

Opinion

Three cortical scene systems and their development

Daniel D. Dilks,^{1,4,*} Frederik S. Kamps,² and Andrew S. Persichetti³

Since the discovery of three scene-selective regions in the human brain, a central assumption has been that all three regions directly support navigation. We propose instead that cortical scene processing regions support three distinct computational goals (and one not for navigation at all): (i) The parahippocampal place area supports scene categorization, which involves recognizing the kind of place we are in; (ii) the occipital place area supports visually guided navigation, which involves finding our way through the immediately visible environment, avoiding boundaries and obstacles; and (iii) the retrosplenial complex supports map-based navigation, which involves finding our way from a specific place to some distant, out-of-sight place. We further hypothesize that these systems develop along different timelines, with both navigation systems developing slower than the scene categorization system.

Carving up human cortical scene processing

The human ability to recognize a place (or **scene**; see [Glossary](#)) forms the bedrock for many of our essential everyday behaviors. In a brief glance, we extract a wealth of information from scenes, such as the category of the scene (e.g., ‘a kitchen’), its identity (e.g., ‘my kitchen’), and other critical properties such as whether it is safe or what behavior is appropriate for the current context. At the same time, we extract information that is vital for navigation, allowing us to effortlessly find our way through the immediately visible environment without running into the kitchen walls or banging into the kitchen table, for example. What’s more, we are able to situate the local visual environment within a broader spatial map, allowing us, for instance, to know where our favorite restaurant is relative to our house. But how do we accomplish these remarkable feats?

One promising strategy to understand human visual scene processing is to characterize the neural systems that accomplish it. Over the past three decades, cognitive neuroscience has revealed a set of three cortical regions that together make up the visual scene processing system in humans: the **parahippocampal place area (PPA)** [1], the **occipital place area (OPA)** [2], and the **retrosplenial complex (RSC)** [3]. However, beyond establishing the general involvement of these regions in human visual scene processing – that is, responding about two to four times more to images of scenes than to images of objects and faces in human fMRI studies – two fundamental questions remain unanswered: (i) What precise role does each region play within the broad domain of adult human cortical scene processing? (ii) How does this functional organization develop from infancy to adulthood?

We review adult fMRI and neuropsychological studies and developmental fMRI studies on human cortical scene processing, and we offer two hypotheses. First, we hypothesize that human cortical scene processing is composed of three distinct systems: One we call the ‘scene categorization’ system, including the PPA, which is involved in recognizing a scene as a kind of place (e.g., a kitchen), but not as a specific place (e.g., my kitchen); another we call the ‘visually guided

Highlights

It is widely believed that cortical scene processing in adult humans – composed of the parahippocampal place area (PPA), the occipital place area (OPA), and the retrosplenial complex (RSC) – is wholly devoted to navigation.

Challenging this pervasive theory, recent research suggests that the PPA is not involved in navigation but instead is involved in ‘scene categorization.’

By contrast, the OPA and RSC are involved in navigation, albeit different kinds (i.e., ‘visually guided navigation’ and ‘map-based navigation,’ respectively).

Thus, we propose that adult human cortical scene processing comprises three dissociable systems.

Furthermore, we propose that the PPA, OPA, and RSC develop along different timelines, and we offer the specific hypothesis that both the visually guided navigation and map-based navigation systems are slower to develop than the scene categorization system.

¹Department of Psychology, Emory University, Atlanta, GA 30322, USA

²Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

³Section on Cognitive Neuropsychology, Laboratory of Brain and Cognition, National Institute of Mental Health, National Institutes of Health, Bethesda, MD 20892, USA

⁴<https://secure.web.emory.edu/psychology/dilks/main.php>

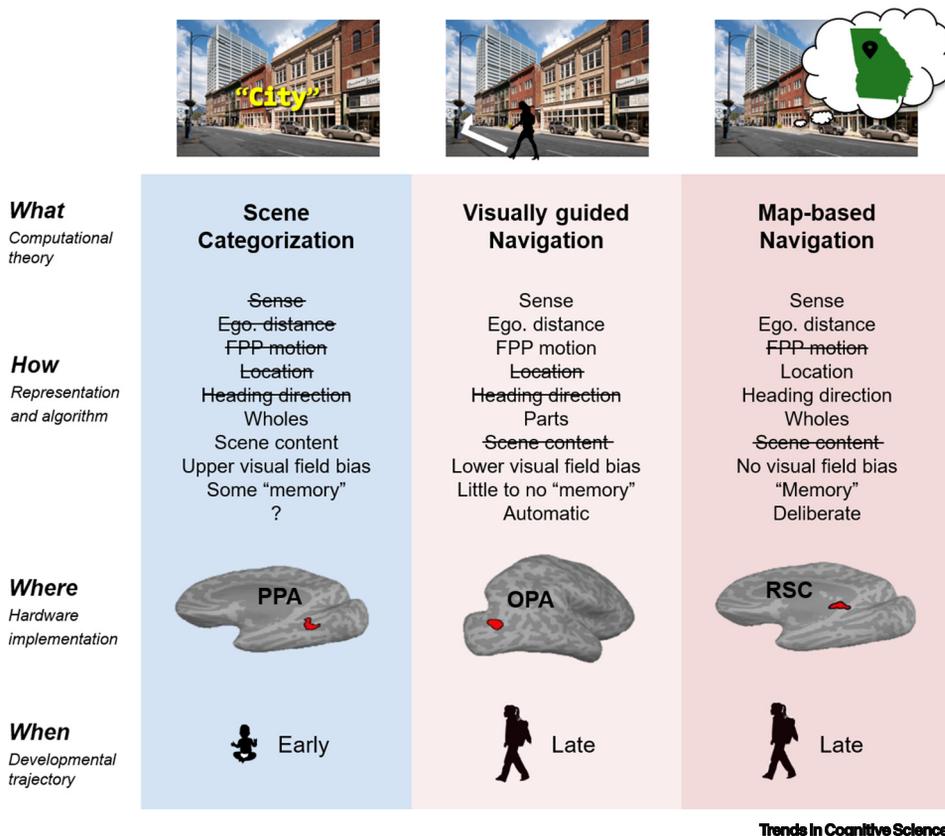
*Correspondence: dilks@emory.edu (D.D. Dilks).



navigation' system, including the OPA, which is involved in finding our way through the immediately visible environment, avoiding boundaries and obstacles; and a third we call the 'map-based navigation' system, including the RSC, which is involved in making our way from a specific place to some distant, out-of-sight place (Figure 1, Key figure). This three scene systems

Key figure

The three scene systems (a la Marr's levels of analysis) and their development



Trends in Cognitive Sciences

Figure 1. The 'scene categorization' system [including the parahippocampal place area (PPA)] supports our ability to recognize the kind of place we are in (e.g., a city versus a beach), not a specific place, and is not involved in navigation. Accordingly, it does not represent information critical for navigation: sense (left/right), egocentric distance (near/far), first-person perspective (FPP) motion, location, or heading direction. Instead, it represents two kinds of information critical for categorizing scenes: (i) spatial layout or 'wholes' (i.e., the spatial arrangement of the major surfaces that make up the overall shape of a scene) and (ii) scene content (e.g., the particular objects and textures that fill that space). By contrast, the 'visually guided' and 'map-based' navigation systems [including occipital place area (OPA) and retrosplenial complex (RSC), respectively] do represent information essential for navigation. However, they do so differently, with the visually guided navigation system representing information critical to guiding navigation through the immediately visible environment (i.e., sense, egocentric distance, FPP motion, and the local 'parts' of a scene that constitute boundaries or obstacles), and the map-based navigation system representing information critical to guiding navigation from a specific place to some distant, out-of-sight place (i.e., sense, egocentric distance, location, heading direction, and the overall shape of a scene). Each system is further dissociated based on visual field bias, memory requirements, and the automaticity of the process they support. Finally, consistent with the distinct computational goals, representations, and cortical regions across the three scene systems, each is further hypothesized to develop along a different timeline. In particular, the scene categorization system is hypothesized to develop earlier than both the visually guided navigation and map-based navigation systems.

Glossary

Landmark recognition: the ability to recognize the locations of particular places or objects in the local environment so they can be used as landmarks to orient ourselves while navigating the broader environment.

Occipital place area (OPA): a scene-selective region located in the parieto-occipital cortex, around the transverse occipital sulcus.

Parahippocampal place area (PPA): the first scene-selective cortical region reliably identified in adult humans using fMRI, the PPA is located medially in the inferior temporo-occipital cortex, between the posterior parahippocampal gyrus and the anterior lingual gyrus, and sometimes extending to the adjacent fusiform gyrus and collateral sulcus.

Reorientation: an organism's ability to find its bearings after being disoriented, using both spatial layout and egocentric information (i.e., sense and egocentric distance).

Retrosplenial complex (RSC): a scene-selective region – more recently called the medial place area [74] – located in the medial parietal cortex, immediately behind the splenium, the most caudal part of the corpus callosum, often including the precuneus and posterior cingulate cortex.

Scene: a large-scale space or environment that a person can recognize as a particular kind of place, navigate through, or situate within a broader environment.

Scene content: the internal features of a scene encompassing objects, textures, colors, and materials.

Scene-selective regions: brain regions, including the PPA, RSC, and OPA, that respond more strongly in fMRI when people view images of scenes (e.g., landscapes, cityscapes, rooms) than when they view other visual stimuli, such as objects and faces.

Spatial layout: the geometry of a space, defined by the spatial arrangement of the large, extended surfaces/planes composing the space.

hypothesis challenges the pervasive theory that human cortical scene processing is entirely in the service of navigation [1,4–18]. Second, we hypothesize that these systems develop along different timelines – more specifically, that both the visually guided navigation and map-based navigation systems are slower to develop than the scene categorization system (Figure 1).

The PPA is involved in scene categorization, not navigation

A central assumption since the discovery of the three **scene-selective regions** has been that all three directly support navigation. For example, in the seminal PPA paper, Epstein and Kanwisher [1] found that the PPA responded significantly more to images of empty rooms than to the same rooms in which the walls, floors, and ceilings had been fractured and rearranged, thus disrupting spatial layout. The authors then argued that the representation of **spatial layout** in the PPA signals its role in navigation because spatial layout information is critical for navigation and **reorientation** [19–23]. Other more recent fMRI studies found that the PPA represents certain aspects of the spatial layout of scenes (e.g., whether a scene is ‘open’ or ‘closed’) [6,7,17], and similarly argued that such encoding of spatial layout information reflects the role of PPA in navigation. But spatial layout information need not be used for navigation only and could also easily facilitate scene categorization (e.g., a beach is ‘open’; a kitchen is ‘closed’). Indeed, several behavioral and computer vision studies have found that spatial layout information (including open/closed information) can be used to categorize scenes and at times is even necessary to do so [24–28]. Thus, we argue that none of the studies claiming that the PPA is involved in navigation provide direct evidence for this claim and that all of the findings are, in fact, open to the alternative interpretation that spatial layout sensitivity reflects the role of the PPA in scene categorization.

To directly test the role that the PPA plays within human cortical scene processing, we conducted an fMRI experiment [29] in which participants were asked to perform either a scene categorization task or a visually guided navigation task on the exact same stimuli. In the scene categorization task, participants imagined standing in a room and indicated whether the scene was a kitchen, a living room, or a bedroom. In the visually guided navigation task, participants imagined walking along a path on the floor that led only to one of three doors and indicated whether they could leave through the door on the left, center, or right wall. We found that the PPA responded significantly more during the scene categorization task than during the visually guided navigation task. Crucially, we further found that the PPA did not respond any more during the visually guided navigation task than during a baseline task (i.e., a one-back task, where participants indicated whether the presented image was exactly the same as or different from the previous one), suggesting a complete lack of task-based modulation in PPA during the visually guided navigation task. By contrast, the OPA showed the exact opposite pattern, responding significantly more during the visually guided navigation task than during both the scene categorization and baseline tasks. These results reveal a double dissociation between the responses in the PPA and the OPA and thus strongly suggest distinct neural systems selectively involved in scene categorization and visually guided navigation, respectively. Meanwhile, the RSC responded similarly during all three tasks, suggesting the lack of RSC involvement in either scene categorization or visually guided navigation, consistent with our hypothesis that the RSC and the OPA play different roles even within navigation – a topic we explore in the next section. Additional evidence against the role of the PPA in visually guided navigation comes from several other fMRI studies showing that the PPA – unlike the OPA – does not represent information necessary for visually guided navigation. Specifically, the PPA does not represent (i) ‘sense’ (left–right) information [30], (ii) egocentric distance (near–far) information [31], (iii) first-person perspective motion information through scenes (i.e., actually mimicking the visual experience of navigating through immediately visible environments) [32,33], or (iv) possible routes through a local scene [34].

Intriguingly, our finding of dissociable neural systems involved in recognizing places (including the PPA in the temporo-occipital cortex) and navigating through them (including the OPA, in the parieto-occipital cortex) suggests that the human cortical scene processing system is similar to the well-characterized ventral/dorsal distinction for human object processing, with one system responsible for recognition (a ‘what’ system) and another for visually guided action (a ‘how’ system), respectively [35,36]. Note, however, that our finding of two distinct systems for scene categorization (including the PPA) and visually guided navigation (including the OPA) does not mean that the two systems cannot and do not interact, as is widely recognized in the case of the dorsal and ventral streams for object processing [35,37]. Indeed, two studies [38,39] found resting-state functional correlations between the PPA and the OPA, suggesting that these regions are connected (at least functionally, maybe anatomically), thereby enabling crosstalk between them. However, the nature of any such interaction remains mysterious, and, in any case, this interaction need not be essential to the operations performed in either system. Just as in object processing, where one need not recognize the category of an object (e.g., ‘cup’) to successfully act on it (and vice versa), in scene processing, one need not recognize the category of a scene (e.g., ‘kitchen’) to successfully navigate through it (and vice versa). Importantly, causal evidence for independent scene categorization and visually guided navigation systems is still lacking. Future studies of patients with focal lesions to either the PPA or the OPA are thus required to provide a definitive answer to this question.

However, perhaps the biggest challenge to our hypothesis that the PPA is not involved in navigation are several studies reporting that the PPA is involved in a key component of map-based navigation, known as **landmark recognition** [8–15]. In other words, although the PPA may not be involved in visually guided navigation – as described previously – it may be involved in map-based navigation. However, these otherwise elegantly designed studies have significant methodological limitations (i.e., not correcting for multiple comparisons and/or ‘double dipping’ [40], not comparing across regions to show neural specificity, and/or not matching on behavioral performance to rule out task difficulty) and hence cannot be taken to show the claimed result. Instead, a recent fMRI multivoxel pattern analysis (MVPA) study [41] – which did not suffer from the previously-described methodological problems – found that PPA does *not* represent the location of particular places (e.g., the restaurant on 10th Street versus the restaurant on Peachtree Street) or the heading direction of the navigator (e.g., the navigator is facing the restaurant in the north part of town versus the south part of town) – but instead represents only the scene category (e.g., a restaurant). The RSC showed the exact opposite pattern of results, representing location and heading direction information, but not category information.

Such a double dissociation strongly suggests distinct neural systems involved in scene categorization (including the PPA) and map-based navigation (including the RSC). Unlike the RSC or the PPA, the OPA represented neither location, nor heading direction, nor category information, again supporting our hypotheses that the RSC and the OPA play different roles within navigation and that the PPA and the OPA play different roles within scene processing more broadly. Additional evidence arguing against the role of the PPA in map-based navigation comes from two other fMRI studies also showing that the PPA – unlike the RSC – does not represent landmark information (i.e., location and/or direction information) [42,43]. Finally, another fMRI study found that, during a landmark recognition task, responses in the parahippocampal cortex (perhaps overlapping with the PPA; although the PPA was never functionally defined) – unlike the RSC – did not vary as a function of one’s ability to navigate [44,45].

Interestingly, support for the idea that the PPA is involved in landmark recognition also comes from several neuropsychological studies reporting that patients with damage to parahippocampal and

lingual gyri (areas encompassing and surrounding the PPA) commonly present with landmark agnosia – an inability to recognize landmarks [46–50]. However, in all of these studies, there is either (i) diffuse damage, including not only parahippocampal cortex but also other navigationally relevant structures including the hippocampus and retrosplenial cortex; (ii) not enough detail about the lesion to determine whether the parahippocampal cortex is even implicated at all; or (iii) a lack of the proper control conditions to rule out alternative hypotheses, especially regarding the role of PPA in scene categorization – making it impossible then to draw any clear conclusions about the function of PPA from these studies. Future work is therefore required to test patients with damage to the PPA. Indeed, another more recent study on two stroke patients with focal damage to the PPA found significant impairment in the encoding of novel scenes, yet sparing of famous landmarks [51], suggesting a role for the PPA in recognizing scene category, but not identity.

So, if PPA is not involved in navigation (visually guided or map based), then what role does it play within human cortical scene processing? We propose that the PPA plays a role in scene categorization for four reasons. First, as previously described [29], the PPA responded more during a scene categorization task than both a visually guided navigation and a baseline task. Second, as previously described [41,43], the PPA represents basic category information about places (e.g., a restaurant), but not landmark information (e.g., the restaurant on 10th Street). Third, two fMRI MVPA studies found that scene category (e.g., beach versus forest, across multiple exemplars per category) – from both color photographs and simple line drawings – is more accurately decoded from the activation patterns in the PPA than the RSC and that the activation patterns in the PPA, but not the RSC, were correlated to behavioral performance on a scene categorization task [52,53]. The OPA was not tested in either of these studies. Fourth, the PPA represents **scene content** information – information undoubtedly used for scene categorization. For example, the PPA responds to objects that are (i) good exemplars of specific scenes (e.g., a bed is found in a bedroom, while a refrigerator is found in a kitchen) [17,54–56], (ii) strongly associated with a given context (e.g., a toothbrush is only found in a bathroom) versus low ‘contextual’ objects (e.g., an apple is found in many places) [57,58], and (iii) large and not portable (e.g., a couch versus a small fan) [59–61]. Similarly, several other studies found that the PPA represents the summary statistics of object ensembles (e.g., the leaves on a tree) and texture information (e.g., brick) [62–65].

Finally, there is some limited evidence that the PPA is not a unified region, but instead can be broken up into posterior and anterior portions [38,39,66], or even by hemisphere [67]. However, despite such suggestions, there is no evidence that these proposed subregions actually support functionally distinct processing; in fact, the best available evidence suggests that these subregions support similar functions, albeit in a graded or hierarchical fashion from posterior to anterior (i.e., with stronger responses or information processing in the anterior portion than the posterior one) [56,68,69]. Critically, even if functionally distinct subregions exist, it is unlikely such a result would challenge the fundamental claim here that at least some portion of the PPA is dedicated to categorization, not navigation. (Indeed, this must be true, given the evidence reviewed here that when treated as a unitary region – i.e., even averaging over subregions that might support other processes, if they exist – the PPA still shows functional responses consistent with scene categorization, not navigation.)

The OPA and RSC are involved in different kinds of navigation

Unlike the PPA, there is clear and consistent evidence for the roles of the OPA and RSC in navigation. Moreover, there is a growing consensus that these two regions play distinct roles even within navigation, with the OPA supporting visually guided navigation and the RSC supporting map-based navigation.

Beginning with the OPA, as previously discussed [29], the OPA responded significantly more during a visually guided navigation task than either a scene categorization task or a baseline task, and similarly during the latter two tasks, strongly suggesting its complete lack of involvement in scene categorization. Furthermore, several other fMRI studies found that the OPA represents at least five kinds of information relevant for visually guided navigation: (i) sense information [30]; (ii) egocentric distance information [31]; (iii) first-person perspective motion information through scenes [32,33]; (iv) local scene elements ('parts'), including boundaries (e.g., walls) and/or individual obstacles (e.g., furniture) that constrain how one can navigate the immediately visible space [54,65,70–72]; and (v) possible routes through a local scene [34]. Additional support for the OPA's involvement in visually guided navigation comes from two fMRI studies showing a lower visual field bias in the OPA – where paths tend to be [73,74]. Finally, causal evidence in support of the OPA's role in visually guided navigation comes from a transcranial magnetic stimulation (TMS) study showing that TMS to the OPA impairs accuracy of navigation to locations in a virtual arena and that this impairment is specific to locations defined by distance to a bounding wall [75].

But is OPA only involved in visually guided navigation and not map-based navigation? The answer appears to be yes. Indeed, several fMRI studies found that the OPA – unlike the RSC – does not represent landmark information (i.e., location and heading direction) [41,43]. Furthermore, the OPA shows little to no memory effects [69] – consistent with its hypothesized role in visually guided navigation, and not map-based navigation, because visually guided navigation, by definition, must operate in the here and now, requiring little to no memory, while map-based navigation requires a 'linking' of the currently visible scene to representations of the broader environment stored in memory.

Next, the RSC does not respond during either a visually guided navigation or scene categorization task [29], strongly suggesting a lack of involvement in either of these processes. Instead, studies of information processing in the RSC point to a role for this region in map-based navigation. Like the OPA, the RSC represents sense and egocentric distance information [30,31] – information critical for any navigation system. However, unlike the OPA, the RSC does not represent information critical for visually guided navigation (e.g., first-person perspective motion information) [32], but rather represents information necessary for map-based navigation more specifically. For example, many fMRI studies have found that the RSC represents landmark information, including both location information and heading direction [41,43,44,61,76–80]. Moreover, several fMRI studies found that the RSC shows strong memory effects – responding two times more strongly to familiar than unfamiliar scenes [18,69,81], and significantly more to personally familiar scenes than famous scenes [82] – as required for a system that must link the currently visible scene to representations of the broader environment stored in memory. Finally, consistent with the hypothesis that map-based navigation involves deliberate processing (i.e., it depends on conscious, intentional control), a recent fMRI study [83] found greater RSC responses during active versus 'passive' navigation through complex (versus simple) virtual mazes. By contrast, the OPA responded strongly to both active and passive navigation, in both complex and simple mazes, suggesting that visually guided navigation instead operates automatically.

A final, tentative piece of evidence for the dissociation between visually guided versus map-based navigation systems comes from the neuropsychological literature on patients with topographical disorientation. For example, Aguirre and D'Esposito [46] 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 [84], and cases of heading disorientation, where participants are unable to derive directional information from the currently visible scene [85]. Egocentric disorientation is related to posterior parietal

damage, whereas heading disorientation is related to retrosplenial and/or posterior cingulate damage, sites that may encompass the OPA and RSC, respectively. However, although the similarity between this neurological dissociation and our proposed division of labor between the OPA and RSC is striking, future experiments are needed to rigorously test such patients in terms of our ‘two-scene navigation systems’ hypothesis and clarify the precise contributions of the OPA and RSC to these deficits.

Differential development across the three cortical scene processing systems?

If the three regions indeed develop along different timelines – as hypothesized here – then this finding would provide strong support for our three scene systems hypothesis. However, no studies to date have tested this hypothesis, for three reasons. First, most developmental fMRI studies only investigated a single scene-selective region (i.e., the PPA), precluding the critical comparisons across regions [86–89]. Second, the few studies that have compared more than one region have failed to find conclusive evidence. For example, one study investigating the OPA, RSC, and PPA found that the RSC is adult-like in scene selectivity (i.e., responding two to four times more to images of scenes than to images of objects and faces) by 7–8 years old, whereas the OPA and PPA are not – consistent with our differential development hypothesis [90]. However, this study did not test for the critical region by age group interaction necessary to support this claim. To that end, a second study investigating the three scene-selective regions in children 5 and 8 years old did test for the critical region by age group interaction and found no differences in scene selectivity in any of the three regions between the two age groups [91]. Thus, although these studies detected scene-selective regions in children as early as age 5 years (Box 1), they did not establish whether these regions are developing at different rates – at least in the age ranges tested – as hypothesized here. Third, and perhaps most important, given the hypothesis that the three scene-selective regions support distinct functions, focusing on the development of scene selectivity alone may be insufficient. Instead, we propose that investigating the information processing necessary for each specific function is needed to understand the developmental trajectory of each of these regions. For example, one recent fMRI study investigated the development of visually guided navigation by asking when in development the OPA represents first-person perspective motion information – information shown to be represented in the OPA in adulthood [32] – in children 5 and 8 years old [91]. The study found that, although the OPA already exhibited scene selectivity by age 5 years, and not different from children age 8 years (as previously discussed), responses to first-person perspective motion were not yet detectable at this

Box 1. The origins of scene-selective cortex

The finding that the PPA, RSC, and OPA are all scene selective by age 5 years, albeit perhaps with subsequent development, raises the question: When does human cortical scene processing first come online? Does it take 5 years (as suggested by the current data), or is it present much earlier in life, within the first few months (or days even) of life? Addressing this question, a recent study [118] found that the overall scene preferences (i.e., significantly greater responses to images of scenes than faces) can already be detected in and around the parahippocampal gyrus (consistent with the location of the PPA) and the lateral occipital cortex (consistent with the location of the OPA) by just 4 to 6 months, but scene selectivity (i.e., greater responses to scenes than objects – what the aforementioned studies find as early as 5 years old) is not yet detectable in these regions at this age. The RSC was not discussed in this study. Moreover, a functional near-infrared spectroscopy (fNIRS) study – using an independently defined region of interest approach, similarly used in fMRI – found that infants between 3 and 12 months of age exhibit scene-preferring regions (i.e., responding significantly more to scenes than faces), like that of adults [119], most likely reflecting the OPA, given that fNIRS is only able to detect signals from lateral regions of the brain (and not the more medial and deep regions, like the PPA and RSC). Finally, another recent resting-state fMRI study found that the functional connectivity underlying the scene-selective cortical system is already intact in as little as 27 days of age – with the PPA showing biased connectivity with the RSC, as well as with the peripheral early visual cortex – suggesting that the connectivity underlying these regions develops prior to, and potentially scaffolds, the development of scene selectivity [120]. Connectivity to the OPA was not discussed. Thus, although the foundations of the scene processing system (in both ‘protofunction’ and connectivity) may be present early in infancy, it is still not clear when the more focal, adult-like PPA, OPA, and RSC emerge.

same age and did not emerge until age 8 years. This protracted development was specific to the OPA, and was not found in either the PPA or the RSC, as expected because prior work in adults found that these regions do not represent first-person perspective motion information [32] – consistent with our three scene systems hypothesis.

The finding that visually guided navigation undergoes protracted development dovetails with a number of behavioral findings. Most directly, a recent study [92] – using a task known to depend on the OPA in adulthood [75] – found that the ability to navigate to locations (defined by distance to a bounding wall) in a virtual arena was still maturing late into childhood, not reaching adult-like status until sometime after 8 years old. Likewise, several other behavioral studies have found that other visually guided navigation abilities, including obstacle avoidance [93,94] and locomotion through a local environment [95,96], also continue to mature well into childhood, not reaching adult-like status until after 7–8 years old. At the same time, however, the idea of a late-developing visually guided navigation system might seem surprising, given that humans begin navigating early in life (e.g., crawling at approximately 9 months) and show remarkably sophisticated navigational ability within the first few years. For example, young children can use boundaries to recover their orientation after becoming disoriented [97], and infants understand whether it is safe to locomote over a ‘visual cliff’ [98]. How, then, can we reconcile these observations with the hypothesis that visually guided navigation undergoes protracted development? One possibility is that these tasks rely on qualitatively different systems. For example, the reorientation task may involve spatial memory systems in the hippocampus, whereas the visual cliff task may depend on basic depth perception. A second possibility is that these tasks *do* rely on the same visually guided navigation system and that, despite development extending well into childhood (as discussed in the preceding text), the foundations of this system are nevertheless intact early and sufficient to support these early navigational behaviors.

Whatever the case – whether early emerging representations present in the OPA are sufficient to support early visually guided navigational abilities and followed by subsequent development, or whether such early navigational abilities depend on qualitatively different systems – all of the aforementioned findings suggest that visually guided navigation is indeed late developing. But what about the development of the scene categorization and map-based navigation systems? Here we suggest that these systems develop along different timelines and offer a specific developmental hypothesis, described next.

Visually guided and map-based navigation develop later than scene categorization?

Reflecting the basic fact that typically developing infants view scenes from the day they are born but only begin to actively navigate through them much later in life, we hypothesize that the visually guided navigation system (including the OPA) is slower to develop than the scene categorization system (including the PPA). Although no pediatric fMRI or behavioral studies to date have directly tested this hypothesis (given the reasons outlined in the previous section), perhaps the work investigating the development of human object processing can provide some clues. For example, building on the two systems for adult human object processing, as described previously, several behavioral studies have found that the ‘vision-for-action’ (dorsal) system is slower to develop than the ‘vision-for-perception’ (ventral) system [99–106], raising the tantalizing possibility that separable systems for recognition and action in scene processing, like object processing, also follow the same developmental trajectories. As such, we hypothesize that the visually guided navigation system (including the OPA, in the dorsal stream) is slower to develop than the scene categorization system (including the PPA, in the ventral stream). Indeed, behavioral work in infants shows that before infants ever independently navigate their surroundings (e.g., by crawling), they can

discriminate the deep versus shallow side of a visual cliff [107]. By contrast, when older infants are encouraged to actually navigate the same environment (e.g., by crawling or even walking), they fail to do so accurately, navigating right over the edge of the cliff, for example [108].

More speculative still, we further hypothesize that the map-based navigation system (including the RSC) develops later than the scene categorization system. Again, no pediatric fMRI or behavioral studies to date have tested this hypothesis (i.e., by directly comparing map-based navigation and scene categorization abilities across childhood). However, partial support for this possibility comes from a large literature showing that map-based navigation ability – especially that relying on allocentric representations of the broader spatial environment – undergoes protracted development late into childhood [92,109–117].

Concluding remarks and future perspectives

Here we provide a fundamental shift in our current understanding of adult human cortical scene processing and highlight the need for research on its development by offering two novel hypotheses for future investigation. First, we propose that human cortical scene processing is composed of three distinct systems: (i) scene categorization (including the PPA), (ii) visually guided navigation (including the OPA), and (iii) map-based navigation (including the RSC) (but see [Outstanding questions](#)). Second, we propose that these systems develop along different timelines and, more specifically, that the visually guided navigation system and the map-based navigation system are slower to develop than the scene categorization system.

Acknowledgments

We thank the Facility for Education and Research in Neuroscience (FERN) Imaging Center in the Department of Psychology, Emory University, Atlanta, GA. We also thank Rebecca Saxe, Nancy Kanwisher, Michael McCloskey, Annie Cheng, and Christopher Jones for helpful comments on the manuscript. This work was supported by grants from the National Eye Institute (R01 EY29724 to D.D.D. and T32 EY007092 to F.S.K.), the National Science Foundation (DGE-1444932 to A.S.P.), the National Institute of Mental Health Intramural Research Program (ZIA MH 002920 09 to A.S.P.), and a grant from the Simons Foundation to the Simons Center for the Social Brain at the Massachusetts Institute of Technology (F.S.K.).

Declaration of interests

The authors have no interests to declare.

References

- Epstein, R. and Kanwisher, N. (1998) A cortical representation of the local visual environment. *Nature* 392, 598–601
- Dilks, D.D. *et al.* (2013) The occipital place area is causally and selectively involved in scene perception. *J. Neurosci.* 33, 1331–1336a
- Maguire, E.A. (2001) The retrosplenial contribution to human navigation: a review of lesion and neuroimaging findings. *Scand. J. Psychol.* 42, 225–238
- Epstein, R.A. and Baker, C.I. (2019) Scene perception in the human brain. *Annu. Rev. Vis. Sci.* 5, 373–397
- Epstein, R.A. *et al.* (2017) The cognitive map in humans: spatial navigation and beyond. *Nat. Neurosci.* 20, 1504–1513
- Park, S. *et al.* (2011) Disentangling scene content from spatial boundary: Complementary roles for the parahippocampal place area and lateral occipital complex in representing real-world scenes. *J. Neurosci.* 31, 1333–1340
- Kravitz, D.J. *et al.* (2011) Real-world scene representations in high-level visual cortex: It's the spaces more than the places. *J. Neurosci.* 31, 7322–7333
- Janzen, G. and van Turenout, M. (2004) Selective neural representation of objects relevant for navigation. *Nat. Neurosci.* 7, 673–677
- Janzen, G. and Jansen, C. (2010) A neural wayfinding mechanism adjusts for ambiguous landmark information. *Neuroimage* 52, 364–370
- Janzen, G. *et al.* (2008) Memory consolidation of landmarks in good navigators. *Hippocampus* 18, 40–47
- Janzen, G. and Weststeijn, C.G. (2007) Neural representation of object location and route direction: An event-related fMRI study. *Brain Res.* 1165, 116–125
- Wegman, J. and Janzen, G. (2011) Neural encoding of objects relevant for navigation and resting state correlations with navigational ability. *J. Cogn. Neurosci.* 23, 3841–3854
- Schinazi, V.R. and Epstein, R.A. (2010) Neural correlates of real-world route learning. *Neuroimage* 53, 725–735
- Sun, L. *et al.* (2021) The parahippocampal place area and hippocampus encode the spatial significance of landmark objects. *Neuroimage* 236, 118081
- Marchette, S.A. *et al.* (2015) Outside looking in: Landmark generalization in the human navigational system. *J. Neurosci.* 35, 14896–14908
- Epstein, R.A. and Higgins, J.S. (2007) Differential parahippocampal and retrosplenial involvement in three types of visual scene recognition. *Cereb. Cortex* 17, 1680–1693
- Harel, A. *et al.* (2013) Deconstructing visual scenes in cortex: Gradients of object and spatial layout information. *Cereb. Cortex* 23, 947–957
- Epstein, R.A. *et al.* (2007) Visual scene processing in familiar and unfamiliar environments. *J. Neurophysiol.* 97, 3670–3683

Outstanding questions

How exactly are scene categories represented in the PPA? For example, like objects, are scenes first categorized at the basic level (e.g., 'kitchen' versus 'room' or 'galley kitchen') in the PPA?

Do nonhuman primates (or even other animals) also possess three systems for scene processing? Although three visual scene-selective regions have been identified in adult macaques, with each region in a similar anatomical location to the human PPA, OPA, and RSC, it is not clear whether there is a homologous functional organization for scene-selective regions in the cortex of human and nonhuman primates.

Can the three-scene-systems hypothesis guide development of better computational models? Training models to perform specific scene processing tasks (i.e., scene categorization, visually guided navigation, or map-based navigation) may yield improved fits between models and the brain and therefore may be an avenue for gaining insight into the computational mechanisms underlying each system, possibly beyond that obtainable from fMRI alone.

19. Ghaem, O. *et al.* (1997) Mental navigation along memorized routes activates the hippocampus, precuneus, and insula. *Neuroreport* 8, 739–744
20. Rosenbaum, R.S. *et al.* (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
21. Cheng, K. and Newcombe, N.S. (2005) Is there a geometric module for spatial orientation? Squaring theory and evidence. *Psychon. Bull. Rev.* 12, 1–23
22. Rauchs, G. *et al.* (2008) Partially segregated neural networks for spatial and contextual memory in virtual navigation. *Hippocampus* 18, 503–518
23. Spelke, E. *et al.* (2010) Beyond core knowledge: Natural geometry. *Cogn. Sci.* 34, 863–884
24. Oliva, A. and Schyns, P.G. (1997) Coarse blobs or fine edges? Evidence that information diagnosticity changes the perception of complex visual stimuli. *Cogn. Psychol.* 34, 72–107
25. Oliva, A. and Torralba, A. (2001) Modeling the shape of a scene: a holistic representation of the spatial envelope. *Int. J. Comput. Vis.* 42, 145–175
26. Greene, M.R. and Oliva, A. (2009) Recognition of natural scenes from global properties: Seeing the forest without representing the trees. *Cogn. Psychol.* 58, 137–176
27. Walther, D.B. and Shen, D. (2014) Nonaccidental properties underlie human categorization of complex natural scenes. *Psychol. Sci.* 25, 851–860
28. Greene, M.R. and Oliva, A. (2010) High-level aftereffects to global scene properties. *J. Exp. Psychol. Hum. Percept. Perform.* 36, 1430–1442
29. Persichetti, A.S. and Dilks, D.D. (2018) Dissociable neural systems for recognizing places and navigating through them. *J. Neurosci.* 38, 10295–10304
30. Dilks, D.D. *et al.* (2011) Mirror-image sensitivity and invariance in object and scene processing pathways. *J. Neurosci.* 31, 11305–11312
31. Persichetti, A.S. and Dilks, D.D. (2016) Perceived egocentric distance sensitivity and invariance across scene-selective cortex. *Cortex* 77, 155–163
32. Kamps, F.S. *et al.* (2016) The occipital place area represents first-person perspective motion information through scenes. *Cortex* 83, 17–26
33. Pitcher, D. *et al.* (2019) A functional dissociation of face-, body- and scene-selective brain areas based on their response to moving and static stimuli. *Sci. Rep.* 9, 8242
34. Bonner, M.F. and Epstein, R.A. (2017) Coding of navigational affordances in the human visual system. *Proc. Natl. Acad. Sci. U. S. A.* 114, 4793–4798
35. Goodale, M.A. and Milner, A.D. (1992) Separate visual pathways for perception and action. *Trends Neurosci.* 15, 20–25
36. Kravitz, D.J. *et al.* (2011) A new neural framework for visuospatial processing. *Nat. Rev. Neurosci.* 12, 217–230
37. Milner, A.D. (2017) How do the two visual streams interact with each other? *Exp. Brain Res.* 235, 1297–1308
38. Baldassano, C. *et al.* (2013) Differential connectivity within the parahippocampal place area. *Neuroimage* 75, 228–237
39. Nasr, S. *et al.* (2013) Spatial encoding and underlying circuitry in scene-selective cortex. *Neuroimage* 83, 892–900
40. Vul, E. and Kanwisher, N. (2010) Begging the question: The non-independence error in fMRI data analysis. In *Foundational Issues in Human Brain Mapping* (Hanson, S.J. and Bunzl, M., eds), pp. 71–91, MIT Press
41. Persichetti, A.S. and Dilks, D.D. (2019) Distinct representations of spatial and categorical relationships across human scene-selective cortex. *Proc. Natl. Acad. Sci. U. S. A.* 116, 21312–21317
42. Vass, L.K. and Epstein, R.A. (2017) Common neural representations for visually guided reorientation and spatial imagery. *Cereb. Cortex* 27, 1457–1471
43. Marchette, S.A. *et al.* (2014) Anchoring the neural compass: Coding of local spatial reference frames in human medial parietal lobe. *Nat. Neurosci.* 17, 1598–1606
44. Auger, S.D. *et al.* (2012) Retrosplenial cortex codes for permanent landmarks. *PLoS One* 7, e43620
45. Auger, S.D. and Maguire, E.A. (2013) Assessing the mechanism of response in the retrosplenial cortex of good and poor navigators. *Cortex* 49, 2904–2913
46. Aguirre, G.K. and D'Esposito, M. (1999) Topographical disorientation: A synthesis and taxonomy. *Brain* 122, 1613–1628
47. Whiteley, A.M. and Warrington, E.K. (1978) Selective impairment of topographical memory – single case-study. *J. Neurol. Neurosurg. Psychiatry* 41, 575–578
48. Hecaen, H. *et al.* (1980) Loss of topographic memory with learning-deficits. *Cortex* 16, 525–542
49. Pallis, C.A. (1955) Impaired identification of faces and places with agnosia for colours – report of a case due to cerebral embolism. *J. Neurol. Neurosurg. Psychiatry* 18, 218–224
50. Derenzi, E. *et al.* (1977) Topographical amnesia. *J. Neurol. Neurosurg. Psychiatry* 40, 498–505
51. Epstein, R. *et al.* (2001) Neuropsychological evidence for a topographical learning mechanism in parahippocampal cortex. *Cogn. Neuropsychol.* 18, 481–508
52. Walther, D.B. *et al.* (2011) Simple line drawings suffice for functional MRI decoding of natural scene categories. *Proc. Natl. Acad. Sci. U. S. A.* 108, 9661–9666
53. Walther, D.B. *et al.* (2009) Natural scene categories revealed in distributed patterns of activity in the human brain. *J. Neurosci.* 29, 10573–10581
54. Kamps, F.S. *et al.* (2016) The occipital place area represents the local elements of scenes. *Neuroimage* 132, 417–424
55. MacEvoy, S.P. and Epstein, R.A. (2009) Decoding the representation of multiple simultaneous objects in human occipitotemporal cortex. *Curr. Biol.* 19, 943–947
56. Bonner, M.F. and Epstein, R.A. (2021) Object representations in the human brain reflect the co-occurrence statistics of vision and language. *Nat. Commun.* 12, 4081
57. Bar, M. and Aminoff, E. (2003) Cortical analysis of visual context. *Neuron* 38, 347–358
58. Bar, M. *et al.* (2008) Scenes unseen: The parahippocampal cortex intrinsically subserves contextual associations, not scenes or places per se. *J. Neurosci.* 28, 8539–8544
59. Mullally, S.L. and Maguire, E.A. (2011) A new role for the parahippocampal cortex in representing space. *J. Neurosci.* 31, 7441–7449
60. Konkle, T. and Oliva, A. (2012) A real-world size organization of object responses in occipitotemporal cortex. *Neuron* 74, 1114–1124
61. Troiani, V. *et al.* (2014) Multiple object properties drive scene-selective regions. *Cereb. Cortex* 24, 883–897
62. Cant, J.S. and Goodale, M.A. (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
63. Cant, J.S. and Xu, Y. (2012) Object ensemble processing in human anterior-medial ventral visual cortex. *J. Neurosci.* 32, 7685–7700
64. Park, J. and Park, S. (2017) Conjoint representation of texture ensemble and location in the parahippocampal place area. *J. Neurophysiol.* 117, 1595–1607
65. Henriksson, L. *et al.* (2019) Rapid invariant encoding of scene layout in human OPA. *Neuron* 103, 161–171 e3
66. Aminoff, E. *et al.* (2007) The parahippocampal cortex mediates spatial and nonspatial associations. *Cereb. Cortex* 17, 1493–1503
67. Stevens, W.D. *et al.* (2012) Hemispheric asymmetry of visual scene processing in the human brain: evidence from repetition priming and intrinsic activity. *Cereb. Cortex* 22, 1935–1949
68. Park, S. *et al.* (2015) Parametric coding of the size and clutter of natural scenes in the human brain. *Cereb. Cortex* 25, 1792–1805
69. Steel, A. *et al.* (2021) A network linking scene perception and spatial memory systems in posterior cerebral cortex. *Nat. Commun.* 12, 2632
70. Dillon, M.R. *et al.* (2018) Places in the brain: Bridging layout and object geometry in scene-selective cortex. *Cereb. Cortex* 28, 2365–2374
71. Park, J. and Park, S. (2020) Coding of navigational distance and functional constraint of boundaries in the human scene-selective cortex. *J. Neurosci.* 40, 3621–3630

72. Cheng, A. *et al.* (2021) Concavity as a diagnostic feature of visual scenes. *Neuroimage* 232, 117920
73. Silson, E.H. *et al.* (2015) A retinotopic basis for the division of high-level scene processing between lateral and ventral human occipitotemporal cortex. *J. Neurosci.* 35, 11921–11935
74. Silson, E.H. *et al.* (2016) Scene-selectivity and retinotopy in medial parietal cortex. *Front. Hum. Neurosci.* 10, 412
75. Julian, J.B. *et al.* (2016) The occipital place area is causally involved in representing environmental boundaries during navigation. *Curr. Biol.* 26, 1104–1109
76. Auger, S.D. *et al.* (2015) A central role for the retrosplenial cortex in de novo environmental learning. *eLife* 4, e09031
77. Wolbers, T. and Buchel, C. (2005) Dissociable retrosplenial and hippocampal contributions to successful formation of survey representations. *J. Neurosci.* 25, 3333–3340
78. Sherrill, K.R. *et al.* (2013) Hippocampus and retrosplenial cortex combine path integration signals for successful navigation. *J. Neurosci.* 33, 19304–19313
79. Morgan, L.K. *et al.* (2011) Distances between real-world locations are represented in the human hippocampus. *J. Neurosci.* 31, 1238–1245
80. Iaria, G. *et al.* (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
81. Sugiyama, M. *et al.* (2005) Cortical representations of personally familiar objects and places: Functional organization of the human posterior cingulate cortex. *J. Cogn. Neurosci.* 17, 183–198
82. Silson, E.H. *et al.* (2019) Distinct subdivisions of human medial parietal cortex support recollection of people and places. *eLife* 8, e47391
83. Suzuki, S. *et al.* (2021) Two scene navigation systems dissociated by deliberate versus automatic processing. *Cortex* 140, 199–209
84. Stark, M. *et al.* (1996) Impairment of an egocentric map of locations: Implications for perception and action. *Cogn. Neuropsychol.* 13, 481–523
85. Takahashi, N. *et al.* (1997) Pure topographic disorientation due to right retrosplenial lesion. *Neurology* 49, 464–469
86. Golarai, G. *et al.* (2007) Differential development of high-level visual cortex correlates with category-specific recognition memory. *Nat. Neurosci.* 10, 512–522
87. Scherf, K.S. *et al.* (2007) Visual category-selectivity for faces, places and objects emerges along different developmental trajectories. *Dev. Sci.* 10, F15–F30
88. Scherf, K.S. *et al.* (2011) 'What' precedes 'which': Developmental neural tuning in face- and place-related cortex. *Cereb. Cortex* 21, 1963–1980
89. Golarai, G. *et al.* (2010) Differential development of the ventral visual cortex extends through adolescence. *Front. Hum. Neurosci.* 3, 80
90. Meissner, T.W. *et al.* (2019) Prolonged functional development of the parahippocampal place area and occipital place area. *Neuroimage* 191, 104–115
91. Kamps, F.S. *et al.* (2020) Late development of navigationally relevant motion processing in the occipital place area. *Curr. Biol.* 30, 544–550 e3
92. Julian, J.B. *et al.* (2019) Dissociable spatial memory systems revealed by typical and atypical human development. *Dev. Sci.* 22, e12737
93. Berard, J.R. and Vallis, L.A. (2006) Characteristics of single and double obstacle avoidance strategies: A comparison between adults and children. *Exp. Brain Res.* 175, 21–31
94. Pryde, K.M. *et al.* (1997) Age-related trends in locomotor ability and obstacle avoidance. *Hum. Mov. Sci.* 16, 507–516
95. Franchak, J.M. and Adolph, K.E. (2010) Visually guided navigation: Head-mounted eye-tracking of natural locomotion in children and adults. *Vis. Res.* 50, 2766–2774
96. Franchak, J.M. *et al.* (2011) Head-mounted eye tracking: A new method to describe infant looking. *Child Dev.* 82, 1738–1750
97. Hermer, L. and Spelke, E.S. (1994) A geometric process for spatial reorientation in young children. *Nature* 370, 57–59
98. Gibson, E.J. and Walk, R.D. (1960) The 'visual cliff'. *Sci. Am.* 202, 64–71
99. Diamond, A. and Goldman-Rakic, P.S. (1989) Comparison of human infants and rhesus monkeys on Piaget's AB task: Evidence for dependence on dorsolateral prefrontal cortex. *Exp. Brain Res.* 74, 24–40
100. Diamond, A. *et al.* (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. *Behav. Neurosci.* 103, 526–537
101. Bertenthal, B.I. (1996) Origins and early development of perception, action, and representation. *Annu. Rev. Psychol.* 47, 431–459
102. Gilmore, R.O. and Johnson, M.H. (1997) Egocentric action in early infancy: Spatial frames of reference for saccades. *Psychol. Sci.* 8, 224–230
103. Gilmore, R.O. and Johnson, M.H. (1997) Body-centered representations for visually-guided action emerge during early infancy. *Cognition* 65, B1–B9
104. Csibra, G. *et al.* (1998) Neural correlates of saccade planning in infants: A high-density ERP study. *Int. J. Psychophysiol.* 29, 201–215
105. Atkinson, J. *et al.* (2003) Neurobiological models of visuospatial cognition in children with Williams syndrome: Measures of dorsal-stream and frontal function. *Dev. Neuropsychol.* 23, 139–172
106. Dilks, D.D. *et al.* (2008) Vision for perception and vision for action: Normal and unusual development. *Dev. Sci.* 11, 474–486
107. Campos, J.J. *et al.* (1970) Cardiac responses on the visual cliff in prelocomotor human infants. *Science* 170, 196–197
108. Adolph, K. (2000) Specificity in learning: Why infants fall over a veritable cliff. *Psychol. Sci.* 11, 290–295
109. Acredolo, L.P. (1977) Developmental changes in ability to coordinate perspectives of a large-scale space. *Dev. Psychol.* 13, 1–8
110. Overman, W.H. *et al.* (1996) Ontogeny of place learning in children as measured in the radial arm maze, Morris search task, and open field task. *Behav. Neurosci.* 110, 1205–1228
111. Lehnung, M. *et al.* (1998) Spatial memory and orientation in healthy and brain-injured children. *Eur. J. Neurosci.* 10, 143
112. Lepow, B. *et al.* (2003) Navigational place learning in children and young adults as assessed with a standardized locomotor search task. *Br. J. Psychol.* 94, 299–317
113. Akers, K.G. and Hamilton, D.A. (2007) Comparison of developmental trajectories for place and cued navigation in the Morris water task. *Dev. Psychobiol.* 49, 553–564
114. Nazareth, A. *et al.* (2018) Charting the development of cognitive mapping. *J. Exp. Child Psychol.* 170, 86–106
115. Newcombe, N.S. (2019) Navigation and the developing brain. *J. Exp. Biol.* 222, jeb186460
116. Allen, G.L. *et al.* (1979) Developmental issues in cognitive mapping: The selection and utilization of environmental landmarks. *Child Dev.* 50, 1062–1070
117. Cousins, J.H. *et al.* (1983) Way finding and cognitive mapping in large-scale environments: A test of a developmental model. *J. Exp. Child Psychol.* 35, 1–20
118. Deen, B. *et al.* (2017) Organization of high-level visual cortex in human infants. *Nat. Commun.* 8, 13995
119. Powell, L.J. *et al.* (2018) Using individual functional channels of interest to study cortical development with fNIRS. *Dev. Sci.* 21, e12595
120. Kamps, F.S. *et al.* (2020) Connectivity at the origins of domain specificity in the cortical face and place networks. *Proc. Natl. Acad. Sci. U. S. A.* 117, 6163–6169