Research Article: New Research | *Cognition and Behavior*

Neural representations of covert attention across saccades: comparing pattern similarity to shifting and holding attention during fixation

https://doi.org/10.1523/ENEURO.0186-20.2021

Cite as: eNeuro 2021; 10.1523/ENEURO.0186-20.2021

Received: 2 May 2020
Revised: 21 January 2021
Accepted: 26 January 2021

---

1. Manuscript Title (50 word maximum)

Neural representations of covert attention across saccades: comparing pattern similarity to shifting and holding attention during fixation

2. Abbreviated Title (50 character maximum)

Neural patterns of covert attention across saccades

3. List all Author Names and Affiliations in order as they would appear in the published article

Xiaoli Zhang (Department of Psychology, The Ohio State University)
Julie D. Golomb (Department of Psychology, The Ohio State University)

4. Author Contributions:

XZ and JDG developed the idea;
XZ and JDG designed the experiment;
XZ collected and analyzed the data;
XZ and JDG wrote the manuscript.

5. Correspondence should be addressed to (include email address)

Xiaoli Zhang (zhang.4734@osu.edu) and Julie D. Golomb (golomb.9@osu.edu)

6. Number of Figures: 13

7. Number of Tables: 6

8. Number of Multimedia: 0

9. Number of words for Abstract: 250

10. Number of words for Significance Statement: 120

11. Number of words for Introduction: 747

12. Number of words for Discussion: 2731

13. Acknowledgements

The authors thank Andrew B. Leber and Jiageng Chen for helpful discussion about the analyses.

14. Conflict of Interest

The authors report no conflict of interest.

15. Funding sources

This research was supported by Alfred P. Sloan Foundation [grant number BR-2014-098] to JDG; National Institutes of Health [grant number R01-EY025648] to JDG; H. Dean and Susan Regis Gibson Research Award to XZ.

Abstract

We can focus visuospatial attention by covertly attending to relevant locations, moving our eyes, or both simultaneously. How does shifting versus holding covert attention during fixation compare with maintaining covert attention across saccades? We acquired human fMRI data during a combined saccade and covert attention task. On Eyes-fixed trials, participants either held attention at the same initial location ("hold attention") or shifted attention to another location midway through the trial ("shift attention"). On Eyes-move trials, participants made a saccade midway through the trial, while maintaining attention in one of two reference frames: The "retinotopic attention" condition involved holding attention at a fixation-relative location but shifting to a different screen-centered location, whereas the "spatiotopic attention" condition involved holding attention on the same screen-centered location but shifting relative to fixation. We localized the brain network sensitive to attention shifts (shift > hold attention), and used multivoxel pattern time-course analyses to investigate the patterns of brain activity for spatiotopic and retinotopic attention across saccades. In the attention shift network, we found transient information about both whether covert shifts were made and whether saccades were executed. Moreover, in this network, both retinotopic and spatiotopic conditions were represented more similarly to shifting than to holding covert attention. An exploratory searchlight analysis revealed additional regions where spatiotopic was relatively more similar to shifting and retinotopic more to holding. Thus, maintaining retinotopic and spatiotopic attention across saccades may involve different types of updating that vary in similarity to covert attention "hold" and "shift" signals across different regions.

Keywords: covert attention shifts, saccades, representational similarity, reference frames, fMRI

Significance Statement:

To our knowledge, this study is the first attempt to directly compare human brain activity patterns of covert attention (to a peripheral spatial location) across saccades and during fixation. We applied fMRI multivoxel pattern time course analyses to capture the dynamic changes of activity patterns, with specific focus on the critical timepoints related to attention shifts and saccades. Our findings indicate that both retinotopic and spatiotopic attention across saccades produce patterns of activation similar to "shifting" attention in the brain, even though both tasks could be interpreted as "holding" attention by the participant. The results offer a novel perspective to understand how the brain processes and updates spatial information under different circumstances to fit the needs of various cognitive tasks.

Introduction

We live in a world with an abundance of visual information, but we have limited visual acuity and cognitive resources. To process visual information across various locations with high sensitivity as needed by daily tasks, we can perform functions like shifting attention allocation covertly or making eye movements. In daily life, covert attention shifts and saccades are often directed to the same to-be-attended location. But we can also covertly attend one location while saccading elsewhere, and the neural mechanisms underlying this case are considerably less explored.

When the eyes are at a stable fixation, covert shifts of attention are often associated with activation in the frontoparietal network (Chica et al., 2013). Specifically, medial superior parietal lobule is activated when covert attention is shifted spatially (Gmeindl et al., 2016; Yantis et al., 2002), between space and feature dimensions (Greenberg et al., 2010), between visual and auditory modalities (Shomstein & Yantis, 2004), and between spatial and nonspatial modalities (Shomstein & Yantis, 2006), suggesting the presence of a general mechanism that mediates shifts of attention.

A number of studies comparing covert attention shifts with overt attention shifts (saccades) further show that these two functions share overlapping brain areas, including intraparietal sulcus (IPS), superior parietal lobule (SPL), and frontal regions like pre-central sulcus/gyrus (Beauchamp et al., 2001; Corbetta et al., 1998; de Haan et al., 2008; Perry & Zeki, 2000). In these neuroimaging studies, a common paradigm is for participants to either shift attention (covert shifts) or make a saccade (overt shifts) between the current fixation point and a target location, with the brain activation in these conditions each contrasted with a baseline condition where no shift happened.

These neuroimaging studies, together with behavioral evidence, suggest a tight coupling between covert spatial attention and eye movements. Covert attentional orientation is an important step preceding saccade execution (Kowler et al., 1995; Peterson et al., 2004). The premotor theory of attention even claims that covert attention simply reflects the central programming of eye movements, just without actual saccade execution (Rizzolatti, Riggio, Dascola, & Umiltá, 1987). However, this theory remains controversial, especially regarding independence between endogenous attention and motor preparation (Smith & Schenk, 2012), and covert spatial attention and saccade target locations can be dissociated in several paradigms, such as anti-saccade tasks (Juan et al., 2004; Smith & Schenk, 2007) and attention in different reference frames, as below.

When attention is allocated to a separate location from the saccade target, the eye movement introduces a discrepancy between retinotopic (eye-centered) and non-retinotopic (e.g. spatiotopic / world-centered) reference frames. Although the spatiotopic reference frame feels more relevant for most behaviors, visual processing starts on our retina in retinotopic coordinates. Behavioral and neural evidence shows that we can allocate attention in both retinotopic and spatiotopic reference frames, though it is debated which is more dominant and whether they differ by brain region (Crespi et al., 2011; Fabius et al., 2016; Fairhall et al., 2017; Golomb et al., 2008; Golomb & Kanwisher, 2012a, 2012b; Melcher & Morrone, 2003; Satel et al., 2012; Shafer-Skelton & Golomb, 2017; Turi & Burr, 2012; Zimmermann et al., 2013).

This ambiguity raises important questions about how our brain processes covert attention across saccades. For example, maintaining covert attention at a stable peripheral real-world location across a saccade (i.e., spatiotopic attention) would be akin to holding attention in spatiotopic coordinates, but shifting attention in retinotopic coordinates. Here, we take a novel approach to understanding the relationship between covert attention and saccades by comparing the neural patterns associated with retinotopic and spatiotopic attention across saccades to holding or shifting covert attention during fixation. We hypothesized that activation patterns should reveal whether maintaining retinotopic or spatiotopic attention is represented more like a stable hold of attention, in functionally relevant regions such as the attention shift network.

We acquired fMRI data during a combined saccade and covert attention task, with four critical conditions. On Eyes-fixed trials, participants either held attention at the same initial peripheral location ("hold attention") or shifted attention to a different location midway through the trial ("shift attention"). On Eyes-move trials, participants made a saccade midway through the trial half of the time, while covertly maintaining either "spatiotopic attention" (hold relative-to-screen, shift relative-to-eyes) or "retinotopic attention" (hold relative-to-eyes, shift relative-to-screen). We used multivoxel pattern time course (MVPTC) analyses to compare whether patterns of brain activity for spatiotopic and retinotopic conditions were more similar to shifting or to holding attention, both in our a priori ROIs, as well as through an exploratory whole-brain searchlight analysis.

Methods

Participants. 12 right-handed subjects participated in the study (7 females, 5 males, mean age 19.08, range 18-25). An additional left-handed subject was also scanned inadvertently, but the data were not included in our analyses. All subjects reported normal or corrected-to-normal vision. They were pre-screened for MRI eligibility, and they gave informed consent. The study protocol was approved by the Ohio State University Biomedical Sciences Institutional Review Board.

Stimuli and task. The paradigm is shown in Figure 1. Eyes-fixed and Eyes-move tasks were done in separate runs.

In the Eyes-fixed task (Figure 1A), subjects fixated their eyes at the fixation cross at the screen center. A letter cue appeared above and below the fixation to indicate the location to be covertly attended (L for left of fixation, R for right of fixation). The stimuli were rapid serial visual presentation (RSVP) streams of random digits (each frame of digits presented for 250ms without gap). Two columns of RSVP streams were located 2.5° to the left and right of the fixation cross, respectively. In each column, the middle stream was the target stream and the upper and lower streams were the flanker streams. Subjects were instructed to attend to the cued side and press the button when they saw a target (the number “5”) in the target stream.

Each trial lasted 8 seconds. The fixation and letter cue alone were presented for 0.75s before the onset of the RSVP streams. On half of the trials, the letter cue changed (e.g., from L to R) midway through the trial (always 3.25s after the onset of the RSVP streams), cueing participants to shift their covert attention to the other side and monitor for the target digit on the new side. Each trial can thus be thought of as containing two task periods, each lasting for 3.25s, separated by a 0.75s gap for the potential shift. (The RSVP streams continued during this potential shift period, but the target number “5” was inhibited.) The task was programmed so that the onset of the first task period was always synced with the scanner pulse (time 0 for each trial). The attended locations of the two periods could either be the same (“Hold-L” and “Hold-R” conditions) or different (“shift-LR” and “shift-RL” conditions), as shown in Figure 2. The four trial types were randomly intermixed in each Eyes-fixed run so that participants could not predict the conditions before each trial.

The RSVP streams were composed of digits; the digit “5” was reserved as the target; other digits were presented randomly in a trial. In the RSVP, for every frame of 250 ms, there was a 1/3 chance that the target “5” would appear on the screen in one of middle (target) streams (when it appeared, it was randomly assigned to one of the target streams, and “5” never appeared in the flanker streams). The target presentation was temporally restricted so that two targets could not appear sequentially within 1s, no matter whether it appeared in the cued or uncued stream.

Stimuli in the Eyes-move task were similar, except that instead of fixating at the screen center, the fixation cross could appear at one of two potential fixation locations at the start of each trial, 2.5° to the left and right of the screen center, and there were three columns of RSVP streams, located at the far left, center, and far right of the screen, each centered 2.5° away from the nearest fixation location (Figure 1B). On half of the trials, the fixation cross remained in the same position for the entire trial (no-saccade trials); on the other half of trials, the fixation cross jumped to the other fixation location halfway through the trial (saccade trials). Subjects were instructed to fixate their eyes on the fixation cross and saccade to the new location whenever it moved.

Each Eyes-move run was subdivided into 4 mini-blocks (8 trials each). Two of these blocks contained the spatiotopic reference frame condition, where subjects were instructed to attend to the central RSVP stream regardless of where their eyes were. This condition was cued at the beginning of the mini-block as “attend screen center”, and the letter cue “C” appeared above and below the current fixation to remind subjects of the attended location. The other two mini-blocks contained the retinotopic reference frame conditions, where subjects were instructed to attend to an RSVP stream defined relative to fixation, i.e., “left of the cross” or “right of the cross”. These conditions were cued as such at the beginning of the mini-block, and with the letters “L” and “R”, respectively, during the trial. The order of these four mini-blocks was randomized in each run. Participants always knew which reference frame condition they were in, but they could not predict either the initial fixation location or whether they would have to make a saccade or not on each trial.

Each trial in the Eyes-move task also lasted 8 seconds. As in the Eyes-fixed task, the scanner pulse was always synced with the onset of the first task period (time 0); the rest of the trial was designed so that the time-course data would be as comparable as possible between Eyes-fixed and Eyes-move tasks. The initial fixation and letter cue alone appeared 1s before the start of the trial (onset of the RSVP streams). The first task period lasted 3s and the second 2.5s, separated by a 1s gap for a potential saccade. (The RSVP streams continued during this potential saccade period, but the target number “5” was inhibited.) There were another 0.5s of blank gap after the second task period before the next trial began.

A summary of all conditions in the Eyes-move task is listed in Figure 2. The conditions were coded based on reference frame, attended location, and fixation location or saccade direction. For example, in spatiotopic blocks, no-saccade trials were coded as SpaC-Rfix (spatiotopic reference frame, attend center stream, fixation on the right cross) and SpaC-Lfix, and saccade trials were coded as SpaC-RLsac (spatiotopic reference frame, attend center stream, saccade from right to left cross) and SpaC-LRsac. In retinotopic blocks, no-saccade trials were coded as RetL-Rfix (retinotopic reference frame, attend stream left of fixation, fixation on the right cross), RetL-Lfix, RetR-Rfix, and RetR-Lfix; however, although our design included both left and right fixation location trials, we aggregated them into RetL-fix and RetR-fix to simplify our analyses. This is because the aggregated conditions did not involve a visual field difference, and any effect coming from pure fixation location difference is beyond the main scope of this study. Retinotopic saccade trials were coded as RetL-RLsac (retinotopic reference frame, attend stream left of fixation, saccade from right to left cross), RetL-LRsac, RetR-RL-sac, and RetR-LRsac. These conditions are all illustrated in Figure 2. In sum, our main MVPA analyses included a total of 10 task conditions. (We also conducted a descriptive univariate analysis with different numbers of conditions; see Results section for details.)

In both Eyes-fixed runs and Eyes-move runs, trial onset times were jittered, with inter-trial intervals (ITIs) of 0s, 2s, and 4s (50%, 35%, and 15% of trials, respectively), in a fast-event related fashion. An additional mini-block (16s) of blank baseline was put in the beginning, middle and end of each run, respectively, where participants were instructed to keep fixated at the fixation cross. Participants completed 4 runs of Eyes-fixed task and 8 runs of Eyes-move task. In addition, they also completed 2 to 4 runs of the standard retinotopic mapping task (see details in the ROI section below).

All stimuli were generated with the Psychtoolbox (Brainard, 1997) in Matlab (MathWorks). Stimuli were displayed with a 3-chips DLP projector onto a screen in the rear of the scanner (resolution 1280×1024 at 60Hz). Participants viewed from a distance of 74cm via a mirror above attached to the head coil.

Eye Tracking. Eye positions were recorded throughout the experiment when the calibration was reliable, using an MRI-compatible Eyelink remote eye-tracker at 500 Hz. Eye position data were used to ensure the participants kept their eyes on the fixation point and made eye movements following the fixation change. When eye position data were not available, the experimenters observed the subjects' eye through the camera and made sure that the participants were making eye movements as intended.

fMRI acquisition. This study was done at the OSU Center for Cognitive and Behavioral Brain Imaging with a Siemens Prisma 3T MRI scanner using a 32-channel phase array receiver head coil. Functional data were acquired using a T2-weighted gradient-echo sequence (TR=2000ms, TE=28ms, flip angle 71°). The slice coverage was oriented about 45° away from the AC-PC plane and placed to prioritize full coverage of occipital and parietal lobes, and then maximize coverage of temporal and frontal lobes (33 slices, 2×2×2mm voxel, 10% gap). We also collected a high-resolution MPRAGE anatomical scan at $1mm^3$ resolution for each participant. Each participant was scanned in one two-hour session.

fMRI preprocessing. The fMRI data were preprocessed with Brain Voyager QX (Brain Innovation). All functional data were corrected for slice acquisition time and head motion, temporally filtered. Runs with abrupt motion greater than 1mm were discarded from later analyses, and the motion correction parameters were logged and input as nuisance variables into the GLM. Spatial smoothing of 4mm FWHM was performed on the preprocessed data for univariate analyses, but not for multivoxel pattern analysis (MVPA). Data of each participant were normalized into Talairach space (Talairach & Tournoux, 1988). We used FreeSurfer to segment the white matter / gray matter boundaries from each participant’s anatomical scan, and imported the images into BrainVoyager for flattening. We extracted each participant’s cortical surface for each hemisphere in Talairach space, and inflated and flattened them into cortical surface space for retinotopic mapping. Other analyses were performed on volume space only.

Regions of Interest. Our analyses focused on two a priori regions of interest (ROIs). These ROIs were our theoretical regions of interest designed to look at attentional representations: bilateral area V4 (considered strongly modulated by attention: McAdams & Maunsell, 2000), and a functionally defined attention shift network (e.g., Yantis et al., 2002).

The attention shift network was functionally defined based on the group-level shift > hold univariate attention contrast in the Eyes-fixed task. For this contrast, we used a whole-brain multi-subject general linear model (GLM) in the Eyes-fixed task with 5 regressors (blank baseline plus the 4 Eyes-fixed conditions) and 6 nuisance regressors from the motion correction processing, with a canonical hemodynamic response function, to calculate beta weights of each condition for each voxel. We then projected the contrasts of shift conditions vs hold conditions onto volume maps. All volume maps were corrected for cluster threshold at $\alpha=0.05$ level, using the BrainVoyager plugin "Cluster-level Statistical Threshold Estimator", after which all significant voxel clusters were picked as the corresponding functional network. The attention shift network is shown in Figure 3 and Table 1. The attention shift network includes inferior parietal lobule (IPL) and temporal gyri, consistent with areas previously found in the literature (Beauchamp et al., 2001; Corbetta et al., 1998; de Haan et al., 2008; Yantis et al., 2002). Due to limited frontal coverage in our scanning protocol, our data only captured more posterior regions.

We used a standard phase-encoded retinotopic mapping localizer (Sereno et al., 1995) to define visual area V4 for each participant. In the retinotopic mapping scans, a rotating wedge with high-contrast radial checkerboard patterns was presented on the screen and flickered at 4 Hz. The 60° wedge stimulus covered eccentricity from 1.6° to 16° and was rotated either clockwise or counter-clockwise for 7 cycles with a period of 24 s per cycle. Participants were instructed to fixate at the center fixation of the screen, and press the button every time when the fixation dot changed color from dark grey to light grey. A pair of clockwise and counterclockwise runs were combined in the analyses. One or two pairs of runs (i.e., 2 to 4 runs) were obtained for each participant. After preprocessing, the brain data were analyzed in custom Matlab code and projected onto the flattened brains as surface maps in Brain Voyager. Bilateral V4 boundaries were defined based on these surface maps. We then used the task > baseline contrast from the Eyes-fixed runs to further constrain the retinotopic ROIs to regions visually activated by this task.

In addition to these a priori regions of interest, we also defined a post-hoc network for exploratory analyses, the "retinotopic-hold" network, based on the cross-task similarity searchlight results (see details below), corrected for cluster threshold in the same way as above. ROI results for this post-hoc network are presented for descriptive purposes only, as the datasets used to define and analyze were not fully independent.

Finally, in the extended data (Table 2-1 and 3-1), we also report results from two additional, comparison ROIs to capture generic visual activation (bilateral area V1) and deactivation (functionally-defined task negative network). Area V1 was defined using the same retinotopic mapping procedure as V4, and the task-negative network was defined based on the group-level baseline > task contrast in the Eyes-fixed task, where task included all 4 Eyes-fixed task conditions.

Multivoxel pattern analyses (MVPA). For all MVPA analyses below, we imported corresponding GLM data to Matlab with BrainVoyager's BVQXtools Matlab toolbox, and all subsequent analyses were done using custom Matlab code.

1) Within-task MVPA (split-half correlation-based analyses)

We first performed MVPA within the Eyes-Fixed and Eyes-Move tasks (e.g., comparing the Eyes-Fixed conditions to each other), using the split-half correlation-based method (Haxby et al., 2001) for each participant and each ROI/network. This split-half procedure is necessary for the within-task analysis to avoid confounds driven by the diagonal cells in the correlation matrix. (The split-half procedure is not necessary for cross-run analyses; see Cross-task pattern similarity analysis in later section.) We obtained GLMs for odd runs and even runs separately for each participant; each GLM had 5 regressors for the Eyes-fixed task (blank baseline plus the 4 Eyes-fixed conditions) and 11 regressors for the Eyes-move task (blank baseline plus the 10 Eyes-move conditions from Figure 2), as well as 6 nuisance regressors from the motion correction processing. For the following analyses, we focused on non-baseline conditions. For each GLM, we normalized the voxel data (beta weights) by subtracting the mean response across all non-baseline conditions from the response of each individual condition, for each voxel. This standard demeaning procedure (Coutanche, 2013) was done within each fold of split-half data. The response patterns (voxel-wise beta weights after de-meaning) for each condition in the even runs were then correlated with the patterns for each condition in the odd runs, generating a correlation matrix for each task. The correlation coefficients were transformed into z-scores using Fisher's r-to-z transform.

We then calculated the following types of information based on the correlation matrix. In the Eyes-fixed task: information about shift execution (holding vs shifting attention), hold attention location (holding left vs holding right), and shift direction (shifting leftward vs shifting rightward). In the Eyes-move task: information about saccade execution (saccade vs no saccade), saccade direction (saccade leftward vs saccade rightward), and reference frame (attend retinotopic task vs attend spatiotopic task). Specifically, we picked out cells in the matrix that reflected the same type of information ("within-category" correlations, e.g., holding attention correlated with holding attention), and cells that reflected the different type of information ("between-category" correlations, e.g., holding attention correlated with shifting attention). The information index was then calculated by subtracting the mean correlation values of "different" cells from those of "the same" cells. A significantly-positive information index value would indicate that there is some amount of information of this type in the ROI.

2) Multivoxel Pattern Time Course (MVPTC) analyses

The first step of analyses described above used regular whole-trial GLMs, which modeled the whole 8 sec (4 TR) trial as a single event. However, since trials contained a potential attention shift or saccade halfway through, the initial analysis might fail to capture some dynamic brain representations. Thus, we also performed timecourse analyses using finite impulse response (FIR) GLM analyses with 10 timepoints, on the same conditions as above. Timepoint zero (TP0) corresponds to the start of the first task period in each trial (i.e., the onset of RSVP stimuli). We fed those FIR GLMs into MVPA analyses (i.e., MVPTC, modified from Chiu et al., 2012). Taking each time point as a separate dataset, we performed similar analyses as above to calculate the information indices. The result figures show all 10 TPs in the FIR, but our statistical analyses focus on three TPs that capture critical time periods in the trial, accounting for BOLD signal lag: TP3 (before the shift/saccade happened), TP4 (capturing the shift/saccade), TP5 (after the shift/saccade). It is also important to clarify that at the behavioral time period corresponding to BOLD signals at TP3, participants did not know yet whether there would be an attentional shift or not (in eyes-fixed task), or a saccade or not (in Eyes-move task), because the trials were intermixed; however, it was predictable that if there would be a shift/saccade, what direction the shift/saccade would be, based on the attention location or the eye location within the first half of a trial.

3) Cross-task pattern similarity analysis

To directly compare the similarity *between* the brain activity patterns of covert attention during Eyes-fixed and Eyes-move tasks, we performed a cross-task pattern similarity analysis for both whole-trial and time-course beta weights. Because the Eyes-fixed and Eyes-move tasks were performed in separate runs, we used GLMs of all runs instead of split-half to increase power; that is, we took Eyes-fixed runs and Eyes-move runs as the two datasets for the correlation analysis. After de-meaning the voxel-wise responses in the same way as above, we calculated the z-scored correlation matrix comparing each condition in the Eyes-fixed task to each saccade condition in the Eyes-move task. We then calculated the pattern similarity between the following four pairings by averaging the z-scored correlation coefficients of corresponding cells in the matrix: retinotopic-to-hold, retinotopic-to-shift, spatiotopic-to-hold, spatiotopic-to-shift. The similarity data were submitted to a 2 (Eyes-move conditions: retinotopic & spatiotopic) by 2 (similarity to Eyes-fixed conditions: similarity-to-hold & similarity-to-shift) ANOVA. In this ANOVA analysis, a main effect of similarity to Eyes-fixed conditions would indicate that both retinotopic and spatiotopic attention (across saccades) are represented more similarly to hold (or shift) attention than shift (or hold); an interaction would indicate relatively greater similarity between retinotopic and holding attention & between spatiotopic and shifting attention (or relatively greater similarity between spatiotopic and holding attention & between retinotopic and shifting attention). To help illustrate the result, we also plot the difference in pattern similarity between spatiotopic-to-shift minus spatiotopic-to-hold correlations, and the difference between retinotopic-to-shift minus retinotopic-to-hold correlations; the difference between these difference scores reflects the interaction term from the ANOVA analysis above.

4) Whole-brain searchlight on cross-task pattern similarity analysis

Finally, we performed MVPA searchlight analyses (Kriegeskorte et al., 2006) to search across the entire slice coverage, for clusters that might show patterns of interest outside our a priori ROIs. The approach is similar to what is described above; instead of taking a priori ROIs, we searched through individual brains iteratively with a “moving” ROI, defined as a sphere of radius 3 voxels. On each iteration, MVPTC analyses were performed as described above on each ROI sphere, and z-scored correlation values were assigned to the center voxel of this ROI sphere to form z-maps for each subject. We generated such searchlight maps for three measures: the difference in similarity between spatiotopic-to-shift and spatiotopic-to-hold correlations, the difference between retinotopic-to-shift and retinotopic-to-hold correlations, and their interaction (i.e. the interaction term in the ANOVA described in the prior section). Specifically, we focused only on TP4, which theoretically captured the timepoint at shift/saccade. To generate these difference maps, we first generated 4 searchlight maps for each individual subject, indexing each pair of correlations: retinotopic-to-hold, spatiotopic-to-hold, retinotopic-to-shift, and spatiotopic-to-shift. We calculated the difference maps by comparing (subtracting) the appropriate similarity maps for each subject accordingly. The resulting searchlight difference and interaction maps for each individual were then spatially smoothed with a 4 mm FWHM kernel to facilitate group analyses. Group t-value maps were constructed using two-tailed t-tests comparing the values for each voxel against zero, correcting for cluster threshold in the same way as above. For the first two difference maps, a positive t-value for a given voxel indicates that spatiotopic/retinotopic attention is represented more similar to shifting than holding attention. For the final interaction map, a positive t-value for a given voxel indicates that retinotopic attention across saccades is represented more similar to holding attention at fixation, and spatiotopic more similar to shifting attention (i.e., the “retinotopic-hold / spatiotopic-shift” pattern); a negative t-value indicates that retinotopic attention across saccades is represented more similar to shifting attention at fixation, and spatiotopic more similar to holding attention (i.e., the “spatiotopic-hold / retinotopic-shift” pattern).

Results

Our main theoretical question of interest is whether maintaining retinotopic or spatiotopic attention across saccades is represented relatively more like holds (or shifts) of attention at fixation. Our primary focus is thus on the cross-task similarity results from our a priori attention-related ROIs (along with an exploratory searchlight analysis). Before presenting these cross-task MVPA results, we first report the behavioral, univariate, and within-task MVPA results to establish the sensitivity of the paradigm and provide context for the cross-task results.

Initial results 1: Behavior

To evaluate participants’ behavioral performance, we defined hits as correctly pressing a button within 1 sec in response to a “5” target at the attended location and false alarms as incorrectly pressing a button when there was no “5” target within 1 sec at the attended location. We calculated the hit rate by dividing the total number of hits in each trial by the total number of targets at the attended location (trials with 0 targets were omitted). We also calculated the false alarm rate by dividing the total number of false alarms in each trial by the total number of frames when there was no target presented in the attended RSVP stream. D-prime was calculated by subtracting z-scored false alarm rates from z-scored hit rates.

Due to a coding mistake for data logging, two subjects did not have reliable behavioral responses logged and were excluded from the analyses of behavioral performance. For the remaining 10 subjects, the mean hit rate in Eyes-fixed task was 66.17% (±5.07% standard deviation), and the mean false alarm rate was 0.52% (±0.14% standard deviation); in Eyes-move task, the mean hit rate was 65.67% (±5.70%) and the mean false alarm rate was 0.50% (±0.18%). These two tasks were designed to be hard to make sure that participants maintained attention on the cued location, so it is reasonable that participants' performance was not at ceiling. The d-prime measurements in both tasks were well above zero, t's$\geq$15.239, p's$\leq$.001, Cohen's d's$\geq$4.819, and there was no significant difference between the two tasks, $t(9)=0.217$, $p=.833$, Cohen's $d=0.069$. In addition, there were no significant differences of d-prime between hold and shift attention in Eyes-fixed task, between saccade and no saccade trials in Eyes-move task, and between spatiotopic and retinotopic attention, all t's$\leq$2.083, p's$\geq$.067, Cohen's d's$\leq$0.659.

Initial results 2: Univariate comparisons

To give a general view of how the brain activity looks like for each condition, Figure 4 plots the percent signal change in the time course as well as the univariate beta weights for our two a priori attention-related ROIs. To better illustrate, we recoded the conditions to plot them according to whether the attended side was ipsilateral/contralateral relative to the ROIs in each hemisphere, and further collapsed across the RL and LR saccade directions in retinotopic saccade trials (that is, only 8 conditions were shown in Eyes-move results). To make it comparable for each condition, we subtracted the percent signal change or beta weights of fixation baseline from all other conditions, in both Eyes-fixed and Eyes-move task. As shown in

Figure 4, there was a separation in the attention shift network between holding and shifting attention around TP4, as well as a clear pattern of contralateral attentional modulation in V4.

Initial results 3: MVPA of shifting vs holding attention (Eyes-fixed)

For the Eyes-fixed task, we examined whether we could decode from the brain patterns information about shift execution (holding vs shifting attention trials), about hold attention location (attending left vs right stream on hold trials), and about shift direction (shift left-right vs shift right-left trials) (Figure 5A). From each of our a priori ROIs/networks, we conducted correlation-based MVPA on the whole-trial GLM beta weights (Figure 5B). We also examined how these three types of information develop over the time course of the trials (MVPTC), by using beta weights from the FIR GLMs (Figure 5C). Table 2 lists t-test statistics for each of these comparisons for the whole-trial betas and critical timepoints TP3, TP4, and TP5, corresponding to the critical behavioral time periods before the shift/saccade happened, around the shift/saccade, and after the shift/saccade was done.

For the information about shift execution (holding vs shifting attention trials), we did not find significant information with the whole trial MVPA analyses in the attention shift network nor in V4. However, recall that trials were 8-seconds long, and the hold and shift trials were designed to be identical for the majority of the trial, except the transient shift occurring midway through the trial. Indeed, when analyzing the time course in the attention shift network, we did find significant information about shift execution at the critical TP4. There was a weak effect at TP5 that did not pass correction for multiple comparisons, and no significant information about shift execution for TP3, before the shift happened. The MVPTC analyses thus successfully captured a transient change in activity pattern around the time when the shifts happened, in the attention shift network (in V4 we found information about shift execution at TP4 that did not pass correction for multiple comparisons). (Note that the attention shift network was defined by the univariate contrast of shift > hold (with the whole-trial betas), so these MVPA results are not completely independent, although a univariate effect alone (linear transform) could not drive a correlation-based MVPA difference. Nonetheless, these MVPTC results are useful as a validity check, and the remaining analyses that we focus on below are fully independent of the ROI definitions.)

For the information about which location was attended on hold trials (holding left vs right stream), we found significant information in the whole-trial MVPA, in both the attention shift network and V4. MVPTC showed that this information was sustained for the duration of the trial and was significant at TP3, TP4 and TP5 in the attention shift network and V4, with the only exception at TP3 in the attention shift network. This is consistent with the behavioral task on these hold trials, in that participants maintained attention in one location throughout the entire trial.

The analogous analysis for the shift attention trials examines information about covert attention shift direction (shift left-right vs shift right-left trials). We did not find significant information in either ROI with whole-trial beta weights. The timecourse analyses may give some insight into why. Interestingly, the MVPTC took a different shape than for the previous analyses; here, instead of peaking at the critical TP4, the information was actually greater at TP3 and TP5 than at TP4 in both ROIs/networks. In V4, the shift direction information was significant at TP3 and TP5 but not TP4. This bimodal pattern also existed in the attention shift network numerically, but all three TPs were significant. It should be noted that in our design, the direction of the shifting was perfectly confounded with the location participants attended to before and after the shift. Thus, the bimodal pattern may reflect a dynamic representation of which location was being attended in the first half of the trial (peaking at TP3), and then after the attention shift in the second half of the trial (peaking at TP5), rather than reflecting information about the shift direction itself.

Initial results 4: MVPA of attention maintained across saccades (Eyes-move)

For the Eyes-move task, we used a similar approach of whole-trial MVPA followed by MVPTC to examine information about saccade execution (saccade vs no-saccade trials), and on saccade trials, about the saccade direction (leftward vs rightward saccade) and reference frame (retinotopic vs spatiotopic attention) (Figure 6, statistics in Table 2).

For the information about saccade execution (saccade vs no-saccade trials), we found significant information in whole-trial MVPA analyses in the attention shift network. When looking at time-course analyses, we found that the information was represented significantly in V4 and the attention shift network at TP4, corresponding to the behavioral time period of saccade execution. In the attention shift network, this information was also significant at both TP3 and TP5. Post hoc t-tests comparing the information indices at TP3/TP5 to TP4 showed that the information at the critical TP4 was significantly greater than at TP3, $t(11)=2.772$, $p=.018$, Cohen's $d=0.800$, but information at TP4 was only numerically larger than at TP5, $t(11)=0.946$, $p=.364$, Cohen's $d=0.273$. It is possible that saccade preparation and saccade execution might have elongated the process and thus blurred the effect temporally in the attention shift network.

For the information about saccade direction (right-left saccade vs left-right saccade), we found weak information that did not pass correction with whole-trial MVPA in V4, but not in the attention shift network. In the MVPTC, the saccade direction information was significant in all three timepoints in V4, and at TP3 and TP5 in the attention shift network. Some of the timecourses appeared to have a similar bimodal shape for information about saccade direction as above for covert attention shift direction, perhaps again driven by information about attended hemisphere over time (Extended data Figure 6-1). Interestingly, although both V4 and the attention shift network represented information on saccade execution and saccade direction information, V4 seems to have more information about saccade direction, whereas the attention shift network had more information about saccade execution.

Finally, we did not find reference frame information (retinotopic-attention vs spatiotopic-attention trials) in whole-trial MVPA analyses in either ROI. In the timecourse analysis, no timepoints were significant in V4 or the attention shift network. Thus, our attentionally-modulated ROIs did not appear to directly differentiate which reference frame participants were maintaining attention in, though as noted above, they contained information about which location was being attended at any given time, and whether saccades were being executed.

Main results: Cross-task similarity analysis of covert attention at fixation and across saccades

The above results demonstrate that brain regions sensitive to attentional modulation (V4 and the attentional shift network) represent information about covert attention shifts and about saccade execution. Now the key question is, how do representations of covert attention during fixation compare to covert attention maintained across saccades? Depending on the reference frame, both spatiotopic and retinotopic attention could be thought of as “hold” or “shift” attention tasks: spatiotopic attention is maintained in the same location relative to the screen, but shifted relative to our eyes, whereas retinotopic attention is the opposite. Is one or both of these tasks represented more similarly to holding attention in some brain regions, and/or more similarly to shifting attention elsewhere in the brain?

To answer these questions, we analyzed the pattern similarity between Eyes-fixed conditions and Eyes-move conditions (Figure 7A). Rather than calculate information indices, in this cross-task MVPA analysis we directly compare the representational similarity scores for each cross-task pair of conditions (i.e., similarity between retinotopic and hold, between spatiotopic and hold, between retinotopic and shift, and between spatiotopic and shift). We also plot the difference scores between spatiotopic-to-shift minus spatiotopic-to-hold correlations, and retinotopic-to-shift minus retinotopic-to-hold correlations. The results of this analysis are shown in Figure 7B & 7C, and statistics from the 2 × 2 ANOVA are reported in Table 3 for each ROI/network at each critical timepoint, as well as for the whole trial data.

In the whole-trial MVPA analysis, there was a significant main effect of similarity-to-shift versus similarity-to-hold in both V4 and the attention shift network, in that the representational similarity scores were generally higher when correlating the Eyes-move conditions with the Eyes-fixed shift attention condition, compared to with the Eyes-fixed hold attention condition. In the MVPTC analysis, this main effect was significant at critical timepoint TP4 in both ROIs, and also at the neighboring timepoints TP3 and TP5 in the attention shift network.

Post-hoc analyses with whole-trial data reveal that in V4, maintaining spatiotopic attention across saccades was represented marginally more similarly to shift than to hold ($t(11)=2.141$, $p=.056$, Cohen’s $d=0.618$), and there was no significant difference between retinotopic-to-shift and retinotopic-to-hold correlations in the whole-trial analysis ($t(11)=0.032$, $p=.975$, Cohen’s $d=0.009$). In the attention shift network, spatiotopic was marginally more similar to shift than to hold ($t(11)=2.056$, $p=.064$, Cohen’s $d=0.594$), and retinotopic was significantly more similar to shift than to hold ($t(11)=3.306$, $p=.007$, Cohen's $d=0.954$). The interaction in the ANOVA was not significant in either ROI in the whole trial analysis. In the MVPTC, post-hoc analyses at the critical timepoint TP4 showed that both retinotopic and spatiotopic attention across saccades were significantly more similar to shift than to hold in both V4 and the attention shift network (t's$\geq$2.387, p's$\leq$.036, Cohen's d's$\geq$0.689). The similarity-to-shift effect also seemed to be numerically greater for the spatiotopic compared to retinotopic attention condition in V4, but again this interaction was not significant, nor was it significant in the attention shift network.

Exploratory results: Whole-brain cross-task similarity searchlight

The above results suggest that both retinotopic and spatiotopic attention across saccades are represented more like shifts than holds of attention at fixation, with no significant interaction in our a priori attention ROIs indicating that one reference frame is represented more strongly than the other. As an exploratory analysis, we next asked: are there other areas in the brain that might show differential similarity patterns? We performed a searchlight analysis for a significant interaction effect at the critical time point TP4, as described in Methods.

In Figure 8A we first show the difference score searchlight maps between spatiotopic-to-shift vs spatiotopic-to-hold, and retinotopic-to-shift vs retinotopic-to-hold. These difference score maps revealed that throughout the brain, both retinotopic and spatiotopic attention across saccades are widely represented as more similar to shifting attention than holding attention, consistent with our ROI findings.

Critically, the interaction map (Figure 8B) allowed us to extract potential regions that significantly differentiate retinotopic and spatiotopic representations via one of two interaction patterns: (1) retinotopic relatively more similar to hold, and spatiotopic relatively more similar to shift; or (2) spatiotopic relatively more similar to hold, and retinotopic relatively more similar to shift. The searchlight revealed four clusters (Figure 8B and Table 4), all with the retinotopic-hold/spatiotopic-shift pattern. The clusters were located in ventral areas and superior parietal regions bilaterally, which were in later visual hierarchy in both ventral and dorsal pathways. No regions with the spatiotopic-hold/retinotopic-shift pattern survived the cluster threshold correction.

For illustration purposes, we plot the cross-task similarities for the regions identified in the searchlight (Figure 8C; plots for separate clusters in Extended data Figure 8-1). Note that this analysis is circular; we show the interaction patterns here for descriptive purposes only. The interaction in these regions seems to be primarily driven by the spatiotopic comparisons, particularly the high similarity between spatiotopic and shifting attention.

Discussion

In summary, we found that both spatiotopic and retinotopic attention across saccades were represented more similarly to shifting compared to holding attention at fixation, especially in the attention shift network. Our a priori attention ROIs did not reveal a significant interaction between retinotopic and spatiotopic similarity, but our exploratory searchlight analysis revealed some brain regions where maintaining spatiotopic attention was represented more similarly to shifting attention and maintaining retinotopic attention was relatively more similar to holding attention (“retinotopic-hold / spatiotopic-shift” regions), with no brain regions displaying the opposite pattern.

In addition to these primary results, we were able to uncover several other signatures of covert attention during fixation and across saccades from the multivoxel activation patterns in various brain regions. First, pattern similarity results from within the Eyes-fixed task support the validity of our design and analyses. In the visual and attention shift areas we could decode which location the participants were holding attention at, even dynamically in the time course, consistent with existing findings that attention modulates the activity in visual areas (Desimone & Duncan, 1995) and pattern activities in shift-related areas can be used to decode attention in the left vs right hemifield (Gmeindl et al., 2016). In the Eyes-move task we could similarly decode which hemifield was being covertly attended both before and after the saccade (Extended data Figure 6-1). We could also reliably decode from the Eyes-fixed task whether a covert attention shift was executed in the middle of the trials, specifically at the critical timepoint TP4 which corresponds to the transient shift, consistent with time-course decoding results about shift execution with SVM in Chiu, Esterman, Gmeindl, & Yantis, 2012. In the Eyes-move task, information about saccade execution emerged at TP4 in V4, and at all of TP3, TP4, and TP5 in the attention shift network. Below we discuss how our study contributes to the existing literature and informs our understanding of the mechanisms of covert attention across saccades.

Representational patterns for covert attention across saccades

How spatial attention is maintained/updated in particular reference frames across saccades has been an open question in the literature, and it is actively debated with various paradigms whether one reference frame is more native or dominant – and thus requires less updating across saccades – than the other (Crespi et al., 2011; Fabius et al., 2016; Fairhall et al., 2017; Golomb et al., 2008; Golomb & Kanwisher, 2012a, 2012b; Melcher & Morrone, 2003; Satel et al., 2012; Shafer-Skelton & Golomb, 2017; Turi & Burr, 2012; Zimmermann et al., 2013). In the case of spatial attention, it has been argued that attention pointers proactively remap to compensate for saccades and maintain spatiotopic attention (Cavanagh et al., 2010; Marino & Mazer, 2018; Rolfs et al., 2011), but also that attention might linger in retinotopic coordinates even after a saccade (Golomb, 2019; Golomb et al., 2008, 2010; Jonikaitis et al., 2012). More generally, spatiotopic remapping signals have been found in several brain regions, including monkeys' lateral intraparietal area (LIP) (Duhamel et al., 1992), superior colliculus (SC) (Walker et al., 1995), frontal eye field (FEF) (Umeno & Goldberg, 1997), and striate and extrastriate cortex (Nakamura & Colby, 2002), and human visual and parietal cortex (Merriam et al., 2003, 2007). Higher-level visual and parietal areas in particular have also been a focus of much debate over dominant reference frames for neuronal receptive fields (Duhamel et al., 1997; Snyder et al., 1998), fMRI adaptation (Baltaretu et al., 2018; Fairhall et al., 2017; McKyton & Zohary, 2007; Zimmermann et al., 2016), functional organization (Crespi et al., 2011; d'Avossa et al., 2007; Golomb & Kanwisher, 2012b; Ward et al., 2010), and attentional modulation (Golomb et al., 2010; Rawley & Constantinidis, 2010).

In the current study, we approached this question from a different angle. As introduced earlier, eye movements distinguish the two reference frames in a way that maintaining retinotopic attention can be considered as "holding" a location relative to the eyes and "shifting" relative to the world, and maintaining spatiotopic attention can be considered as "shifting" relative to the eyes and "holding" relative to the world. To our knowledge, the current study is the first attempt to directly compare the brain activity patterns of covert attention maintained/updated in the periphery across saccades and during fixation. We found that in the pre-defined attention shift network, maintaining both retinotopic and spatiotopic attention across saccades evoked more similar representational patterns to covertly shifting attention than to covertly holding attention at fixation. Perhaps this is not surprising, given that both retinotopic and spatiotopic trials involved an eye movement, which is expected to engage attentional shifts as discussed below. In that sense, it is less notable that both retinotopic and spatiotopic resembled shifts more than holds per se; but the lack of a *relative* difference in representational similarity is intriguing. If attention were represented more natively in one reference frame, we may have predicted the other condition to show relatively more similarity to the shift condition. Our exploratory searchlight analysis did reveal some regions where maintaining spatiotopic attention across saccades was relatively more similar to shifting attention and retinotopic relatively more to holding, but no regions with the opposite pattern.

Why were both retinotopic and spatiotopic attention represented like covert attention shifts?

Why did saccade trials of both reference frames have greater representational similarity to the covert shift attention trials than hold attention trials? We suggest that the answer may be related to our within-task similarity analyses finding that information about both covert attention shifts (in Eyes-Fixed runs) and saccade execution (in Eyes-Move runs) could be decoded from our attention shift network. As mentioned in the introduction, overt and covert attention have been found to involve overlapping brain areas (Beauchamp et al., 2001; Corbetta et al., 1998; de Haan et al., 2008; Nobre et al., 2000; Perry & Zeki, 2000). Our study differed from these studies in that the paradigm used in these previous studies typically involved overt and covert attention shifts aiming at the same target. In our design, we tried to disentangle the saccade execution from the allocation of top-down task-directed attention, by using top-down covert retinotopic and spatiotopic tasks. There are several possible interpretations of this overlap between representations of saccades and covert attention shifts in our task that could account for why the saccade and no-saccade trials may have produced differentiable activation patterns in the attention shift network – and why in these areas, saccade trials of both reference frames may have had greater representational similarity to the covert shift-attention trials in the cross-task similarity analysis.

One reason could be that covert shifting of attention is directly involved in making a saccade; i.e., the execution of the saccade required a pre-saccadic shift of attention towards the saccade target, and this initial covert shift was what was driving the representational similarity to the covert shift-attention trials. It has been widely shown that shifts of covert attention precede saccade execution (Godijn & Pratt, 2002; Peterson et al., 2004), and pre-saccadic attention is considered critical for determining the saccade endpoints to execute accurate saccades and enhancing perceptual representations of the saccade target (Gersch et al., 2004; Zhao et al., 2012). Even when the task is designed as attending to peripheral locations other than the saccade target, there is evidence that attention is still pre-saccadically shifted to the saccade target (Kowler et al., 1995). In our experiment, the information about saccade vs no saccade in the attention shift network emerged fairly early (around TP3), which could be related to the preparation stage (pre-saccadic shift stage) before the saccade was executed, potentially providing indirect support for this account.

Another potential account is that the Eyes-move task involved a covert shift of attention not related to execution of the saccade per se, but due to perisaccadic updating or remapping of the peripheral focus of attention, on both retinotopic and spatiotopic saccade trials. Previous studies involving spatiotopic remapping have found anticipatory remapping signals in the lateral intraparietal sulcus in monkeys (Duhamel et al., 1992), which could overlap with our parietal attention shift regions in humans. As described earlier, maintaining retinotopic attention can be seen as shifting attention relative to the screen/world, and maintaining spatiotopic attention can be seen as shifting attention relative to the eyes. It is possible that both types of attention in our task involved some updating process across saccades that engaged an attentional shift signal in this brain region, which would be consistent with our cross-task correlation results that both spatiotopic and retinotopic attention were more similar to shifting compared to holding attention in the attention shift network.

A third possibility could be that our Eyes-move task may have triggered a more generic temporary disengaging/reengaging of top-down attention; i.e. a transient change or shift of attention on saccade trials that might have occurred independently of saccade planning, executing, or remapping processes. For example, although our task and instructions were designed to encourage continuous attention, we cannot rule out the possibility that participants may have approached the task as a serial attention task (attend the relevant stream, then disengage to execute saccade, then reengage again on the relevant stream), instead of attending continuously on the relevant stream. Or the abrupt onset of the saccade cue might have captured attention and caused an involuntary shift of attention away from the to-be-attended location. In cases like these examples, a transient shift in attention may have evoked representationally similar patterns of activity in this region on saccade trials to the goal-directed shifts of covert attention on fixation trials, without being directly related to the saccade itself. We found that both maintaining retinotopic and spatiotopic attention are represented as more similar to shifting than holding attention widely in dorsal and ventral areas (Figure 8A), possibly revealing this generic representation of dynamic change. However, it is unlikely that this scenario could have accounted for our full pattern of results, particularly the searchlight findings of the interaction.

Why didn't we see greater differences between retinotopic and spatiotopic representations?

In general, we found less of a difference between retinotopic and spatiotopic conditions than what might have been expected. In analyses directly comparing the two reference frames, we did not reveal any representational difference between retinotopic and spatiotopic conditions in the whole-trial MVPA in the attention shift network or other ROIs. In the MVPTC analyses, significant information about retinotopic vs spatiotopic attention was only found in V1 at TP4, but not in other pre-defined ROIs/networks or timepoints (Extended data Table 2-1). We further probed for retinotopic vs spatiotopic differences with whole-brain MVPTC searchlight (Extended data Figure 6-2) and a whole-brain univariate contrast (Extended data Figure 6-3), but only small scattered regions were found outside of our a priori ROIs.

The behavioral performance confirms that participants were allocating attention properly, but why didn't we find greater differences in retinotopic vs spatiotopic patterns in our attention-related ROIs? One important consideration is that our task was designed to equate visual input across these two conditions. Both conditions contained constant, dynamic stimulation (RSVP streams) in the same three locations; the only difference was which of the streams, depending on which reference frame, was *attended* at any moment in time. This design is in contrast to a design commonly used in prior studies probing other aspects of reference frames across saccades, where only one stimulus is presented at a time, and retinotopic and spatiotopic conditions differ in terms of both stimulus-driven visual input and attentional locus (e.g., Baltaretu et al., 2018; Crespi et al., 2011; d'Avossa et al., 2007; Fairhall et al., 2017; Gardner et al., 2008; Golomb & Kanwisher, 2012; McKyton & Zohary, 2007; Pertzov et al., 2011; Rawley & Constantinidis, 2010; Zimmermann et al., 2016).

Moreover, our analysis was designed to look for representational signatures associated with attending in a retinotopic or spatiotopic reference frame across saccades (i.e., how shift- or hold-like they were); not to ask whether we could decode which particular retinotopic or spatiotopic locations were being attended. Early visual areas are known to be retinotopically organized (Crespi et al., 2011; d'Avossa et al., 2007; Gardner et al., 2008; Golomb & Kanwisher, 2012b; Merriam et al., 2013; Sereno et al., 1995), and we would expect that at least in these areas, attending to a particular retinotopic location across a saccade would look more similar to holding covert attention at that same retinotopic location during fixation than to shifting attention to a different retinotopic location (i.e., the brain activity pattern of RetL-RLsac would be more similar to Hold-L compared to Shift-LR, for example). Indeed, we could decode which hemifield(s) were attended on saccade trials (Extended data Figure 6-1), but this was not the goal of our study. Instead, the primary goal of the current study was to ask more broadly, whether the neural processes associated with maintaining attention in retinotopic (or spatiotopic) coordinates across saccades evoked more similar representational patterns to holding compared to shifting covert attention (at fixation). Thus, our analysis included correlations of conditions with both hemifields (e.g., similarity between retinotopic and hold attention includes correlations between RetL (with both RL and LR saccades) vs Hold-L, RetL vs Hold-R, RetR vs Hold-L, and RetR vs Hold-R; same for other cross-task correlations; see Extended data Figure 7-1 for a more detailed comparison between matching and not matching hemifields). This likely explains why we did not find a Retinotopic-Hold / Spatiotopic-Shift effect with the cross-task similarity searchlight analysis in early visual areas.

Instead, our cross-task pattern similarity analysis was better suited to reflect potential connections between the representations of covert attention across saccades and during fixations, independent of potential confounds from visual stimulation and hemifield-based attentional effects. Thus, it is telling that our pre-defined ROIs – particularly the attention shift network – did not show a difference in representational similarity between the retinotopic and spatiotopic reference frames in the cross-task similarity analysis, such that both were more representationally similar to shifting attention; but the exploratory searchlight analysis revealed some potential regions where maintaining spatiotopic attention was relatively more shift-like than retinotopic attention but not vice versa. This asymmetry may reflect the idea that retinotopic attention is the more “native” coordinate system for spatial attention (Golomb et al., 2008) and suggest potential regions for differentiating retinotopic and spatiotopic attention across saccades, though it is interesting that neither this pattern nor the opposite pattern was found within the attention shift network itself.

Interactions across brain regions

Our findings suggest that maintaining spatiotopic and retinotopic attention across saccades may involve different types of updating that might be represented with “hold” and “shift” signals combined across different sets of regions. Some regions might be involved in both reference frames in a similar way (e.g., the attention shift network), and some other regions might use “shift” signals to further differentiate these two updating processes. That these other areas include bilateral anterior ventral areas and superior parietal regions, located in later visual hierarchy in both ventral and dorsal pathways, may hold further clues for understanding this complex process.

Our results support a close link between the neural mechanisms associated with covert attention shifts during fixation and maintaining retinotopic/spatiotopic attention across saccades in V4 and the attention shift network. In comparing the relative amounts and types of information present in the attention shift network versus area V4 patterns, we found an intriguing parallel; the attention shift network had relatively more information about the execution of covert attention shifts and saccades, while V4 had more information about the location of covert attention and the direction of saccades. This pattern aligns with the general understanding that the attention shift network is more involved in the *execution* of shifting spatial attention, and V4 in the *modulation* of spatial attention (Yantis et al., 2002). Outside the domain of perisaccadic processing, previous literature has shown that the attention shift network is associated with broad, domain-independent brain activity for transient shifts of attention (Chica et al., 2013; Gmeindl et al., 2016; Greenberg et al., 2010; Shomstein & Yantis, 2004, 2006; Yantis et al., 2002). Our findings comparing covert attention shifts with attention updating across saccades further indicate that the brain activity patterns associated with covert attention shifts may be widely and reliably involved in various domains, contexts, and tasks.

In summary, coordination between different brain networks/regions may support more flexible updating of attention across saccades in different contexts, raising interesting follow-up questions regarding how and when this process might be achieved mechanistically, and how it is related to behavior, development, and clinical implications.

References

Baltaretu, B. R., Dunkley, B. T., Monaco, S., Chen, Y., & Crawford, J. D. (2018). Transsaccadic feature interactions in multiple reference frames: An fMRIa study. *BioRxiv*, 413815. https://doi.org/10.1101/413815

Beauchamp, M. S., Petit, L., Ellmore, T. M., Ingeholm, J., & Haxby, J. V. (2001). A Parametric fMRI Study of Overt and Covert Shifts of Visuospatial Attention. *NeuroImage*, *14*(2), 310–321. https://doi.org/10.1006/nimg.2001.0788

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436. https://doi.org/10.1163/156856897X00357

Cavanagh, P., Hunt, A. R., Afraz, A., & Rolfs, M. (2010). Visual stability based on remapping of attention pointers. *Trends in Cognitive Sciences*, *14*(4), 147–153. https://doi.org/10.1016/j.tics.2010.01.007

Chica, A. B., Bartolomeo, P., & Lupiáñez, J. (2013). Two cognitive and neural systems for endogenous and exogenous spatial attention. *Behavioural Brain Research*, *237*, 107–123. https://doi.org/10.1016/j.bbr.2012.09.027

Chiu, Y.-C., Esterman, M. S., Gmeindl, L., & Yantis, S. (2012). Tracking cognitive fluctuations with multivoxel pattern time course (MVPTC) analysis. *Neuropsychologia*, *50*(4), 479–486. https://doi.org/10.1016/j.neuropsychologia.2011.07.007

Corbetta, M., Akbudak, E., Conturo, T. E., Snyder, A. Z., Ollinger, J. M., Drury, H. A., Linenweber, M. R., Petersen, S. E., Raichle, M. E., Van Essen, D. C., & Shulman, G. L. (1998). A Common Network of Functional Areas for Attention and Eye Movements. *Neuron*, *21*(4), 761–773. https://doi.org/10.1016/S0896-6273(00)80593-0

Coutanche, M. N. (2013). Distinguishing multi-voxel patterns and mean activation: Why, how, and what does it tell us? *Cognitive, Affective, & Behavioral Neuroscience*, *13*(3), 667–673. https://doi.org/10.3758/s13415-013-0186-2

Crespi, S., Biagi, L., d'Avossa, G., Burr, D. C., Tosetti, M., & Morrone, M. C. (2011). Spatiotopic Coding of BOLD Signal in Human Visual Cortex Depends on Spatial Attention. *PLOS ONE*, *6*(7), e21661. https://doi.org/10.1371/journal.pone.0021661

d'Avossa, G., Tosetti, M., Crespi, S., Biagi, L., Burr, D. C., & Morrone, M. C. (2007). Spatiotopic selectivity of BOLD responses to visual motion in human area MT. *Nature Neuroscience*, *10*(2), 249–255. https://doi.org/10.1038/nn1824

de Haan, B., Morgan, P. S., & Rorden, C. (2008). Covert orienting of attention and overt eye movements activate identical brain regions. *Brain Research*, *1204*, 102–111. https://doi.org/10.1016/j.brainres.2008.01.105

Desimone, R., & Duncan, J. (1995). Neural Mechanisms of Selective Visual Attention. *Annual Review of Neuroscience*, *18*(1), 193–222. https://doi.org/10.1146/annurev.ne.18.030195.001205

Duhamel, J.-R., Bremmer, F., Ben Hamed, S., & Graf, W. (1997). Spatial invariance of visual receptive fields in parietal cortex neurons. *Nature*, *389*(6653), 845–848. https://doi.org/10.1038/39865

Duhamel, J.-R., Colby, C. L., & Goldberg, M. E. (1992). The Updating of the Representation of Visual Space in Parietal Cortex by Intended Eye Movements. *Science*, *255*(5040), 90.

Fabius, J. H., Fracasso, A., & Van der Stigchel, S. (2016). Spatiotopic updating facilitates perception immediately after saccades. *Scientific Reports*, *6*, 34488. https://doi.org/10.1038/srep34488

Fairhall, S. L., Schwarzbach, J., Lingnau, A., Van Koningsbruggen, M. G., & Melcher, D. (2017). Spatiotopic updating across saccades revealed by spatially-specific fMRI adaptation. *NeuroImage*, *147*, 339–345. https://doi.org/10.1016/j.neuroimage.2016.11.071

Gardner, J. L., Merriam, E. P., Movshon, J. A., & Heeger, D. J. (2008). Maps of Visual Space in Human Occipital Cortex Are Retinotopic, Not Spatiotopic. *Journal of Neuroscience*, *28*(15), 3988–3999. https://doi.org/10.1523/JNEUROSCI.5476-07.2008

Gersch, T. M., Kowler, E., & Dosher, B. (2004). Dynamic allocation of visual attention during the execution of sequences of saccades. *Vision Research*, *44*(12), 1469–1483. https://doi.org/10.1016/j.visres.2003.12.014

Gmeindl, L., Chiu, Y.-C., Esterman, M. S., Greenberg, A. S., Courtney, S. M., & Yantis, S. (2016). Tracking the will to attend: Cortical activity indexes self-generated, voluntary shifts of attention. *Attention, Perception, & Psychophysics*, *78*(7), 2176–2184. https://doi.org/10.3758/s13414-016-1159-7

Godijn, R., & Pratt, J. (2002). Endogenous saccades are preceded by shifts of visual attention: Evidence from cross-saccadic priming effects. *Acta Psychologica*, *110*(1), 83–102. https://doi.org/10.1016/S0001-6918(01)00071-3

Golomb, J. D. (2019). Remapping locations and features across saccades: A dual-spotlight theory of attentional updating. *Current Opinion in Psychology*, *29*, 211–218. https://doi.org/10.1016/j.copsyc.2019.03.018

Golomb, J. D., Chun, M. M., & Mazer, J. A. (2008). The Native Coordinate System of Spatial Attention Is Retinotopic. *Journal of Neuroscience*, *28*(42), 10654–10662. https://doi.org/10.1523/JNEUROSCI.2525-08.2008

Golomb, J. D., & Kanwisher, N. (2012a). Retinotopic memory is more precise than spatiotopic memory. *Proceedings of the National Academy of Sciences*, *109*(5), 1796–1801. https://doi.org/10.1073/pnas.1113168109

Golomb, J. D., & Kanwisher, N. (2012b). Higher Level Visual Cortex Represents Retinotopic, Not Spatiotopic, Object Location. *Cerebral Cortex*, *22*(12), 2794–2810. https://doi.org/10.1093/cercor/bhr357

Golomb, J. D., Nguyen-Phuc, A. Y., Mazer, J. A., McCarthy, G., & Chun, M. M. (2010). Attentional Facilitation throughout Human Visual Cortex Lingers in Retinotopic Coordinates after Eye Movements. *Journal of Neuroscience*, *30*(31), 10493–10506. https://doi.org/10.1523/JNEUROSCI.1546-10.2010

Greenberg, A. S., Esterman, M., Wilson, D., Serences, J. T., & Yantis, S. (2010). Control of Spatial and Feature-Based Attention in Frontoparietal Cortex. *Journal of Neuroscience*, *30*(43), 14330–14339. https://doi.org/10.1523/JNEUROSCI.4248-09.2010

Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and Overlapping Representations of Faces and Objects in Ventral Temporal Cortex. *Science*, *293*(5539), 2425–2430. https://doi.org/10.1126/science.1063736

Jonikaitis, D., Szinte, M., Rolfs, M., & Cavanagh, P. (2012). Allocation of attention across saccades. *Journal of Neurophysiology*, *109*(5), 1425–1434. https://doi.org/10.1152/jn.00656.2012

Juan, C.-H., Shorter-Jacobi, S. M., & Schall, J. D. (2004). Dissociation of spatial attention and saccade preparation. *Proceedings of the National Academy of Sciences*, *101*(43), 15541–15544. https://doi.org/10.1073/pnas.0403507101

Kowler, E., Anderson, E., Dosher, B., & Blaser, E. (1995). The role of attention in the programming of saccades. *Vision Research*, *35*(13), 1897–1916. https://doi.org/10.1016/0042-6989(94)00279-U

Kriegeskorte, N., Goebel, R., & Bandettini, P. (2006). Information-based functional brain mapping. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(10), 3863–3868. https://doi.org/10.1073/pnas.0600244103

Marino, A. C., & Mazer, J. A. (2018). Saccades Trigger Predictive Updating of Attentional Topography in Area V4. *Neuron*, *98*(2), 429-438.e4. https://doi.org/10.1016/j.neuron.2018.03.020

McAdams, C. J., & Maunsell, J. H. R. (2000). Attention to Both Space and Feature Modulates Neuronal Responses in Macaque Area V4. *Journal of Neurophysiology*, *83*(3), 1751–1755.

McKyton, A., & Zohary, E. (2007). Beyond Retinotopic Mapping: The Spatial Representation of Objects in the Human Lateral Occipital Complex. *Cerebral Cortex*, *17*(5), 1164–1172. https://doi.org/10.1093/cercor/bhl027

Melcher, D., & Morrone, M. C. (2003). Spatiotopic temporal integration of visual motion across saccadic eye movements. *Nature Neuroscience*, *6*(8), 877–881. https://doi.org/10.1038/nn1098

Merriam, E. P., Gardner, J. L., Movshon, J. A., & Heeger, D. J. (2013). Modulation of Visual Responses by Gaze Direction in Human Visual Cortex. *Journal of Neuroscience*, *33*(24), 9879–9889. https://doi.org/10.1523/JNEUROSCI.0500-12.2013

Merriam, E. P., Genovese, C. R., & Colby, C. L. (2003). Spatial Updating in Human Parietal Cortex. *Neuron*, *39*(2), 361–373. https://doi.org/10.1016/S0896-6273(03)00393-3

Merriam, E. P., Genovese, C. R., & Colby, C. L. (2007). Remapping in Human Visual Cortex. *Journal of Neurophysiology*, *97*(2), 1738–1755. https://doi.org/10.1152/jn.00189.2006

Nakamura, K., & Colby, C. L. (2002). Updating of the visual representation in monkey striate and extrastriate cortex during saccades. *Proceedings of the National Academy of Sciences*, *99*(6), 4026–4031. https://doi.org/10.1073/pnas.052379899

Nobre, A. C., Gitelman, D. R., Dias, E. C., & Mesulam, M. M. (2000). Covert Visual Spatial Orienting and Saccades: Overlapping Neural Systems. *NeuroImage*, *11*(3), 210–216. https://doi.org/10.1006/nimg.2000.0539

Perry, R. J., & Zeki, S. (2000). The neurology of saccades and covert shifts in spatial attention. An event-related fMRI study. *Brain*, *123*(11), 2273–2288. https://doi.org/10.1093/brain/123.11.2273

Pertzov, Y., Avidan, G., & Zohary, E. (2011). Multiple Reference Frames for Saccadic Planning in the Human Parietal Cortex. *Journal of Neuroscience*, *31*(3), 1059–1068. https://doi.org/10.1523/JNEUROSCI.3721-10.2011

Peterson, M. S., Kramer, A. F., & Irwin, D. E. (2004). Covert shifts of attention precede involuntary eye movements. *Perception & Psychophysics*, *66*(3), 398–405. https://doi.org/10.3758/BF03194888

Rawley, J. B., & Constantinidis, C. (2010). Effects of task and coordinate frame of attention in area 7a of the primate posterior parietal cortex. *Journal of Vision*, *10*(1), 12–12. https://doi.org/10.1167/10.1.12

Rizzolatti, G., Riggio, L., Dascola, I., & Umiltá, C. (1987). Reorienting attention across the horizontal and vertical meridians: Evidence in favor of a premotor theory of attention. *Neuropsychologia*, *25*(1, Part 1), 31–40. https://doi.org/10.1016/0028-3932(87)90041-8

Rolfs, M., Jonikaitis, D., Deubel, H., & Cavanagh, P. (2011). Predictive remapping of attention across eye movements. *Nature Neuroscience*, *14*(2), 252–256. https://doi.org/10.1038/nn.2711

Satel, J., Wang, Z., Hilchey, M. D., & Klein, R. M. (2012). Examining the dissociation of retinotopic and spatiotopic inhibition of return with event-related potentials. *Neuroscience Letters*, *524*(1), 40–44. https://doi.org/10.1016/j.neulet.2012.07.003

Sereno, M. I., Dale, A. M., Reppas, J. B., Kwong, K. K., & al, et. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science*, *268*(5212), 889.

Shafer-Skelton, A., & Golomb, J. D. (2017). Memory for retinotopic locations is more accurate than memory for spatiotopic locations, even for visually guided reaching. *Psychonomic Bulletin & Review*, 1–11. https://doi.org/10.3758/s13423-017-1401-x

Shomstein, S., & Yantis, S. (2004). Control of Attention Shifts between Vision and Audition in Human Cortex. *Journal of Neuroscience*, *24*(47), 10702–10706. https://doi.org/10.1523/JNEUROSCI.2939-04.2004

Shomstein, S., & Yantis, S. (2006). Parietal Cortex Mediates Voluntary Control of Spatial and Nonspatial Auditory Attention. *Journal of Neuroscience*, *26*(2), 435–439. https://doi.org/10.1523/JNEUROSCI.4408-05.2006

Smith, D. T., & Schenk, T. (2007). Enhanced probe discrimination at the location of a colour singleton. *Experimental Brain Research*, *181*(2), 367–375. https://doi.org/10.1007/s00221-007-0937-9

Smith, D. T., & Schenk, T. (2012). The Premotor theory of attention: Time to move on? *Neuropsychologia*, *50*(6), 1104–1114. https://doi.org/10.1016/j.neuropsychologia.2012.01.025

Snyder, L. H., Grieve, K. L., Brotchie, P., & Andersen, R. A. (1998). Separate body- and world-referenced representations of visual space in parietal cortex. *Nature*, *394*(6696), 887–891. https://doi.org/10.1038/29777

Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain. 3-Dimensional proportional system: An approach to cerebral imaging*. Thieme.

Turi, M., & Burr, D. (2012). Spatiotopic perceptual maps in humans: Evidence from motion adaptation. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1740), 3091–3097. https://doi.org/10.1098/rspb.2012.0637

Umeno, M. M., & Goldberg, M. E. (1997). Spatial Processing in the Monkey Frontal Eye Field. I. Predictive Visual Responses. *Journal of Neurophysiology*, *78*(3), 1373–1383.

Walker, M. F., Fitzgibbon, E. J., & Goldberg, M. E. (1995). Neurons in the monkey superior colliculus predict the visual result of impending saccadic eye movements. *Journal of Neurophysiology*, *73*(5), 1988–2003.

Ward, E. J., MacEvoy, S. P., & Epstein, R. A. (2010). Eye-centered encoding of visual space in scene-selective regions. *Journal of Vision*, *10*(14), 6–6. https://doi.org/10.1167/10.14.6

Yantis, S., Schwarzbach, J., Serences, J. T., Carlson, R. L., Steinmetz, M. A., Pekar, J. J., & Courtney, S. M. (2002). Transient neural activity in human parietal cortex during spatial attention shifts. *Nature Neuroscience*, *5*(10), 995–1002. https://doi.org/10.1038/nn921

Zhao, M., Gersch, T. M., Schnitzer, B. S., Dosher, B. A., & Kowler, E. (2012). Eye movements and attention: The role of pre-saccadic shifts of attention in perception, memory and the control of saccades. *Vision Research*, *74*, 40–60. https://doi.org/10.1016/j.visres.2012.06.017

Zimmermann, E., Morrone, M. C., Fink, G. R., & Burr, D. (2013). Spatiotopic neural representations develop slowly across saccades. *Current Biology*, *23*(5), R193–R194. https://doi.org/10.1016/j.cub.2013.01.065

Zimmermann, E., Weidner, R., Abdollahi, R. O., & Fink, G. R. (2016). Spatiotopic Adaptation in Visual Areas. *Journal of Neuroscience*, *36*(37), 9526–9534. https://doi.org/10.1523/JNEUROSCI.0052-16.2016

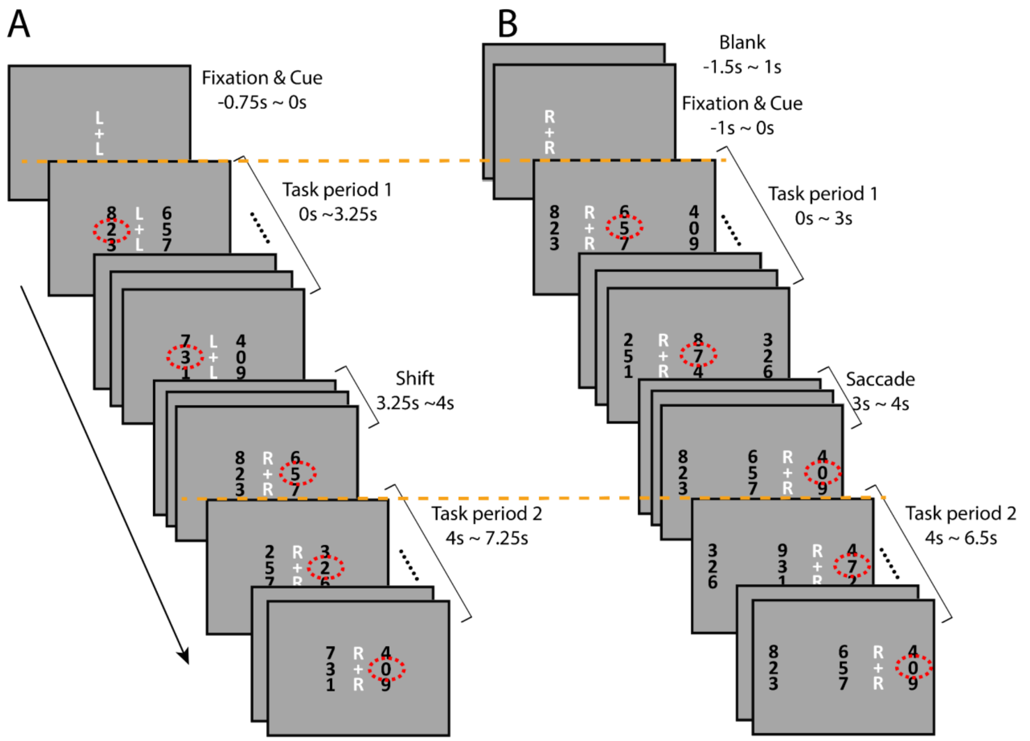


Figure 1 Paradigms of the Eyes-fixed and Eyes-move tasks. (A) An example of an Eyes-Fixed, Shift-attention trial, where covert attention is shifted from the left stream to the right stream; the letter cues "L" and "R" above and below the fixation cross indicate "left" and "right". (B) An example of an Eyes-Move, maintain Retinotopic-attention trial, where covert attention is maintained on the stream located to the right of fixation across the saccade; here the letter cues "L" and "R" indicate "left of fixation" and "right of fixation", and "C" would indicate "center of screen" for maintain Spatiotopic-attention trials (see Figure 2 for examples). Red dotted circles (not shown in the actual experiment) indicate the digit stream that participants should attend to according to the letter cue. Time 0s is taken as the onset of each trial, and orange dotted lines are to show that the onsets of task period 1 and 2 were synced with scanner pulse in both Eyes-fixed and Eyes-move tasks.

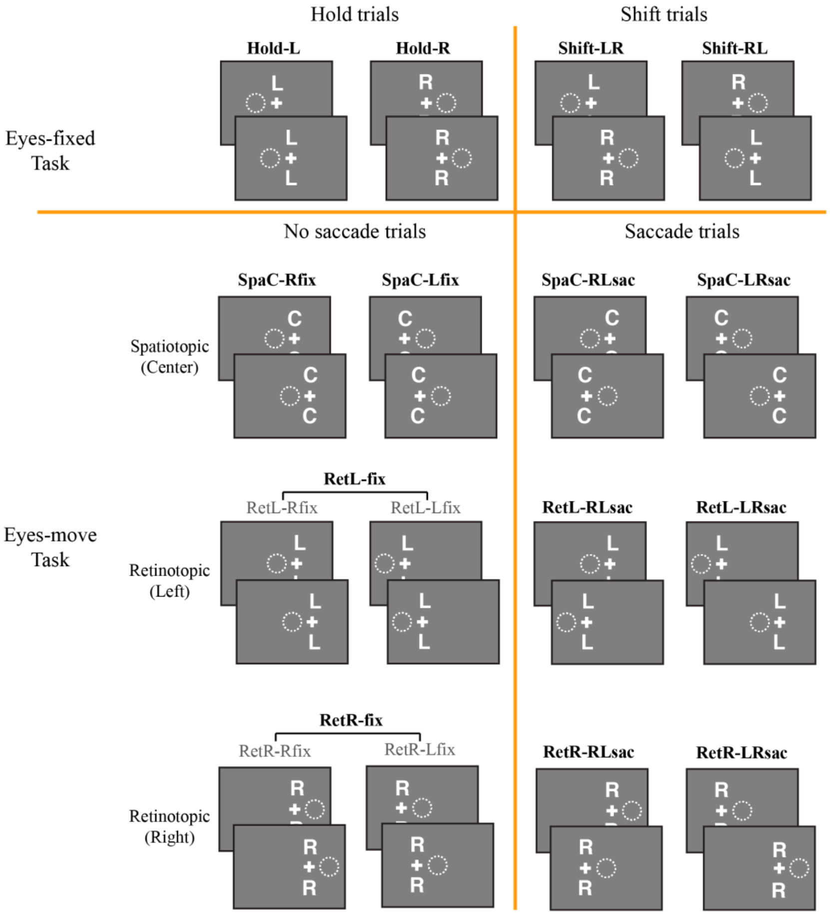


Figure 2 Diagrams of all conditions. Each condition was separated into the first half (before shift/saccade) and the second half (after shift/saccade), shown as the top and bottom panel for each condition. White crosses indicate the fixation location, and white dotted circles indicate the attention location on the screen, corresponding to the letter cues above and below the fixation. Note that in our analyses, we did not separate the left and right fixation for retinotopic no-saccade conditions; that is, only the bolded conditions were included in the GLMs.

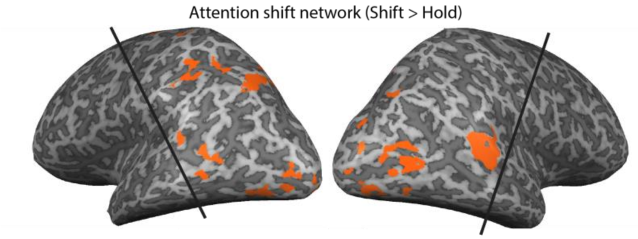


Figure 3 The attention shift network. Result is based on group-level GLM contrasts on the volume space, after being cluster-threshold corrected at p<.05. The volume maps were projected onto an inflated brain only for visualization purpose. The black lines demonstrate the approximate coverage (slightly different for each subject).

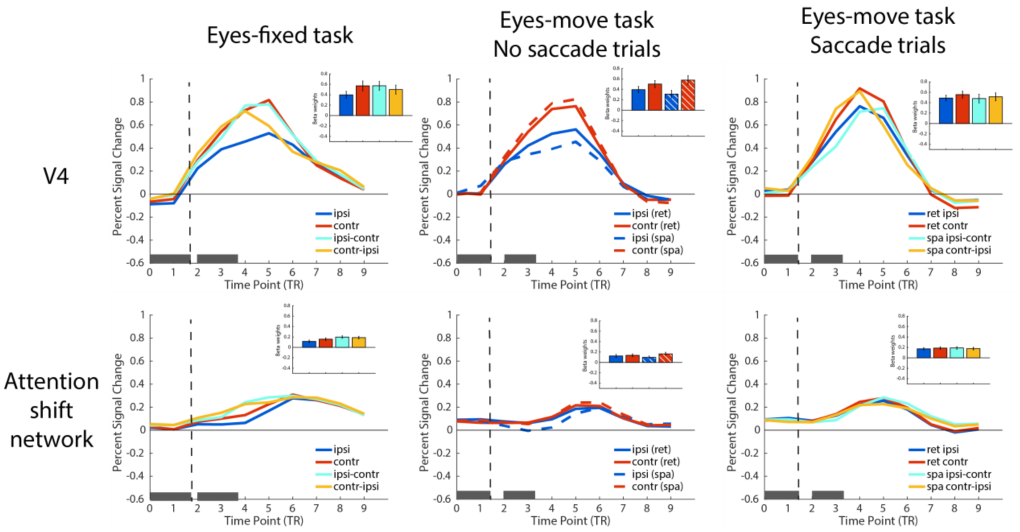


Figure 4 FIR results of Eyes-fixed task (left column), Eyes-move task with no saccade trials (middle column), and Eyes-move task with saccade trials (right column). The pair of gray boxes along the x-axis in each plot indicates the time duration of the two task periods in the trial, and the vertical dashed lines indicate the onset of shift or saccade cues. Inset bar plots show the whole-trial beta weights for each condition in each ROI/network, color-coded in the same way as the corresponding FIR plots.

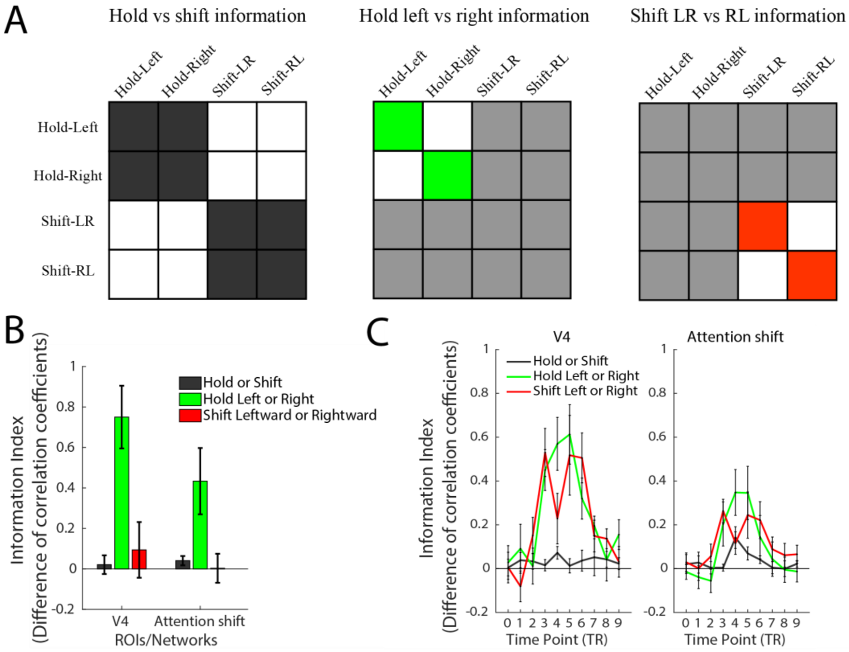


Figure 5 MVPA and MVPTC analyses and results of Eyes-fixed tasks. (A) Hypothetical matrices for hold vs shift, hold left vs right, and shift LR vs RL information. Cells colored in dark grey, green and red are the within-group correlations, and white cells are the between-group correlations. Light grey cells are not used in the corresponding analysis. The index value of each type of information is calculated by subtracting the z-scored between-group correlation coefficients from the z-scored within-group correlation coefficients. (B) The index value of each type of information in each ROI/network. (C) The index value of each type of information at 10 time points, in each ROI/network. Error bars represent SEM.

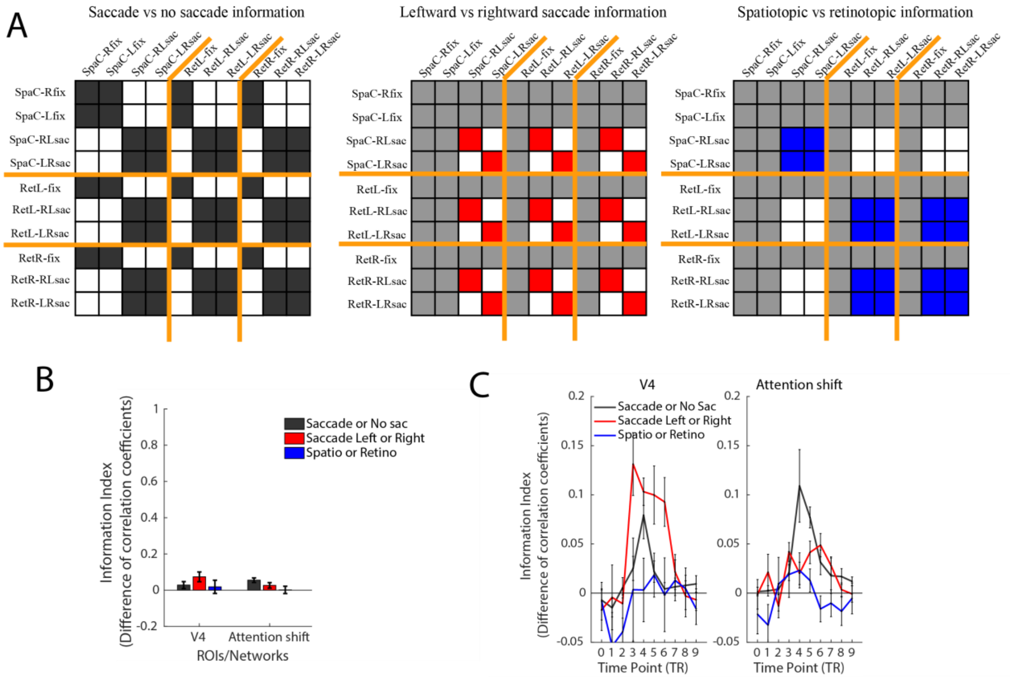


Figure 6 MVPA and MVPTC analyses and results of the Eyes-move task. (A) Hypothetical matrices for information about: saccade vs no saccade, leftward vs rightward saccade, and spatiotopic vs retinotopic attention (across saccades). Orange lines separate conditions in spatiotopic ("attend center"), retinotopic left ("attend left of cross"), and retinotopic right ("attend right of cross") blocks. Cells colored in dark grey, red, and blue are the within-group correlations, and white cells are the between-group correlations. Light grey cells are not used in the corresponding analysis. The index value of each type of information is calculated by subtracting the z-scored between-group correlation coefficients from the z-scored within-group correlation coefficients. (B) The index value of each type of information in each ROI/network; the scale is the same as Figure 5B. (C) The index value of each type of information at 10 time points, in each ROI/network; the scale is different from Figure 5B, 5C, and 6B. Error bars represent SEM. Extended Analyses are shown in Figure 6-1, 6-2, and 6-3.

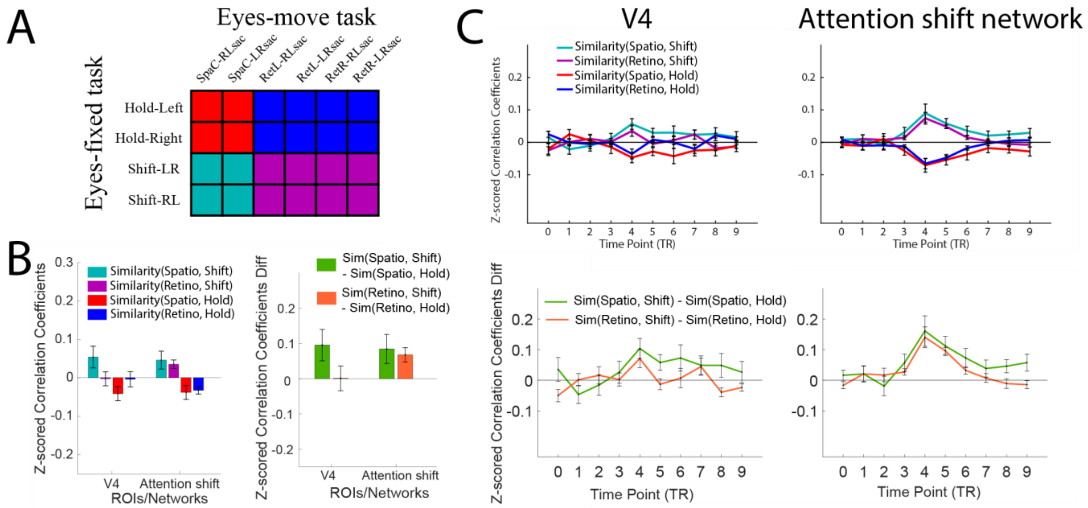


Figure 7 Cross-task similarity analyses in a priori ROIs/networks. (A) A hypothetical matrix indicating each combination of similarity: Retinotopic-to-hold (blue), retinotopic-to-shift (magenta), spatiotopic-to-hold (red), and spatiotopic-to-shift (cyan). (B-C) Pattern similarity (z-scored correlation coefficients) for each combination of conditions, for each ROI/network. (B) Pattern similarity based on whole-trial beta weights. Left: Similarity for each of the 4 cross-task pairings. Right: Pattern similarity difference scores, showing [Spatio-to-shift minus Spatio-to-hold] and [Retino-to-shift minus Retino-to-hold]. (C) Pattern similarity timecourses based on FIR beta weights for each of 10 time points. Top row: For each of the 4 cross-task pairings; Bottom two: Pattern similarity difference scores as in B. Error bars represent SEM. Note that the roughly symmetrical patterns of the timecourse plots are likely due to the de-meaning step of subtracting the grand mean activity across conditions for each timepoint's MVPA analysis, but it does not influence the interpretation for the main effects and interactions. (see Methods).

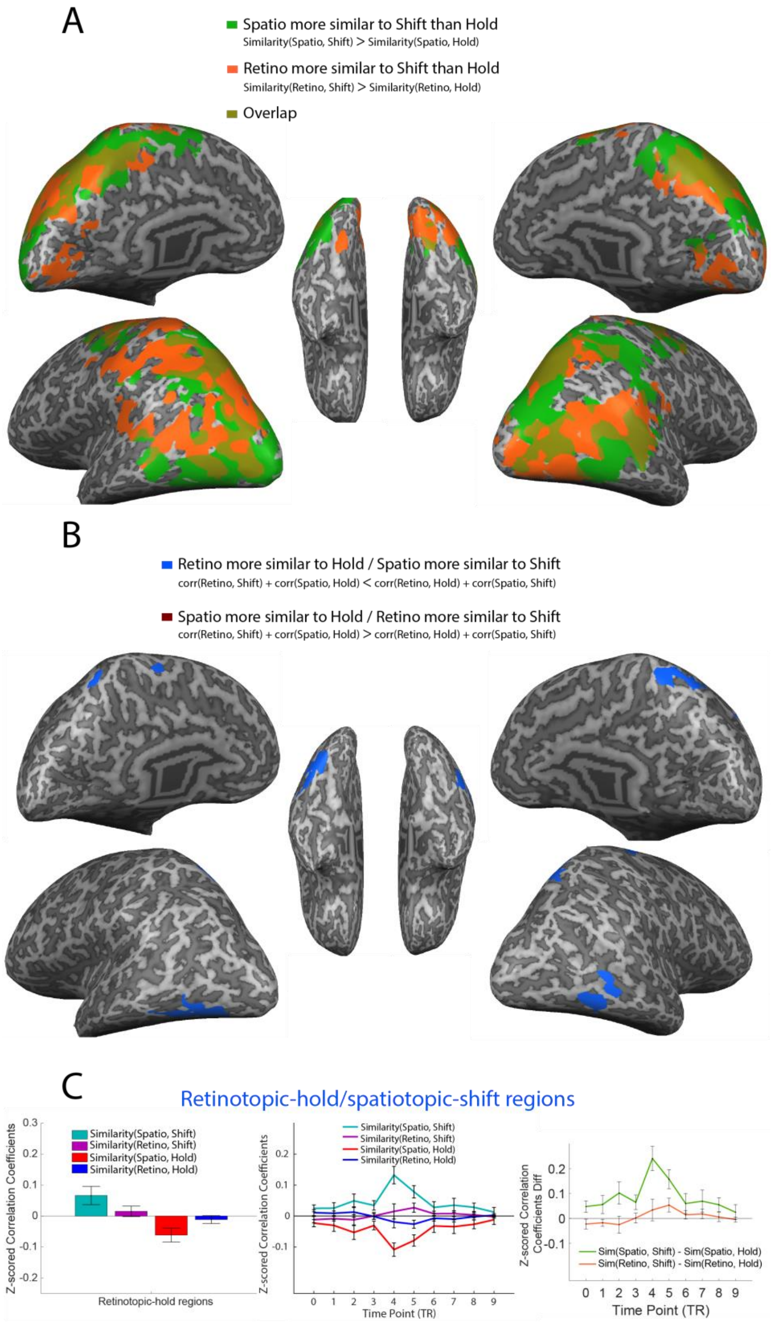


Figure 8 Cross-task pattern similarity, whole-brain searchlight analyses. (A) Regions showing significant difference between retinotopic-shift similarity and retinotopic-hold similarity (orange), and regions showing significant difference between spatiotopic-shift similarity and spatiotopic-hold similarity (green). Overlapping regions shown in brown. Note, no regions showing higher similarity to holding than shifting (for either comparison) survived the cluster threshold correction. (B) Regions

showing a significant interaction effect. Regions exhibiting a significant retino-hold/spatio-shift pattern shown in blue; Regions exhibiting a significant spatio-hold/retino-shift pattern shown in scarlet (no clusters passed significance threshold for this contrast). All searchlights are based on cross-task MVPTC, using the pattern correlation difference at TP4, with direction of contrast as indicated in the legends. The searchlight maps were corrected for cluster-threshold in the same way as other brain maps. Searchlight analyses were conducted on the volume maps and projected onto an inflated brain for visualization purpose. (C) Pattern similarity in the whole-trial (left) and in timecourses (middle) for each combination of conditions, and the difference scores for similarity-to-shift and similarity-to-hold (right), shown for the retino-hold/spatio-shift areas extracted from (B) (all voxels averaged into single network; for separate plots for each individual area, see Extended data Figure 8-1). Plots are for illustrative purposes only to explore the specific pattern driving the significant interaction. Error bars represent SEM.

Table 1 Description of clusters in the attention shift network, including Talairach coordinates of the peak voxel, number of voxels, and t values.

Networks	Areas	Hemisphere	TAL coordinates of peak voxel				
			x	y	z	# of voxels	t value (df=11)
Attention shift network	Superior Temporal Gyrus	R	63	-39	14	398	6.5813
		L	-53	-53	12	180	7.4536
	Middle Temporal Gyrus	R	41	-61	6	389	6.5224
	Inferior Occipital Gyrus	R	31	-81	-4	254	6.8871
		L	-37	-77	-4	226	5.7041
	Inferior Parietal Lobule	L	-37	-37	42	1388	5.5973
	Lingual Gyrus (posterior)	L	-11	-87	-14	163	5.2757
	Lingual Gyrus (anterior)	L	-27	-61	4	133	4.6952
	Superior Frontal Gyrus	L	-17	-13	74	146	4.9916

Table 2 Statistical tests of information indices in each ROI/network, separately for whole-trial analyses and time points of interest in the time-course analyses. N=12.

	V4	Attention shift network
Hold or Shift	*t*(11)=0.451, *p*=.661, *d*=0.130 TP3: *t*(11)=0.608, *p*=.555, *d*=0.176 TP4: *t*(11)=2.507, *p*=.029, *d*=0.724* TP5: *t*(11)=0.394, *p*=.701, *d*=0.114	*t*(11)=1.755, *p*=.107, *d*=0.507 TP3: *t*(11)=0.278, *p*=.787, *d*=0.080 TP4: *t*(11)=2.853, *p*=.016, *d*=0.823** TP5: *t*(11)=2.316, *p*=.041, *d*=0.668*
Hold L or Hold R	*t*(11)=4.843, *p*<.001, *d*=1.380** TP3: *t*(11)=4.818, *p*<.001, *d*=1.391** TP4: *t*(11)=4.709, *p*<.001, *d*=1.359** TP5: *t*(11)=4.521, *p*<.001, *d*=1.305**	*t*(11)=2.645, *p*=.023, *d*=0.764** TP3: *t*(11)=2.025, *p*=.068, *d*=0.585 TP4: *t*(11)=3.326, *p*=.007, *d*=0.960** TP5: *t*(11)=2.834, *p*=.016, *d*=0.818**
Shift leftward or rightward	*t*(11)=0.682, *p*=.510, *d*=0.197 TP3: *t*(11)=4.840, *p*<.001, *d*=1.397** TP4: *t*(11)=1.975, *p*=.074, *d*=0.570 TP5: *t*(11)=2.839, *p*=.016, *d*=0.820**	*t*(11)=0.040, *p*=.969, *d*=0.012 TP3: *t*(11)=4.903, *p*<.001, *d*=1.415** TP4: *t*(11)=2.273, *p*=.044, *d*=0.656** TP5: *t*(11)=2.310, *p*=.041, *d*=0.667*
Saccade or no saccade	*t*(11)=1.452, *p*=.175, *d*=0.419 TP3: *t*(11)=2.056, *p*=.064, *d*=0.594 TP4: *t*(11)=3.305, *p*=.007, *d*=0.954** TP5: *t*(11)=2.014, *p*=.069, *d*=0.581	*t*(11)=4.432, *p*=.001, *d*=1.279** TP3: *t*(11)=2.598, *p*=.025, *d*=0.750** TP4: *t*(11)=2.956, *p*=.013, *d*=0.853** TP5: *t*(11)=7.249, *p*<.001, *d*=2.093**
Saccade leftward or rightward	*t*(11)=2.730, *p*=.020, *d*=0.788* TP3: *t*(11)=4.113, *p*=.002, *d*=1.187** TP4: *t*(11)=7.401, *p*<.001, *d*=2.136** TP5: *t*(11)=3.370, *p*=.006, *d*=0.973**	*t*(11)=1.771, *p*=.104, *d*=0.511 TP3: *t*(11)=4.420, *p*=.001, *d*=1.276** TP4: *t*(11)=1.775, *p*=.104, *d*=0.512 TP5: *t*(11)=3.615, *p*=.004, *d*=1.044**
Retinotopic or spatiotopic	*t*(11)=0.504, *p*=.625, *d*=0.145 TP3: *t*(11)=0.067, *p*=.948, *d*=0.019 TP4: *t*(11)=0.101, *p*=.921, *d*=0.029 TP5: *t*(11)=0.819, *p*=.430, *d*=0.237	*t*(11)=0.074, *p*=.943, *d*=0.021 TP3: *t*(11)=0.816, *p*=.432, *d*=0.236 TP4: *t*(11)=1.295, *p*=.222, *d*=0.374 TP5: *t*(11)=1.095, *p*=.297, *d*=0.316

* indicate statistical significance at *p*<.05

** indicate statistical significance at *p*<.05 (Holm-Bonferroni corrected for multiple post hoc comparisons, separately across ROIs/networks for whole-trial MVPA, and across three TPs for MVPTC)

Table 3 Statistics of 2×2 repeated-measure ANOVAs for each ROI at TP3, TP4, and TP5 respectively, on pattern similarity between Eyes-fixed conditions (hold & shift attention) and Eyes-move conditions (spatiotopic & retinotopic attention), separately for whole-trial analyses and time points of interest.

	V4	Attention shift network
Main effect of similarity to Eyes-fixed conditions (hold & shift)	F=8.367, p=.015, η_p^2=.432** TP3: F=2.549, p=.139, η_p^2=.188 TP4: F=13.113, p=.004, η_p^2=.544** TP5: F=4.269, p=.063, η_p^2=.280	F=18.892, p=.001, η_p^2=.632** TP3: F=15.604, p=.002, η_p^2=.587** TP4: F=15.293, p=.002, η_p^2=.582** TP5: F=29.311, p<.001, η_p^2=.727**
Main effect of Eyes-move conditions (spatiotopic & retinotopic)	F=0.486, p=.500, η_p^2=.042 TP3: F=0.510, p=.490, η_p^2=.044 TP4: F=0.922, p=.358, η_p^2=.077 TP5: F=0.184, p=.676, η_p^2=.016	F=0.291, p=.601, η_p^2=.026 TP3: F=0.291, p=.600, η_p^2=.026 TP4: F=5.514, p=.039, η_p^2=.334* TP5: F=0.827, p=.383, η_p^2=.070
Interaction between similarity to Eyes-fixed conditions and similarity to Eyes-move conditions	F=1.672, p=.223, η_p^2=.132 TP3: F=0.182, p=.678, η_p^2=.016 TP4: F=0.727, p=.412, η_p^2=.062 TP5: F=3.407, p=.092, η_p^2=.236	F=0.091, p=.768, η_p^2=.008 TP3: F=0.640, p=.441, η_p^2=.055 TP4: F=0.348, p=.567, η_p^2=.031 TP5: F=0.351, p=.566, η_p^2=.031

* indicate statistical significance at p<.05

** indicate statistical significance at p<.05 (Holm-Bonferroni corrected for multiple post hoc comparisons, separately across ROIs/networks for whole-trial beta weights, and across three TPs for time-course beta weights)

Table 4 Description of clusters in regions with the retinotopic-hold pattern, including Talairach coordinates of the peak voxel, number of voxels, and t values.

Networks	Areas	Hemisphere	TAL coordinates of peak voxel				
			x	y	z	# of voxels	t values (df=11)
Retinotopic-hold	Parahippocampal Gyrus	R	39	-49	4	691	4.0403
	Fusiform Gyrus	L	-37	-59	-12	708	4.9133
	Precuneus	R	11	-60	67	884	3.9127
	Paracentral Lobule	L	-3	-41	60	472	3.4624

Extended data

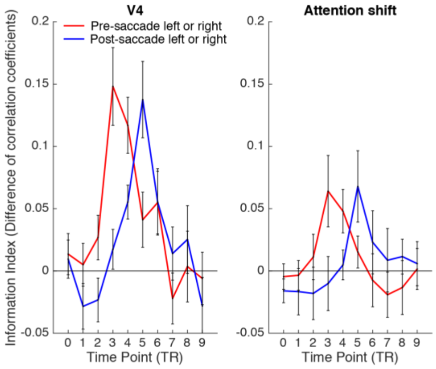


Figure 6-1 Extended Data showing MVPTC results of information about the hemifield attended (left or right) before and after the saccade separately (in the Eyes-move task). The index values of each type of information at 10 time points are plotted for each ROI/network. Error bars represent SEM. Results show that we could decode which hemifield was being covertly attended both before and after the saccade.

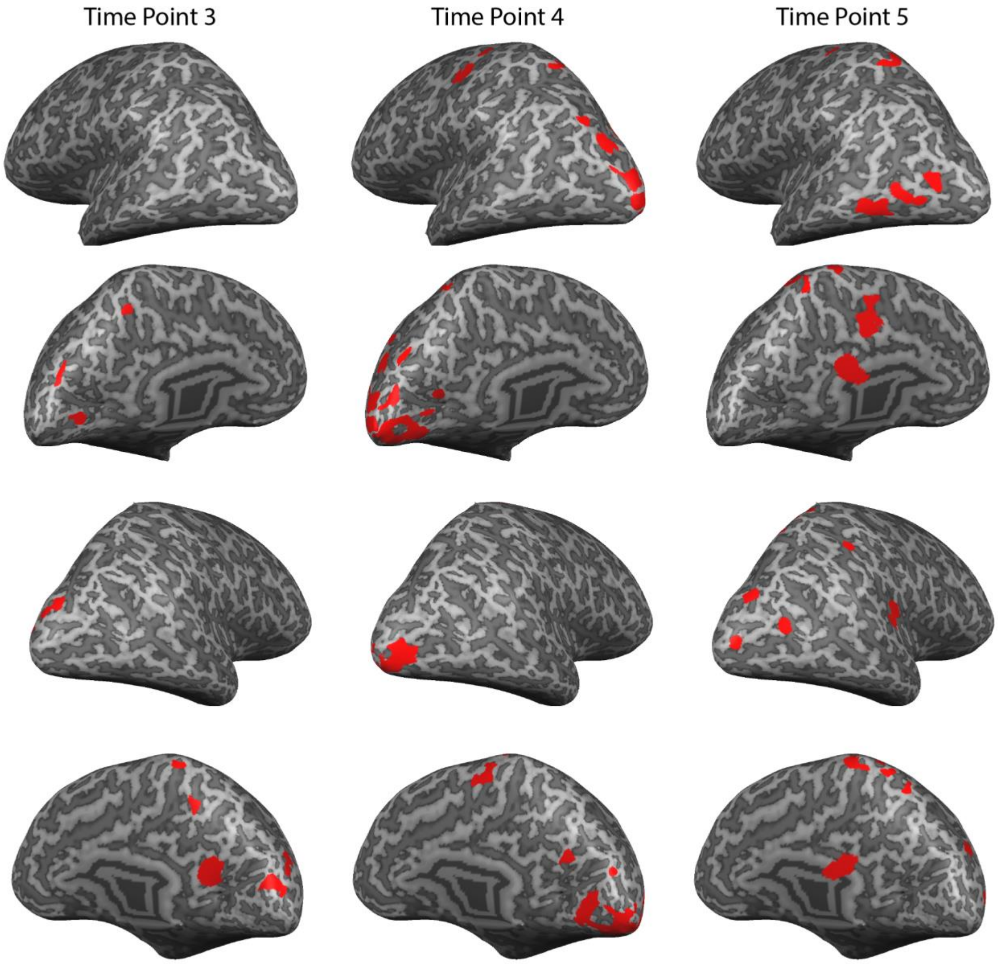
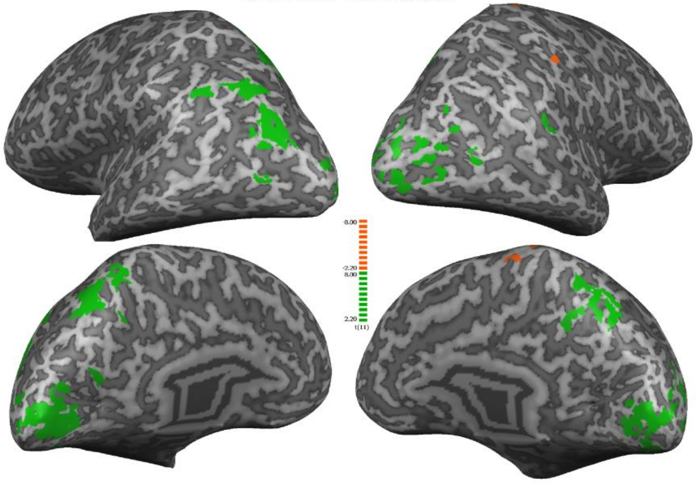


Figure 6-2 Extended Data showing information of retinotopic vs spatiotopic attention in searchlight analyses, for time point 3, 4, and 5 separately. This whole-brain analysis is analogous to Figure 6, blue condition (information about spatiotopic vs retinotopic). Red areas show significant information after cluster-threshold correction at p<.05. The viewing angle for each row is left lateral, left medial, right lateral, and right medial, respectively.

Saccade > No saccade



Spatiotopic > Retinotopic

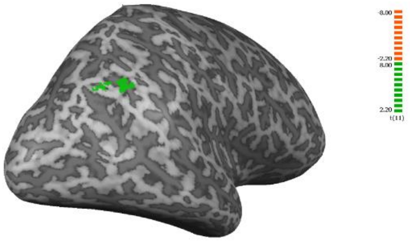
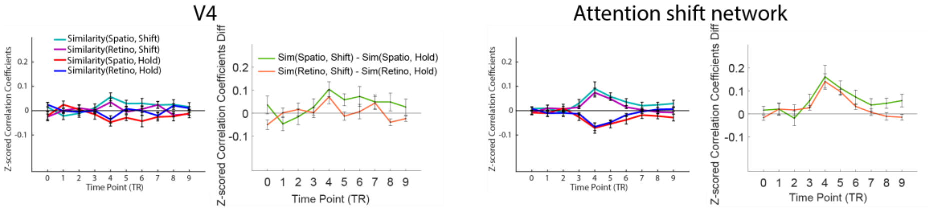
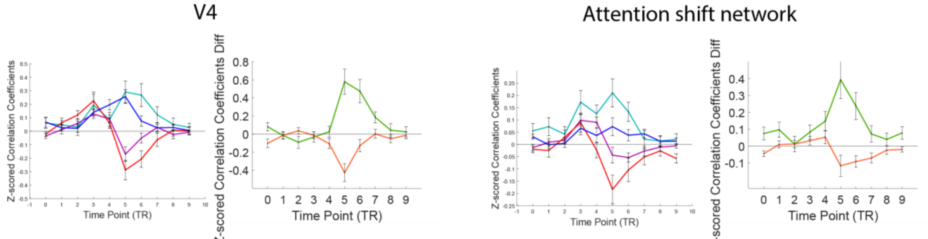


Figure 6-3 Extended Data showing univariate differences, based on whole-trial betas, between saccade and no saccade conditions in the Eyes-move task (top), and between retinotopic and spatiotopic conditions in the Eyes-move task (bottom). For each contrast, significant clusters in the positive direction are shown in green and negative in orange. Maps were cluster threshold corrected at p<.05. For Spatiotopic > Retinotopic contrast, the only significant cluster found was located in the left hemisphere, so only the left lateral viewing angle is shown here.

A. Manuscript Figure 7



B. Match based on first half



C. Match based on second half

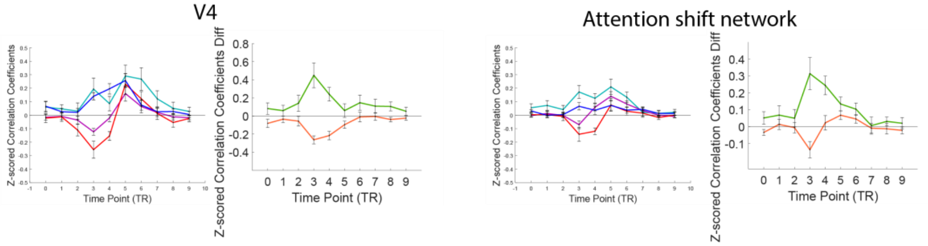


Figure 7-1 Extended data showing an alternative way to analyze cross-task similarity – matching the hemispheric locations of covert attention, e.g., only correlating RetL and HoldL (RetR and HoldR) to calculate similarity between retinotopic and hold. Compared with the analysis in the paper (Panel A), here we perform this alternative analysis based on matching the first half of the trial (before shift/saccade) (Panel B) and matching the second half of the trial (after shift/saccade) (Panel C).

In Panel B, we can see that in both ROIs, at time point 3 when BOLD signals corresponds to the first half of trial, all four pairs of correlations are positive, because we have explicitly matched the retinotopic location of attention for all. In the second half of the trial, now the correlations between retinotopic and hold (blue) and between spatiotopic and shift (cyan) are greater than those between retinotopic and shift (magenta) and spatiotopic and hold (red). This looks like the retinotopic attention condition has more representational similarity to holding attention, and the spatiotopic is more similar to shift, as reflected in the difference score plots. Panel C can be interpreted in a similar way. But again, this interpretation would be biased because we explicitly defined the conditions in terms of their retinotopic locations; thus, it is an unsurprising result. (Note, though, that the difference score plots are not symmetrical around zero, especially for the attention shift network; if the ROIs coded attention in a purely retinotopic manner, we would expect the difference curves to be of equal magnitude in opposite directions. Thus, even this retinotopically-biased analysis still reveals a pattern consistent with our original conclusions: that both retinotopic and spatiotopic attention in saccade trials carry some similarity to shifting attention.)

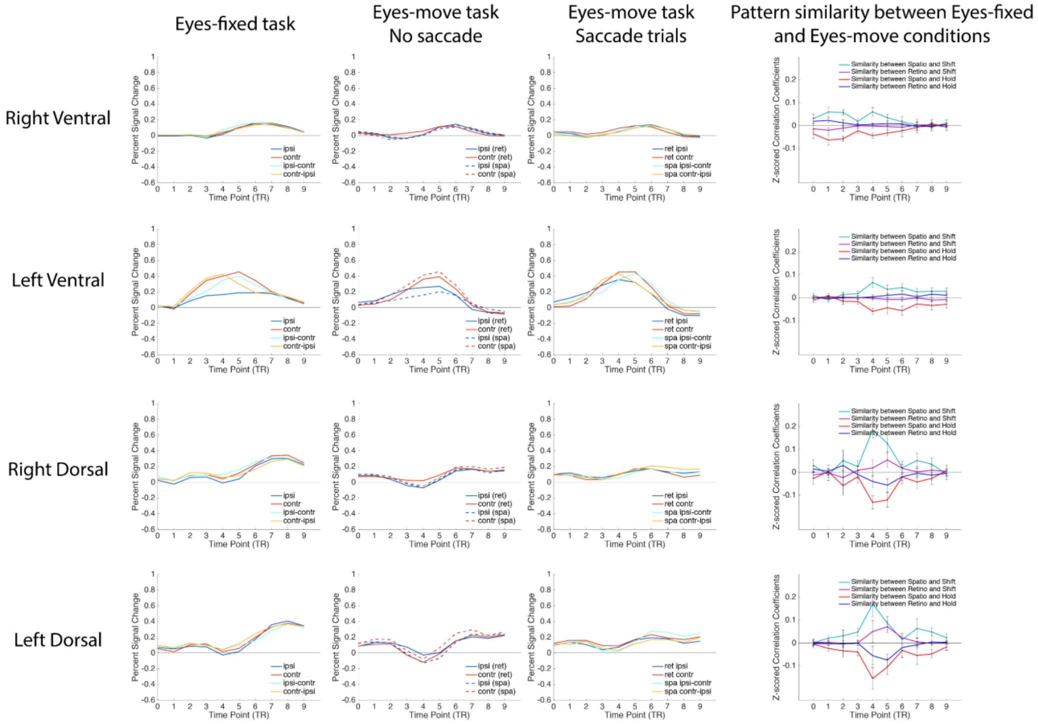
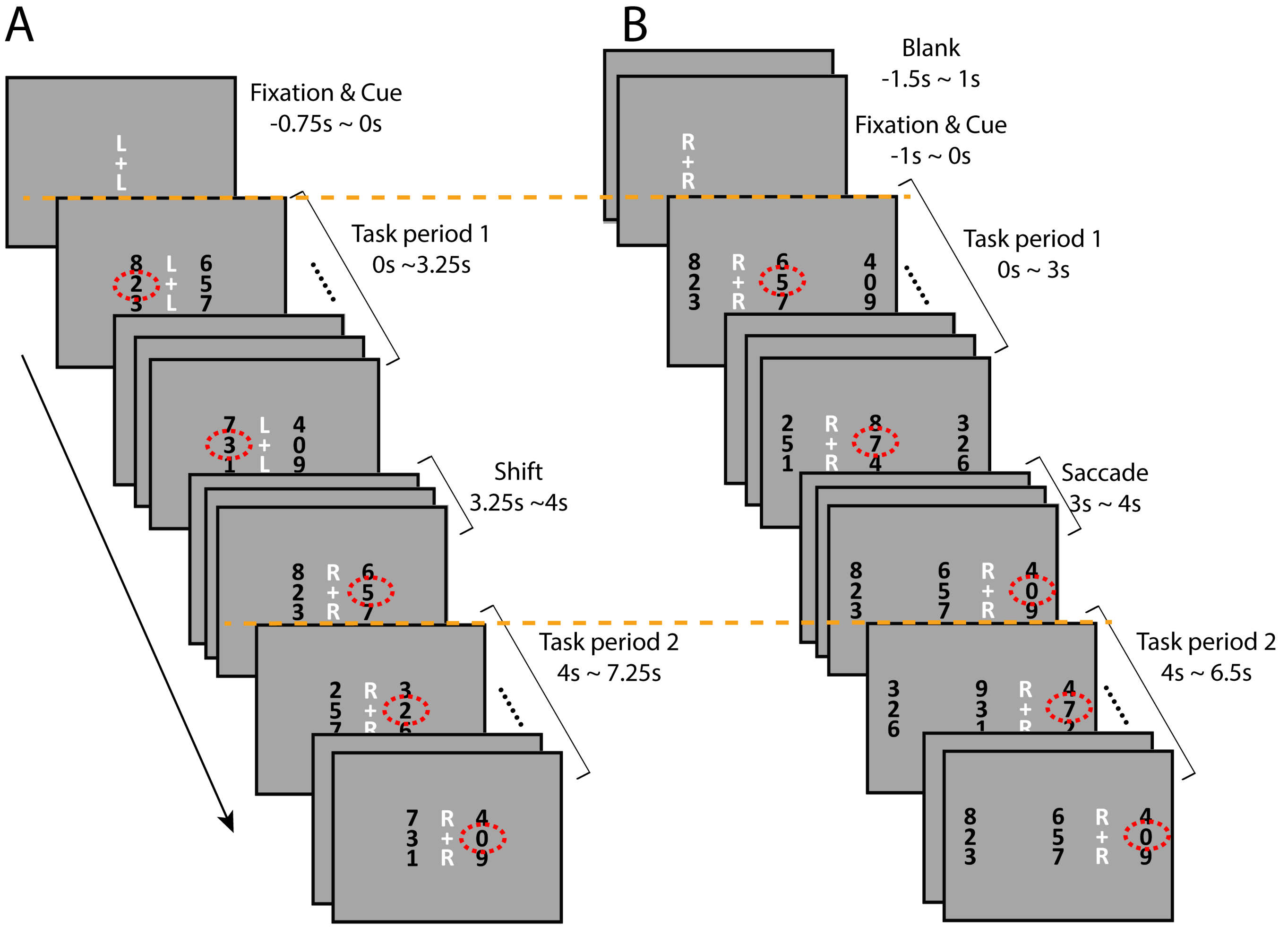
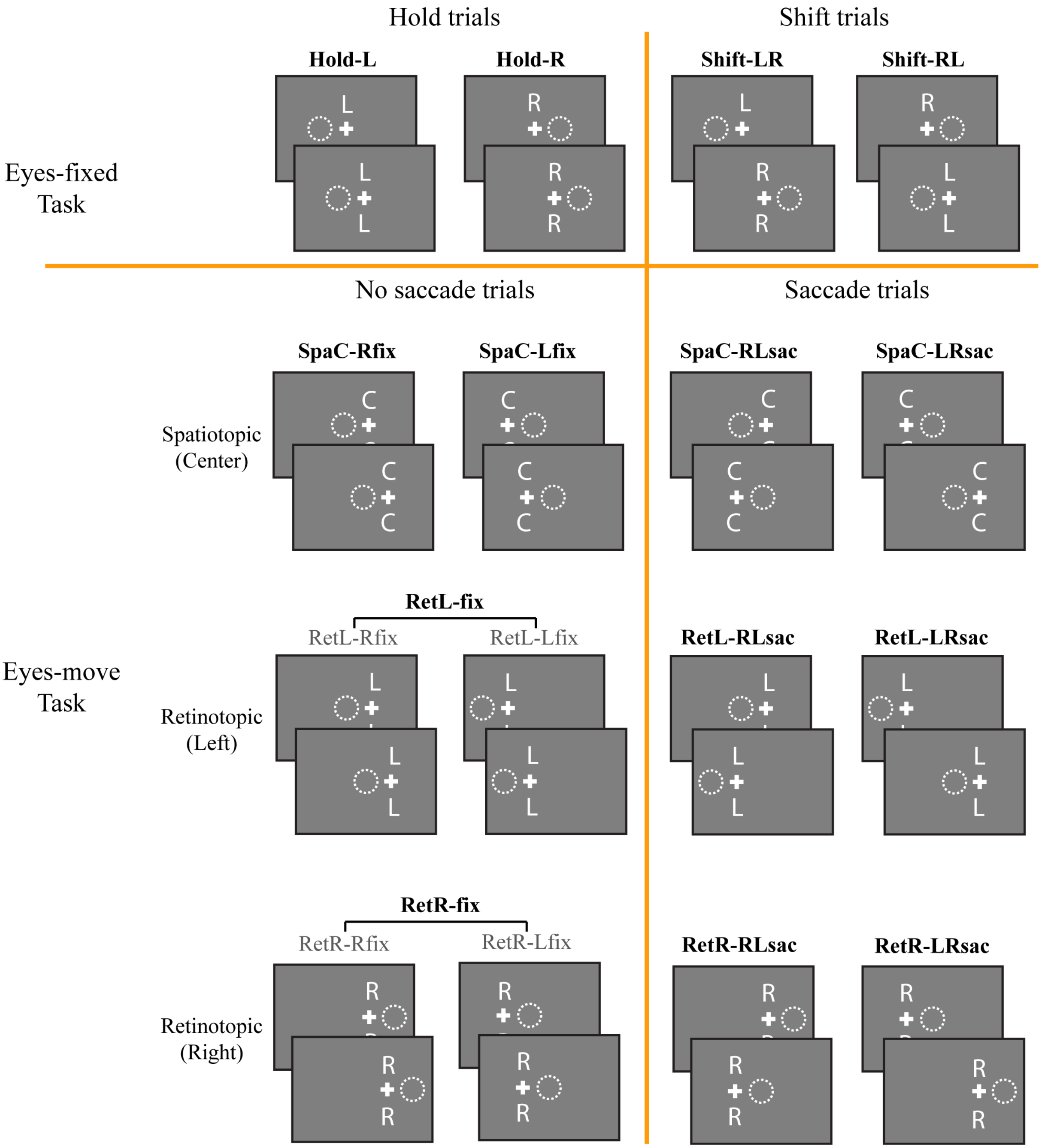
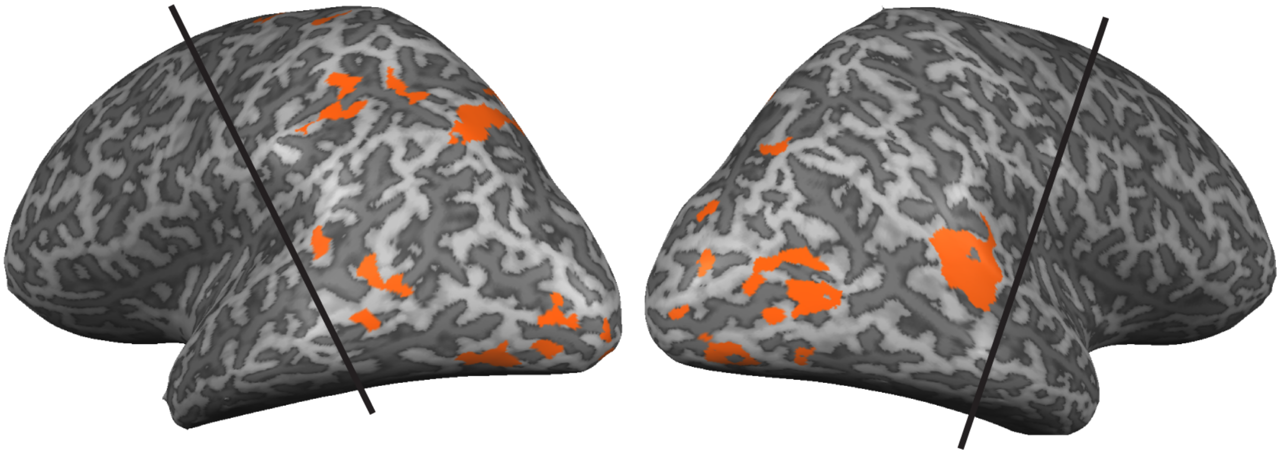


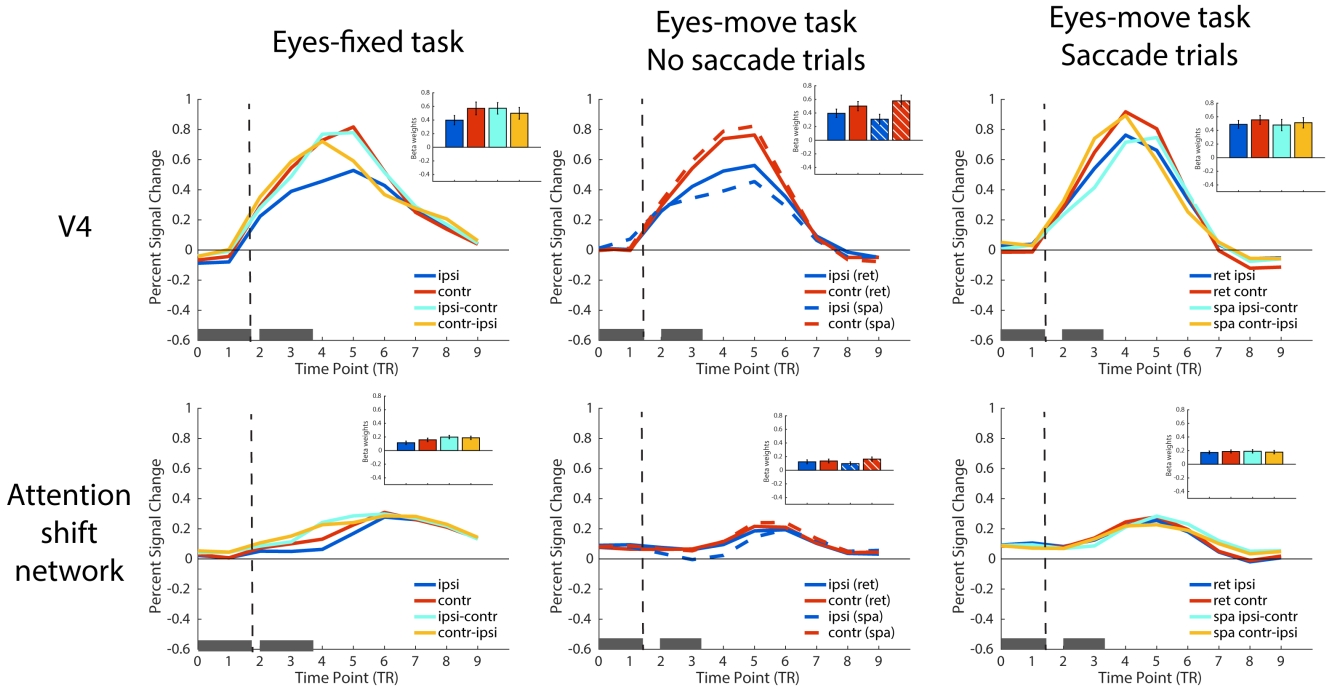
Figure 8-1 Extended Data showing univariate activation (the first three columns) and cross-task pattern similarities (the last column), separately for each cluster of the retinotopic-hold regions from the exploratory searchlight analyses. The univariate activation plots were comparable to Figure 4, and the pattern similarity plots to Figure 7C & 8B.

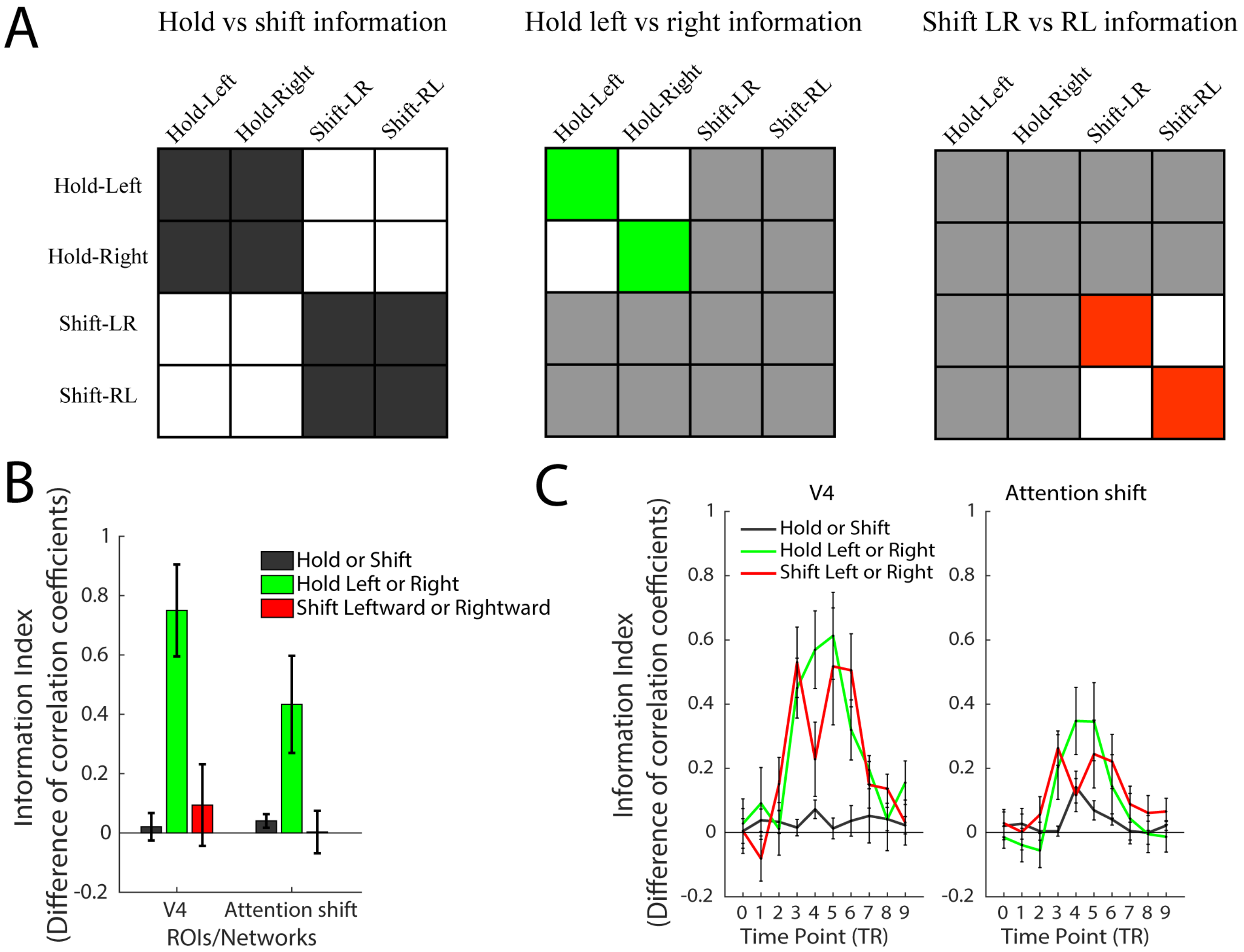

A
Fixation & Cue
-0.75s ~ 0s
Task period 1
0s ~3.25s
Shift
3.25s ~4s
Task period 2
4s ~ 7.25s
B
Blank
-1.5s ~ 1s
Fixation & Cue
-1s ~ 0s
Task period 1
0s ~ 3s
Saccade
3s ~ 4s
Task period 2
4s ~ 6.5s

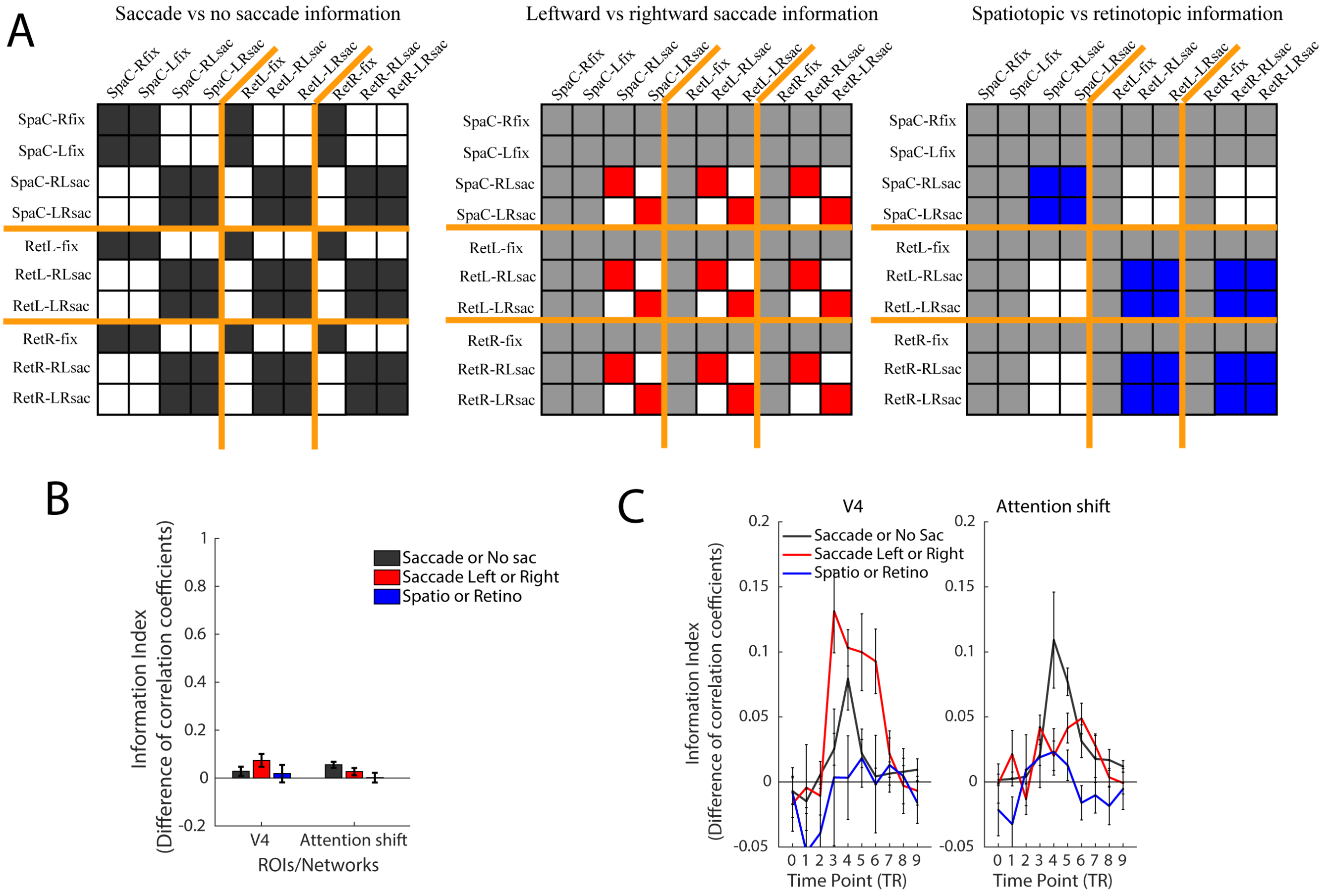

Hold trials
Shift trials
Hold-L
Hold-R
Shift-LR
Shift-RL
Eyes-fixed Task
No saccade trials
Saccade trials
SpaC-Rfix
SpaC-Lfix
SpaC-RLsac
SpaC-LRsac
Spatiotopic (Center)
RetL-fix
RetL-Rfix
RetL-Lfix
RetL-RLsac
RetL-LRsac
Eyes-move Task
Retinotopic (Left)
RetR-fix
RetR-Rfix
RetR-Lfix
RetR-RLsac
RetR-LRsac
Retinotopic (Right)

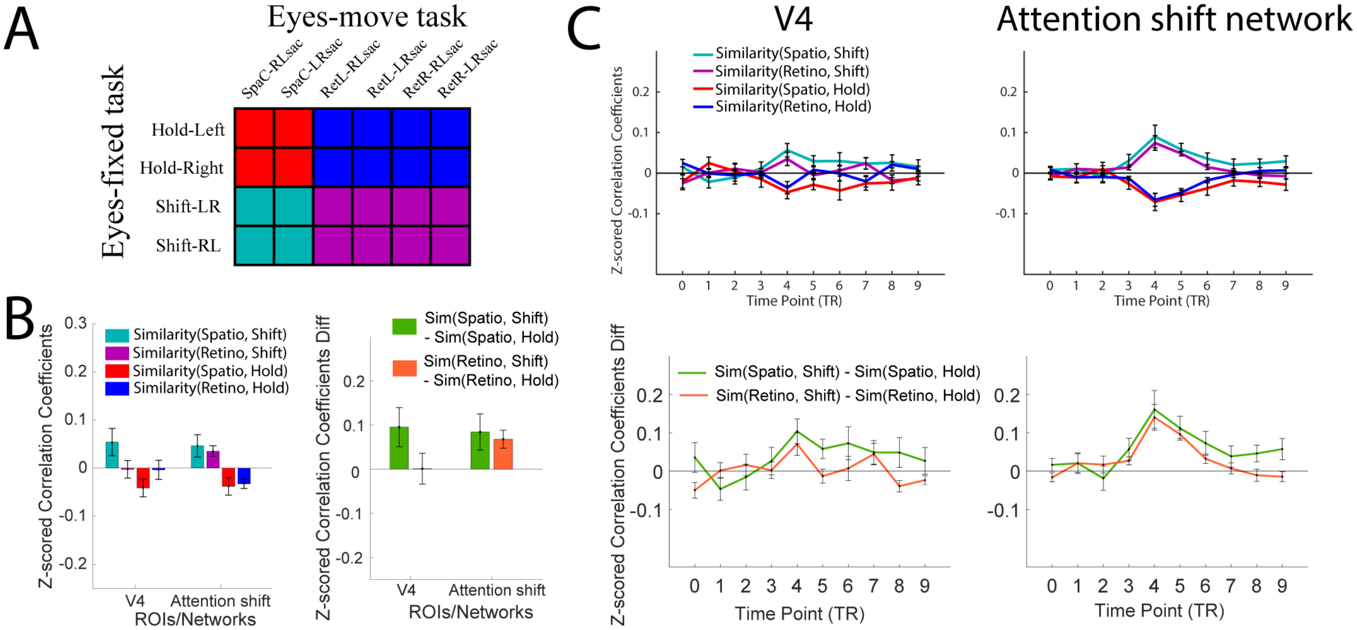
Attention shift network (Shift > Hold)

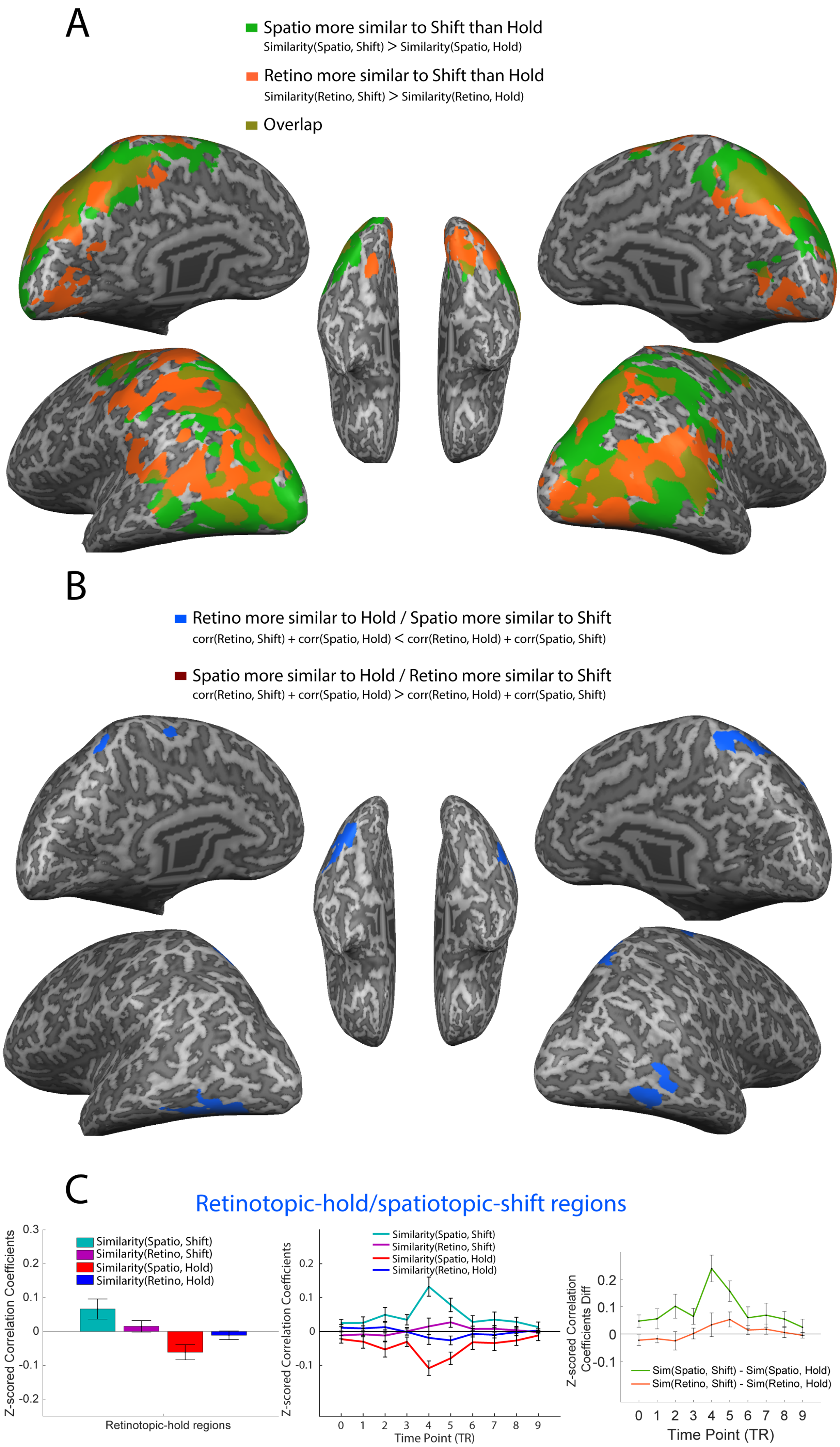



Eyes-fixed task
Eyes-move task
No saccade trials
Eyes-move task
Saccade trials
V4
Attention shift network
Percent Signal Change
Time Point (TR)
Beta weights
ipsi
contr
ipsi-contr
contr-ipsi
ipsi (ret)
contr (ret)
ipsi (spa)
contr (spa)
ret ipsi
ret contr
spa ipsi-contr
spa contr-ipsi
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
0 1 2 3 4 5 6 7 8 9


A
Hold vs shift information
Hold left vs right information
Shift LR vs RL information
Hold-Left
Hold-Right
Shift-LR
Shift-RL
B
Information Index
(Difference of correlation coefficients)
1
0.8
0.6
0.4
0.2
0
-0.2
V4
Attention shift
ROIs/Networks
Hold or Shift
Hold Left or Right
Shift Leftward or Rightward
C
V4
Attention shift
Hold or Shift
Hold Left or Right
Shift Left or Right
0 1 2 3 4 5 6 7 8 9
Time Point (TR)


A
Saccade vs no saccade information
Leftward vs rightward saccade information
Spatiotopic vs retinotopic information
SpaC-Rfix
SpaC-Lfix
SpaC-RLsac
SpaC-LRsac
RetL-fix
RetL-RLsac
RetL-LRsac
RetR-fix
RetR-RLsac
RetR-LRsac
B
Information Index
(Difference of correlation coefficients)
1
0.8
0.6
0.4
0.2
0
-0.2
V4
Attention shift
ROIs/Networks
Saccade or No sac
Saccade Left or Right
Spatio or Retino
C
V4
Attention shift
Saccade or No Sac
Saccade Left or Right
Spatio or Retino
Information Index
(Difference of correlation coefficients)
0.2
0.15
0.1
0.05
0
-0.05
0 1 2 3 4 5 6 7 8 9
Time Point (TR)


A
Eyes-move task
SpaC-RLsac
SpaC-LRsac
RetL-RLsac
RetL-LRsac
RetR-RLsac
RetR-LRsac
Eyes-fixed task
Hold-Left
Hold-Right
Shift-LR
Shift-RL
B
Similarity(Spatio, Shift)
Similarity(Retino, Shift)
Similarity(Spatio, Hold)
Similarity(Retino, Hold)
Z-scored Correlation Coefficients
0.3
0.2
0.1
0
-0.1
-0.2
V4
Attention shift
ROIs/Networks
Sim(Spatio, Shift) - Sim(Spatio, Hold)
Sim(Retino, Shift) - Sim(Retino, Hold)
Z-scored Correlation Coefficients Diff
0.2
0.1
0
-0.1
-0.2
V4
Attention shift
ROIs/Networks
C
V4
Attention shift network
Similarity(Spatio, Shift)
Similarity(Retino, Shift)
Similarity(Spatio, Hold)
Similarity(Retino, Hold)
Z-scored Correlation Coefficients
0.2
0.1
0
-0.1
0 1 2 3 4 5 6 7 8 9
Time Point (TR)
Sim(Spatio, Shift) - Sim(Spatio, Hold)
Sim(Retino, Shift) - Sim(Retino, Hold)
Z-scored Correlation Coefficients Diff
0.2
0.1
0
-0.1
0 1 2 3 4 5 6 7 8 9
Time Point (TR)


A
Spatio more similar to Shift than Hold
Similarity(Spatio, Shift) > Similarity(Spatio, Hold)
Retino more similar to Shift than Hold
Similarity(Retino, Shift) > Similarity(Retino, Hold)
Overlap
B
Retino more similar to Hold / Spatio more similar to Shift
corr(Retino, Shift) + corr(Spatio, Hold) < corr(Retino, Hold) + corr(Spatio, Shift)
Spatio more similar to Hold / Retino more similar to Shift
corr(Retino, Shift) + corr(Spatio, Hold) > corr(Retino, Hold) + corr(Spatio, Shift)
C
Retinotopic-hold/spatiotopic-shift regions
Z-scored Correlation Coefficients
0.3
0.2
0.1
0
-0.1
-0.2
Similarity(Spatio, Shift)
Similarity(Retino, Shift)
Similarity(Spatio, Hold)
Similarity(Retino, Hold)
Retinotopic-hold regions
Z-scored Correlation Coefficients
0.2
0.1
0
-0.1
Similarity(Spatio, Shift)
Similarity(Retino, Shift)
Similarity(Spatio, Hold)
Similarity(Retino, Hold)
0 1 2 3 4 5 6 7 8 9
Time Point (TR)
Z-scored Correlation Coefficients Diff
0.2
0.1
0
-0.1
Sim(Spatio, Shift) - Sim(Spatio, Hold)
Sim(Retino, Shift) - Sim(Retino, Hold)
0 1 2 3 4 5 6 7 8 9
Time Point (TR)