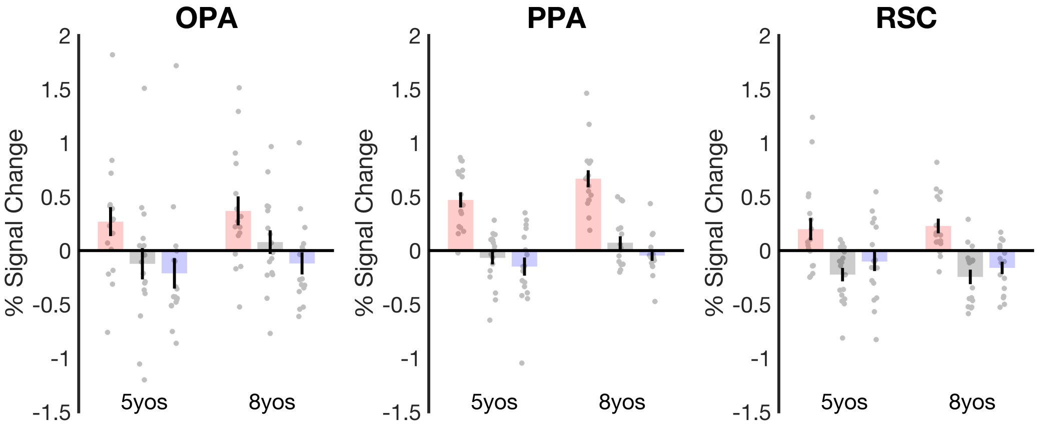

Figure 3. *Scene selectivity is present in all three scene-selective regions by age 5.* Responses are shown for static stimuli only, following standard contrasts used to measure scene selectivity in adults. For 5-year-olds and 8-year-olds, OPA, PPA, and RSC each responded significantly more to scenes than both objects and faces (all *p*'s < 0.001). Further, no region showed a significant age group (5-year-olds, 8-year-olds) x condition (scenes, objects, faces) interaction (all *p*'s > 0.23), suggesting that scene selectivity does not develop across this age range. Error bars depict the standard error of the mean.

(Meissner et al., 2019), while another study found that scene selectivity is *not* present at 5 months old (Deen et al., 2017), leaving open the question of whether scene-selectivity is present as early as age 5. To test for scene selectivity, we compared responses in OPA to Static Scenes with those to Static Objects – following the standard contrast used to define OPA in adults – as well as to Static Faces (note that responses to dynamic stimuli are analyzed extensively below). A 2 (group: 5-year-olds, 8-year-olds) x 3 (condition: Static Scenes, Static Objects, Static Faces) mixed-model ANOVA revealed a main effect of condition ($F_{(2,60)} = 34.30$, $p < 0.001$, $\eta_p^2 = 0.53$), with stronger responses to scenes than both objects (pairwise comparison, $p < 0.001$) and faces ($p < 0.001$). However, this analysis did not reveal a significant group x condition interaction ($F_{(2,60)} = 1.21$, $p = 0.31$, $\eta_p^2 = 0.04$) (Figure 3). These findings show that scene selectivity is present in OPA by 5 years old, and already of similar magnitude to that observed by 8 years old.

For completeness, we also investigated scene selectivity in two additional scene-selective regions involved in other aspects of scene processing and navigation, including the parahippocampal place area (PPA) and the retrosplenial complex (RSC). For both PPA and RSC, a 2 (group: 5-year-olds, 8-year-olds) x 3 (condition: Static Scenes, Static Objects, Static Faces) mixed-model ANOVA revealed a main effect of condition (PPA: $F_{(2,60)} = 222.06$, $p < 0.001$, $\eta_p^2 = 0.88$; RSC: $F_{(2,60)} = 114.60$, $p < 0.001$, $\eta_p^2 = 0.79$), with stronger responses to scenes than both objects (PPA: $p < 0.001$; RSC: $p < 0.001$) and faces (PPA: $p < 0.001$; RSC: $p < 0.001$), but no significant group x condition interaction (PPA: $F_{(2,60)} = 1.51$, $p = 0.23$, $\eta_p^2 = 0.05$; RSC: $F_{(2,60)} = 1.36$, $p = 0.26$, $\eta_p^2 = 0.04$) (Figure 3). These results indicate that all three regions show scene selectivity by 5 years old, with no changes in scene selectivity across ages 5 to 8. Notably, given that we defined and tested responses in the scene regions using independent halves of the data, this analysis reveals the within-subject replicabilty of these results, suggesting that the data are high quality. Further, given the similar magnitude of scene selectivity between 5- and 8-year-olds, these results also provide a further indication that data quality was well matched between the two child groups

OPA responses to first-person perspective motion develop from age 5 to age 8

Having established that OPA is scene selective by 5 years old, we next turned to our central question: Does first-person perspective motion processing develop in OPA within the first few years of life, or does such development extend well into childhood? To address this question, we calculated a “scene motion” difference score by subtracting the response in OPA to Static Scenes from that to Dynamic Scenes for each participant. Strikingly, we found significantly greater scene motion responses in the 8-year-olds than the 5-year-olds ($t_{(30)} = 2.17$, $p = 0.04$, $d = 0.77$), with a significant response to scene motion (i.e., scene motion difference score > 0) in OPA for

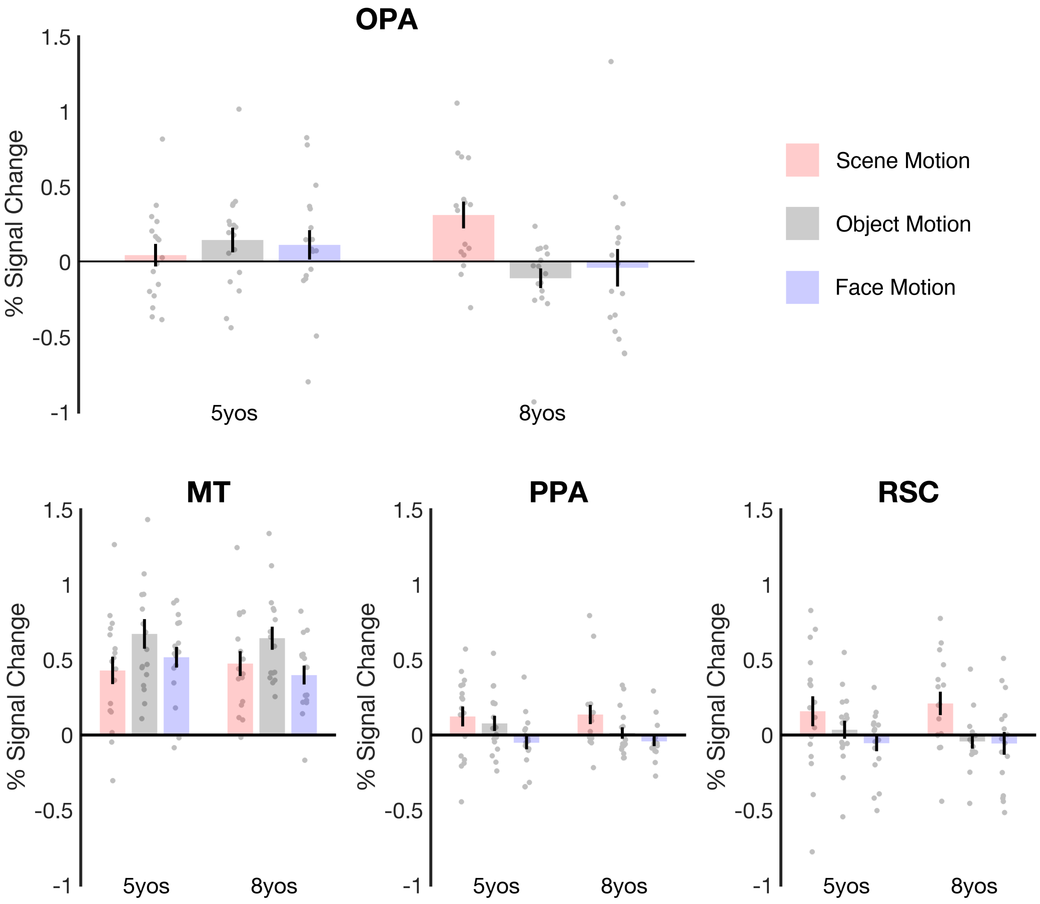

Figure 4. *First-person perspective motion processing develops in OPA from 5 to 8 years old.* Scene motion, object motion, and face motion difference scores were calculated by subtracting responses to the Static stimuli from responses to the Dynamic stimuli separately for each condition. OPA responded strongly to Scene Motion in the 8-year-olds ($p = 0.003$), but not the 5-year-olds ($p = 0.52$). This increase in scene motion processing from 5 to 8 years was greater than that for either object or face motion processing (both p's < 0.05), indicating that this developmental effect was specific to motion through scenes. Further, no developmental changes were found in a motion selective region (MT) ($p = 0.64$), or in other scene-selective regions (PPA or RSC) (both p's > 0.68). These findings suggest that selective visually-guided navigational function is still developing in OPA across childhood. Error bars depict the standard error of the mean.

the 8-year-olds ($t_{(15)} = 3.50$, $p = 0.003$, $d = 0.88$), but not the 5-year-olds ($t_{(15)} = 0.66$, $p = 0.52$, $d = 0.32$) (Figure 4). These findings suggest that first-person perspective motion processing in OPA develops slowly across childhood, and only emerges by age 8.

However, does the developmental increase in first-person perspective motion processing in OPA reflect increasing sensitivity to any kind of visual motion information, rather than motion information in scenes, in particular? To address this question, we compared OPA responses to scene motion with those to object motion (i.e., calculated by subtracting the responses to Static Objects from that to Dynamic Objects) and face motion (i.e., calculated by subtracting the responses to Static Faces from that to Dynamic Faces). A 2 (group: 5-year-olds, 8-year-olds) x 3 (motion domain: scene motion, object motion, face motion) mixed-model ANOVA revealed a significant group x motion domain interaction ($F_{(2,60)} = 4.73$, $p = 0.01$, $\eta_p^2 = 0.14$), with the 5- and 8-year-olds showing a significantly greater change in scene motion responses than in object motion (interaction contrast, $p = 0.006$) or face motion responses (interaction contrast, $p = 0.04$) (Figure 4).

Given that motion information was not precisely matched across our stimuli, it is still possible that OPA shows increasing sensitivity to any kind of visual motion information, and responds more to scene motion than object or face motion by age 8 simply because more motion information was present in the Dynamic Scene stimuli than the Dynamic Object or Dynamic Face stimuli. To address this possibility, we compared responses in OPA with those in the middle temporal area (MT), a domain-general motion-selective region. For MT, a 2 (group: 5-year-olds, 8-year-olds) x 3 (motion domain: scene motion, object motion, face motion) mixed-model ANOVA did not reveal a significant group x motion domain interaction ($F_{(2,60)} = 0.45$, $p = 0.64$, $\eta_p^2 = 0.01$), but rather revealed a main effect of motion domain ($F_{(2,60)} = 11.83$, $p < 0.001$, $\eta_p^2 = 0.28$), with greater responses to object motion than both face and scene motion (pairwise comparisons, both *p's* < 0.002) (Figure 4). These findings suggest that unlike OPA, MT does not respond more to scene motion than face and object motion, and does not develop from 5 to 8

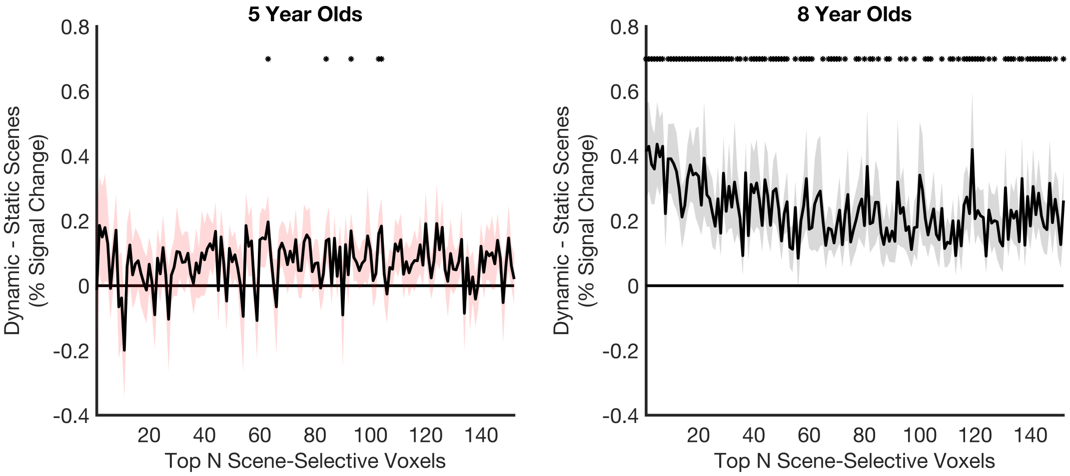

Figure 5. *No evidence of first-person perspective motion processing in the 5-year-old OPA.* Volume-selectivity function analyses were conducted to explore responses to scene motion (Dynamic Scenes minus Static Scenes) in the top 152 most scene-selective OPA voxels in each group. For the 5-year-olds, only 5 out of 152 voxels responded significantly greater than 0 (asterisks indicate significant voxels), fewer than would be expected by chance; by contrast, for the 8-year-olds, 105 out of 152 responded significantly greater than 0, far more than would be expected by chance. Thus, we found no evidence of even a small population of voxels responding significantly to scene motion in the 5-year-old OPA, despite strong evidence of first-person perspective motion processing at the individual-voxel level in the 8-year-old OPA. Shaded regions depict the standard error of the mean.

years old. Testing this claim directly, a 2 (region: OPA, MT) x 2 (group: 5-year-olds, 8-year-olds) x 3 (motion domain: scene motion, object motion, face motion) mixed-model ANOVA revealed a significant region x age x motion domain interaction ($F_{(2,60)} = 4.52$, $p = 0.02$, $\eta_p^2 = 0.13$). These results rule out the possibility that the Dynamic Scenes simply contained more motion information than the Dynamic Faces or Objects, and therefore reveal a remarkably specific developmental increase in scene motion selectivity in OPA from age 5 to 8.

Finally, we asked whether the development of first-person perspective motion processing in OPA was driven by differences in attention between the 5- and 8-year-olds, with the 5-year-olds simply paying less attention to the Dynamic Scenes than the 8-year-olds? While this possibility is unlikely, given that MT responses to scene motion did not differ between the two groups ($t_{(30)} = 0.002$, $p > 0.99$, $d < 0.001$), we nevertheless further addressed this concern by comparing the responses in OPA to those in PPA and RSC. For both PPA and RSC, a 2 (group: 5-year-olds, 8-year-olds) x 3 (motion domain: scene motion, object motion, face motion) mixed-model ANOVA failed to reveal a significant group x motion domain interaction (PPA: $F_{(2,62)} = 0.35$, $p = 0.71$, $\eta_p^2 = 0.01$; RSC: $F_{(2,62)} = 0.39$, $p = 0.68$, $\eta_p^2 = 0.01$) (Figure 4). These findings suggest that unlike OPA, PPA and RSC show no change in motion processing from ages 5 to 8 (consistent with previous work in adults showing that these regions never develop the strong, first-person perspective motion responses found in OPA (Kamps et al., 2016a)). Testing this claim directly, a 3 (region: OPA, PPA, RSC) x 2 (group: 5-year-olds, 8-year-olds) x 3 (motion domain: scene motion, object motion, face motion) mixed-model ANOVA revealed a significant region x group x motion domain interaction ($F_{(2,62)} = 2.82$, $p = 0.03$, $\eta_p^2 = 0.09$). This three-way interaction provides strong evidence against the possibility that the 5-year-olds simply paid less attention to the Dynamic Scenes than the 8-year-olds, since a group difference in attention would predict a main effect of group (with the 8-year-olds paying more attention to all Dynamic Stimuli than the 5-year-olds, causing all regions in the 8-year-olds to respond more to all kinds of motion information than the 5-year-olds), or a group x motion domain interaction (with 8-year-olds paying more attention than 5-year-olds to the Dynamic Scenes in particular, causing all regions in the 8-year-olds to respond more to scene motion than face or object motion, relative to the 5-year-olds), but crucially *not* a region x group x motion domain. Rather, our results reveal differential development within the cortical scene processing system, with first-person perspective motion processing developing specifically in OPA, not PPA or RSC.

Does the 5-year-old OPA show any sensitivity to first-person perspective motion?

The findings above present strong evidence that first-person perspective motion processing develops in OPA between 5 and 8 years old. However, is it truly the case that the 5-year-old OPA does not represent first-person perspective motion whatsoever? For example, it could be the case that a small population of voxels in the peak of the 5-year-old OPA respond strongly to scene motion, but were not detected in the analysis above because they were averaged together with the surrounding voxels that did not respond strongly to scene motion. To test this possibility, we performed a volume-selectivity function (VSF) analysis (Norman-Haignere et al., 2016; Saygin et al., 2016; Kamps et al., 2019a), which allowed us to explore scene motion responses in individual voxels across the volume of OPA extending from the peak scene-selective response outward. Individual OPA voxels were ranked in each participant from most-to-least scene selective using one half of the runs, and the response to scene motion in the top 152 individual voxels of each participant was then averaged across participants in the remaining, independent half of the runs, yielding the average scene motion responses in the top 152 individual voxels (our analysis focused on the top 152 voxels, since all participants had at least 152 voxels in OPA). For the 5-year-olds, only 5 out of 152 voxels (3.29%) showed a significant response to scene motion (i.e., scene motion responses > 0). Given that we ran 152 statistical tests with alpha = 0.05, this number is similar to the number of false-positive results expected by chance (5%, or 8 voxels) (Figure 5). By contrast, for the 8-year-olds, 105/152 voxels (69.08%) showed a significant response to scene motion – well beyond the number expected by chance (Figure 5). Thus, we found no evidence of even a small population of scene motion-selective voxels in the 5-year-old OPA, at least using the methods employed here.

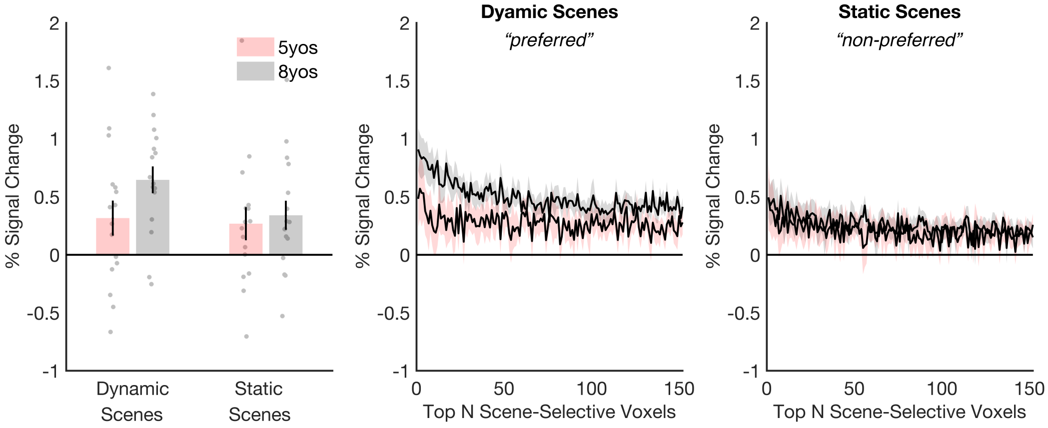

Figure 6. *The development of first-person perspective motion processing in OPA occurs via construction of preferred responses, not pruning of non-preferred responses.* To investigate how development of first-person perspective motion processing occurs in OPA, we compared responses between the 5- and 8-year-olds to the "preferred" stimuli (i.e., Dynamic Scenes) and "non-preferred" stimuli (i.e., Static Scenes) separately. A) Increasing responses to the preferred stimuli were observed from 5 to 8 years old, with no change in responses to the nonpreferred stimuli ($p = 0.03$). B) Further inspection of VSF confirms confirmed this effect across the volume of the OPA. These results suggest that the development of first-person perspective motion processing occurs via construction of preferred responses, with no evidence of pruning of non-preferred responses. Error bars depict the standard error of the mean.

The development of first-person perspective motion processing in OPA reflects construction of preferred responses, not pruning of non-preferred responses

Having established that first-person perspective motion processing develops in OPA in childhood, we next asked *how* such development occurs. Does the development of first-person perspective motion processing occur by construction of preferred responses, by pruning of non-preferred responses, or by a combination of these two mechanisms? To address this question, we directly compared responses to "preferred" stimuli (i.e., the Dynamic Scenes) between 5- and 8-year-olds, as well as responses to "non-preferred" stimuli (i.e., Static Scenes) between 5- and 8-year-olds. A 2 (response: preferred, non-preferred) x 2 (group: 5-year-olds, 8-year-olds) mixed-model ANOVA revealed a significant response x group interaction ($F_{(1,31)} = 5.27$, $p = 0.03$, $\eta_p^2 = 0.15$), with responses to the preferred stimuli (i.e., Dynamic Scenes) showing a marginally significant increase from 5 to 8 (pairwise comparison, $p = 0.06$), and with responses to the non-preferred stimuli (i.e., Static Scenes) showing no significant difference between the 5- and 8-year-olds (pairwise comparison, $p = 0.61$) (Figure 6a). To further explore how this development occurred across the volume of OPA, and whether any smaller population of voxels might show evidence of pruning, we next conducted a VSF analysis. This analysis confirmed that the increasing response to the preferred stimulus was found consistently in individual voxels extending from the peak scene-selective response well into the surrounding cortex, while no discernable changes were found anywhere across the volume of OPA for the non-preferred stimulus (Figure 6b). Taken together, these results suggest that the development of first-person perspective motion processing occurs primarily via construction of preferred responses, rather than pruning of non-preferred responses.

Discussion

Here we explored how visually-guided navigation develops by testing the development of OPA, a key region supporting visually-guided navigation in adulthood. Five- and 8-year-old children were scanned while OPA responses were measured to first-person perspective motion, a proxy for the visual experience of navigation through scenes. Strikingly, we found that although OPA already responded selectively to scenes (relative to objects and faces) by age 5, OPA did not respond to first-person perspective motion at this same age, and only responded to such navigationally-relevant information by age 8. This increase in first-person perspective motion processing is specific to motion on scenes, not motion on faces or objects, and is found only in OPA, not in motion-selective cortex (i.e., MT) or in other scene-selective regions (i.e., PPA and

RSC). Our results therefore present novel neural evidence for the hypothesis that visually-guided navigation undergoes protracted development, with key signatures of visually-guided navigational function still emerging in childhood.

The hypothesis that visually-guided navigation is late developing dovetails with a number of findings showing late development of visually-guided navigation ability. Perhaps most directly, a recent behavioral study of navigation in a virtual environment found that spatial memory for locations defined relative to boundary walls, but not landmark objects, is still developing in children ages 6-10 (Julian et al., 2019), while a transcranial magnetic stimulation (TMS) study using the same task in adults found that disruption of OPA impaired performance for boundary-related locations, but not landmark-related locations (Julian et al., 2016). The present findings unite these studies by finding direct evidence that selective navigational function is still emerging in OPA across childhood, and suggest that late development of boundary-based spatial memory may therefore be mediated in part by late development of OPA. The hypothesis of a late developing visually-guided navigation system is also consistent with developmental studies of locomotion and obstacle avoidance, which show that children ages 4-10 are slower and make more errors than adults when navigating past obstacles in a cluttered terrain (Pryde et al., 1997; Berard and Vallis, 2006) and are less adept than adults at using peripheral vision to guide locomotion (Franchak and Adolph, 2010b).

At the same time, however, the hypothesis that visually-guided navigation undergoes protracted development across childhood might seem incompatible with other findings showing that sophisticated navigational ability is present within the first few years of life. For example, children as young as 18-24 months are able to use the geometry of local boundaries to recall previously learned locations (Hermer and Spelke, 1994; Hermer and Spelke, 1996), while infants as young as 6-14 months understand whether it is safe to navigate a “visual cliff” (Gibson and Walk, 1960). Although the relationship between the navigational behaviors probed by these tasks and those supported by OPA is unknown, there are two possibilities. First, it is possible that these tasks do not depend on OPA, and rather are supported by distinct neural systems. For example, the reorientation task may assess the ability to recall *orientation* relative to boundaries – as supported by extra-hippocampal structures (Winter and Taube, 2014) – more so than the ability to recall *location* relative to boundaries – as supported by OPA (Julian et al., 2019). Likewise, the visual cliff task may be supported by a more general visual ability to perceive depth, rather than the visually-guided navigation system in particular. Second, it is possible that these tasks *do,* in fact, depend on the visually-guided navigation system in OPA, but can be solved using representations that emerge earlier in this region than the representations of first-person perspective motion processing studied here. Indeed, a preference for scenes relative to faces is already detectable in the OPA by just 4-6 months (Deen et al., 2017), with scene-selectivity emerging by 5 years or earlier, as revealed in the present study. While the precise nature of these early-emerging scene representations is unknown, it is possible that they are sufficient, for example, to determine which of two walls is closer (i.e., in the case of the reorientation task) or which of two cliffs is steeper (i.e., in the case of the visual cliff task). Under this account, the later emergence of first-person perspective motion processing may reflect increasingly sophisticated navigational function (e.g., allowing the navigator to dynamically plan and update possible navigational paths through a space). Intriguingly, a similar developmental progression has been found in studies of the developing rodent medial temporal lobe. For example, the rudiments of head direction cells, grid cells, and place cells are detectable as soon as rat pups make their first movements outside of the nest, while the stability and precision of the spatial coding in these cells continues to be refined well into juvenility (Langston et al., 2010; Wills et al., 2010).

Beyond the realm of visually-guided navigation, our results are also consistent with the broad hypothesis that the dorsal visual system, which supports visually-guided action, is later to develop than the ventral visual system, which supports visual perception (Bertenthal, 1996; Atkinson et al., 2003; Braddick et al., 2003; Dilks et al., 2008). For example, studies using the classic posting/matching task developed by Milner and Goodale have shown that the ability to accurately post a card into an oriented slot (a measure of visually-guided action) is later developing than the ability to match a card to the same slot's orientation (a measure of visual perception) (Dilks et al., 2008). Although most studies have focused exclusively on action and perception in object processing, we recently proposed that a similar functional division of labor is found for scene processing, with the more dorsal OPA supporting visually-guided navigation through scenes, and the more ventral PPA supporting categorization of scenes (Dilks et al., 2011; Kamps et al., 2016b; Kamps et al., 2016a; Persichetti and Dilks, 2016; Persichetti and Dilks, 2018a). The present findings therefore provide initial support for the novel hypothesis that the scene-processing system undergoes differential development analogous to that found in object processing, with the dorsal visually-guided navigation system maturing later than the ventral scene categorization system. Critically, however, our study did not test the development of scene categorization in the PPA, and future work will therefore be required to provide a complete test of this hypothesis.

Finally, we found that all three known scene-selective regions (i.e., OPA, PPA, and RSC) respond selectively to scenes (i.e., relative to objects, as well as faces) by age 5. To our knowledge, this is the earliest demonstration of scene selectivity in this system to date, given that previous work has not tested children younger than 7 (Golarai et al., 2007; Pelphrey et al., 2009; Scherf et al., 2011b; Jiang et al., 2014; Meissner et al., 2019), with the possible exceptions of two studies of 5- to 8-year-olds (although it is unclear whether the scene selectivity observed in those studies was driven by the older children only, given that the mean age of the children in each sample was approximately 7 years old) (Scherf et al., 2007; Cohen et al., 2019). This finding therefore places new constraints on theories of how scene selectivity develops, since the selective response to scenes must be established relatively rapidly, within the first few years of life.

In conclusion, we present the first study of the development of visually-guided navigational function in the cortical scene processing system. Strikingly, although OPA is already scene selective by age 5, responses to first-person perspective motion information in OPA were not present at this same age, and only emerged by age 8. These results therefore present novel neural evidence for the hypothesis that human visually-guided navigation undergoes protracted development.

General Discussion

The present thesis sought to address two fundamental questions about human visual scene processing: i) How is the adult human visual scene processing system functionally organized? and ii) How does that functional organization develop? Addressing these questions, I proposed two novel hypotheses. First, I hypothesized that the adult human visual scene processing system contains functionally dissociable systems for navigation (including OPA and RSC) versus scene categorization (including PPA). Second, I hypothesized that these two systems develop independently, with the navigation system maturing later than the scene categorization system. The results of three Papers supported both of these hypotheses. Papers 1 and 2 provided evidence for dissociable visually-guided navigation and scene categorization systems by showing that these systems represent different information about scenes, and are differentially impaired in the case of WS. Papers 2 and 3 provided evidence that these systems develop along independent timelines by showing that the visually-guided navigational ability matures later than scene categorization ability in children ages 4 to 7, and that late development of visually-guided navigational function occurs specifically in OPA across the same age range. Taken together then, these studies provide strikingly consistent support for the hypotheses that human visual scene processing depends on two independently developing systems: an earlier developing scene categorization system, and a later developing visually-guided navigation system.

The hypothesis that OPA and PPA are independent systems supporting visually-guided navigation and scene categorization, respectively, echos the classic division of labor in object processing between the dorsal "action" and ventral "perception" systems (Ungerleider and Mishkin, 1982; Goodale and Milner, 1992). By extending this division of labor to scene processing, the present results suggest that the division of labor between action and perception systems is a general organizing principle for the visual system, not just a principle of the object processing system in particular. Critically, however, the studies in this thesis present only a single dissociation. As such, it could still be the case that a single mechanism (e.g., for scene processing in general) underlies both visually-guided navigation and scene categorization, and that visually-guided navigation is simply a more difficult or complex case than scene categorization. Thus, it will be critical for future work to find the complementary case of impaired scene categorization (i.e., damage to PPA) coupled with spared visually-guided navigation (i.e., sparing of OPA). Indeed, the analogous case in object processing was found in patient D.F., who famously was unable to recognize objects, despite successful visually-guided reaching for those same objects (Goodale et al., 1991).

Importantly, while we have drawn a broad division between systems for categorizing scenes versus navigating through them, a key question for future work will be to understand precisely how scene categorization and visually-guided navigation are accomplished *within* each system. For example, although the currently available evidence rules out a role for PPA in navigation, it does not provide strong constraints on precisely how PPA supports scene categorization. For example, it could be the case that PPA represents superordinate categories (indoor versus outdoor), basic categories (a beach versus a forest), subordinate categories (a temperate forest versus a tropical forest), or all of the above. At very least, recent evidence suggests that PPA does support scene categorization per se, and not scene individuation: patterns of activity in PPA reliability discriminated coffee shops from hardware stores, but could not distinguish one coffee shop from another (Persichetti and Dilks, in revision). Likewise, although we present strong evidence for the role of OPA in visually-guided navigation, there are many possibilities as to how, precisely, OPA contributes to this process. For example, OPA might contribute to the selection of a navigational goal within a scene, the visual identification of possible paths (or "navigational affordances") through the scene toward that goal, or the online guidance of locomotion along the chosen path. Critically, while the precise role of OPA in visually-guided navigation has yet to be discovered, our results do nevertheless help establish that navigation supported by OPA (i.e., what I have termed "visually-guided navigation") is distinct from that supported in RSC (i.e., what I have termed "memory-guided navigation") (Kamps et al., 2016b; Kamps et al., 2016a), placing an initial limit on the hypothesis spaces for the function role of each region.

The present thesis also found converging behavioral and neural evidence for the hypothesis that the visually-guided navigation and scene categorization systems develop independently, with the visually-guided navigation system maturing later than the scene categorization system (at least in the case of the visually-guided navigation system supported by OPA). These differential developmental trajectories provide a striking complement to related work in object processing showing that the action system is slower to develop than the perception system (Dilks et al., 2008), and suggest the overarching hypothesis that the dorsal stream in general is later developing than the ventral stream. Based on this hypothesis of a late-developing action system (i.e., relative to the perception system), it has further been suggested that this system is also more susceptible to breakdown in cases of developmental disorder (Atkinson et al., 1997; Braddick et al., 2003; Dilks et al., 2008). Consistent with this possibility, we found that individuals with WS (a genetic, developmental disorder) are selectively impaired in visually-guided navigation, not scene categorization. An interesting further question, however, is whether individuals with any developmental disorder will show this pattern (Spencer et al., 2000), or whether this pattern is specific to WS. Indeed, given that WS is caused by a specific genetic deletion (Ewart et al., 1993), involves cortical thinning in the parietal lobe (the location of the dorsal action system) (Meyer-Lindenberg et al., 2004; Meyer-Lindenberg et al., 2006), and that many genes are expressed in targeted regions of cortex (for example, 70% of gene signals in mice localize to fewer than 20% of all brain cells, indicating that most genes localize to small brain regions) (Lein et al., 2007), it is possible that this pattern of results is a specific consequence of the genetic deletion in WS, targeting the dorsal stream in particular, and not the ventral stream. Beyond revealing the more precise nature of the genetic impairment in WS, such a finding would suggest the intriguing hypothesis that the division between the dorsal and ventral visual systems has a genetic basis.

Notably, although we present neural evidence that the visually-guided navigation system in OPA is late emerging, and found behaviorally that scene categorization ability develops earlier than visually-guided navigation ability, we did not test the complementary neural prediction that the scene categorization system in PPA develops early, since Paper 3 did not test for a specific signature of scene categorization function. Future work will therefore be required to understand how scene categorization function develops in PPA. Moreover, given that we limited our scope to the visually-guided navigation system (as supported by OPA), none of the papers in the present theses tested the development of memory-guided navigation (as supported by RSC). Given the overarching hypothesis that navigation develops later than scene categorization, we predict that memory-guided navigation function in RSC will mature later than scene categorization function in PPA. However, it is unclear how regions within the navigation system will develop. One possibility is that the navigation system as a whole develops in tandem. However, given that OPA and RSC are hypothesized to support distinct navigational functions, it may also be the case that these systems will develop independently as well.

A final intriguing question concerns the developmental origins of the distinction between the navigation and categorization systems. Do the scene categorization and navigation systems emerge slowly over the first few years of life, consistent with a role for experience in shaping their development? Or are the distinct foundations of these systems detectable from the earliest days of life, allowing less room for experience in shaping their development? One clue comes from a recent finding that the scene processing system already shows differential connectivity from the face-processing system in as little as 27 days on average (Kamps et al., 2019b). Thus, an intriguingly possibility is that regions within the scene processing system – while sharing some connections that drive their overall selectivity for scenes – also show distinct patterns of connectivity from one another, accounting for the later divergence in their functions within scene processing. For example, by adulthood, OPA shows stronger connectivity than PPA to areas of parietal cortex that putatively support visually-guided action and egocentric spatial processing (Baldassano et al., 2013). If present in infancy, these connections could bias the OPA (but not the PPA) toward visually-guided action processing on scenes (i.e., visually-guided navigation). Beyond this connectivity hypothesis, it is also possible that cellular and molecular properties of these systems bias their functions toward action versus perception. For example, although the relationship between cytoarchitecture and function is far from established (Weiner et al., 2017), OPA and PPA reside in distinct cytoarchitectonic regions (Zilles and Amunts, 2010), consistent with this possibility.

Whatever the ultimate answers to these questions, the present thesis provides converging evidence for the hypothesis that human scene processing depends on dissociable systems for scene processing and navigation, and further, that these systems develop independently, with the navigation system developing later than the categorization system.

References

Aguirre GK, D'Esposito M (1999) Topographical disorientation: a synthesis and taxonomy. Brain 122 (Pt 9):1613-1628.

Aguirre GK, Zarahn E, D'Esposito M (1998) An area within human ventral cortex sensitive to "building" stimuli: Evidence and implications. Neuron 21:373-383.

Atkinson J, King J, Braddick O, Nokes L, Anker S, Braddick F (1997) A specific deficit of dorsal stream function in Williams' syndrome. Neuroreport 8:1919-1922.

Atkinson J, Braddick O, Anker S, Curran W, Andrew R, Wattam-Bell J, Braddick F (2003) Neurobiological models of visuospatial cognition in children with Williams syndrome: measures of dorsal-stream and frontal function. Dev Neuropsychol 23:139-172.

Auger SD, Mullally SL, Maguire EA (2012) Retrosplenial cortex codes for permanent landmarks. Plos One 7:e43620.

Auger SD, Zeidman P, Maguire EA (2015) A central role for the retrosplenial cortex in de novo environmental learning. Elife 4.

Baldassano C, Beck DM, Fei-Fei L (2013) Differential connectivity within the Parahippocampal Place Area. NeuroImage 75:228-237.

Bar M (2004) Visual objects in context. Nature reviews Neuroscience 5:617-629.

Bar M, Aminoff E (2003) Cortical analysis of visual context. Neuron 38:347-358.

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.

Bellugi U, Bihrle AM, Neville HJ, Doherty S, Jernigan TL (1992) Language, cognition, and brain organization in a neurodevelopmental disorder. In: Developmental Behavioral Neuroscience: The Minnesota Symposia on Child Psychology (Gunner M, Nelson C, eds), pp 201-232. Hillsdale, NJ: Erlbaum Press.

Berard JR, Vallis LA (2006) Characteristics of single and double obstacle avoidance strategies: a comparison between adults and children. Exp Brain Res 175:21-31.

Bertenthal BI (1996) Origins and early development of perception, action, and representation. Annu Rev Psychol 47:431-459.

Bettencourt KC, Xu YD (2013) The Role of Transverse Occipital Sulcus in Scene Perception and Its Relationship to Object Individuation in Inferior Intraparietal Sulcus. J Cognitive Neurosci 25:1711-1722.

Bonner MF, Epstein RA (2017) Coding of navigational affordances in the human visual system. Proc Natl Acad Sci U S A 114:4793-4798.

Bostelmann M, Fragniere E, Costanzo F, Di Vara S, Menghini D, Vicari S, Lavenex P, Lavenex PB (2017) Dissociation of spatial memory systems in Williams syndrome. Hippocampus 27:1192-1203.

Braddick O, Atkinson J, Wattam-Bell J (2003) Normal and anomalous development of visual motion processing: motion coherence and 'dorsal-stream vulnerability'. Neuropsychologia 41:1769-1784.

Brainard DH (1997) The Psychophysics Toolbox. Spatial Vision 10:433-436.

Britten KH, van Wezel RJ (1998) Electrical microstimulation of cortical area MST biases heading perception in monkeys. Nature neuroscience 1:59-63.

Broadbent HJ, Farran EK, Tolmie A (2014) Egocentric and allocentric navigation strategies in Williams syndrome and typical development. Developmental science 17:920-934.

Burgess N (2008) Spatial cognition and the brain. Annals of the New York Academy of Sciences 1124:77-97.

Burgess N, Becker S, King JA, O'Keefe J (2001) Memory for events and their spatial context: models and experiments. Philos T R Soc B 356:1493-1503.

Byrne P, Becker S, Burgess N (2007) Remembering the past and imagining the future: A neural model of spatial memory and imagery. Psychol Rev 114:340-375.

Cain DP, Boon F, Corcoran ME (2006) Thalamic and hippocampal mechanisms in spatial navigation: a dissociation between brain mechanisms for learning how versus learning where to navigate. Behavioural brain research 170:241-256.

Cant JS, Goodale MA (2011) Scratching Beneath the Surface: New Insights into the Functional Properties of the Lateral Occipital Area and Parahippocampal Place Area. J Neurosci 31:8248-8258.

Cavina-Pratesi C, Goodale MA, Culham JC (2007) FMRI Reveals a Dissociation between Grasping and Perceiving the Size of Real 3D Objects. Plos One 2.

Chai XJ, Ofen N, Jacobs LF, Gabrieli JDE (2010) Scene complexity: influence on perception, memory, and development in the medial temporal lobe. Front Hum Neurosci 4.

Cheng K, Newcombe NS (2005) Is there a geometric module for spatial orientation? Squaring theory and evidence. Psychon Bull Rev 12:1-23.

Ciaramelli E, Rosenbaum RS, Solcz S, Levine B, Moscovitch M (2010) Mental Space Travel: Damage to Posterior Parietal Cortex Prevents Egocentric Navigation and Reexperiencing of Remote Spatial Memories. J Exp Psychol Learn 36:619-634.

Cohen MA, Dilks DD, Koldewyn K, Weigelt S, Feather J, Kell AJ, Keil B, Fischl B, Zollei L, Wald L, Saxe R, Kanwisher N (2019) Representational similarity precedes category selectivity in the developing ventral visual pathway. NeuroImage 197:565-574.

Cooper BG, Manka TF, Mizumori SJ (2001) Finding your way in the dark: the retrosplenial cortex contributes to spatial memory and navigation without visual cues. Behavioral neuroscience 115:1012-1028.

Cox RW, Jesmanowicz A (1999) Real-time 3D image registration for functional MRI. Magn Reson Med 42:1014-1018.

Csibra G, Tucker LA, Johnson MH (1998) Neural correlates of saccade planning in infants: a high-density ERP study. Int J Psychophysiol 29:201-215.

Culham JC, Danckert SL, DeSouza JF, Gati JS, Menon RS, Goodale MA (2003) Visually guided grasping produces fMRI activation in dorsal but not ventral stream brain areas. Exp Brain Res 153:180-189.

Deen B, Richardson H, Dilks DD, Takahashi A, Keil B, Wald LL, Kanwisher N, Saxe R (2017) Organization of high-level visual cortex in human infants. Nat Commun 8:13995.

Diamond A (2006) The Early Development of Executive Functions. Lifespan cognition: Mechanisms of change 210:70-95.

Diamond A, Goldman-Rakic PS (1989) Comparison of human infants and infant rhesus monkeys on Piaget's AB task: evidence for dependence on dorsolateral prefrontal cortex. Exp Brain Res 74:24-40.

Diamond A, Zola-Morgan S, Squire LR (1989) Successful performance by monkeys with lesions of the hippocampal formation on AB and object retrieval, two tasks that mark developmental changes in human infants. Behavioral neuroscience 103:526– 537.

Dilks DD, Hoffman JE, Landau B (2008) Vision for perception and vision for action: normal and unusual development. Developmental science 11:474-486.

Dilks DD, Julian JB, Paunov AM, Kanwisher N (2013) The occipital place area is causally and selectively involved in scene perception. J Neurosci 33:1331-1336a.

Dilks DD, Julian JB, Kubilius J, Spelke ES, Kanwisher N (2011) Mirror-image sensitivity and invariance in object and scene processing pathways. The Journal of neuroscience : the official journal of the Society for Neuroscience 31:11305-11312.

Dillon MR, Persichetti AS, Spelke ES, Dilks DD (2018) Places in the Brain: Bridging Layout and Object Geometry in Scene-Selective Cortex. Cerebral cortex 28:2365-2374.

Doeller CF, King JA, Burgess N (2008) Parallel striatal and hippocampal systems for landmarks and boundaries in spatial memory. Proc Natl Acad Sci U S A 105:5915-5920.

Epstein (2008) Parahippocampal and retrosplenial contributions to human spatial navigation. Trends in cognitive sciences 12:388-396.

Epstein, Kanwisher (1998a) A cortical representation of the local visual environment. Nature 392:598-601.

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

Epstein R, Graham KS, Downing PE (2003) Viewpoint-specific scene representations in human parahippocampal cortex. Neuron 37:865-876.

Epstein R, Harris A, Stanley D, Kanwisher N (1999) The parahippocampal place area: recognition, navigation, or encoding? Neuron 23:115-125.

Epstein R, Deyoe EA, Press DZ, Rosen AC, Kanwisher N (2001) Neuropsychological evidence for a topographical learning mechanism in parahippocampal cortex. Cogn Neuropsychol 18:481-508.

Epstein RA, Ward EJ (2010) How Reliable Are Visual Context Effects in the Parahippocampal Place Area? Cerebral cortex 20:294-303.

Epstein RA, Vass LK (2015) Neural systems for landmark-based wayfinding in humans. Philos T R Soc B 370.

Epstein RA, Higgins JS, Thompson-Schill SL (2005) Learning places from views: variation in scene processing as a function of experience and navigational ability. J Cogn Neurosci 17:73-83.

Epstein RA, Parker WE, Feiler AM (2007a) Where am I now? Distinct roles for parahippocampal and retrosplenial cortices in place recognition. J Neurosci 27:6141-6149.

Epstein RA, Higgins JS, Jablonski K, Feiler AM (2007b) Visual scene processing in familiar and unfamiliar environments. Journal of neurophysiology 97:3670-3683.

Ewart AK, Morris CA, Atkinson D, Jin W, Sternes K, Spallone P, Stock AD, Leppert M, Keating MT (1993) Hemizygosity at the elastin locus in a developmental disorder, Williams syndrome. Nature genetics 5:11-16.

Farran EK, Blades M, Boucher J, Tranter LJ (2010) How do individuals with Williams syndrome learn a route in a real-world environment? Developmental science 13:454-468.

Fedorenko E, Hsieh PJ, Nieto-Castanon A, Whitfield-Gabrieli S, Kanwisher N (2010) New method for fMRI investigations of language: defining ROIs functionally in individual subjects. Journal of neurophysiology 104:1177-1194.

Ferrara K, Landau B (2015) Geometric and featural systems, separable and combined: Evidence from reorientation in people with Williams syndrome. Cognition 144:123-133.

Foti F, Petrosini, L., Cutuli, D., Menghini, D., Chiarotti, F., Vicari, S., & Mandolesi, L. (2011) Explorative function in Williams syndrome analyzed through a large-scale task with multiple rewards. Research in developmental disabilities 32:972-985.

Franchak JM, Adolph KE (2010a) Visually guided navigation: head-mounted eye-tracking of natural locomotion in children and adults. Vision research 50:2766-2774.

Franchak JM, Adolph KE (2010b) Visually guided navigation: Head-mounted eye-tracking of natural locomotion in children and adults. Vision research 50:2766-2774.

Gallivan JP, Goodale MA (2018) The dorsal "action" pathway. Handb Clin Neurol 151:449-466.

Ganaden RE, Mullin CR, Steeves JK (2013) Transcranial magnetic stimulation to the transverse occipital sulcus affects scene but not object processing. J Cogn Neurosci 25:961-968.

Ghaem O, Mellet E, Crivello F, Tzourio N, Mazoyer B, Berthoz A, Denis M (1997) Mental navigation along memorized routes activates the hippocampus, precuneus, and insula. Neuroreport 8:739-744.

Gibson EJ, Walk RD (1960) The "visual cliff". Sci Am 202:64-71.

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

Gilmore RO, Johnson MH (1997a) Egocentric action in early infancy: spatial frames of reference for saccades. Psychological science 8:224-230.

Gilmore RO, Johnson MH (1997b) Body-centered representations for visually-guided action emerge during early infancy. Cognition 65:B1-B9.

Golarai G, Liberman A, Yoon JM, Grill-Spector K (2010) Differential development of the ventral visual cortex extends through adolescence. Front Hum Neurosci 3:80.

Golarai G, Ghahremani DG, Whitfield-Gabrieli S, Reiss A, Eberhardt JL, Gabrieli JD, Grill-Spector K (2007) Differential development of high-level visual cortex correlates with category-specific recognition memory. Nature neuroscience 10:512-522.

Goodale MA, Milner AD (1992) Separate visual pathways for perception and action. Trends in Neuroscience 15:20-25.

Goodale MA, Milner AD, Jakobson LS, Carey DP (1991) A neurological dissociation between perceiving objects and grasping them. Nature 349:154-156.

Greene MR, Oliva A (2009) Recognition of natural scenes from global properties: seeing the forest without representing the trees. Cognitive psychology 58:137-176.

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 (2012) Deconstructing visual scenes in cortex: gradients of object and spatial layout information. J Mol Neurosci 48:S50-S50.

Hasson U, Harel M, Levy I, Malach R (2003) Large-scale mirror-symmetry organization of human occipito-temporal object areas. Neuron 37:1027-1041.

Haxby JV, Ungerleider LG, Clark VP, Schouten JL, Hoffman EA, Martin A (1999) The effect of face inversion on activity in human neural systems for face and object perception. Neuron 22:189-199.

Hecaen H, Tzortzis C, Rondot P (1980) Loss of Topographic Memory with Learning-Deficits. Cortex; a journal devoted to the study of the nervous system and behavior 16:525-542.

Henderson JM, Larson CL, Zhu DC (2007) Cortical activation to indoor versus outdoor scenes: an fMRI study. Exp Brain Res 179:75-84.

Henderson JM, Larson CL, Zhu DC (2008) Full Scenes produce more activation than Close-up Scenes and Scene-Diagnostic Objects in parahippocampal and retrosplenial cortex: An fMRI study. Brain and cognition 66:40-49.

Hermer L, Spelke ES (1994) A geometric process for spatial reorientation in young children. Nature 370:57-59.

Hermer L, Spelke E (1996) Modularity and development: the case of spatial reorientation. Cognition 61:195-232.

Iaria G, Chen JK, Guariglia C, Ptito A, Petrides M (2007) Retrosplenial and hippocampal brain regions in human navigation: complementary functional contributions to the formation and use of cognitive maps. Eur J Neurosci 25:890-899.

Ino T, Inoue Y, Kage M, Hirose S, Kimura T, Fukuyama H (2002) Mental navigation in humans is processed in the anterior bank of the parieto-occipital sulcus. Neuroscience letters 322:182-186.

Janzen G, van Turennout M (2004) Selective neural representation of objects relevant for navigation. Nature neuroscience 7:673-677.

Jarrold C, Baddeley AD, Hewes AK (1998) Verbal and Nonverbal Abilities in the Williams Syndrome Phenotype: Evidence for Diverging Developmental Trajectories. J Child Psychol Psychiat 39:511-523.

Jiang P, Tokariev M, Aronen ET, Salonen O, Ma YY, 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.

Johnson MR, Johnson MK (2014) Decoding individual natural scene representations during perception and imagery. Front Hum Neurosci 8.

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.

Julian JB, Ryan J, Hamilton R, Epstein R (2015) The Occipital Place Area is causally involved in representing environmental boundaries during navigation. Journal of vision 15:512-512.

Julian JB, Ryan J, Hamilton RH, Epstein RA (2016) The occipital place area is causally involved in representing environmental boundaries during navigation. Curr Biol 26:1104-1109.

Julian JB, Kamps FS, Epstein RA, Dilks DD (2019) Dissociable spatial memory systems revealed by typical and atypical human development. Developmental science 22:e12737.

Kamps FS, Lall V, Dilks DD (2016a) Occipital place area represents first-person perspective motion through scenes. Cortex; a journal devoted to the study of the nervous system and behavior 83:17-26.

Kamps FS, Morris EJ, Dilks DD (2019a) A face is more than just the eyes, nose, and mouth: fMRI evidence that face-selective cortex represents external features. NeuroImage 184:90-100.

Kamps FS, Hendrix CL, Brennan PA, Dilks DD (2019b) Connectivity at the origins of domain specificity: the case of the cortical face network. Vision Sciences Society Annual Meeting.

Kamps FS, Julian JB, Kubilius J, Kanwisher N, Dilks DD (2016b) The occipital place area represents the local elements of scenes. NeuroImage 132:417-424.

Kamps FS, Pincus JE, Radwan SF, Wahab S, Dilks DD (Submitted) Neural evidence for the protracted development of human visually-guided navigation.

Kamps FS, Julian JB, Battaglia P, Landau B, Kanwisher N, Dilks DD (2017) Dissociating intuitive physics from intuitive psychology: Evidence from Williams syndrome. Cognition 168:146-153.

Kanwisher N, Dilks D (2013) The functional organization of the ventral visual pathway in humans. In: The New Visual Neurosciences (Chalupa L, Werner J, eds). Cambridge, MA: The MIT Press.

Karmiloff-Smith A (1997) Crucial differences between developmental cognitive neuroscience and adult neuropsychology. Developmental Neuropsychology 13:513-524.

Kauffmann L, Ramanoel S, Guyader N, Chauvin A, Peyrin C (2015) Spatial frequency processing in scene-selective cortical regions. NeuroImage 112:86-95.

Kaufman AS, & Kaufman, N. L., ed (1990) Manualforthe Kaufman Brief Intelligence Test. Circle Pines, MN: American Guidance Service.

Kobayashi Y, Amaral DG (2003) Macaque monkey retrosplenial cortex: II. Cortical afferents. J Comp Neurol 466:48-79.

Kobayashi Y, Amaral DG (2007) Macaque monkey retrosplenial cortex: III. Cortical efferents. J Comp Neurol 502:810-833.

Konkle T, Oliva A (2012) A Real-World Size Organization of Object Responses in Occipitotemporal Cortex. Neuron 74:1114-1124.

Kravitz DJ, Baker CI (2011) Toward a new model of scientific publishing: discussion and a proposal. Frontiers in computational neuroscience 5:55.

Kravitz DJ, Peng CS, Baker CI (2011a) Real-world scene representations in high-level visual cortex: it's the spaces more than the places. The Journal of neuroscience : the official journal of the Society for Neuroscience 31:7322-7333.

Kravitz DJ, Saleem KS, Baker CI, Mishkin M (2011b) A new neural framework for visuospatial processing. Nature reviews Neuroscience 12:217-230.

Kriegeskorte N, Formisano E, Sorger B, Goebel R (2007) Individual faces elicit distinct response patterns in human anterior temporal cortex. Proc Natl Acad Sci U S A 104:20600-20605.

Kriegeskorte N, Simmons WK, Bellgowan PS, Baker CI (2009) Circular analysis in systems neuroscience: the dangers of double dipping. Nature neuroscience 12:535-540.

Lakusta L, Dessalegn B, Landau B (2010) Impaired geometric reorientation caused by genetic defect. Proc Natl Acad Sci U S A 107:2813-2817.

Landau B, Hoffman JE (2012) Spatial representation : from gene to mind. Oxford ; New York: Oxford University Press.

Landau B, Ferrara K (2013) Space and language in Williams syndrome: insights from typical development. Wires Cogn Sci 4:693-706.

Langston RF, Ainge JA, Couey JJ, Canto CB, Bjerknes TL, Witter MP, Moser EI, Moser MB (2010) Development of the Spatial Representation System in the Rat. Science 328:1576-1580.

Lein ES et al. (2007) Genome-wide atlas of gene expression in the adult mouse brain. Nature 445:168-176.

MacEvoy SP, Epstein RA (2007) Position selectivity in scene- and object-responsive occipitotemporal regions. Journal of neurophysiology 98:2089-2098.

MacEvoy SP, Epstein RA (2011) Constructing scenes from objects in human occipitotemporal cortex. Nature neuroscience 14:1323-1329.

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

Malouin F, Richards CL, Jackson PL, Dumas F, Doyon J (2003) Brain activations during motor imagery of locomotor-related tasks: a PET study. Human brain mapping 19:47-62.

Marchette SA, Vass LK, Ryan J, Epstein RA (2014) Anchoring the neural compass: coding of local spatial reference frames in human medial parietal lobe. Nature neuroscience 17:1598-1606.

Marchette SA, Vass LK, Ryan J, Epstein RA (2015a) Anchoring the neural compass: coding of local spatial reference frames in human medial parietal lobe (vol 17, pg 1598, 2014). Nature neuroscience 18:926-926.

Marchette SA, Vass LK, Ryan J, Epstein RA (2015b) Outside Looking In: Landmark Generalization in the Human Navigational System. J Neurosci 35:14896-14908.

McKone E, Crookes K, Jeffery L, Dilks DD (2012) A critical review of the development of face recognition: Experience is less important than previously believed. Cogn Neuropsychol 29:174-212.

Megevand P, Groppe DM, Goldfinger MS, Hwang ST, Kingsley PB, Davidesco I, Mehta AD (2014) Seeing Scenes: Topographic Visual Hallucinations Evoked by Direct Electrical Stimulation of the Parahippocampal Place Area. J Neurosci 34:5399-5405.

Meissner TW, Nordt M, Weigelt S (2019) Prolonged functional development of the parahippocampal place area and occipital place area. NeuroImage 191:104-115.

Menghini D, Addona F, Costanzo F, Vicari S (2010) Executive functions in individuals with Williams syndrome. J Intellect Disabil Res 54:418-432.

Meyer-Lindenberg A, Mervis CB, Berman KF (2006) Neural mechanisms in Williams syndrome: a unique window to genetic influences on cognition and behaviour. Nature Reviews Neuroscience 7:380-393.

Meyer-Lindenberg A, Kohn P, Mervis CB, Kippenhan JS, Olsen RK, Morris CA, Berman KF (2004) Neural basis of genetically determined visuospatial construction deficit in Williams syndrome. Neuron 43:623-631.

Miyake A, Friedman NP, Emerson MJ, Witzki AH, Howerter A, Wager TD (2000) The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: a latent variable analysis. Cognitive psychology 41:49-100.

Mobbs D, Eckert MA, Mills D, Korenberg J, Bellugi U, Galaburda AM, Reiss AL (2007) Frontostriatal dysfunction during response inhibition in Williams syndrome. Biological psychiatry 62:256-261.

Morgan LK, Macevoy SP, Aguirre GK, Epstein RA (2011) Distances between real-world locations are represented in the human hippocampus. J Neurosci 31:1238-1245.

Mullally SL, Maguire EA (2011) A new role for the parahippocampal cortex in representing space. J Neurosci 31:7441-7449.

Musolino J, Landau B (2012) Genes, language, and the nature of scientific explanations: The case of Williams syndrome. Cogn Neuropsychol 29:123-148.

Nakamura K, Kawashima R, Sato N, Nakamura A, Sugiura M, Kato T, Hatano K, Ito K, Fukuda H, Schormann T, Zilles K (2000) Functional delineation of the human occipito-temporal areas related to face and scene processing. A PET study. Brain : a journal of neurology 123 (Pt 9):1903-1912.

Nasr S, Tootell RBH (2012) A Cardinal Orientation Bias in Scene-Selective Visual Cortex. J Neurosci 32:14921-14926.

Nasr S, Echavarria CE, Tootell RBH (2014) Thinking Outside the Box: Rectilinear Shapes Selectively Activate Scene-Selective Cortex. J Neurosci 34:6721-6735.

Nasr S, Liu N, Devaney KJ, Yue X, Rajimehr R, Ungerleider LG, Tootell RB (2011) Scene-selective cortical regions in human and nonhuman primates. The Journal of neuroscience : the official journal of the Society for Neuroscience 31:13771-13785.

Newcombe NS (2019) Navigation and the developing brain. J Exp Biol 222.

Norman-Haignere SV, Albouy P, Caclin A, McDermott JH, Kanwisher NG, Tillmann B (2016) Pitch-Responsive Cortical Regions in Congenital Amusia. J Neurosci 36:2986-2994.

O'Craven KM, Kanwisher N (2000) Mental imagery of faces and places activates corresponding stiimulus-specific brain regions. J Cogn Neurosci 12:1013-1023.

O'Hearn K, Courtney S, Street W, Landau B (2009) Working memory impairment in people with Williams syndrome: effects of delay, task and stimuli. Brain and cognition 69:495-503.

O'hearn K, Roth JK, Courtney SM, Luna B, Street W, Terwillinger R, Landau B (2011) Object recognition in Williams syndrome: uneven ventral stream activation. Developmental science 14:549-565.

O'Keefe J, Nadel L (1978) The hippocampus as a cognitive map. Oxford: Clarendon Press.

Ofen N, Chai XQJ, Schuil KDI, Whitfield-Gabrieli S, Gabrieli JDE (2012) The Development of Brain Systems Associated with Successful Memory Retrieval of Scenes. J Neurosci 32:10012-10020.

Oliva A, Schyns PG (1997) Coarse blobs or fine edges? Evidence that information diagnosticity changes the perception of complex visual stimuli. Cognitive psychology 34:72-107.

Oliva A, Torralba A (2001) Modeling the shape of a scene: a holistic representation of the spatial envelope. International journal in computer vision 42:145-175.

Oliva A, Torralba A (2002) Scene-centered description from spatial envelope properties. Lect Notes Comput Sc 2525:263-272.

Pallis CA (1955) Impaired Identification of Faces and Places with Agnosia for Colours - Report of a Case Due to Cerebral Embolism. J Neurol Neurosur Ps 18:218-224.

Palomares M, Landau B, Egeth H (2009) Orientation perception in Williams Syndrome: discrimination and integration. Brain and cognition 70:21-30.

Park S, Chun MM (2009) Different roles of the parahippocampal place area (PPA) and retrosplenial cortex (RSC) in panoramic scene perception. NeuroImage 47:1747-1756.

Park S, Brady TF, Greene MR, Oliva A (2011) Disentangling scene content from spatial boundary: complementary roles for the parahippocampal place area and lateral occipital complex in representing real-world scenes. The Journal of neuroscience : the official journal of the Society for Neuroscience 31:1333-1340.

Pelphrey KA, Lopez J, Morris JP (2009) Developmental continuity and change in responses to social and nonsocial categories in human extrastriate visual cortex. Front Hum Neurosci 3:25.

Persichetti A, Dilks DD (2018a) Dissociable neural systems for recognizing places and navigating through them. J Neurosci 38.

Persichetti AS, Dilks DD (2016) Perceived egocentric distance sensitivity and invariance across scene-selective cortex Cortex; a journal devoted to the study of the nervous system and behavior 77:155-163.

Persichetti AS, Dilks DD (2018b) Dissociable Neural Systems for Recognizing Places and Navigating through Them. J Neurosci 38:10295-10304.

Persichetti AS, Dilks DD (in revision) Distinct representations of spatial and categorical relationship across human scene-selective cortex. Proceedings of the National Academy of Sciences.

Pitcher D, Dilks DD, Saxe RR, Triantafyllou C, Kanwisher N (2011) Differential selectivity for dynamic versus static information in face-selective cortical regions. NeuroImage 56:2356-2363.

Pryde KM, Roy EA, Patla AE (1997) Age-related trends in locomotor ability and obstacle avoidance. Hum Movement Sci 16:507-516.

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

Rauchs G, Orban P, Balteau E, Schmidt C, Degueldre C, Luxen A, Maquet P, Peigneux P (2008) Partially segregated neural networks for spatial and contextual memory in virtual navigation. Hippocampus 18:503-518.

Rosenbaum RS, Ziegler M, Winocur G, Grady CL, Moscovitch M (2004) "I have often walked down this street before": fMRI studies on the hippocampus and other structures during mental navigation of an old environment. Hippocampus 14:826-835.

Saxe R, Brett M, Kanwisher N (2006) Divide and conquer: a defense of functional localizers. NeuroImage 30:1088-1096; discussion 1097-1089.

Saygin ZM, Osher DE, Norton ES, Youssoufian DA, Beach SD, Feather J, Gaab N, Gabrieli JDE, Kanwisher N (2016) Connectivity precedes function in the development of the visual word form area. Nature neuroscience 19:1250-1255.

Scherf KS, Behrmann M, Humphreys K, Luna B (2007) Visual category-selectivity for faces, places and objects emerges along different developmental trajectories. Developmental science 10:F15-30.

Scherf KS, Luna B, Avidan G, Behrmann M (2011a) "What" precedes "which": developmental neural tuning in face- and place-related cortex. Cerebral cortex 21:1963-1980.

Scherf KS, Luna B, Avidan G, Behrmann M (2011b) "What" Precedes "Which": Developmental Neural Tuning in Face- and Place-Related Cortex. Cerebral cortex 21:1963-1980.

Schinazi VR, Epstein RA (2010) Neural correlates of real-world route learning. NeuroImage 53:725-735.

Schindler A, Bartels A (2016) Motion parallax links visual motion areas and scene regions. NeuroImage 125:803-812.

Schwarzlose RF, Swisher JD, Dang S, Kanwisher N (2008) The distribution of category and location information across object-selective regions in human visual cortex. Proc Natl Acad Sci U S A 105:4447-4452.

Sergent J, Ohta S, MacDonald B (1992) Functional neuroanatomy of face and object processing. A positron emission tomography study. Brain 115 Pt 1:15-36.

Silson EH, Chan AW, Reynolds RC, Kravitz DJ, Baker CI (2015) A Retinotopic Basis for the Division of High-Level Scene Processing between Lateral and Ventral Human Occipitotemporal Cortex. J Neurosci 35:11921-11935.

Smith SM, Nichols TE (2009) Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. NeuroImage 44:83-98.

Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TEJ, Johansen-Berg H, Bannister PR, De Luca M, Drobnjak I, Flitney DE, Niazy RK, Saunders J, Vickers J, Zhang YY, De Stefano N, Brady JM, Matthews PM (2004) Advances in functional and structural MR image analysis and implementation as FSL. NeuroImage 23:S208-S219.

Spelke E, Lee SA, Izard V (2010) Beyond core knowledge: Natural geometry. Cogn Sci 34:863-884.

Spencer J, O'Brien J, Riggs K, Braddick O, Atkinson J, Wattam-Bell J (2000) Motion processing in autism: evidence for a dorsal stream deficiency. Neuroreport 11:2765-2767.

Spiers HJ, Maguire EA (2007) A navigational guidance system in the human brain. Hippocampus 17:618-626.

Stark M, Coslett HB, Saffran EM (1996) Impairment of an egocentric map of locations: Implications for perception and action. Cogn Neuropsychol 13:481-523.

Sugiura M, Shah NJ, Zilles K, Fink GR (2005) Cortical representations of personally familiar objects and places: functional organization of the human posterior cingulate cortex. J Cogn Neurosci 17:183-198.

Suzuki WA, Amaral DG (1994) Perirhinal and Parahippocampal Cortices of the Macaque Monkey - Cortical Afferents. J Comp Neurol 350:497-533.

Takahashi N, Kawamura M, Hirayama K, Tagawa K (1989) [Non-verbal facial and topographic visual object agnosia--a problem of familiarity in prosopagnosia and topographic disorientation]. No to shinkei= Brain and nerve 41:703-710.

Takahashi N, Kawamura M, Shiota J, Kasahata N, Hirayama K (1997) Pure topographic disorientation due to right retrosplenial lesion. Neurology 49:464-469.

Tootell RB, Reppas JB, Kwong KK, Malach R, Born RT, Brady TJ, Rosen BR, Belliveau JW (1995) Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. J Neurosci 15:3215-3230.

Troiani V, Stigliani A, Smith ME, Epstein RA (2014) Multiple Object Properties Drive Scene-Selective Regions. Cerebral cortex 24:883-897.

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 Press.

Valyear KF, Culham JC, Sharif N, Westwood D, Goodale MA (2006) A double dissociation between sensitivity to changes in object identity and object orientation in the ventral and dorsal visual streams: A human fMRI study. Neuropsychologia 44:218-228.

van Assche M, Kebets V, Vuilleumier P, Assal F (2016) Functional Dissociations Within Posterior Parietal Cortex During Scene Integration and Viewpoint Changes. Cerebral cortex 26:586-598.

Vann SD, Aggleton JP, Maguire EA (2009) What does the retrosplenial cortex do? Nature Reviews Neuroscience 10:792-U750.

Vass LK, Epstein RA (2013) Abstract Representations of Location and Facing Direction in the Human Brain. J Neurosci 33:6133-6142.

Vul E, Kanwisher N (2010) Begging the question: The non-independence error in fMRI data analysis. In: Foundations and Philosophy for Neuroimaging (Hanson S, Bunzl M, eds), pp 71-91.

Vuontela V, Jiang P, Tokariev M, Savolainen P, Ma YY, Aronen ET, Fontell T, Liiri T, Ahlstrom M, Salonen O, Carlson S (2013) Regulation of brain activity in the fusiform face and parahippocampal place areas in 7-11-year-old children. Brain and cognition 81:203-214.

Walther DB, Caddigan E, Fei-Fei L, Beck DM (2009) Natural scene categories revealed in distributed patterns of activity in the human brain. J Neurosci 29:10573-10581.

Walther DB, Chai B, Caddigan E, Beck DM, Fei-Fei L (2011) Simple line drawings suffice for functional MRI decoding of natural scene categories. P Natl Acad Sci USA 108:9661-9666.

Wang L, Mruczek REB, Arcaro MJ, Kastner S (2015) Probabilistic Maps of Visual Topography in Human Cortex. Cerebral cortex 25:3911-3931.

Weigelt S, Koldewyn K, Dilks DD, Balas B, McKone E, Kanwisher N (2014) Domain-specific development of face memory but not face perception. Developmental science 17:47-58.

Weiner KS, Barnett MA, Lorenz S, Caspers J, Stigliani A, Amunts K, Zilles K, Fischl B, Grill-Spector K (2017) The Cytoarchitecture of Domain-specific Regions in Human High-level Visual Cortex. Cerebral cortex 27:146-161.

Welsh MC, Pennington, B. F., & Groisser, D. B. (1991) A normative‐developmental study of executive function: A window on prefrontal function in children. Developmental neuropsychology 7:131-149.

Whiteley AM, Warrington EK (1978) Selective impairment of topographical memory: a single case study. Journal of neurology, neurosurgery, and psychiatry 41:575-578.

Wills TJ, Cacucci F, Burgess N, O'Keefe J (2010) Development of the Hippocampal Cognitive Map in Preweanling Rats. Science 328:1573-1576.

Wilson BA, Berry E, Gracey F, Harrison C, Stow I, Macniven J, Weatherley J, Young AW (2005) Egocentric disorientation following bilateral parietal lobe damage. Cortex; a journal devoted to the study of the nervous system and behavior 41:547-554.

Winkler AM, Ridgway GR, Webster MA, Smith SM, Nichols TE (2014) Permutation inference for the general linear model. NeuroImage 92:381-397.

Winter SS, Taube JS (2014) Head Direction Cells: From Generation to Integration. Space, Time and Memory in the Hippocampal Formation:83-106.

Wolbers T, Buchel C (2005) Dissociable retrosplenial and hippocampal contributions to successful formation of survey representations. J Neurosci 25:3333-3340.

Wolbers T, Klatzky RL, Loomis JM, Wutte MG, Giudice NA (2011) Modality-Independent Coding of Spatial Layout in the Human Brain. Curr Biol 21:984-989.

Zelazo PD, Muller U, Frye D, Marcovitch S, Argitis G, Boseovski J, Chiang JK, Hongwanishkul D, Schuster BV, Sutherland A (2003) The development of executive function in early childhood. Monogr Soc Res Child Dev 68:vii-137.

Zilles K, Amunts K (2010) Centenary of Brodmann's map--conception and fate. Nature reviews Neuroscience 11:139-145.