Late development of navigationally-relevant motion processing in the occipital place area

Frederik S. Kamps, **Jordan E. Pincus**, **Samaher F. Radwan**, **Stephanie Wahab**, **Daniel D. Dilks**[*]

Department of Psychology, Emory University, 36 Eagle Row, Atlanta, GA, 30322, USA

Summary

Human adults flawlessly and effortlessly navigate boundaries and obstacles in the immediately visible environment, a process we refer to as "visually-guided navigation". Neuroimaging work in adults suggests this ability involves the occipital place area (OPA) [1, 2] – a scene selective region in the dorsal stream that selectively represents information necessary for visually-guided navigation [3-9]. Despite progress in understanding the neural basis of visually-guide navigation, however, little is known about how this system develops. Is navigationally-relevant information processing present in the first few years of life? Or does this information processing only develop after many years of experience? Although a handful of studies have found selective responses to scenes (relative to objects) in OPA in childhood [10-13], no study has explored how more specific navigationally-relevant information processing emerges in this region. Here we do just that by measuring OPA responses to first-person perspective motion information – a proxy for the visual experience of actually navigating the immediate environment – using fMRI in 5- and 8-year-old children. We found that although OPA already responded more to scenes than objects by age 5, responses to first-person perspective motion were not yet detectable at this same age, and rather only emerged by age 8. 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). These findings therefore suggest that navigationally-relevant information processing in OPA undergoes prolonged development across childhood.

[*]Corresponding author and lead contact: Daniel D. Dilks, Department of Psychology, Emory University, 36 Eagle Row, Atlanta, GA 30322, 404-727-2980, dilks@emory.edu.

Declaration of Interests

The authors declare no competing interests.

Keywords

Navigation; Scene perception; Development; fMRI; Occipital Place Area (OPA); Transverse Occipital Sulcus (TOS); Parahippocampal Place Area (PPA); Retrosplenial Complex (RSC); High-level Vision

Results

To study how navigationally-relevant information processing develops in OPA, we measured responses in OPA across childhood to first-person perspective motion – a proxy for the visual experience of actually navigating through an immediately visible scene. Given previous findings that boundary-based navigation and other dorsal stream processes are still developing in childhood [14-16], we scanned two groups of children: 5-year-olds and 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 Faces (Figure 1). If navigationally-relevant information processing develops early in OPA, 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 navigationally-relevant information processing develops late in OPA, then stronger selectivity for first-person perspective motion information will be observed in 8-year-olds than 5-year-olds. Importantly, fMRI data quality was closely matched between the two groups (Figure 2; STAR Methods), ensuring that data quality could not explain any observed developmental differences.

OPA is scene selective by 5 years old

Before testing first-person perspective motion processing in OPA, we asked whether scene selectivity could be detected in OPA at age 5 by comparing responses in OPA to Static Scenes and Static Objects – following the standard contrast used to define OPA in adults – as well as to Static Faces (responses to dynamic stimuli are analyzed extensively below). A 2 (group: 5-year-olds, 8-year-olds) × 3 (condition: Static Scenes, Static Objects, Static Faces) mixed-model ANOVA revealed a main effect of condition ($F_{(2,60)} = 21.78$, $p < 0.001$, $\eta_p^2 = 0.42$), with stronger responses to scenes than both objects (pairwise comparison, $p < 0.001$) and faces ($p < 0.001$), but no significant group × condition interaction ($F_{(2,60)} = 0.61$, $p = 0.54$, $\eta_p^2 = 0.02$) (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 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) × 3 (condition: Static Scenes, Static Objects, Static Faces) mixed-model ANOVA revealed a main effect of condition (PPA: $F_{(2,60)} = 149.86$, $p < 0.001$, $\eta_p^2 = 0.83$;

RSC: $F_{(2,60)} = 41.73$, $p < 0.001$, $\eta_p^2 = 0.58$), 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 × condition interaction (PPA: $F_{(2,60)} = 1.09$, $p = 0.34$, $\eta_p^2 = 0.04$; RSC: $F_{(2,60)} = 0.61$, $p = 0.54$, $\eta_p^2 = 0.02$) (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.

To confirm that the findings above were driven by scene selectivity, rather than stimulus complexity or low-level visual information, we also investigated responses in V1. Unlike the three scene-selective regions, V1 responded strongly to all three conditions, and a 2 (group: 5-year-olds, 8-year-olds) × 3 (condition: Static Scenes, Static Objects, Static Faces) mixed-model ANOVA did not reveal a main effect of condition ($F_{(2,60)} = 1.22$, $p = 0.30$, $\eta_p^2 = 0.04$), nor a significant group × condition interaction ($F_{(2,60)} = 1.49$, $p = 0.23$, $\eta_p^2 = 0.05$) (Figure 3). Further, comparing V1 to each scene-selective region directly, a 2 (region: V1, each scene region) × 3 (condition: Static Scenes, Static Objects, Static Faces) repeated-measures ANOVA revealed a significant region × condition interaction for all three regions (all F's > 18.82, all p's < 0.001, all $\eta_p^2 > 0.38$). Thus, responses in OPA, PPA, and RSC reflect scene selectivity, not stimulus complexity or low-level visual information.

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

Having established that OPA is scene selective by age 5, 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 each participant's response in OPA to Static Scenes from that to Dynamic Scenes. 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 OPA response to scene motion (i.e., scene motion difference score > 0) for the 8-year-olds ($t_{(15)} = 3.50$, $p = 0.003$, $d = 0.88$), but not the 5-year-olds ($t_{(15)} = 0.65$, $p = 0.52$, $d = 0.16$) (Figure 4). Further analysis investigating scene motion responses across the entire volume of OPA (voxel by voxel) failed to find any evidence of scene motion responses in 5 year olds, even in the peak scene-selective voxels (Supplemental Experimental Procedures; Figure S1). These findings suggest that first-person perspective motion processing in OPA develops slowly across childhood, and only emerges by age 8.

We next asked whether the developmental increase in first-person perspective motion processing in OPA reflects increasing sensitivity to any kind of visual motion information, or motion information in scenes, in particular, by comparing OPA responses to scene motion with those to object and face motion (i.e., calculated for objects and faces separately by subtracting responses to Static conditions from those to Dynamic conditions). A 2 (group: 5-year-olds, 8-year-olds) × 3 (motion domain: scene motion, object motion, face motion) mixed-model ANOVA revealed a significant group × 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 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 other 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) × 3 (motion domain: scene motion, object motion, face motion) mixed-model ANOVA did not reveal a significant group × 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 $p's < 0.003$) (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 years old. Testing this claim directly, a 2 (region: OPA, MT) × 2 (group: 5-year-olds, 8-year-olds) × 3 (motion domain: scene motion, object motion, face motion) mixed-model ANOVA revealed a significant region × age × motion domain interaction ($F_{(2,60)} = 4.52$, $p = 0.01$, $\eta_p^2 = 0.13$) (full ANOVA results are listed in Table S1). 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.

We also asked whether differences in first-person perspective motion processing could be explained by 5-year-olds simply paying less attention to the Dynamic Scenes than 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 responses in OPA with those in PPA and RSC. For both PPA and RSC, a 2 (group: 5-year-olds, 8-year-olds) × 3 (motion domain: scene motion, object motion, face motion) mixed-model ANOVA failed to reveal a significant group × motion domain interaction (PPA: $F_{(2,60)} = 0.34$, $p = 0.71$, $\eta_p^2 = 0.01$; RSC: $F_{(2,60)} = 0.28$, $p = 0.76$, $\eta_p^2 = 0.01$) (Figure 4), suggesting 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 [5]). Testing this claim directly, a 3 (region: OPA, PPA, RSC) × 2 (group: 5-year-olds, 8-year-olds) × 3 (motion domain: scene motion, object motion, face motion) mixed-model ANOVA revealed a significant region × group × motion domain interaction ($F_{(4,120)} = 2.82$, $p = 0.03$, $\eta_p^2 = 0.09$) (full ANOVA results are listed in Table S2). This three-way interaction provides strong evidence against the possibility group differences in attention drove the results. Moreover, this evidence of differential development within the cortical scene processing system is consistent with proposals that these regions support dissociable aspects of scene navigation and categorization [3, 5, 6, 17-19].

Finally, we explored whether developmental effects in OPA occur via construction of preferred responses or pruning of non-preferred responses [20-22]. Two analyses revealed that the development of scene motion processing in OPA depends on increasing responses to Dynamic Scenes (consistent with construction of preferred responses), rather than on decreasing responses to static scenes (as predicted by a pruning account) (Figure S2).

Discussion

Taken together, we found that although OPA already responds selectively to scenes (relative to objects and faces) by age 5, OPA does not respond to first-person perspective motion at this same age, and only responds 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 support the hypothesis that navigationally-relevant information processing in OPA undergoes protracted development, dovetailing with a number of behavioral findings. Most directly, spatial memory for locations defined relative to environmental boundaries – an ability that depends on OPA in adulthood [9] – is still developing in children ages 6-10 [14]. The present finding that navigationally-relevant information processing is still emerging in OPA across childhood therefore suggests that this late development of boundary-based spatial memory may be mediated in part by late development of scene processing in OPA. The hypothesis of late-developing navigationally-relevant information processing is also consistent with evidence that obstacle avoidance [23, 24], locomotion (especially that relying on peripheral vision) [25, 26], and reorientation ability [27, 28] continue to mature across childhood (although the role of OPA in these tasks has not been directly established).

At the same time, however, the idea that navigationally-relevant information processing develops late in OPA might seem incompatible with evidence that sophisticated navigational ability emerges within the first few years of life [29, 30]. For example, children use the geometry of local boundaries to recall previously learned locations by 18-24 months [31, 32], and understand whether it is safe to navigate a "visual cliff" by 6-14 months [33]. Although the role of OPA in these tasks is unknown, there are two possibilities. First, these tasks may not depend on OPA, and other neural systems may instead mediate early-developing navigational abilities. For example, reorientation tasks may assess the ability to recall *orientation* relative to boundaries – as supported by extra-hippocampal structures [34] – rather than *location* relative to boundaries; likewise, visual cliff tasks may assess general depth perception, rather than visually-guided navigation. Second, it is possible that these tasks *do,* in fact, depend on OPA, but can be solved using earlier-emerging representations in this region. Indeed, OPA shows a preference for scenes relative to faces by 4-6 months [10, 13], with scene selectivity emerging by 5 years or earlier, as revealed here. It is possible then that these early emerging scene responses are sufficient to represent the approximate location of boundaries, in order to determine, for example, which of two walls is closer (in the reorientation task) or which of two cliffs is steeper (in 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 one to dynamically extract and update possible navigational paths, or represent boundary and obstacle locations with increasing precision). Intriguingly, an analogous developmental progression is found in the developing rodent medial temporal lobe, where the rudiments of head direction, grid, and place cells are detectable as soon as rat pups make their first movements outside of the nest,

yet the stability and precision of the spatial coding in these cells continues to be refined well into juvenility [35, 36].

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, the earliest demonstration of scene selectivity in this system to date ([11, 37-40], but see [12, 41]). This finding suggests that all three regions are earmarked to support scene processing within the first few years of life.

In conclusion, we found that OPA is already scene selective by age 5, yet does not develop selective responses to first-person perspective motion until age 8. These findings raise important new questions. For example, what kinds of scene representations does OPA encode early in life? Which late-emerging navigational behaviors [42, 43] depend on and drive OPA development? And what factors (e.g., inputs, outputs, or internal computations) ensure that OPA always represents visual information relevant to navigation, and not other functions, even so late into childhood? Whatever the answers to these questions, our findings provide novel evidence of protracted development in OPA across childhood.

STAR Methods

LEAD CONTACT AND MATERIALS AVAILABILITY

Further information requests and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Daniel Dilks (dilks@emory.edu). This study did not generate new unique reagents.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Participants—Sixteen 5-year-olds (mean age: 63 months; range: 60-72 months; 11 females) and 16 8-year-olds (mean age: 101 months; range: 93-112 months; 10 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). The mean age of excluded participants (65 months) did not significantly differ from that of included participants ($t_{(26)} = 0.87$, $p = 0.39$); however, a larger proportion of the excluded 5 year old participants were male (6/12 excluded participants, compared with 5/16 included participants). 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.

METHOD DETAILS

Design and Stimuli—We used a region of interest (ROI) approach in which we used one set of runs to localize scene-selective, motion-selective, and primary visual cortex (V1), 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 [44, 45], as detailed in the Data Analysis section. In addition to the standard ROI analyses, we conducted a

"volume-selectivity function" (VSF) analysis [46-48], 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). These stimuli consisted of 16 3-sec video clips depicting first-person perspective motion (as would be experienced during locomotion through an immediately visible environment), and subtended approximately 21×16 degrees of visual angle. 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 90.3×58.9 degrees) held at eye level. As such, the Dynamic Scenes contained rich, naturalistic visual information relevant to guiding locomotion past boundaries and obstacles in a current scene. Most videos depicted only a straight walking path, while some clips included turning events (e.g., walking around a corner), and therefore primarily depicted radial optic flow motion. We were confident that the Dynamic Scenes were sufficient to allow perception of self motion (as distinct from object or face motion) in 5 and 8 year old children for four reasons. First, previous work in infants and newborns finds that optic flow stimuli (but not other kinds of moving stimuli) elicit locomotive responses [49] and postural changes (including in the absence of vestibular or somatosensory cues) [50-53], and that infants' sensitivity to this information is relatively sophisticated, such that this information can be used to discriminate heading direction [54], or be distinguished from object motion [55]. Second, these self motion responses occur rapidly; even neonates show postural responses to visual optic flow (again, in the absence of concurrent vestibular or somatosensory information) within 1.90 (+- 1.75) seconds [50], well within the duration of a single 3 s video. Third, the 3s clips were presented in the context of 15 second blocks (i.e., 5 consecutive 3s clips of scene motion). It is likely then that even if 5 year olds take longer than 8 year olds to perceive the clips as specifying self-motion (i.e., on the first video), they should nevertheless be primed to perceive self-motion in the remaining videos for that block. Fourth, our findings revealed that area MT, which responds strongly to self-motion in adults [56-58], responded equally strongly to scene motion in 5 and 8 year olds, providing indirect support for the idea that 5 and 8 year olds did not differ in their ability to perceive self-motion. Next, 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 21×16 degrees of visual angle.

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, also presented at 21×16 degrees of visual angle. 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 motion of the toys involved a mixture of translations (laterally, vertically, and in depth) and rotations (in depth and in the vertical plane), sometimes occurring in a periodic fashion (e.g., for swinging or spinning objects). The Dynamic Face stimuli depicted only the faces of 4 children against a black background as they smiled, laughed, and looked around. The dynamic face information therefore involved eye, lip, and mouth movements; emotional expressions; and small head movements (including roll, pitch, and yaw rotations). 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 × 1.5 × 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).

QUANTIFICATION AND STATISTICAL 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) [60] 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 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). As a result of these procedures, the final groups of 5- and 8-year-olds were matched on temporal signal to noise ratio (tSNR) in all ROIs (all t's < 0.64, all p's > 0.52, all d's < 0.23), motion (average absolute frame displacement for all usable runs) ($t_{(30)} = 0.75$, $p = 0.46$, $d = 0.26$), 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). The number of usable runs from each subject was also comparable: for the final sample of 5 year olds, 10 participants had 4 runs, 5 participants had 3 runs, and 1 participant had 2 runs; for the final sample of 8 year olds, 14 participants had 4 runs, 1 participant had 3 runs and 1 participant had 2 runs.

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 [44, 45]. 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 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 [46-48], 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).

DATA CODE AND AVAILABILITY

The dataset generated during this study is available at https://osf.io/ztbuq/.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Dilks DD, Julian JB, Paunov AM, and Kanwisher N (2013). The occipital place area is causally and selectively involved in scene perception. J Neurosci 33, 1331–1336a. [PubMed: 23345209]

2. Ganaden RE, Mullin CR, and Steeves JK (2013). Transcranial magnetic stimulation to the transverse occipital sulcus affects scene but not object processing. J Cogn Neurosci 25, 961–968. [PubMed: 23410031]

3. Dilks DD, Julian JB, Kubilius J, Spelke ES, and 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. [PubMed: 21813690]

4. Kamps FS, Julian JB, Kubilius J, Kanwisher N, and Dilks DD (2016). The occipital place area represents the local elements of scenes. NeuroImage 132, 417–424. [PubMed: 26931815]

5. Kamps FS, Lall V, and Dilks DD (2016). 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. [PubMed: 27474914]

6. Persichetti AS, and 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. [PubMed: 26963085]

7. Persichetti AS, and Dilks DD (2018). Dissociable Neural Systems for Recognizing Places and Navigating through Them. J Neurosci 38, 10295–10304. [PubMed: 30348675]

8. Bonner MF, and Epstein RA (2017). Coding of navigational affordances in the human visual system. Proc Natl Acad Sci U S A 114, 4793–4798. [PubMed: 28416669]

9. Julian JB, Ryan J, Hamilton RH, and Epstein RA (2016). The occipital place area is causally involved in representing environmental boundaries during navigation. Curr Biol 26, 1104–1109. [PubMed: 27020742]

10. Deen B, Richardson H, Dilks DD, Takahashi A, Keil B, Wald LL, Kanwisher N, and Saxe R (2017). Organization of high-level visual cortex in human infants. Nat Commun 8, 13995. [PubMed: 28072399]

11. Meissner TW, Nordt M, and Weigelt S (2019). Prolonged functional development of the parahippocampal place area and occipital place area. NeuroImage 191, 104–115. [PubMed: 30763610]

12. Cohen MA, Dilks DD, Koldewyn K, Weigelt S, Feather J, Kell AJ, Keil B, Fischl B, Zollei L, Wald L, et al. (2019). Representational similarity precedes category selectivity in the developing ventral visual pathway. NeuroImage 197, 565–574. [PubMed: 31077844]

13. Powell LJ, Deen B, and Saxe R (2017). Using individual functional channels of interest to study cortical development with fNIRS. Developmental science 21, e12595. [PubMed: 28944612]

14. Julian JB, Kamps FS, Epstein RA, and Dilks DD (2019). Dissociable spatial memory systems revealed by typical and atypical human development. Developmental science 22, e12737. [PubMed: 30176106]

15. Dilks DD, Hoffman JE, and Landau B (2008). Vision for perception and vision for action: normal and unusual development. Developmental science 11, 474–486. [PubMed: 18576955]

16. Dekker TM, Ban H, van der Velde B, Sereno MI, Welchman AE, and Nardini M (2015). Late Development of Cue Integration Is Linked to Sensory Fusion in Cortex. Curr Biol 25, 2856–2861. [PubMed: 26480841]

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

18. Dillon MR, Persichetti AS, Spelke ES, and Dilks DD (2018). Places in the Brain: Bridging Layout and Object Geometry in Scene-Selective Cortex. Cerebral cortex 28, 2365–2374. [PubMed: 28633321]

19. Persichetti AS, and Dilks DD (2019). Distinct representations of spatial and categorical relationships across human scene-selective cortex. Proc Natl Acad Sci U S A.

20. Cantlon JF, Pinel P, Dehaene S, and Pelphrey KA (2011). Cortical representations of symbols, objects, and faces are pruned back during early childhood. Cerebral cortex 21, 191–199. [PubMed: 20457691]

21. Quartz SR (1999). The constructivist brain. Trends in cognitive sciences 3, 48–57. [PubMed: 10234227]

22. Changeux JP, and Danchin A (1976). Selective stabilisation of developing synapses as a mechanism for the specification of neuronal networks. Nature 264, 705–712. [PubMed: 189195]

23. Berard JR, and Vallis LA (2006). Characteristics of single and double obstacle avoidance strategies: a comparison between adults and children. Exp Brain Res 175, 21–31. [PubMed: 16761138]

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

25. Franchak JM, and Adolph KE (2010). Visually guided navigation: Head-mounted eye-tracking of natural locomotion in children and adults. Vision research 50, 2766–2774. [PubMed: 20932993]

26. Franchak JM, Kretch KS, Soska KC, and Adolph KE (2011). Head-mounted eye tracking: a new method to describe infant looking. Child development 82, 1738–1750. [PubMed: 22023310]

27. Nardini M, Thomas RL, Knowland VC, Braddick OJ, and Atkinson J (2009). A viewpoint-independent process for spatial reorientation. Cognition 112, 241–248. [PubMed: 19501349]

28. Gianni E, De Zorzi L, and Lee SA (2018). The developing role of transparent surfaces in children's spatial representation. Cognitive psychology 105, 39–52. [PubMed: 29920399]

29. Kretch KS, and Adolph KE (2017). The organization of exploratory behaviors in infant locomotor planning. Developmental science 20.

30. Adolph KE, Eppler MA, Gibson EJ (1993). Development of perception of affordances. Advances in infancy research.

31. Hermer L, and Spelke E (1996). Modularity and development: the case of spatial reorientation. Cognition 61, 195–232. [PubMed: 8990972]

32. Hermer L, and Spelke ES (1994). A geometric process for spatial reorientation in young children. Nature 370, 57–59. [PubMed: 8015605]

33. Gibson EJ, and Walk RD (1960). The "visual cliff". Sci Am 202, 64–71. [PubMed: 13827949]

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

35. Wills TJ, Cacucci F, Burgess N, and O'Keefe J (2010). Development of the Hippocampal Cognitive Map in Preweanling Rats. Science 328, 1573–1576. [PubMed: 20558720]

36. Langston RF, Ainge JA, Couey JJ, Canto CB, Bjerknes TL, Witter MP, Moser EI, and Moser MB (2010). Development of the Spatial Representation System in the Rat. Science 328, 1576–1580. [PubMed: 20558721]

37. Jiang P, Tokariev M, Aronen ET, Salonen O, Ma YY, Vuontela V, and 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.

38. Golarai G, Ghahremani DG, Whitfield-Gabrieli S, Reiss A, Eberhardt JL, Gabrieli JD, and Grill-Spector K (2007). Differential development of high-level visual cortex correlates with category-specific recognition memory. Nature neuroscience 10, 512–522. [PubMed: 17351637]

39. Scherf KS, Luna B, Avidan G, and Behrmann M (2011). "What" Precedes "Which": Developmental Neural Tuning in Face- and Place-Related Cortex. Cerebral cortex 21, 1963–1980. [PubMed: 21257673]

40. Pelphrey KA, Lopez J, and Morris JP (2009). Developmental continuity and change in responses to social and nonsocial categories in human extrastriate visual cortex. Front Hum Neurosci 3, 25. [PubMed: 19826492]

41. Scherf KS, Behrmann M, Humphreys K, and Luna B (2007). Visual category-selectivity for faces, places and objects emerges along different developmental trajectories. Developmental science 10, F15–30. [PubMed: 17552930]

42. Lourenco SF, Frick A (2013). Remembering where: The origins and early development of spatial memory. The Wiley Handbook on the Development of Children's Memory 1, 361–393.

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

44. Fedorenko E, Hsieh PJ, Nieto-Castanon A, Whitfield-Gabrieli S, and Kanwisher N (2010). New method for fMRI investigations of language: defining ROIs functionally in individual subjects. Journal of neurophysiology 104, 1177–1194. [PubMed: 20410363]

45. Julian JB, Fedorenko E, Webster J, and Kanwisher N (2012). An algorithmic method for functionally defining regions of interest in the ventral visual pathway. Neuroimage 60, 2357–2364. [PubMed: 22398396]

46. Norman-Haignere SV, Albouy P, Caclin A, McDermott JH, Kanwisher NG, and Tillmann B (2016). Pitch-Responsive Cortical Regions in Congenital Amusia. J Neurosci 36, 2986–2994. [PubMed: 26961952]

47. Saygin ZM, Osher DE, Norton ES, Youssoufian DA, Beach SD, Feather J, Gaab N, Gabrieli JDE, and Kanwisher N (2016). Connectivity precedes function in the development of the visual word form area. Nature neuroscience 19, 1250–1255. [PubMed: 27500407]

48. Kamps FS, Morris EJ, and Dilks DD (2019). A face is more than just the eyes, nose, and mouth: fMRI evidence that face-selective cortex represents external features. NeuroImage 184, 90–100. [PubMed: 30217542]

49. Barbu-Roth M, Anderson DI, Despres A, Provasi J, Cabrol D, and Campos JJ (2009). Neonatal Stepping in Relation to Terrestrial Optic Flow. Child development 80, 8–14. [PubMed: 19236388]

50. Jouen F, Lepecq JC, Gapenne O, and Bertenthal BI (2000). Optic flow sensitivity in neonates. Infant Behav Dev 23, 271–284.

51. Stoffregen TA, Schmuckler MA, and Gibson EJ (1987). Use of Central and Peripheral Optical-Flow in Stance and Locomotion in Young Walkers. Perception 16, 113–119. [PubMed: 3671034]

52. Lee DN, and Aronson E (1974). Visual Proprioceptive Control of Standing in Human Infants. Perception & psychophysics 15, 529–532.

53. Bertenthal BI, and Bai DL (1989). Infants Sensitivity to Optical-Flow for Controlling Posture. Developmental Psychology 25, 936–945.

54. Gilmore RO, Baker TJ, and Grobman KH (2004). Stability in young infants' discrimination of optic flow. Developmental Psychology 40, 259–270. [PubMed: 14979765]

55. Kellman PJ, Gleitman H, and Spelke ES (1987). Object and Observer Motion in the Perception of Objects by Infants. J Exp Psychol Human 13, 586–593.

56. Cardin V, and Smith AT (2010). Sensitivity of human visual and vestibular cortical regions to egomotion-compatible visual stimulation. Cerebral cortex 20, 1964–1973. [PubMed: 20034998]

57. Uesaki M, and Ashida H (2015). Optic-flow selective cortical sensory regions associated with self-reported states of vection. Front Psychol 6.

58. Kovacs G, Raabe M, and Greenlee MW (2008). Neural correlates of visually induced self-motion illusion in depth. Cerebral cortex 18, 1779–1787. [PubMed: 18063566]

59. Pitcher D, Dilks DD, Saxe RR, Triantafyllou C, and Kanwisher N (2011). Differential selectivity for dynamic versus static information in face-selective cortical regions. NeuroImage 56, 2356–2363. [PubMed: 21473921]

60. Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TEJ, Johansen-Berg H, Bannister PR, De Luca M, Drobnjak I, Flitney DE, et al. (2004). Advances in functional and structural MR image analysis and implementation as FSL. NeuroImage 23, S208–S219. [PubMed: 15501092]

61. Wang L, Mruczek REB, Arcaro MJ, and Kastner S (2015). Probabilistic Maps of Visual Topography in Human Cortex. Cerebral cortex 25, 3911–3931. [PubMed: 25452571]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript
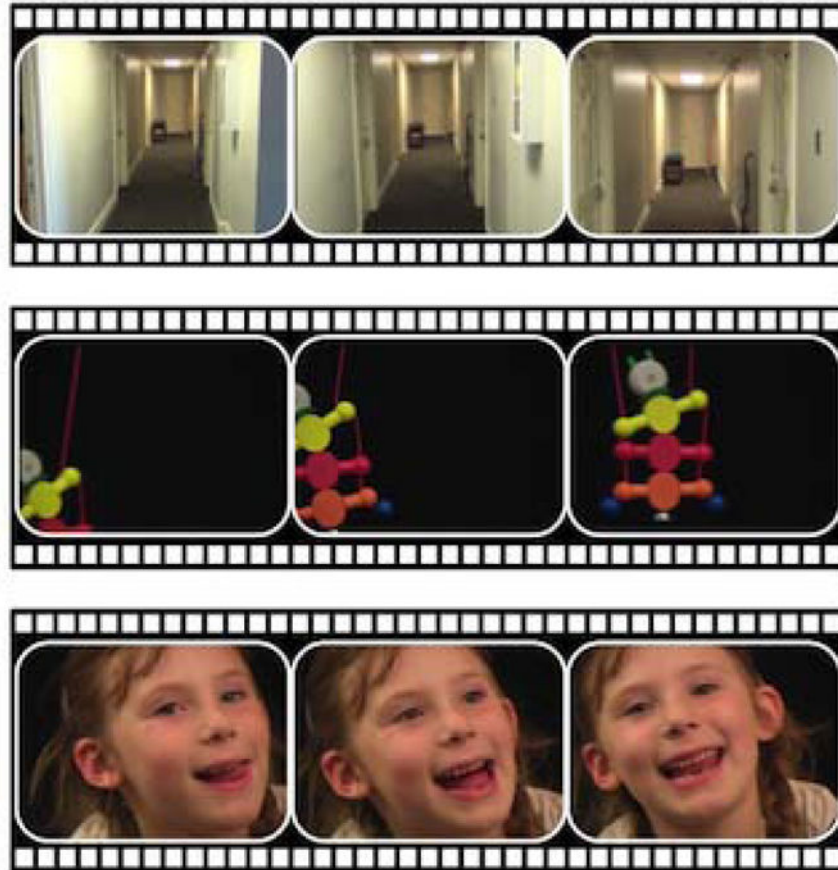
Figure 1. Example experimental stimuli.
The conditions included Dynamic and Static Scenes (top row), Dynamic and Static Objects (middle row), and Dynamic and Static Faces (bottom row). The Dynamic Scene stimuli consisted of 3-s video clips of first-person perspective motion through scenes. The Static 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 Object stimuli consisted of 3-s video clips of moving toys presented against a black background, while 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 Static Object and Static Face stimuli were created following the same procedure described for the Static Scene stimuli.
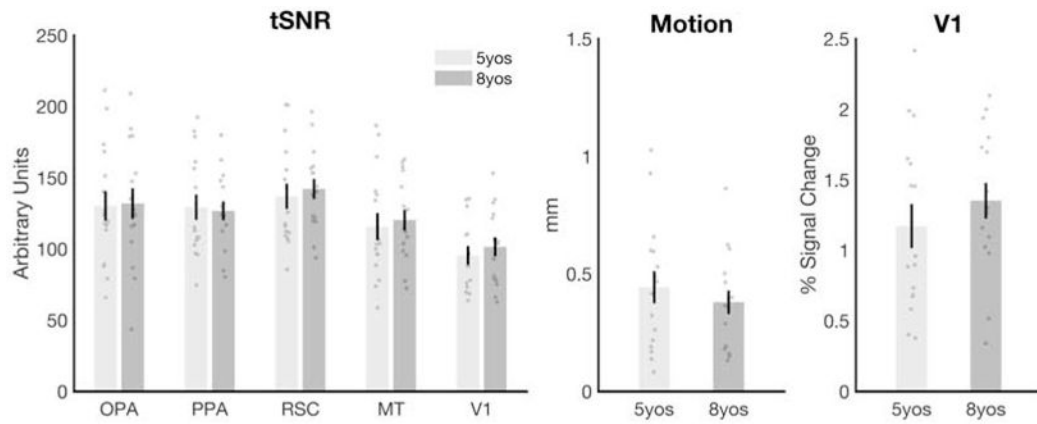
**Figure 2. Data quality did not differ between the 5- and 8-year-old children.**
No significant differences were found between the 5- and 8-year-olds for any of the following measures: A) the temporal signal to noise ratio (tSNR) in any region of interest (all t's < 0.64, all p's > 0.52); B) participant 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.
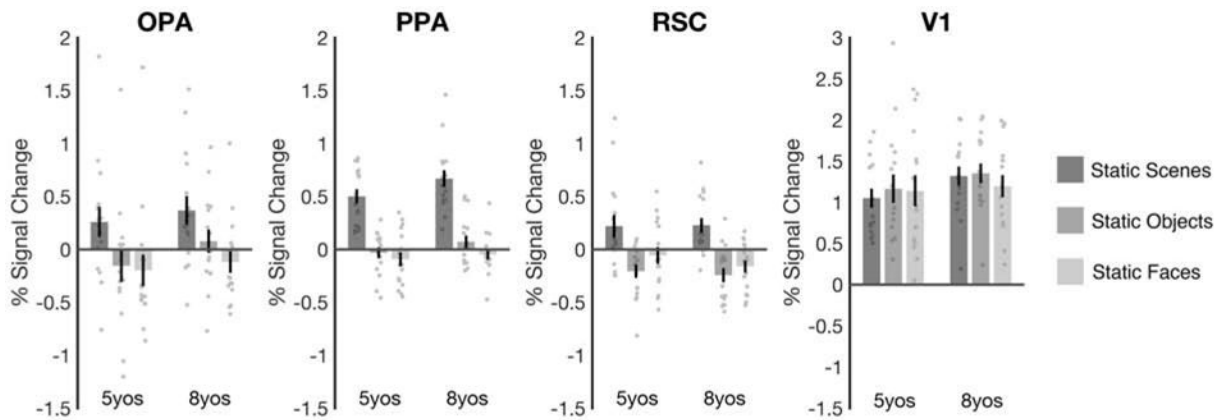
Figure 3. Scene selectivity is present in all three scene-selective regions by age 5.
Responses are shown for static stimuli only, following standard contrasts used to measure scene selectivity in adults. For 5-year-olds and 8-year-olds, OPA, PPA, and RSC each responded significantly more to scenes than both objects and faces (all p's < 0.001). 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. Finally, V1 responded similarly across the three conditions ($p = 0.30$), ruling out the possibility that the scene-selective responses in OPA, PPA, and RSC were driven by stimulus complexity or low-level visual features. Error bars depict the standard error of the mean.
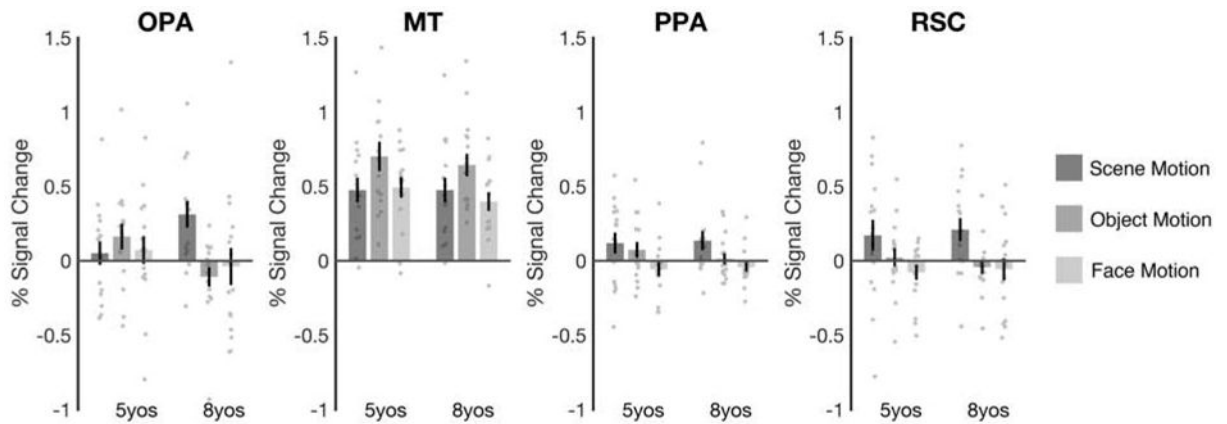
Figure 4. First-person perspective motion processing develops in OPA from 5 to 8 years old.
Scene motion, object motion, and face motion difference scores were calculated by
subtracting responses to the Static stimuli from responses to the Dynamic stimuli separately
for each condition. OPA responded 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) ($p = 0.64$; Table S1),
or in other scene-selective regions (PPA or RSC) (both p's > 0.68; Table S2). These findings
suggest that navigationally-relevant motion processing is still developing in OPA across
childhood. Error bars depict the standard error of the mean. See also Figures S1 and S2.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited Data		
ROI response data	This paper	https://osf.io/ztbuq/
ROI search spaces	[48]	https://web.mit.edu/bcs/nklab/GSS.shtml
Software and Algorithms		
MATLAB 2019b	MathWorks	https://www.mathworks.com
SPSS Statistics	IBM	https://www.ibm.com/products/spss-statistics
FMIRB Software Library (FSL)	Analysis Group, FMRIB	https://www.fmrib.ox.ac.uk/fsl