Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

______________ ___________

Shosuke Suzuki Date

Retrosplenial complex represents the broader spatial environment during the active experience of navigation.

By

Shosuke Suzuki

Master of Arts

Psychology

Michael T. Treadway, Ph.D.

Daniel D. Dilks, Ph.D.

Scott O. Lilienfeld, Ph.D.

Accepted:

Lisa A. Tedesco, Ph.D.

Dean of the James T. Laney School of Graduate Studies

Date

Retrosplenial complex represents the broader spatial environment during the active experience of navigation.

By

Shosuke Suzuki

B.S., University of California Los Angeles, 2014

Advisor: Michael T. Treadway, Ph.D.

An abstract of

A thesis submitted to the Faculty of the

James T. Laney School of Graduate Studies of Emory University

in partial fulfillment of the requirements for the degree of

Master of Arts

in Psychology

2020

Abstract

Retrosplenial complex represents the broader spatial environment during the active experience of navigation.

By Shosuke Suzuki

Successfully navigating the world requires not only being able to find one's way through the immediately visible environment, but also situating the immediately visible environment to a broader spatial environment. Two scene systems are thought to support these different navigational processes, with the occipital place area (OPA) supporting navigation through the immediate environment ("visually-guided navigation") and the retrosplenial complex (RSC) supporting navigation of the broader environment ("map-based navigation"). However, the precise roles of these systems are still not well understood, especially since most studies use only static displays and/or passive-viewing tasks, rather than studying responses within these regions during the active, first-person perspective experience of navigating. Moreover, a widespread and intuitive hypothesis in the spatial learning literature is that such active experience is critical for spatial processing, raising the possibility that active processing may also be critical for scene navigation; although the role of active vs. passive viewing has never been tested in the cortical navigation regions. Here we address these two hypotheses using fMRI and a novel maze-navigation paradigm. While in the scanner, participants navigated either simple or complex mazes (testing representation of the broader spatial environment) and did so while either controlling their own movement via button presses (i.e., "active") or by being ferried through the maze (i.e., "passive") (testing the active experience hypothesis). Consistent with the hypothesized division of labor within navigation between OPA and RSC, we indeed found that RSC responded significantly more to complex than simple mazes, while OPA did not. Surprisingly, however, we further found a specific effect of active processing in RSC only, showing greater processing of the broader environment in the active than passive condition (no effect was found in OPA). Taken together, these findings suggest the novel hypothesis that map-based navigation in RSC is relatively deliberate, whereas visually-guided navigation in OPA is relatively automatic, shedding new light on the dissociable functions of these systems.

Retrosplenial complex represents the broader spatial environment during the active experience of navigation.

By

Shosuke Suzuki

B.S., University of California Los Angeles, 2014

Advisor: Michael T. Treadway, Ph.D.

A thesis submitted to the Faculty of the

James T. Laney School of Graduate Studies of Emory University

in partial fulfillment of the requirements for the degree of

Master of Arts

in Psychology

2020

Table of Contents

List of Figures

Introduction

To successfully navigate our spatial world, we must both rapidly and effortlessly move about the immediately visible environment, avoiding boundaries and obstacles (e.g., walking around one's bedroom, not bumping into the walls or furniture), and more slowly and consciously situate the immediately visible environment within the broader spatial environment, ultimately enabling us to find our way from the current place to another distant place (e.g., knowing that the bathroom is down the hall and to the right of one's bedroom). Neuroimaging studies have begun to identify neural systems that may support these two navigational processes—herein referred to as "visually-guided navigation" and "map-based navigation", respectively. Critically, however, given the primary reliance of prior work on using passive-viewing tasks to examine navigation-related neural responses, the precise roles of these neural systems during the active, dynamic experience of first-person navigation remains unclear.

Two distinct neural systems have thus far been implicated in supporting map-based and visually-guided navigation. We define map-based navigation as a set of processes that allows an individual to learn a map by integrating the immediately visible scene with the broader spatial structure, localize their current position within the map, and choose a course of actions given the current position (Filliat & Meyer, 2003). Specifically, one scene-selective region—the retrosplenial complex (RSC) (Maguire, 2001)—has been shown to represent information necessary for navigating the broader environment (i.e., sense, egocentric distance, and heading direction and location of a navigator) (Dilks, Julian, Kubilius, Spelke, & Kanwisher, 2011; Marchette, Vass, Ryan, & Epstein, 2014; Persichetti & Dilks, 2016, 2019). Moreover, RSC encodes landmarks and their spatial relationships in the service of forming a cognitive map when learning a virtual environment (Iaria, Chen, Guariglia, Ptito, & Petrides, 2007), and activation in

RSC is positively correlated with gained map expertise (Sherrill et al., 2013; Wolbers & Büchel, 2005). By contrast, we define visually-guided navigation as a set of processes that allows individuals to use visual information within the immediate scene to find one's way through the environment by avoiding boundaries and obstacles. Another scene-selective region—the occipital place area (OPA) (Dilks, Julian, Paunov, & Kanwisher, 2013)—has been shown to represent information necessary for navigating the currently visible environment (i.e., "sense" or direction, egocentric distance, boundaries and obstacles, first-person perspective motion, relative length and angle of surfaces, as well as potential paths for movement in one's immediate surroundings) (Bonner & Epstein, 2017; Dilks et al., 2011; Dillon, Persichetti, Spelke, & Dilks, 2018; Julian, Ryan, Hamilton, & Epstein, 2016; Kamps, Julian, Kubilius, Kanwisher, & Dilks, 2016; Kamps, Lall, & Dilks, 2016; Persichetti & Dilks, 2016). Further, OPA responds to imagined visually-guided navigation (Persichetti & Dilks, 2018).

While neuroimaging studies have begun to identify the neural correlates of map-based and visually-guided navigation, how these neural systems interact with the active-experience of navigation have largely been overlooked. This is critical given that a widespread and intuitive hypothesis in the spatial learning literature is that such active experience is critical for spatial processing (Chrastil & Warren, 2012). For example, scene recognition is disrupted when objects in a scene are moved with respect to a stationary observer, but not when the observer moves with respect to the scene (Simons & Wang, 1998). However, no study to our knowledge has examined the role of active compared to passive experiences of navigation in cortical navigation regions.

Importantly, the distinct neural machinery supporting map-based and visually-guided navigation are expected to interact differently to the active experience of navigation: A map-based navigation system which is concerned with navigation of the broader environment should exhibit greater responses to the active experience of navigation compared to passive viewing, given that integrating a mental spatial representation most likely involves cognitive processes that are relatively deliberate. By contrast, a visually-guided navigation system concerned with navigating the immediate environment should respond similarly regardless of whether navigation is actively experienced or passively-viewed. That is, a visually-guided navigation should be relatively indifferent to whether the individual is approaching a wall, or the wall is approaching the individual, so as long as the information required to prevent bumping into the wall (i.e., sense and distance information) is encoded. Thus, here we hypothesize that RSC supports a map-based navigation system that is relatively deliberate, whereas OPA supports a visually-guided navigation system that is relatively automatic.

To directly test this hypothesis, we used fMRI and a novel maze-navigation paradigm in adult humans. Specifically, on each trial, participants were required to virtually navigate through a 3D maze environment. Maze structures were either simple (with one turn) or complex (with two turns), and participants navigated through them by controlling their own movements (i.e., “active”) or by simply watching a prerecorded animation (i.e., “passive”). If RSC is involved in map-based navigation, then it should respond significantly more while participants are navigating through the complex mazes as compared to simple mazes, which would suggest that RSC represents information about the broader spatial environment. Importantly, we further predict that this representation should be significantly greater during the active than passive navigation condition, which would suggest that processing of navigational information in RSC is relatively deliberate. By contrast, if OPA is indeed involved in visually-guided navigation, then it should respond similarly when participants are navigating through both the simple or complex mazes. Such a pattern would be consistent with our hypothesis that OPA encodes navigational information from the immediately visible environment, but not the broader spatial environment. Further, OPA responses should be similar during the active and passive navigation conditions, demonstrating that encoding of navigational information in OPA is relatively automatic.

Materials and Methods

Participants

We recruited 20 healthy adults from the Atlanta community through fliers and online advertisements. Following prior work on our regions of interest, we expected relatively large within-subject effect sizes for two-way (η^2~.58) and three-way (η^2~.16) interactions, which we estimated by averaging previously-reported effect sizes (Dilks et al., 2013; Kamps, Lall, et al., 2016; Persichetti & Dilks, 2016, 2018). Power analyses using the Power Analysis for General Anova designs program (PANGEA)(Westfall, 2016) indicated that sample sizes of 4 and 15 would afford 80% power to detect such interaction effects, respectively, at alpha=0.05. Our recruited sample of 20 participants corresponded to a power of 96.4% to detect a significant three-way interaction.

Eligibility was determined using an online pre-screening survey. Eligible participants were right-handed, English-speaking individuals between the ages of 18-35. Individuals were excluded if they: (1) had any contraindications for magnetic resonance imaging (e.g., claustrophobia, metallic implants, central nervous system disorders, pregnancy in females); (2) were currently taking psychoactive medications, investigational drugs, or drugs that affect blood flow (e.g., for hypertension); or (3) reported current medical, neurological, or psychiatric illnesses. All participants gave informed consent and had normal or corrected-to-normal vision.

One participant was excluded because we could not localize any of his scene-selective cortical regions from the functional localizer. Thus, the final sample included 19 participants (M_{age}=25.36; SD_{age}=5.55; 13 females). Upon completion of study procedures, participants were compensated for their participation. All procedures were approved by the Emory Institutional Review Board.

Experimental Design

We used a region of interest (ROI) approach, in which we localized scene-selective cortical regions via a functional localizer run. Then, we used an independent set of experimental runs to investigate the responses of these regions while participants completed the maze-navigation task. Prior to scanning, all participants underwent a 15-minute training procedure on the maze-navigation task (see *Training and Calibration*).

For the functional localizer run, ROIs were identified using a standard method (Epstein & Kanwisher, 1998). Specifically, in a blocked design, participants saw pictures of scenes and objects for a total of 336s, consisting of 8 blocks per stimulus category. Each block was 16s long and contained 20 pictures from the same category. Each picture was presented for 300ms, followed by a 500ms interstimulus interval (ISI). We also included five 16s fixation blocks: 1 at the beginning; 3 in the middle interleaved after every 4 blocks; and 1 at the end of the run. Participants performed a one-back task, responding with a button press every time the same picture was presented twice in a row.

For the experimental runs (i.e., the maze-navigation task), participants were required to virtually navigate through single-path mazes via a 4-button box with the right (i.e., dominant) hand. Specifically, one button moved the participant forward (r=2.2units/s), and two buttons each rotated the participant clockwise and counterclockwise ($\omega=0.5\pi$ rad/s). Holding down the buttons continuously applied the translational/rotational effects. Acceleration was applied to participant motion to mimic real-life motion. Only one type of motion was allowed at any given moment (i.e., pressing multiple buttons resulted in no motion).

Each trial was associated with one of four navigation conditions, including i) an *active navigation* condition, in which participants were required to advance through the maze using button presses to move and rotate, and ii) a *passive navigation* condition, in which participants were required to simply view moving through the maze without making any action. Two other navigation conditions were included for the purpose of another study, including an *effortful navigation* condition, in which participants were required to repeatedly press the button to advance through the maze, and a *no navigation* condition, in which participants waited for the approximate duration of the maze until they were teleported to the end of the maze. All conditions were equated in navigation time (~4.75s). In addition to the navigation condition, we also varied the structure of the mazes. Specifically, there were “simple” mazes which required a single 90° turn (left or right) and “complex” mazes which required two 90° turns (left-then-right or right-then-left). Regardless of structure, each maze was comprised of 1x1 unit^2 floors placed adjacently to form a path, bounded by 1x1 unit^2 walls. The participant view was set at 0.6 units above the floor, with a -10° tilt along the surface axes.

Each trial proceeded as follows (Figure 1): (1) Start phase: At the beginning of each trial, the participant’s position was initialized to the beginning of the maze, immediately followed by a cue phase. (2) Cue phase: The floor tile immediate to the participant’s view changed color, which represented a cue informing the participants to the navigation condition for that trial. The colored cue lasted for 2s before returning to its original floor texture, and the participant was rendered immobile during the cue presentation. (3) Jittered interstimulus interval (ISI): The cue was immediately followed by a jittered fixation period, whereby a ‘+’ was rendered on top of the current scene for a Poisson-distributed duration with a mean of 2.5s. The participant was still restricted from moving during this period. (4) Navigation phase: After the fixation cross disappeared, participants completed navigation depending on the navigation condition, as detailed above. (5) Jittered ISI: Once participants successfully reached the door at the end of the maze (the “goal”), the player was again rendered immobile for an ISI jittered around 2.5s. (6) Reward phase (used for another study): Following the ISI, participants were presented with an animation of the door opening followed by a monetary reward, represented by a coin with the dollar amount rendered on its surface. Each trial was associated with one of 4 bins of reward magnitudes ($0, $1.68-2.78, $2.79-$3.89, $3.90-5.00), from which an amount was randomly selected. (7) Rating phase (used for another study): Once every 4 trials, the participant was asked to make a mood rating on a Likert-scale between 1 (not happy at all) to 4 (very happy) using a button-press. (8) Jittered inter-trial interval (ITI): Finally, participants were presented with a ‘+’ rendered on a grey screen for a duration jittered around 3s.

All navigation conditions, maze structures, and reward bins were equally distributed and balanced across trials and runs. Each participant completed 3 runs with 32 trials each (~11min/run), and trials were presented in a fixed-randomized order.

Training and Calibration

Participants completed a 15-min training procedure of the maze-navigation task prior to the scanning session, to ensure that they understood the instructions and could complete the task. First, participants were told that they will be navigating through virtual mazes to obtain monetary rewards. They were introduced to the player controls and all navigation conditions. To incentivize the participants, they were also told that a proportion of the reward they obtain on each trial will be added to their payment as a bonus. Once participants indicated that they understand the task and can follow the instructions, they completed 16 practice trials on a laptop computer. To maximize the effect of practice, participants used the same hand and fingers used to perform the task in the MRI scanner.

Data Acquisition & Analysis

The maze-navigation task was programmed using Unity 3D (Unity Technologies ApS). Stimuli were presented via back-projection mirror, and participants completed the maze-navigation task and functional localizer runs using an MR-compatible 4-button box (Current Designs Inc). Foam pads placed around participants' heads were used to minimize motion.

Participants were scanned in a 3-Tesla Siemens TIM Trio scanner (Siemens AG) with a 32-channel head-coil. We used multiband functional and structural imaging (Feinberg et al., 2010; Feinberg & Setsompop, 2013; Xu et al., 2013), similarly used by the Human Connectome Project consortium (Van Essen et al., 2013). Each session began with a 3-plane localizer scan for slice alignment, and a single-shot, high-resolution structural MPRAGE sequence (TR/TE=1900/2.27ms; flip angle=9°; FoV=250x250mm; 192x1.0mm slices). BOLD functional images are acquired with $T2^{*}$-weighted EPI sequences with a multiband acceleration factor of 4 (TR/TE=1000/30.0ms; flip angle=65°; FoV=220xx220mm; 52x3.00mm slices).

Functional images were preprocessed using SPM12 scripts through NeuroElf v1.1, following best practice guidelines described by the Human Connectome Project for multiband data analysis (Glasser et al., 2013). Specifically, images were co-registered to the structural image, motion-corrected, warped to the Montreal Neurological Institute template, and smoothed using a Gaussian filter (6mm full width-half maximum). Raw and preprocessed data were subjected to multiple tests for quality assurance and inspected for spiking and motion. Volumes were discarded if the root mean square of motion parameters exceeded a single voxel dimension (3mm), or if striping was identified through visual inspection. Subject-level modeling of trial events was conducted using robust regression to reduce the influence of strong outliers.

To identify scene-selective regions, the scene and object blocks in the functional localizer run were included in a subject-level GLM as regressors of interest. Motion parameters and their squares, as well as high-pass filter parameters were included as additional nuisance regressors. For each participant, fixed-effects contrasts (Scenes>Objects) were generated to individually define ROIs. Specifically, we identified scene-selective ROIs (bilateral OPA, RSC, and a third scene-selective control region – the parahippocampal place area, PPA) at a voxelwise threshold of $p<10^{-6}$, uncorrected (Figure 2).

For the maze-navigation task, the Cue, Navigation, and Reward phases were included in a subject-level GLM as regressors of interest. A separate regressor was included for each navigation condition at each phase, and additionally for each maze structure during the Navigation phase. We also included the Start and Rating phases as well as the ISIs in the model to omit their influences on the implicit baseline. In addition, motion parameters and their squares, as well as high-pass filter parameters were again included as nuisance regressors. We then extracted beta parameters for the navigation conditions for each maze structure from each scene-selective ROI for each participant, and conducted repeated-measures ANOVAs on the neural response for each ROI.

Results

Given our hypothesis that RSC deliberately integrates navigational information about the immediate environment to the broader spatial environment, we predicted greater RSC responses to more complex mazes, and especially during active, rather than passive, navigation. Consistent with these predictions, a 2 (navigation condition: active, passive) x 2 (maze structure: simple, complex) repeated measures ANOVA revealed significant main effects of navigation condition ($F_{(1,18)}=23.21$, $p<.001$, $\eta^2=.56$), maze structure ($F_{(1,18)}=19.58$, $p<.001$, $\eta^2=.52$), and their interaction ($F_{(1,18)}=6.11$, $p=.02$, $\eta^2=.25$). As predicted, post-hoc analyses revealed that this interaction was driven by RSC's heightened response to maze structure (complex > simple) when actively navigating the maze ($t_{(18)}=4.90$, $p<.001$, d=1.12), compared to passive navigation ($t_{(18)}=1.95$, $p=.07$) (Figure 3). Given RSC's implicated role in navigation, we also repeated the analysis comparing the passive and no navigation conditions to examine the effect of navigation within RSC, and found a significant effect of navigation condition ($F_{(1,18)}=14.18$, $p<.001$, $\eta^2=.44$).

In contrast to RSC, given our hypothesis that OPA encodes navigational information about the immediate environment rather than the broader spatial environment, and does so automatically, we predicted that OPA will respond similarly to navigation through simple and complex structures, and regardless of whether navigation is performed actively or passively. Indeed, a 2 (navigation condition: active, passive) x 2 (maze structure: simple, complex) repeated-measures ANOVA revealed no significant main effects of navigation condition or maze structure (navigation condition: $F_{(1,18)}=4.09$, $p=.06$, $\eta^2=.19$; maze structure: $F_{(1,18)}=3.19$, $p=.09$, $\eta^2=.15$; Note while the p-values are near significant, the effect sizes are low). Crucially, we found no navigation condition by maze structure interaction ($F_{(1,18)}=0.01$, $p=.93$, $\eta^2=.00$), consistent with our hypothesis that OPA encodes navigational information, regardless of the environmental structure, and does so automatically (Figure 3). But does this lack of effect mean that we cannot detect any significant effects? As with RSC, we repeated the analysis comparing the passive and no navigation conditions given OPA's implicated role in navigation. We found a significant effect of navigation condition ($F_{(1,18)}$=24.76, p<.001, η^2=.58), which not only suggests that we are able to detect differences, but also replicates a prior finding showing that OPA represents first-person perspective motion through scenes (Kamps, Lall, et al., 2016).

The above results suggest differential navigational processing across RSC and OPA. To directly test this, we conducted a 2 (region: OPA, RSC) x 2 (navigation condition: active, passive) x 2 (maze structure: simple, complex) repeated-measures ANOVA. We found a significant main effect of region ($F_{(1,18)}$=10.58, p=.004, η^2=.37), navigation condition ($F_{(1,18)}$=12.81, p=.002, η^2=.42), and maze structure ($F_{(1,18)}$=11.16, p=.004, η^2=.38), as well as a significant interaction between region and maze structure ($F_{(1,18)}$=14.78, p=.001, η^2=.45) and a three-way interaction between region, navigation condition, and structure ($F_{(1,18)}$=14.41, p=.001, η^2=.45; observed power=94.8%). These interactions were driven by the effects of maze structure and interaction between navigation condition and maze structure that were uniquely observed in RSC as described above, providing evidence of dissociable navigational processing between RSC and OPA.

Next, we also examined the effects of navigation condition and maze structure in RSC to PPA, another scene-selective region thought to be involved in the categorization of scenes (e.g., recognizing a kitchen versus a beach) (Persichetti & Dilks, 2018), but not navigation (Epstein, Harris, Stanley, & Kanwisher, 1999). A 2 (region: RSC and PPA) x 2 (navigation condition: active, passive) x 2 (maze structure: simple, complex) repeated-measures ANOVA revealed a significant three-way interaction between region, navigation condition, and maze structure ($F_{(1,18)}$=6.97, p=.02, η^2=.28). Whereas PPA, like RSC, also exhibited significant main effects of navigation condition ($F_{(1,18)}$=16.82, p=.001, η^2=.48) and maze structure ($F_{(1,18)}$=6.25, p=.02, η^2=.26), it did not show a significant interaction between maze structure and navigation condition ($F_{(1,18)}$=1.08, p=.31, η^2=.06). Additionally, we found a significant interaction between region and maze structure ($F_{(1,18)}$=13.46, p=.002, η^2=.43), such that RSC more strongly differentiated between maze structures compared to PPA.

Though not the focus of the current paper, our paradigm also included an effortful navigation condition, and presentation of reward upon completion of navigation. For completeness, we report the response of scene regions for effort and reward, predicting that neither of these conditions should differentially affect either the OPA or RSC. Indeed, a 2 (region: OPA, RSC) x 2 (condition: effortful, active) repeated-measures ANOVA revealed no significant region by condition interaction ($F_{(1,18)}$=0.00, p=.99, η^2=.00), suggesting similar patterns of activity between the regions with respect to the level of effort required for navigation. Similarly, a 2 (region: OPA, RSC) x 2 (reward: reward, no reward) repeated-measures ANOVA revealed no significant region by condition interaction ($F_{(1,18)}$=1.55, p=.23, η^2=.08), suggesting that the regions do not differentially respond to reward outcome following navigation.

Discussion

The current study examined the neural systems that support visually-guided navigation and map-based navigation, by examining neural activity in scene-selective regions during active and passive navigation through varying maze structures. We provide two pieces of evidence that were necessary to test in order to support RSC's role in map-based navigation. First, we demonstrated RSC’s sensitivity to the maze structure by showing greater responses in complex compared to simple mazes, suggesting that RSC represents information about the broader spatial environment. Second, RSC’s response to the maze structure was heightened when actively experiencing the navigation, compared to passively viewing the navigation, which may reflect the deliberate nature of the processes associated with integrating spatial information. In contrast to RSC, OPA behaved similarly when navigating through simple and complex mazes, suggesting that OPA is relatively indifferent to the broader structure. Additionally, OPA did not distinguish between active and passive navigation, suggesting that OPA encodes navigational information relatively automatically, regardless of whether the individual has active control over the navigation. Thus, taken together, we provide direct evidence that RSC is disproportionately involved in map-based navigation compared to other scene-selective regions. Our results are also consistent with OPA’s hypothesized role in visually-guided navigation, to the extent that OPA was relatively indifferent about the broader spatial structure or the active experience of navigation.

Our finding that RSC is sensitive to maze structure is consistent with prior research implicating RSC in anchoring a cognitive map (Epstein, Patai, Julian, & Spiers, 2017). Importantly, extending prior work, we show that this sensitivity is elicited especially during active, rather than passive, navigation. Thus, our data suggests that RSC integrates navigational information about the immediate environment with the broader structure relatively deliberately. This finding provides a possible neural correlate to the earliest works on cognitive maps, which suggested that individuals actively process spatial information rather than simply operating on a stimulus-response relationship (Blodgett, 1929; Tolman, 1948). Interestingly, RSC responses did not differ from OPA when increasing the effort level required to navigate through the spatial environment, suggesting that RSC may be sensitive to the act of navigation, but not necessarily to the level of physical engagement. These observations are consistent with our hypothesis that RSC supports map-based navigation. Importantly, our findings are unlikely due to non-task factors like attention, given the observed difference in RSC activity between complex and simple maze structures during active navigation.

Our finding that OPA does not differentiate between varying navigation conditions or maze structures extends prior research associating OPA with visually-guided navigation. Specifically, prior work has reported OPA activation during imagined navigation (Persichetti & Dilks, 2018) and "passive" first-person perspective motion (Kamps, Lall, et al., 2016), but no studies to our knowledge have examined OPA activity during actual (or virtual) navigation. Here, we find that OPA is indifferent to whether virtual navigation is being executed actively or passively, suggesting that OPA guides navigation automatically. Additionally, prior work has reported that OPA represents information about the local environment (Bonner & Epstein, 2017; Julian et al., 2016). We confirm and extend these results by showing that OPA is not sensitive to the broader structure (i.e., maze complexity) of the navigated environment. Thus, our findings provide novel evidence suggesting that OPA supports a visually-guided navigation system that automatically and selectively encodes information about the immediate, not broader, environment.

Still, we suggest that these results within OPA are interpreted with caution, given near-significant effects of both the active experience of navigation and maze structure despite low effect sizes. Here, we offer alternative explanations to these potential effects whilst maintaining our hypothesis that OPA supports visually-guided navigation. First, a potentially increased OPA response to active navigation, compared to passive navigation, may be explained by a generalized increase in neural signaling due to attentional demands. Although participants were instructed to focus on the task even during the passive-viewing condition, assuming active control of movement may have engaged additional attentional processes that may have distributed downstream effects. Consistent with this, we also observed increased responses to the active experience of navigation even within PPA, a region which is not believed to support navigation. Nonetheless, we observed a differential pattern of responses between the regions, suggesting that RSC represents navigational information above and beyond any generalized effect of attention. Second, a potentially increased OPA response to complex mazes, compared to simple mazes, may be explained by slight differences in the immediately visible scenes while navigating the maze structures. Specifically, the complex mazes included two turns/corners, while the simple mazes included a single turn/corner. If OPA supports visually-guided navigation which allows an individual to avoid boundaries, and if OPA not only represents environmental boundaries (Julian et al., 2016) but also the *type* of boundaries, then one may expect OPA to respond greater to corners compared to straight paths; however, this remains to be tested.

Although our task required first-person virtual navigation through visuospatial environments, to what extent the observed effects generalize to navigation through the physical world remains an open question. Specifically, the restrictions of MRI measurement required participants to remain motionless on the scanner bed while they controlled task movement using button presses. Despite this limitation, to our knowledge this is the first study to examine evidence for neural systems supporting map-based and visually-guided navigation using a virtually-navigable environment, and we hope that future studies will follow-up using increasingly ecological and innovative designs. Importantly, to the extent that these neural systems responded to the active experience of navigation, we were able to confirm the role of these neural regions in representing navigation information. This is critical given that prior work had primarily used static images and passive-viewing tasks. Additionally, our findings point to the importance of incorporating an active condition when studying the neural underpinnings of navigation. More broadly, it has been suggested that studying neural processes under naturalistic and dynamic conditions is critical to achieving a deeper understanding of real-world brain function (Sonkusare, Breakspear, & Guo, 2019).

In conclusion, using a novel 3D maze-navigation task and fMRI, we found evidence that the adult human navigation is composed of two systems: a map-based navigation system, including RSC, that processes information relatively deliberately, and a visually-guided navigation system, including OPA, that processes information relatively automatically. This finding helps refine our understanding of the brain's navigation system and offers further evidence to suggest functional specialization across scene-selective cortex.

References

Blodgett, H. C. (1929). The effect of the introduction of reward upon the maze performance of rats. *University of California publications in psychology*.

Bonner, M. F., & Epstein, R. A. (2017). Coding of navigational affordances in the human visual system. *Proceedings of the National Academy of Sciences*, *114*(18), 4793-4798.

Chrastil, E. R., & Warren, W. H. (2012). Active and passive contributions to spatial learning. *Psychonomic bulletin & review, 19*(1), 1-23.

Dilks, D. D., Julian, J. B., Kubilius, J., Spelke, E. S., & Kanwisher, N. (2011). Mirror-image sensitivity and invariance in object and scene processing pathways. *Journal of Neuroscience, 31*(31), 11305-11312.

Dilks, D. D., Julian, J. B., Paunov, A. M., & Kanwisher, N. (2013). The occipital place area is causally and selectively involved in scene perception. *Journal of Neuroscience*, *33*(4), 1331-1336.

Dillon, M. R., Persichetti, A. S., Spelke, E. S., & Dilks, D. D. (2018). Places in the brain: bridging layout and object geometry in scene-selective cortex. *Cerebral Cortex, 28*(7), 2365-2374.

Epstein, R., Harris, A., Stanley, D., & Kanwisher, N. (1999). The parahippocampal place area: Recognition, navigation, or encoding? *Neuron, 23*(1), 115-125.

Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature, 392*(6676), 598-601.

Epstein, R., Patai, E. Z., Julian, J. B., & Spiers, H. J. (2017). The cognitive map in humans: spatial navigation and beyond. *Nature neuroscience, 20*(11), 1504.

Feinberg, D. A., Moeller, S., Smith, S. M., Auerbach, E., Ramanna, S., Glasser, M. F., . . . Yacoub, E. (2010). Multiplexed echo planar imaging for sub-second whole brain FMRI and fast diffusion imaging. *PloS one, 5*(12), e15710.

Feinberg, D. A., & Setsompop, K. (2013). Ultra-fast MRI of the human brain with simultaneous multi-slice imaging. *Journal of magnetic resonance, 229*, 90-100.

Filliat, D., & Meyer, J.-A. (2003). Map-based navigation in mobile robots:: I. a review of localization strategies. *Cognitive Systems Research, 4*(4), 243-282.

Glasser, M. F., Sotiropoulos, S. N., Wilson, J. A., Coalson, T. S., Fischl, B., Andersson, J. L., . . . Polimeni, J. R. (2013). The minimal preprocessing pipelines for the Human Connectome Project. *Neuroimage, 80*, 105-124.

Iaria, G., Chen, J. K., Guariglia, C., Ptito, A., & Petrides, M. (2007). Retrosplenial and hippocampal brain regions in human navigation: complementary functional contributions to the formation and use of cognitive maps. *European Journal of Neuroscience, 25*(3), 890-899.

Julian, J. B., Ryan, J., Hamilton, R. H., & Epstein, R. A. (2016). The occipital place area is causally involved in representing environmental boundaries during navigation. *Current Biology, 26*(8), 1104-1109.

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

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

Maguire, E. (2001). The retrosplenial contribution to human navigation: a review of lesion and neuroimaging findings. *Scandinavian journal of psychology, 42*(3), 225-238.

Marchette, S. A., Vass, L. K., Ryan, J., & Epstein, R. A. (2014). Anchoring the neural compass: coding of local spatial reference frames in human medial parietal lobe. *Nature neuroscience, 17*(11), 1598.

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

Persichetti, A. S., & Dilks, D. D. (2018). Dissociable neural systems for recognizing places and navigating through them. *Journal of Neuroscience, 38*(48), 10295-10304.

Persichetti, A. S., & Dilks, D. D. (2019). Distinct representations of spatial and categorical relationships across human scene-selective cortex. *Proceedings of the National Academy of Sciences, 116*(42), 21312-21317.

Sherrill, K. R., Erdem, U. M., Ross, R. S., Brown, T. I., Hasselmo, M. E., & Stern, C. E. (2013). Hippocampus and retrosplenial cortex combine path integration signals for successful navigation. *Journal of Neuroscience, 33*(49), 19304-19313.

Simons, D. J., & Wang, R. F. (1998). Perceiving real-world viewpoint changes. *Psychological Science, 9*(4), 315-320.

Sonkusare, S., Breakspear, M., & Guo, C. (2019). Naturalistic Stimuli in Neuroscience: Critically Acclaimed. *Trends in Cognitive sciences, 23*(8), 699-714.

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological review, 55*(4), 189.

Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E., Yacoub, E., Ugurbil, K., & Consortium, W.-M. H. (2013). The WU-Minn human connectome project: an overview. *Neuroimage, 80*, 62-79.

Westfall, J. (2016). Power ANalysis for GEneral Anova designs (PANGEA, v0.2), from https://jakewestfall.shinyapps.io/pangea/

Wolbers, T., & Büchel, C. (2005). Dissociable retrosplenial and hippocampal contributions to successful formation of survey representations. *Journal of Neuroscience, 25*(13), 3333-3340.

Xu, J., Moeller, S., Auerbach, E. J., Strupp, J., Smith, S. M., Feinberg, D. A., . . . Uğurbil, K. (2013). Evaluation of slice accelerations using multiband echo planar imaging at 3 T. *Neuroimage, 83*, 991-1001.

Figures

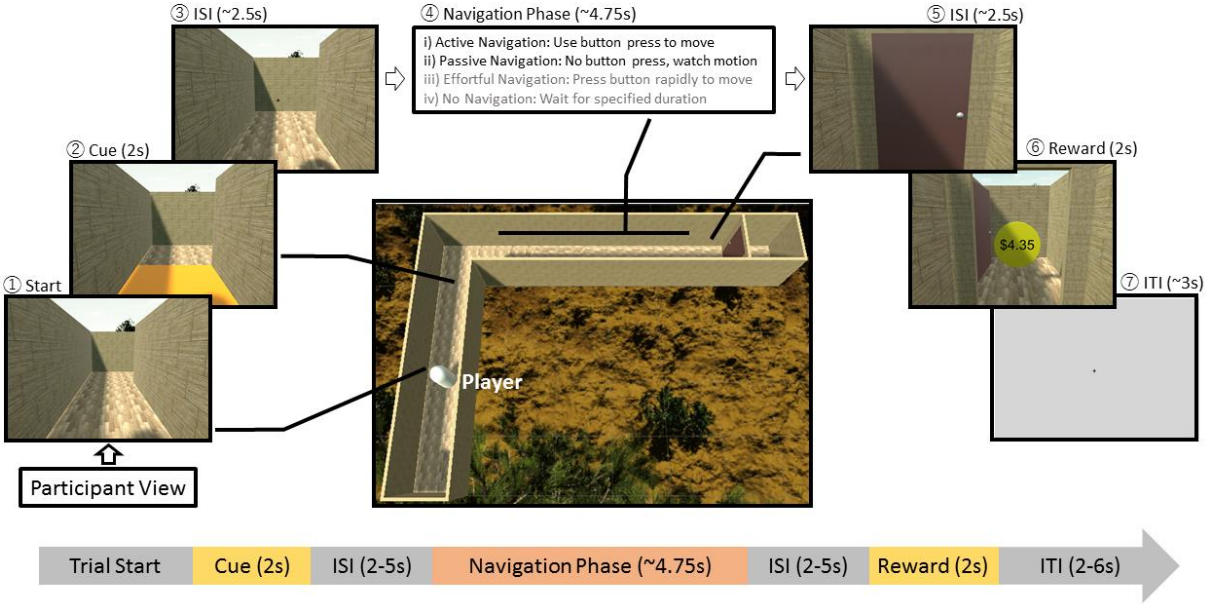

Figure 1. Schematic representation of a single trial of the maze-navigation task. On each trial, participants (labeled as Player) navigates through a maze structure in first-person viewpoint. First, participants make an approach to the cue location, which triggers the floor color to change for 2 seconds. This color indicates the navigation-condition on each trial. Specifically, each trial was associated with one of four navigation conditions: (i) the active navigation condition requires individuals to move using button pressing; (ii) the passive-navigation condition requires individuals to watch navigation without pressing any buttons; (iii) the effortful-navigation condition requires individuals to rapidly press buttons in order to move; and (iv) the no-navigation condition requires individuals to wait for a specified duration without pressing any buttons before being 'teleported' to the goal. Center image shows the top view of one of the four maze structures (single right turn, single left turn, right-then-left turn, left-then-right turn).

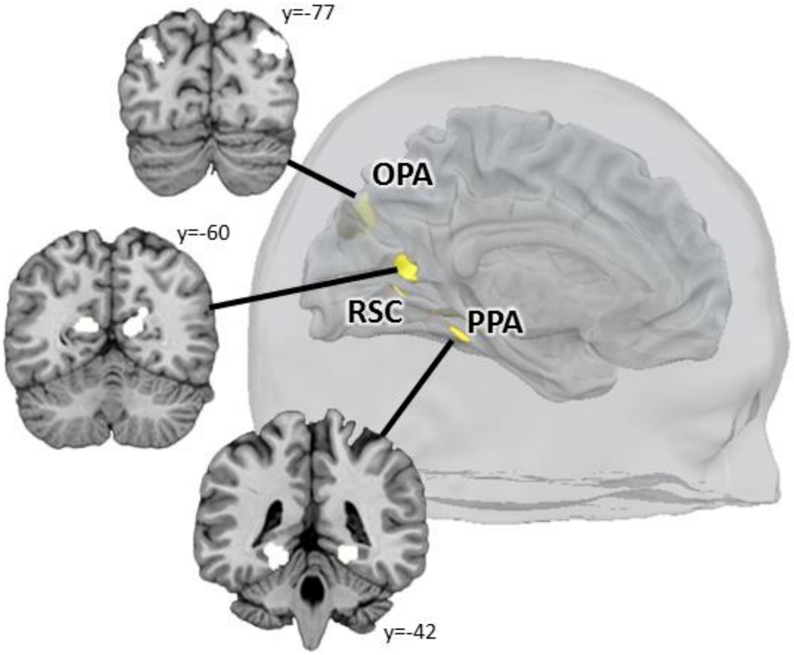

Figure 2. Localization of scene-selective regions (OPA, RSC, and PPA) in a sample participant. Each region was individually localized for each participant using the Scene>Object contrast in the functional localizer task, using a voxel-wise threshold of $p<10^{-6}$.

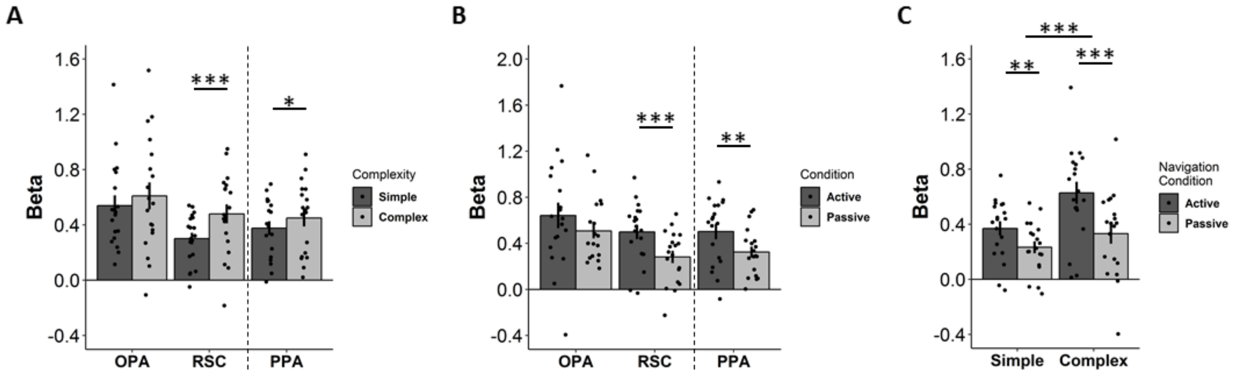

Figure 3. Effect of navigation condition and maze structure in OPA, RSC, and PPA. The three regions showed differential patterns of responses to (A) navigation condition (active vs passive) and (B) maze structure (simple vs complex), with (C) RSC responding greater to complex compared to simple maze structures, especially during the active experience of navigation.