

Dartmouth College

Dartmouth Digital Commons

Dartmouth College Ph.D Dissertations

Theses and Dissertations

2024

Target Selection and Enhancement During Attentional Tracking

Marvin R. Maechler

Dartmouth College, marvin.r.maechler.gr@dartmouth.edu

Follow this and additional works at: <https://digitalcommons.dartmouth.edu/dissertations>



Part of the [Cognition and Perception Commons](#), and the [Cognitive Neuroscience Commons](#)

Recommended Citation

Maechler, Marvin R., "Target Selection and Enhancement During Attentional Tracking" (2024). *Dartmouth College Ph.D Dissertations*. 184.

<https://digitalcommons.dartmouth.edu/dissertations/184>

This Thesis (Ph.D.) is brought to you for free and open access by the Theses and Dissertations at Dartmouth Digital Commons. It has been accepted for inclusion in Dartmouth College Ph.D Dissertations by an authorized administrator of Dartmouth Digital Commons. For more information, please contact dartmouthdigitalcommons@groups.dartmouth.edu.

TARGET SELECTION AND ENHANCEMENT DURING ATTENTIONAL TRACKING

A Thesis
Submitted to the Faculty
in partial fulfillment of the requirements for the
degree of

Doctor of Philosophy

in

Cognitive Neuroscience

by Marvin Rainer Mächler

Guarini School of Graduate and Advanced Studies
Dartmouth College
Hanover, New Hampshire

June 2023

Examining Committee:

(chair) Peter U. Tse, Ph.D.

Viola S. Störmer, Ph.D.

Caroline E. Robertson, Ph.D.

Martin Rolfs, Ph.D.

F. Jon Kull, Ph.D.

Dean of The Guarini School of Graduate and Advanced Studies

Copyright © 2023 Marvin Mächler (marvin.maechler@gmail.com).

Abstract

At any waking moment, we are bombarded with more sensory information than we can fully process. Attention is necessary to deal with the dynamic world we live in. One fundamental function of vision and attention is to keep track of moving objects, but what are the targets of attention during tracking?

One of the first theories of attentional tracking predicted that targets would be selected at early processing stages. By employing the double-drift illusion, which dissociates physical and perceived positions of moving objects, we investigated which of these positions is selected for tracking. Contrary to earlier theories and in line with newer findings, targets were selected rather late in visual processing, at least after the construction of illusory percepts, for both covert (Chapter One) and overt tracking (Chapter Two).

Furthermore, capacity and speed limits to attentional tracking are hemifield specific. Brain activity in many areas is known to covary with the number of tracked targets, but it was previously unknown whether this effect would also show hemifield bias. Only targets presented in the contralateral hemifield influenced activity in earlier visual areas, while both contralateral and ipsilateral targets affected activity in parietal and frontal areas associated with attention (Chapter Three). Due to the hemifield specific nature of the capacity limit, we conjecture that it should emerge where load dependent activity is strongly contralateral.

Overall, the studies presented in this dissertation illuminate two different aspects of attentional tracking. While selection happens late in the visual hierarchy, capacity and speed limits appear to emerge early in visual processing.

Preface

This dissertation would not have been completed without the help of many people. I am immensely thankful to everyone who has helped me succeed throughout the last six years.

First, I want to thank my advisor Peter Tse for inviting me to his lab and for his guidance and support. Peter's bird's eye view of our field and science at large are an inspiration. Even though some of our more ambitious projects did not work out, I am thankful for the freedom that allowed me to try out so many different ideas. Beyond all that, I am deeply appreciative of our friendship. Thank you, Peter!

I also want to thank my co-advisor Patrick Cavanagh for sharing his vast knowledge of vision science. I am immensely grateful for Patrick's help with all the minutiae of science. His calmness and his wisdom are things I aspire to achieve in the future. Thank you, Patrick!

Additional thanks go to my dissertation committee: Viola Störmer, Caroline Robertson, and Martin Rolfs. Their insights into my projects and our discussions about them have greatly enriched my thinking. Thank you, Viola, Caroline, and Martin!

I feel privileged to have worked alongside my current and former lab mates Peter Kohler, Sebastian Frank, Liwei Sun, Kevin Hartstein, Sirui Liu, Marie-Luise Kieseler, Nate Heller, Sharif Saleki, Mert Özkan, Cody Plante, and Eunhye Choe. Thank you for your camaraderie.

Finally, I thank my parents Ruth and Rainer for all their love, support, and encouragement. I would not have made it here without them. Danke Mama und Papa!

Table of Contents

Abstract.....	ii
Preface.....	iii
Table of Contents	iv
List of Figures.....	vi
General Introduction	1
Chapter 1: Attentional Tracking Takes Place Over Perceived Rather Than Veridical Positions	9
Abstract.....	10
Introduction.....	11
Method	16
Results	21
Discussion	24
Supplementary Information	29
Chapter 2: Smooth Pursuit Operates Over Perceived not Physical Positions of the Double-Drift Stimulus	30
Abstract.....	31
Introduction.....	32
Method	38
Results	41
Discussion	48
Supplementary Information	54
Chapter 3: Hemifield Bias in Load Dependent Activity During Multiple Object Tracking.....	56
Abstract.....	57
Introduction.....	58
Method	63
Results	70

Discussion	85
Conclusions	91
General Discussion.....	92
Conclusions	97
References.....	98

List of Figures

Figure 1. Schematic of the gain map metaphor.	4
Figure 2. Stimulus schematic.	14
Figure 3. Results from Pre-Test and Tracking Experiment.	23
Figure 4. Schematic of the double-drift.	33
Figure 5. Schematic of the stimulus.	37
Figure 6. Average individual and group eye movement traces.	42
Figure 7. Differences in direction of eye traces with and without internal drift.	43
Figure 8. Scatter plot of the x-component of saccade vectors.	47
Figure 9. Visualization of catch-up saccades during smooth pursuit.	55
Figure 10. Task schematic showing a single trial.	62
Figure 11. ROIs projected onto the inflated fsaverage brain.	69
Figure 12. Performance in the tracking task by condition.	71
Figure 13. Average fixation heatmaps for each condition.	73
Figure 14. Statistical map comparing tracking to passive viewing.	74
Figure 15. Statistical map of the canonical load-dependent ARF.	76
Figure 16. Statistical map of parametric modulation with contralateral load.	78
Figure 17. Statistical map of parametric modulation with ipsilateral load.	79
Figure 18. Statistical map comparing modulation with contralateral load to ipsilateral load.	80
Figure 19. Determining the effect of adding targets separately to each hemifield.	83
Figure 20. Change in BOLD signal caused by increasing load by one target.	84

General Introduction

Sensory systems provide the brain with more information than can fully be processed at once. Prioritization, selection, and filtering of the available sensory signals are fundamental functions of the brain, which enable flexible behavior in dynamic environments. These functions, collectively referred to as attention, have been studied in psychology and neuroscience for decades (Carrasco, 2011; Cavanagh & Alvarez, 2005). One common task that participants in such studies perform, is to keep track of the locations of one or more moving objects, sometimes among similar or identical distractors (Scholl, 2009; Holcombe, 2023). An important question in this area of research has to do with what is being tracked. This dissertation consists of three studies that investigated the nature of tracked object representations.

Attentional tracking

Object tracking is a fundamental visual process that allows us to monitor and follow multiple moving objects in dynamic visual environments. In a typical attentional tracking task, participants are presented with several identical objects and instructed to keep track of a subset of them while they all move. The number of to-be-tracked targets can vary from one to eight or more (e.g., Alvarez & Franconeri, 2007), and the task becomes increasingly difficult the more targets are attentionally tracked. Initially, it was believed that there was a fixed capacity limit of around four targets that could be successfully tracked (Pylyshyn & Storm, 1988; Holcombe, 2023). Challenges to this notion have revealed that the capacity limit of tracking is not a fixed number. Instead, many factors affect tracking performance, including the number, speed, and spatial

separation of targets and distractors (Franconeri, Lin, Enns, Pylyshyn & Fisher, 2008; Holcombe & Chen, 2012; Intriligator & Cavanagh, 2001; Störmer, Alvarez & Cavanagh, 2014).

Because of the combined effects of these factors, tracking serves as a valuable paradigm for studying multiple aspects of attention simultaneously, such as its resolution, speed, and capacity (Scholl, 2009). Instead of a fixed capacity of four targets, the currently available data are most consistent with a resource limitation model for tracking (Holcombe, 2023). For example, if the speed of objects is high, the trackable number of targets is lower, and vice versa, because faster targets require more resources to be tracked than slower ones (Franconeri, Lin, Enns, Pylyshyn & Fisher, 2008; Holcombe & Chen 2012; Holcombe, 2023). Surprisingly, the availability of resources almost doubles when targets are spread out across the left and right hemifield, rather than being confined to just one hemifield (Alvarez & Cavanagh, Holcombe & Chen, 2012; Störmer, Alvarez & Cavanagh, 2014).

Spatial attention as a spotlight on gain maps

So-called “gain maps” serve as a commonly used metaphor for attention. Such maps implement increased salience or attentional control at certain hotspots and project to corresponding locations in other areas, where they increase or suppress gain (Itti & Koch, 2001; Cavanagh, Caplovitz, Lytchenko, Maechler, Tse & Sheinberg, 2023). According to this framework, attention can be likened to a spotlight that leads to either enhancement or suppression (i.e., changes in gain) in specific regions of these maps. This is not just spatial, as gain maps even exist for features. Feature-based attention operates

over the entire visual field, enhancing relevant features and suppressing similar features while leaving more dissimilar features unchanged (Störmer & Alvarez, 2014). In the case of spatial attention, these spotlights act as focal points, akin to the fovea, where more detailed processing of visual information occurs but is limited to a small portion of the visual field. One question in attention research pertains to the layout and coordinate system of gain maps, which play a crucial role in determining the targets that attention selects for tracking.

Several coordinate systems have been described in the literature. Firstly, retinotopy is a reference frame that is based on the physical location of objects relative to the retina. Retinotopic maps exist throughout cortex, with some of them being additionally constrained to a single hemifield, especially in earlier visual areas (Wang, Mruczek, Arcaro & Kastner, 2015). Secondly, spatiotopy is a coordinate system, which represents objects in absolute locations in space that are independent of the viewer's gaze. Lastly, the perceptual coordinate system encodes the inferred locations of objects rather than being solely determined by their detected physical positions. This might seem counterintuitive, but even ostensibly basic properties, such as an object's position, can be subject to illusory shifts which can make them appear to be in radically different locations than their retinotopic positions (Cavanagh & Anstis, 2013; Cavanagh & Tse, 2019; Cavanagh et al., 2022; De Valois & De Valois, 1991; Tse & Hsieh, 2006).

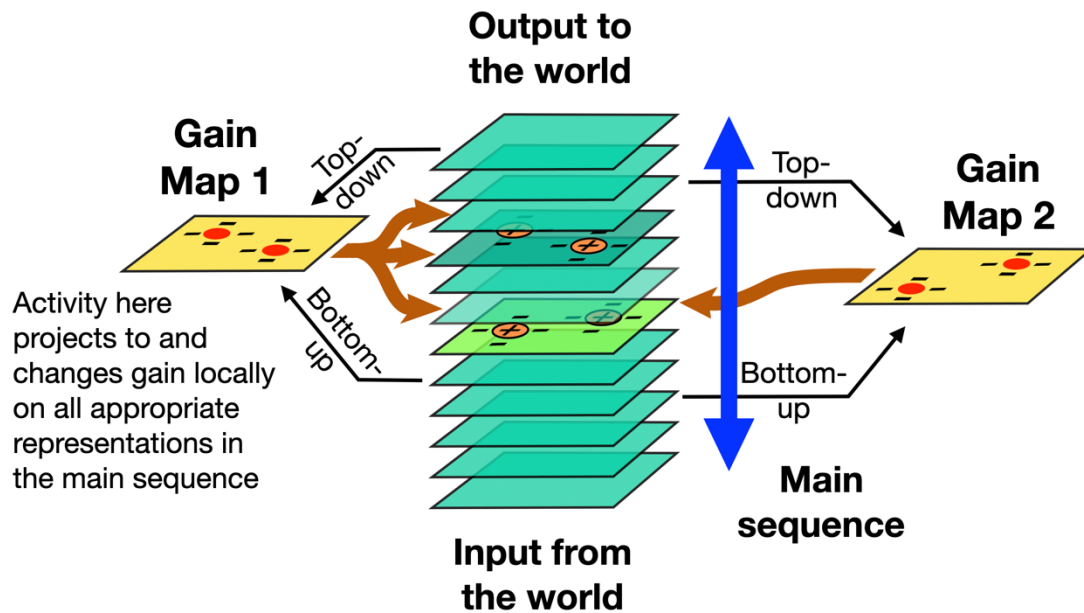


Figure 1. Schematic of the gain map metaphor. Peaks at attended locations (with suppressive surround) project throughout the visual system to change gain. It is unclear how many such gain maps exist or how many coordinate systems there might be. Reproduced from Cavanagh and colleagues (2023) with permission.

Understanding the relationship between attention and the different coordinate systems of gain maps is crucial for elucidating how attention selects targets for tracking. Spatial attention could hypothetically operate at any level of the visual processing stream. However, different maps and coordinate systems impose different challenges and bottlenecks for attention. Ideally, attention would track spatial locations and objects based on those representations that are most helpful for the task at hand, but previous research as well as the studies described in this dissertation show that this is not always the case. Exogenous attention to the location of peripheral cues is often found to be in retinotopic coordinates (Golomb, Nguyen-Phuc, Mazer, McCarthy & Chun, 2010;

Golomb, Chun & Mazer, 2008; Golomb, Pulido, Albrecht, Chun & Mazer, 2010). On the other hand, the locus of voluntary spatial attention can be preserved in spatiotopic coordinates via remapping from the current retinal location of the target to its future location after an upcoming eye movement (Rolfs, Jonikaitis, Deubel & Cavanagh, 2011).

Summary of the Studies

Chapters One and Two are concerned with the coordinate systems of tracking. In a study investigating tracking with regards to spatiotopic and retinotopic coordinates (Howe, Drew, Pinto & Horowitz, 2011) participants were required to execute saccades while tracking multiple objects. Notably, when the display shifted with the eye movement, preserving retinotopic locations, the tracking performance of participants was disrupted. However, intriguingly, when the spatiotopic locations of the objects were preserved across saccades, meaning that nothing shifted on the screen during the eye movement, participants exhibited improved tracking performance.

Chapter One extends the finding that attentional tracking operates on spatiotopic coordinates to the question of whether attentional tracking selects targets from perceived or physical coordinates. Several illusions have provided evidence that although positions are explicitly encoded on the retina, perceived object positions are inferred based on various factors, including motion (Cavanagh & Anstis, 2013; Cavanagh & Tse, 2019; Cavanagh et al., 2022; De Valois & De Valois, 1991; Tse & Hsieh, 2006). We dissociated perceived and physical coordinates of tracking targets and distractors using the double-drift illusion, in which the motion direction of a Gabor patch gets combined with the motion direction of its internal grating to cause large mislocalizations of the

Gabor's perceived position. Targets and distractors coming too close to each other is one source of errors during attentional tracking (Holcombe, 2023; Cavanagh & Alvarez, 2005; Intrilligator & Cavanagh, 2001). In the gain map metaphor, this would mean that when a target and a distractor coinhabit the space that is enhanced by a spotlight, participants lose track of that target. To avoid this, a minimum target distractor spacing must be maintained, which depends among other things on eccentricity (Intrilligator & Cavanagh, 2001). The experiment described in Chapter One shows that the crucial target-distractor distance is in perceptual (illusory) coordinates rather than physical coordinates. While keeping physical distances between targets and distractors constant, we manipulated their perceived distances with illusory position shifts. Despite equal physical target distractor spacing in both conditions, performance was better with increased perceived distance and worse with decreased perceived distances.

Chapter Two is concerned with the coordinate system of overt tracking using smooth pursuit. Eye movements are clearly and deeply linked to spatial attention (Carrasco, 2011). Most studies linking oculomotor systems to attention use saccades and compare them to ballistic shifts of attention (e.g., Zirnsak & Moore, 2017). Fewer studies have investigated the link between smooth pursuit and continuous attentional shifts like attentional tracking, but there is some evidence that they are also linked. It has been shown that a spotlight of spatial attention moves with the pursued object, slightly ahead of the fovea during pursuit (Chen, Valsecchi & Gegenfurtner, 2017; Lovejoy, Fowler & Krauzlis, 2009). Additionally, demanding secondary tasks that deplete attentional resources can interfere with smooth pursuit, suggesting that they rely on a shared resource pool (Hutton & Tegally, 2005).

In this study, we used the same illusion as in Chapter One and investigated whether parafoveal smooth pursuit (i.e., tracking of the midpoint between two stimuli) would use perceived or physical locations to select the pursuit target. We found that smooth pursuit, like attention, selects targets from perceived coordinates. By showing that attentional tracking and smooth pursuit use the same coordinate system, we provide more circumstantial evidence for a link between pursuit and attentional tracking.

Chapter Three describes an investigation into brain activity during attentional tracking with a focus on hemisphere- and hemifield-specific effects. Participants performed a multiple object tracking task, while the number of targets (load) was independently varied in each hemifield. While attentional tracking becomes harder when the number of targets is increased, blood-oxygen-level-dependent (BOLD) activity increases correspondingly in many brain areas (Culham, Cavanagh & Kanwisher, 2001; Jovicich, Peters, Koch, Braun, Chang & Ernst, 2001; Shim, Alvarez, Vickery & Jiang, 2010; Jahn, Wendt, Lotze, Papenmeier & Huff, 2012; Alnæs, Sneve, Espeseth, Endestad, van de Pavert & Laeng, 2014; Nummenmaa, Oksama, Glerean & Hyönä, 2017). Hemisphere specificity has been established as a characteristic of attentional tracking resources. The number of trackable targets (Alvarez & Cavanagh, 2005) and their speed (Störmer, Alvarez & Cavanagh, 2014) show significantly greater capacity when targets are distributed across both visual hemifields rather than confined to one. Furthermore, tracking performance declines when two targets cross the vertical meridian simultaneously (Strong & Alvarez, 2020). Collectively, these findings provide substantial behavioral evidence supporting the existence of neural processes during tracking that are exclusive to a single hemifield and hemisphere. We show that the modulation of BOLD

activity in earlier visual areas is also driven more by contralateral than ipsilateral load, showing hemifield specificity. In contrast, later areas in the visual hierarchy in the parietal and frontal lobes associated with attention, modulate their activity significantly with load in both hemifields.

Chapter 1: Attentional Tracking Takes Place Over Perceived Rather Than Veridical Positions

Authors: Marvin R. Maechler, Patrick Cavanagh & Peter U. Tse

Author Contributions: M.R.M and P.U.T. designed the study, M.R.M. wrote the experiment code, and collected and analyzed the data. P.U.T. and P.C. oversaw analyses, M.R.M., P.C., and P.U.T wrote the paper. All authors provided critical feedback and helped shape the research, analysis, and manuscript.

Published: Maechler, M. R., Cavanagh, P., & Tse, P. U. (2021). Attentional tracking takes place over perceived rather than veridical positions. *Attention, Perception, & Psychophysics*, 83(4), 1455-1462. <https://doi.org/10.3758/s13414-020-02214-9>

The author holds copyright over the included publication.

Abstract

Illusions can induce striking differences between perception and retinal input. For instance, a static Gabor with a moving internal texture appears to be shifted in the direction of its internal motion, a shift that increases dramatically when the Gabor itself is also in motion. Here we ask whether attention operates on the perceptual or physical location of this stimulus. To do so we generated an attentional tracking task where participants ($N=15$) had to keep track of a single target among three Gabors that rotated around a common center in the periphery. During tracking, the illusion was used to make three Gabors appear either shifted away from or toward one another while maintaining the same physical separation. Because tracking performance depends in part on target to distractor spacing, if attention selects targets from perceived positions, performance should be better when the Gabors appear further apart and worse when they appear closer together. We find that tracking performance is superior with greater perceived separation, implying that attentional tracking operates over perceived rather than physical positions.

Introduction

Visual percepts are radically different from the pattern of light that activates photoreceptors in the retina. For example, we perceive a 3D world of objects that are in motion relative to ourselves and to each other. However, the image on the retina is 2D, with no explicit representation of motion, objects, materials or depth. Motion, material, shape, scene layout and causation must be constructed on the basis of computations over the sequence of 2D images in the two eyes guided by assumptions about the image-to-world mapping.

Even seemingly low-level properties, such as position are not simply detected; an object's perceived position can appear radically shifted from its retinotopic position (Cavanagh & Anstis, 2013; Ramachandran & Anstis, 1990; De Valois & De Valois, 1991; Tse & Hsieh, 2006). In particular, when the internal grating of a peripherally moving Gabor patch drifts in a direction orthogonal to its envelope's path, an observer perceives the Gabor's trajectory to be tilted in the direction of the internal drift. Such a doubly drifting Gabor might appear to be moving in a direction 45 degrees offset from its actual path. This illusion has been known variously as the infinite regress illusion (Tse & Hsieh, 2006), the curveball illusion (Shapiro, Lu, Huang, Knight, & Ennis, 2010; Gurnsey & Biard, 2012; Kwon, Tadin & Knill, 2015; Ueda, Abekawa, & Gomi, 2018), and the double-drift illusion (Lisi & Cavanagh, 2015, Liu, Tse & Cavanagh, 2018; Liu, Yu, Tse, & Cavanagh, 2019; Cavanagh & Tse, 2019). Here we will use the latter name.

The double-drift illusion allows a dissociation of perceived and physical locations. By 'perceived position', we mean where an object appears to be in the world as consciously experienced by the observer. By 'physical position' or 'veridical position' we

mean the position of the object in the world or the stimulus on the screen. The visual system generally represents object positions correctly, namely we see them at their physical location in the world, but illusions like the double-drift can induce differences between physical positions and perceived positions.

In this study we used the double-drift illusion to induce large differences between perceived and physical positions in order to investigate whether attentional tracking operates over perceptual or physical position representations. Recent results have ruled out early visual cortex as the origin of the double-drift illusion (Liu, Tse & Cavanagh, 2018; Cavanagh & Tse, 2019) and even suggest that it might arise outside of visual cortex entirely, perhaps in frontoparietal regions (Liu, Yu, Tse, & Cavanagh, 2019).

A similar paradigm, albeit with static Gabor envelopes, was used by Maus and colleagues (Maus, Fischer & Whitney, 2011) as well as Dakin and colleagues (Dakin, Greenwood, Carlson & Bex, 2011) to investigate the effect of perceived position shifts on crowding. When reporting the orientation of a Gabor patch that is flanked by other Gabor patches, crowding will reduce performance when the distance to the flankers is less than about $\frac{1}{2}$ of the eccentricity of the target. In these experiments, distances were fixed in physical coordinates, but the perceived distance was altered by drifting the internal textures of the Gabors toward or away from the target (De Valois & De Valois, 1991). When the perceived spacing was decreased due to the internal drift, crowding increased, suggesting that the region of crowding is defined over perceived positions (Dakin, Greenwood, Carlson & Bex, 2011; Maus, Fischer & Whitney, 2011).

We build on these results here in order to examine the spatial representation accessed by attention in a tracking task using the double-drift illusion to introduce a

dissociation between perceived and physical target distractor spacing (Fig. 1).

Performance in multiple object tracking tasks depends on, among other factors, the distance between target and distractors (review, Cavanagh & Alvarez, 2005). If the tracking of target locations during attentional tracking operates over representations in physical coordinates, no difference between the conditions should be observed. However, if attentional tracking operates on perceived object positions, tracking should be easier in the condition with increased perceived distance than in the condition with decreased perceived distance.

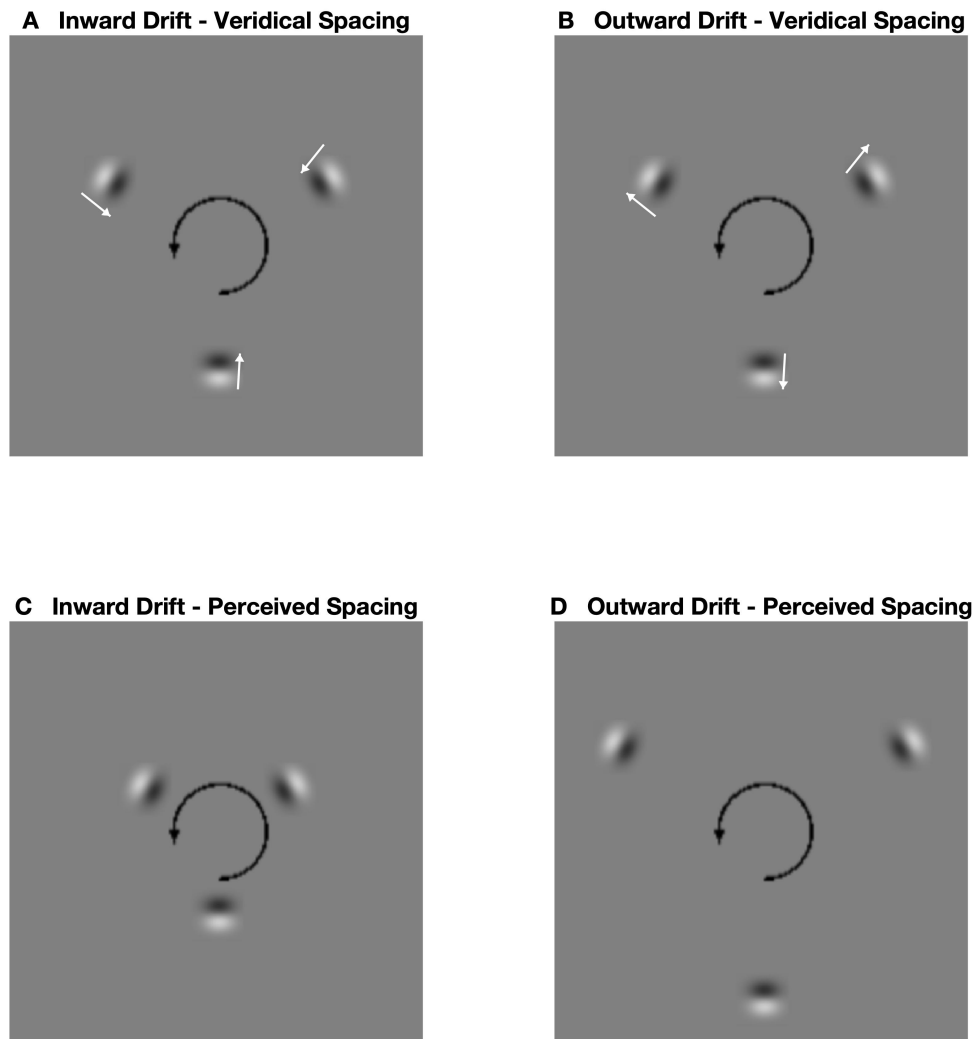


Figure 2. Stimulus schematic. The Gabors rotated around a stationary point in the periphery. Internal drift (white arrows) made them appear closer to each other in the inward drift case (A & C) and farther from each other in the outward drift case (B & D). White arrows show internal drift, while black arrows indicate the Gabors' envelope motion. After the staircase procedure to determine a baseline spacing for each participant, the physical distance was the same on all trials, while the perceptual distance varied according to the drift condition.

When a target and a distractor come too close to each other, it is not possible to individuate them, and the observer is more likely to lose track of the target (Cavanagh & Alvarez, 2005). This zone of interaction or pooling has been linked to the general phenomenon of crowding (He, Cavanagh, & Intriligator, 1996; Pelli, 2008). If two objects fall within the radius of the pooling region, their features are mixed and cannot be further individuated.

The present study addresses the question of whether the distance that limits target selection in attentional tracking is based on physical distance or perceptual distance. We find that perceptual, and not physical, target distractor spacing underlies tracking performance, demonstrating that attentional tracking operates in perceptual coordinates. This could either occur at a stage prior to visual consciousness, if illusory shifts have already emerged at an earlier level, or at the stage of visual consciousness itself. In either case, attentional tracking operates on object locations after they have been converted to perceptual positions.

Method

Participants

We recruited 15 participants (10 women, 5 men, age range: 18-30, mean age: 22.4 years \pm 4 years) from the Dartmouth community and reimbursed them with \$10. Their vision was normal or corrected-to-normal. Participants volunteered and gave informed consent. The experimental protocol was approved by the Institutional Review Board at Dartmouth College.

Apparatus

Participants sat alone in a dark testing room, facing an LCD screen (15'' wide, 1280 x 1024, 60 Hz). A chinrest was used in order to hold the distance to the screen constant at 57 cm. Stimuli were created in MATLAB (The MathWorks, Natick, MA, USA) using the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997).

Eye movements and pupil dilation of the right eye were monitored with a head-mounted eye-tracker (EyeLink 2, SR Research, Oakville, ON, Canada; 500 Hz sampling rate). This eye-tracking system is unable to measure the pupil diameter in absolute units. Instead the values it returns depend on individual factors like pupil position and the distance between camera and eye. In order to normalize the pupil dilation values across subjects, they were z-scored within each recording session.

Stimuli

The display consisted of a white fixation point in the middle of the screen (0.2 dva diameter) and Gabor patches on a uniform gray background. The Gabor patches – sinusoidal gratings with a Gaussian envelope, ($\sigma_{\text{env}} = 0.1$ degrees visual angle [dva]), and

a spatial frequency of 2 cycles per degree – served as stimuli for all experiments described below.

Gabors were presented at the three corners of an imaginary equilateral triangle (compare Figure 2). The Gabor patches were oriented, such that their internal grating was orthogonal to a hypothetical line from each Gabor's center to the triangle's center. This orientation was chosen to maximize the effect of internal drift on the perceived distance between them. The internal drift of the grating (4 Hz) of all three Gabors was moving towards their common center in the *inward* condition and directly away from it in the *outward* condition. In all experiments, the Gabor triangles rotated around their center, which was 8 dva away from fixation, with an angular velocity of 180°/s. Distances between Gabors here always refer to center to center distance, not to the gray space between visible parts of the Gabor.

Pre-test

In order to determine by how much internal drift would change the perceived distance between Gabors in our specific configuration, we devised a quick perceptual pre-test utilizing the method of constant stimuli. Participants were asked to compare two of the triangular Gabor configurations described above. One of them consisted of Gabors with drift toward the virtual triangle center (*inward*), while the other Gabors drifted away from their virtual center (*outward*). While participants fixated in the middle of the screen, one triangle of Gabors appeared in the top right and the other in the bottom left quadrant of the screen. *Inward* and *outward* drifting Gabor triangles were presented pseudo-randomly interleaved on the left and right side of the screen. Participants were tasked to report whether the spacing of the left or right triangle was wider in a two-alternative

forced choice (2AFC) design. Their responses were then recoded to mean whether the *outward* or the *inward* drifting triangle was perceived as more widely spaced.

All triangles were equilateral, but the physical distances between *inward* drifting Gabors were larger than those between the *outward* drifting Gabors by either 0, 0.264, 0.498, 0.762, 1.025, or 1.26 dva. The spacing between Gabors for the *outward* drifting Gabors was always 1.2 dva. Participants completed 144 trials (i.e. 24 trials per spacing difference). A psychometric function of these spacing differences can be fit to the frequency with which the *outward* drifting triangle was perceived to be farther apart. The point of subjective equality (PSE) on this psychometric curve is then equivalent to the average perceived distance shift caused by the internal drift of the Gabors.

Tracking Experiment

We presented participants with one triangular configuration of drifting Gabors, which appeared pseudo-randomly in one of the four quadrants of the screen. Participants were tasked with tracking one Gabor, while ignoring the other two. In contrast to most other tracking paradigms, where targets and distractors move in seemingly random directions, here they always rotated smoothly around the center of the equilateral triangle (Alvarez & Cavanagh, 2005; Holcombe, Chen & Howe, 2014; Störmer, Alvarez & Cavanagh, 2014). In order to discourage tracking strategies involving only the start and end positions of the Gabors rather than actual tracking, the duration of motion varied from trial to trial. The physical distance between the three Gabor patches was first set to give an 80% tracking accuracy for the *outward* drifting condition using a staircase procedure and so it varied from participant to participant. This distance was then used to determine performance with both inward and outward internal drift.

The experiment consisted of a total of 120 trials, split evenly into three blocks. All trials were initiated by a button press, making the procedure entirely self-paced. At the beginning of a trial, the to-be-tracked Gabor was indicated by its offset and onset prior to motion onset. The indicated Gabor flashed for 1 s, marking it as the to-be-tracked target, and after another 250 ms pause, all three Gabors moved (i.e., rotating around their center) for a random duration of between 4.5 and 5.5 seconds. At the end of each trial, one of the Gabors flashed. The subject responded whether or not the same Gabor had flashed in the beginning and end of the trial. The guessing rate was 50% rather than 33.33% because the initially cued object was probed at the end on half the trials, while one of the two non-target Gabors was probed at the end of the other half of the trials. Participants also had to keep their eyes fixated on the central spot in the middle of the screen while they covertly tracked the cued target. Additionally, written instructions were provided at the beginning of every block. After participants had given their response, they received feedback and were informed about the number of trials they had answered correctly so far. The task was completely self-paced as subjects could decide to take a break between any two trials.

The first block of 40 trials was dedicated to a standard one-up/one-down staircasing procedure with the *outward* drifting Gabors where the distance between the Gabor patches was adjusted before every trial depending on the participant's performance in the preceding trial. If they tracked the target correctly, the distance between Gabors decreased by 0.111 dva; if not, it increased by 0.147 dva. The starting distance was 1.5 dva and the minimum distance between Gabors was 0.7 dva to avoid overlap. At the end of the first block, the reversal points of the staircasing procedure were averaged to give

the distance at which a participant should perform with roughly 80% accuracy (Kaernbach, 1991; García-Pérez, 1998). The three Gabor patches were held at this distance in all remaining trials for both *inward* and *outward* drifting Gabors. The following two blocks had an equal number of trials of both conditions (i.e., 20 per block and condition).

The double drift illusion works best in the visual periphery (Tse & Hsieh, 2006). In order to prevent subjects from looking directly at the rotating Gabors, which would disrupt the illusion and help tracking performance, we monitored participants' eye movements during the experiment. Eye tracking data were used to exclude all trials in which participants looked away from fixation by more than 2 dva during tracking (an average 14.5% of trials). Performance was collapsed across blocks two and three, while the staircasing data from block one was analyzed separately.

Results

Pre-Test

If internal drift had no impact on perceived distance between Gabors, participants should have reported equally spaced sets of Gabors as equally far apart for both the inward and outward drift conditions. However, internal drift caused participants to perceive Gabors as farther apart in the *outward* drift case and closer together in the *inward* drift case. At equal spacings, participants on average reported the *inward* drifting Gabors to be farther from each other only 10.3% of the time, which is significantly below the 50% level expected if there were no illusion ($t(14)=16.786$, $p<0.001$, Cohen's $D=6.130$). At the largest difference in Gabor spacings, when *inward* drifting Gabors were 1.26 dva farther apart than *outward* drifting Gabors, participants on average reported the *inward* drifting Gabors to be farther apart 96.9% of the time.

In order to find the point at which these two different sets of Gabors appeared equally spaced for each participant, we fit a Weibull function to the proportion of “*inward* appears farther” responses (see Figure 3 for a representative participant's data). We then used the fitted function to estimate the point of subjective equality (PSE) by finding the spacing at which the function predicts 50% of responses to fall either way. The average PSE across participants was 0.46 dva, which was significantly different from 0 ($t(14)=14.02$, $p<0.001$). Thus, the internal direction of the Gabors shifted the apparent spacing by about 38% so that a set of Gabors with *inward* motion needed to be spaced by 1.66 dva to appear to have the same spacing as the set with *outward* drift that had 1.20 dva spacing.

Tracking Experiment

If attentional tracking operates on perceived positions, tracking should be more difficult when the Gabors' internal grating drifts inward, because the three Gabors are then perceived to be ~38% closer to one another. Indeed, on average, participants had 12.73% more correct trials in the *outward* than in the *inward* condition, which was highly significant ($t(14)=4.415$, $p<.001$, Cohen's $D=1.077$). There were no floor or ceiling effects in either condition. Mean tracking accuracy in the *outward* condition was 79.66% ($SD=12\%$), quite close to the baseline of 80% targeted by the staircase procedure of the first session, while the mean tracking accuracy in the *inward* condition was 66.93% ($SD=12\%$). Compare Figure 3 for more details.

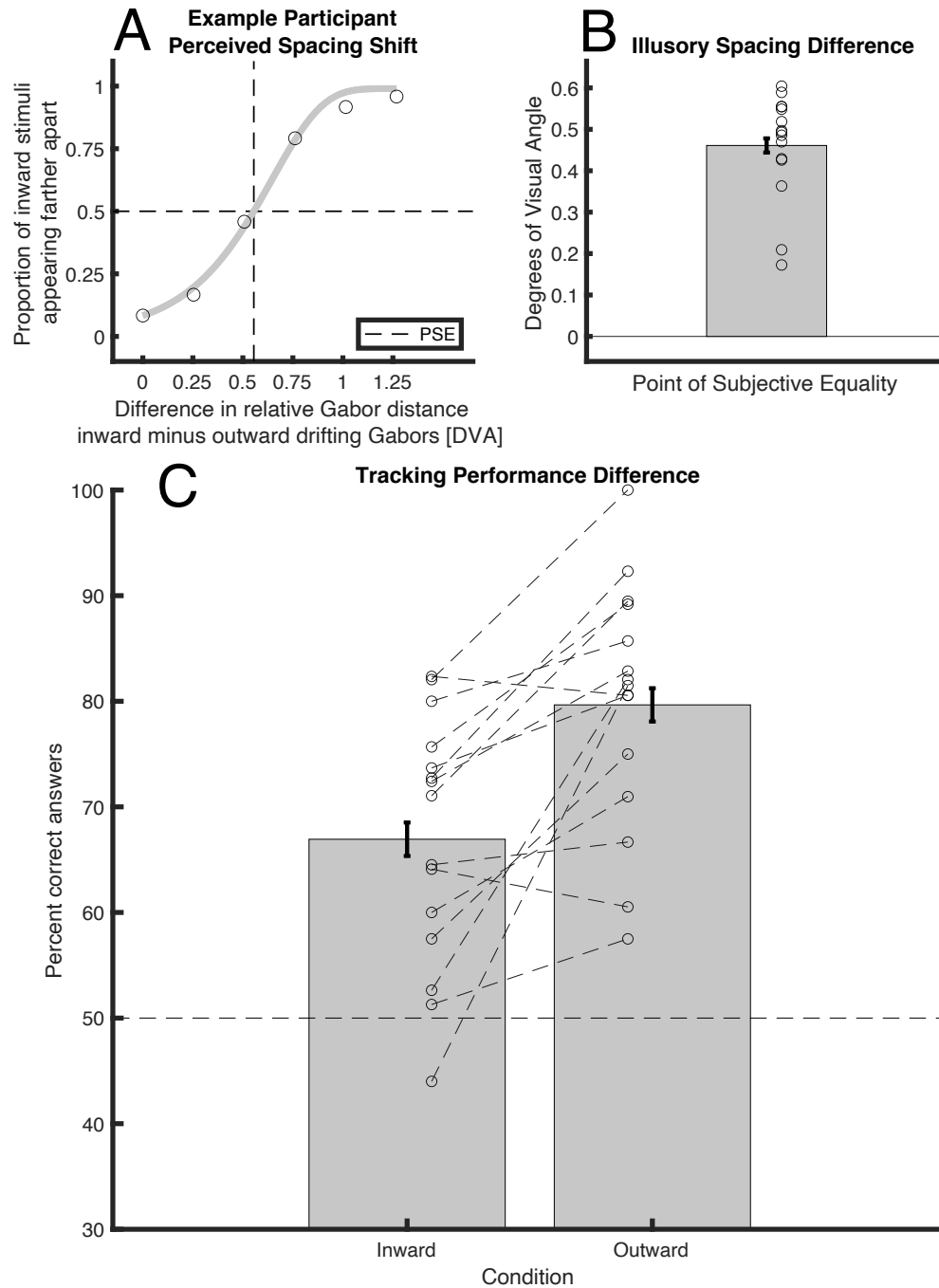


Figure 3. Results from Pre-Test and Tracking Experiment. A: Sample fit of psychometric Weibull function with estimated PSE. B: Average and individual PSEs. C: Average and individual tracking performance for *inward* and *outward* drift conditions. All error bars correspond to the standard error of the mean.

Discussion

Attentional tracking was notably impaired when the double-drift illusion appeared to move the target closer to the two distractors. Since physical target distractor spacing was equal in both conditions, this effect could only have been driven by the perceived distance between targets and distractors. Consequently, our results suggest that the attentional tracking system, which determines the locus of attentional selection, is influenced at some stage by the perceptual representations of target location. Our results support the view that attentional selection and tracking occur late in the visual hierarchy, after the conversion to perceptual representations of object location (He & Nakayama, 1992; Hochstein & Ahissar, 2002; Suzuki & Cavanagh, 1995; Özkan, Tse, & Cavanagh, 2020). Our data do not support models of attentional tracking that describe tracking solely as an encapsulated, low-level visual process (e.g., the ‘Fingers of Instantiation’ theory, Pylyshyn & Storm, 1988). This aligns with data that instead support higher level mechanisms of attentional tracking (Cavanagh & Alvarez, 2005; Oksama & Hyönä, 2004). Our account of attentional tracking as operating on high-level perceptual representations is also in line with recent findings linking attentional tracking ability with other higher cognitive processes (Tullo, Faubert & Bertone, 2018).

Our results suggest that attentional tracking operates primarily on representations in the perceptual coordinates of conscious vision, and quite plausibly on the content of conscious perception itself. It seems that attentional tracking is unable to operate solely on the retinotopic representations that are used to construct conscious percepts (He & Nakayama, 1992; Hochstein & Ahissar, 2002, Suzuki & Cavanagh, 1995). This is consistent with the recent finding that pop-out in a visual search paradigm happens only

among perceptual rather than stimulus-level double-drift oddballs (Özkan, Tse & Cavanagh, 2020). Similarly, it has been shown that crowding does not affect representations in early visual areas (He, Cavanagh & Intriligator, 1996) and instead operates over perceived positions (Dakin, Greenwood, Carlson & Bex, 2011; Maus, Fischer & Whitney, 2011).

Our results are also consistent with attentional tracking performance that is seen when saccades are involved (Howe et al., 2011). Participants in this study had to execute multiple saccades while attentionally tracking. When the display shifted with the eye movement, preserving retinotopic locations, tracking was disrupted. However, when spatiotopic locations of the objects were preserved across saccades (i.e., nothing shifted on the screen during the saccade) participants performed better. These results show that it is possible to execute saccades while tracking, provided that the spatiotopic target locations are preserved. This suggests that attention tracks targets in their spatiotopic locations, corrected for eye movements, ruling out retinotopic locations. Our results here go farther and demonstrate that attentional tracking selects from target representations that include illusory perceptual shifts.

Although our results clearly show that tracking in our task is influenced by perceptual coordinates, two other findings with the double-drift stimulus have suggested that attention might be in physical coordinates. First, Lisi and Cavanagh (2015) showed that saccades to the drifting Gabor land along a line parallel to the physical path of the Gabor rather than its perceived path. To the extent that spatial attention is linked to saccades (e.g., Awh, Armstrong & Moore, 2006), we might expect attention to also be unaffected by the double-drift illusion. Second, in an fMRI study, Liu and colleagues

(2019) found that activity in early visual areas allowed the decoding of the physical but not the perceptual positions. Liu and colleagues suggested that the attentional feedback went to the physical location, not the perceived location, because attention, like saccades would be immune to the illusion.

The saccade results of Lisi and Cavanagh (2015) were replicated by Nakayama and Holcombe (2020). However, they also found that irrelevant transients (e.g., a flash of light that would grab attention) also reset the illusion, bringing the perceived location back to the physical location. According to these authors, saccades targeted the physical location not because saccades were immune to the illusion but because the attention drawn by the saccade had reset the illusion. Due to this, the perceived location would be the same as the physical location at the time of the saccade. Their result challenges the claim that attention might be in physical coordinates. It also raises a question about whether the effects of attention during tracking in our task might eliminate the illusion. We can reject that notion on two grounds. First, attentional tracking is more like smooth pursuit than a saccade (Howe, Drew, Pinto & Horowitz, 2011; Howe, Pinto & Horowitz, 2010), and smooth pursuit has been shown not to affect the illusion (Cavanagh & Tse, 2019). Second, we did measure a significant effect of the illusion on the apparent size of the rotating trajectory and on the performance in tracking, which would have been impossible if attention had reset the illusion. Although ballistic attentional shifts may reset the illusion (Nakayama & Holcombe, 2020), we demonstrate that smooth attentional shifts do not.

Why do our results demonstrate that attentional tracking operates in perceptual coordinates while Liu and colleagues (2019) suggest it operates in physical coordinates?

There are a number of possibilities, and we outline three here. First, it may be that when tracking is disrupted, the participants must rely on memory to recover the target, no matter which coordinate system attention operates on during tracking. Memory of location is most likely in perceptual coordinates, at least, we know that memory saccades are influenced by the double-drift illusion (Massendari, Lisi, Collins & Cavanagh, 2018) even if immediate saccades are not. In this event, the remembered (perceptual) location in the outward motion case would be less crowded than in the inward motion case.

Second, the illusory shifts produced by our stimulus may be in a smaller range that does affect saccades and so by inference, attention. Specifically, the perceptual shifts we find with rotating Gabors are smaller than have been reported in other double-drift studies using linear trajectories. Possibly the presence of continual rotation or curved trajectories leads to a saturated offset or limited accumulation of offset errors. Whatever the reason, the offset we find is closer to that seen for static Gabors with internal drift (De Valois & De Valois, 1991). Notably, for offsets of this size with stationary Gabors with internal drift, saccades target the perceived positions (Shafer & Moore, 2006; Kosovicheva, Wolfe & Whitney, 2016). Indeed, in Lisi and Cavanagh's (2015) study, as well as in Nakayama and Holcombe's (2020), although saccades did not follow the increasing offset of the illusory path, they did show a constant, small shift at all locations. This constant offset in the direction of the internal motion was similar in magnitude to that seen for static Gabors. In contrast to the double-drift, the effects of this smaller position shift can be observed in V1 using fMRI (Whitney et al. 2003; Schneider, Marquardt, Sengupta, de Martino, & Goebel, 2019), although it is paradoxically in the opposite direction from the perceived shift.

Finally, there is a hybrid alternative. Attention may be guided to the physical location but then must select the target from a map that carries perceived locations. The two locations are not very far apart, and the perceived location would fall within the selection zone of attention (Intriligator & Cavanagh, 2001) so that it can be tracked. Nevertheless, the perceived locations are influenced by the illusion and so are closer together in the inward condition leading to a higher probability of interference. In the end, our data do not resolve which of these accounts is correct. These unanswered questions need to be addressed by future research.

A number of motion illusions demonstrate that perceived object positions are not merely detected but rather constructed depending on – among other things – an object's motion (Tse & Hsieh, 2006; Cavanagh & Anstis, 2013; De Valois & De Valois, 1991; Cavanagh & Tse, 2019). Positions are, however, explicitly encoded on the retina. At some stage between these two extremes of the visual processing hierarchy there must be a conversion of position information from retinotopic to perceptual object locations. We do not fully understand where or how this occurs but our data show that attentional tracking does not operate only on the early representations that would still be in physical coordinates. Instead, we show that its processing bottleneck is at a later stage where positions are coded in perceptual coordinates. This, together with the evidence from visual search (Özkan et al., 2020), suggests that attention accesses perceptual representations at a high level, perhaps at the level of consciousness.

Supplementary Information

A control analysis of the staircase data confirmed that changes in physical target-distractor spacing have the same effect on tracking performance as the illusion. The physical spacing of the *outward* drifting stimulus was adjusted during the staircase after each trial. For each participant, we separated the trials of the staircase procedure into two groups: one where spacing was larger than the median and one where it was smaller than or equal to the median. The average spacing difference between these groupings was 0.4 dva, while the effect of the illusion was estimated to be 0.46 dva. Average performance in the two kinds of staircase trials was 79.34% and 66.32% for the wider and smaller spacings respectively. This was very similar to the performance in the main experiment (79.66% and 66.93%, for *outward* and *inward* respectively).

Chapter 2: Smooth Pursuit Operates Over Perceived not Physical Positions of the Double-Drift Stimulus

Authors: Marvin R. Maechler, Nathan H. Heller, Matteo Lisi, Patrick Cavanagh, Peter U. Tse

Author Contributions: ML and PC conceived of the presented idea, MRM, NHH, ML, PUT, and PC designed the study, ML and MRM wrote the stimulus code, MRM collected the data, MRM analyzed the data, MRM and NH visualized the data, ML, NHH, PC, and PUT oversaw analyses, MRM and NHH wrote the first draft of the paper. All authors provided critical feedback and helped shape the research, analysis, and manuscript.

Published: Maechler, M. R., Heller, N. H., Lisi, M., Cavanagh, P., & Tse, P. U. (2021). Smooth pursuit operates over perceived not physical positions of the double-drift stimulus. *Journal of Vision*, 21(11), 6. <https://doi.org/10.1167/jov.21.11.6>

The author holds copyright over the included publication.

Abstract

The double-drift illusion produces a large deviation in perceived direction that strongly dissociates physical position from perceived position. Surprisingly, saccades do not appear to be affected by the illusion (Lisi & Cavanagh, 2015). When targeting a double-drift stimulus, the saccade system is driven by retinal rather than perceived position. Here, using paired double-drift targets, we test whether the smooth pursuit system is driven by perceived or physical position. Participants ($n=7$) smoothly pursued the inferred midpoint (Steinbach, 1976) between two horizontally aligned Gabor patches that were separated by 20° and moving on parallel, oblique paths. On the first half of each trial, the Gabors' internal textures were static while both drifted obliquely downward. On the second half of each trial, while the envelope moved obliquely upward, the internal texture drifted orthogonally to the envelope's motion, producing a large perceived deviation from the downward path even though the upward and downward trajectories always followed the same physical path but in opposite directions. We find that smooth pursuit eye movements accurately followed the non-illusory downward path of the midpoint between the two Gabors, but then followed the illusory rather than the physical trajectory on the upward return. Thus, virtual targets for smooth pursuit are derived from perceived rather than retinal coordinates.

Introduction

Motion-induced position shifts are a class of illusion where the presence of motion causes the misperception of position (Cavanagh & Anstis, 2013; Dunker, 1929; Eagleman & Sejnowski, 2007; Wallach et al., 1978; Whitney, 2002). These illusions have been used to investigate ways in which different sources of low-level stimulus information are combined in the visual system prior to the formation of a conscious representation. A particularly striking example of this class of illusion is the double-drift stimulus (Lisi & Cavanagh, 2015; also known as ‘the curveball illusion’, Kwon et al., 2015; Shapiro et al., 2010; and originally the ‘infinite regress illusion’, Tse & Hsieh, 2006). The double-drift stimulus contains two sources of motion information: an envelope that translates across the screen (i.e., the ‘external drift’) and a moving visual texture that is confined within the envelope (i.e., the ‘internal drift’). When these two sources of motion are oriented orthogonally, the visual system combines them to produce an intermediate motion percept. This results in a dramatic misperception of both the position of the stimulus and the direction of its motion when viewed peripherally (see Figure 4). The dissociation between an object’s physical position (i.e., where it is encoded on the retina) and its perceptual position (i.e., where it is represented in consciousness) provides a probe that can be used to determine whether physical or perceived position drives downstream cognitive operations and motor actions.

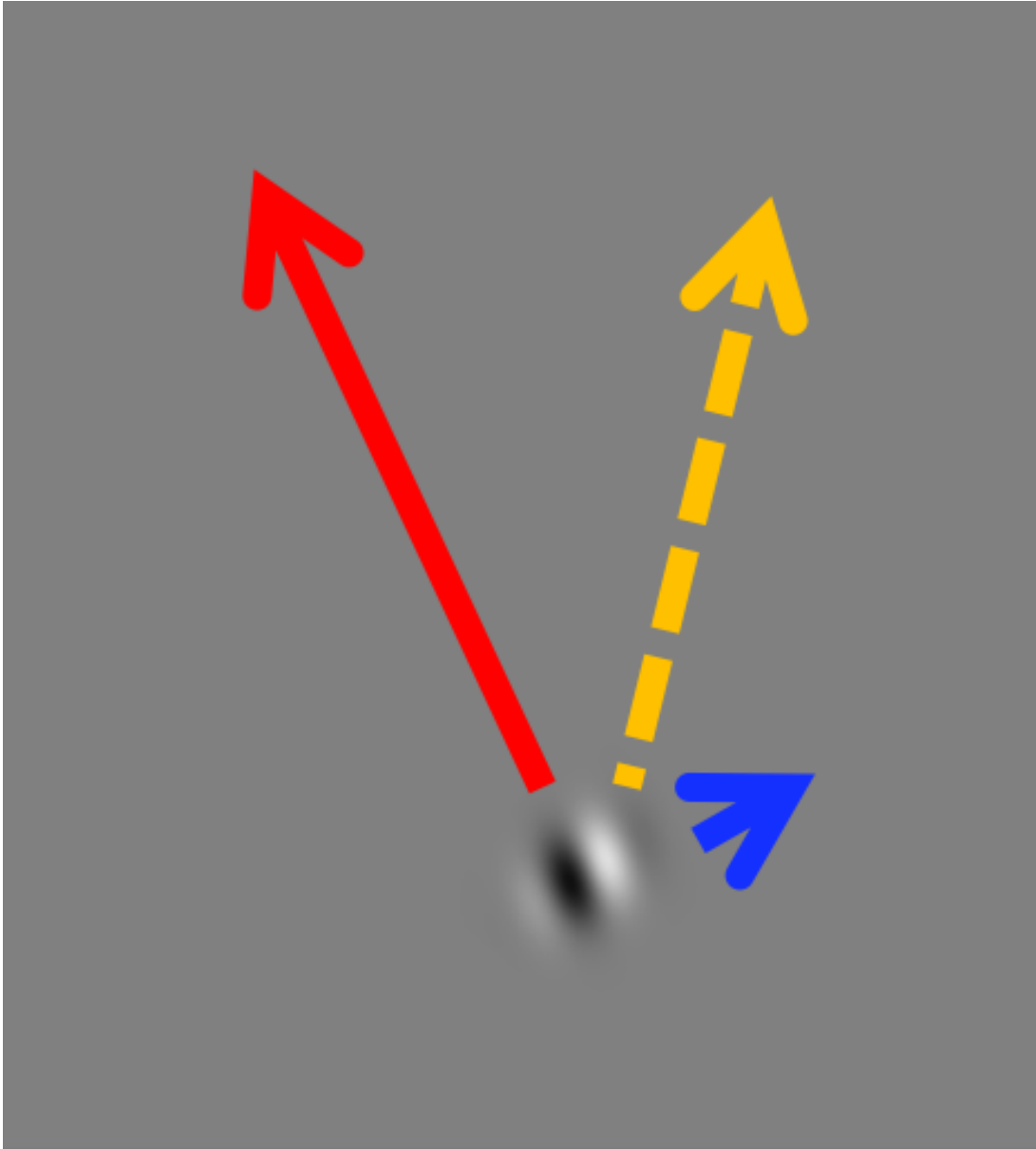


Figure 4. Schematic of the double-drift. When viewed peripherally, a Gabor patch moving over a gray background will appear to travel in the direction (yellow arrow) that is a combination of its external direction (red arrow) and the direction of its internal grating (blue arrow).

Using the double-drift stimulus as a probe, Lisi and Cavanagh (2015) found that saccades were made not to the probe's perceived location, but rather to its physical location. This is a dramatic example of an action process that is uncoupled from conscious experience (Goodale & Milner, 1992). Subsequent results have shown that this dissociation between action and consciousness may be unique to saccades that are programmed while the stimulus is present on the screen. For example, when the stimulus is removed from the screen and the saccade system must target a memory representation of the stimulus instead, saccades are driven toward the perceived position (Massendari, Lisi, Collins, & Cavanagh, 2018; Ueda, Abekawa, & Gomi, 2018) suggesting that memories are stored in a perceptual format. Likewise, motor actions involving hand motion, such as pointing movements and tracking with a stylus, are guided by the perceived position and not the physical position ('t Hart, Henriques, & Cavanagh, 2021; Lisi & Cavanagh, 2017). Moreover, it has been shown that a number of higher order perceptual and cognitive operations, including attentional tracking (Maechler, Cavanagh, and Tse, 2021), pop-out effects (Özkan, Tse, & Cavanagh, 2020), and the computation of apparent motion trajectories (Hui et al., 2020), are all driven by the perceptual representation of the stimulus.

In this study, we investigate a second eye movement system that can act while the stimulus is present — smooth pursuit — and ask whether it is driven by the physical or the perceived position of the double-drift probe. It was recently shown that the illusion itself survives smooth pursuit. Cavanagh and Tse (2019) had participants pursue a smoothly moving fixation point while the Gabor patch, located 18.75 degrees away from fixation, followed the same trajectory as the fixation point. In this way, the envelope was

approximately stabilized on the retina during pursuit. This condition nulls retinal sources of motion information about the envelope's trajectory, leaving primarily information generated by the pursuit system. Cavanagh and Tse (2019) found that the illusion persists without loss during smooth pursuit. Thus, efference copy from the pursuit system is combined with local motion information generated by the texture (i.e., the internal drift) to produce offsets in both position and direction. Therefore, the illusion must be computed after the recovery of motions in the world based on eye movement signals.

In their experiment, Cavanagh and Tse (2019) had participants smoothly pursue a high contrast fixation spot while the double drift patch was in the periphery. Here we test whether smooth pursuit of a virtual target derived from an illusory stimulus is affected by the illusion. To do so, we adapted a paradigm in which participants pursued the inferred mid-point between two objects (Steinbach, 1976; Beutter & Stone, 2000; Hafed, & Krauzlis, 2008). Two Gabors moved in tandem down and then up the screen, while participants tracked the virtual midpoint halfway between the two Gabors. On the way down, the Gabors had no internal drift, but on the way up the internal grating drifted orthogonally to the external motion direction, creating the classic double-drift effect. If smooth pursuit is driven by physical position as is the case for saccades, the gaze trajectories would have the same angle in both downward and upward segments, because the Gabors in fact traversed the same path on the way down as on the way up. However, if perception determines smooth pursuit targets, then the trajectories and their angles should diverge, because the double-drift illusion was only present during the upward trajectory.

Many other studies have investigated whether smooth pursuit follows the physical or perceived path of a stimulus (see Spering & Montagnini, 2011 for a review). Overall, the existing literature is split on the issue. For instance, two previous studies have looked at whether smooth pursuit is based on retinal or perceived target trajectories using the Duncker illusion to make perceived path deviate from the physical path (Wyatt & Pola, 1979; Zivotofsky, 2005). The two articles came to opposite conclusions. Previous research (Zhang, Yeh, & DeValois, 1993; Hughes, 2018) has revealed that a Gabor patch appears to move faster if both internal and external motions have the same direction and slower if the two motions have opposite directions. A study examining smooth pursuit of Gabor patterns with internal motion in the same or opposite direction as the envelope motion found that the pursuit gains were influenced by the drift in the same way as perception (Hughes, 2018). However, in the same experiment the perceptual error and the pursuit error were not correlated, suggesting a dissociation of perception and action in smooth pursuit. In contrast, two studies reported smooth pursuit of a static stimulus that appeared to move because of motion aftereffects, linking perception and smooth pursuit (Braun, Praceus & Gegenfurtner, 2006; Matsumiya & Shioiri, 2015). Other investigations into conflicts between action and perception using visual illusions have also found mixed results (Bruno, 2001). Here, the smooth pursuit that we observed for a virtual target derived from an illusory stimulus showed a strong effect of the illusion on eye movement trajectories.

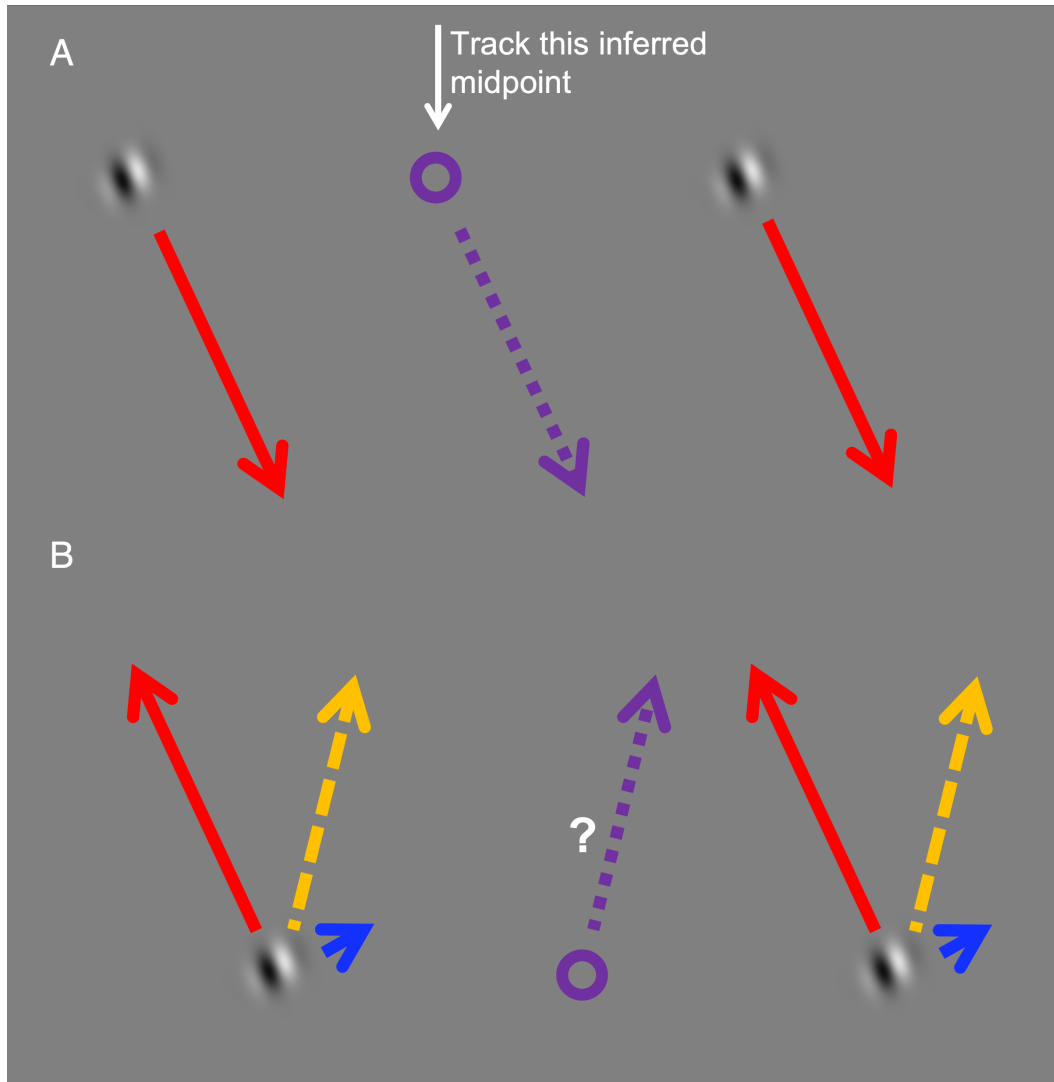


Figure 5. Schematic of the stimulus. A) During the first part of the trial the physical path (red arrows) of both Gabors is downward and the fixated midpoint (purple circle) follows the same trajectory (purple dotted arrow). B) When the Gabors begin to return upward along the same physical path, the orthogonal internal motion (small blue arrow) creates an illusorily perceived trajectory (yellow arrow). The fixated midpoint follows this illusorily perceived trajectory.

Method

Participants

We recruited seven people as participants (six men and one woman, mean age: 32 years, six right-handed) two of whom are authors of this article. The experiment was approved by the Institutional Review Board at Dartmouth to ensure compliance with ethical standards. Additionally, all procedures were approved by Dartmouth's Environmental Health and Safety department regarding COVID safe experimentation.

Stimuli

Two Gabor patches – sinusoidal luminance gratings with a Gaussian envelope – served as stimuli for this experiment. The visible diameter of a Gabor was ~ 0.5 degrees of visual angle [dva], with a sigma of 0.1 dva and spatial frequency of 2 cycles per degree. They were aligned on a horizontal line and separated by 20 dva (10 dva in each direction from their midpoint). At the beginning of each trial, there was also a black fixation spot with 0.2 dva diameter in the middle between the two Gabors, which disappeared once the trial started.

Participants initiated a trial by fixating their gaze on the fixation spot for half a second. After the fixation spot disappeared, the two Gabors, neither of which had internal motion, started moving downwards toward either the left or right lower corner of the screen along parallel paths, at an angle of plus or minus 22° relative to vertical. Participants were asked to fixate and smoothly pursue the inferred (but invisible) midpoint between the two Gabors. To facilitate smooth pursuit eye movements and diminish the need to make catch-up saccades, the Gabors accelerated and decelerated in a

sinusoidal fashion over the whole path. After traversing a 2 dva long path within 1 second (i.e., their average movement speed was 2 dva/s), the Gabors reversed direction and moved back along the same path at the same speed. On the way back, however, their internal grating continuously drifted at 2 dva/s. The drift direction of the internal grating was orthogonal to the Gabors' external motion paths, creating the double-drift stimulus. Although the Gabors moved up the screen along the same path that they traversed on the way down, they looked as if they were following a V-shaped trajectory. See video S1 in the supplementary material for a demonstration of the stimulus.

Procedure

Participants initiated the motion of the Gabors by fixating the black dot between the Gabors, which disappeared after fixating for half of a second. They were then tasked with keeping their gaze fixated on the moving, invisible midpoint between the two Gabors. If smooth pursuit follows the image of an object on the retina, there should be no difference between pursuing drifting or non-drifting Gabors. However, if smooth pursuit follows the perceived position of an object, then the added internal drift should make the participants' gaze deviate from the veridical midpoint of the Gabors. Using an eye tracker (EyeLink 1000, SR Research, Oakland ON, Canada, 1000 Hz sampling rate) we recorded and monitored participants' gaze during the experiment. When their gaze deviated from the Gabors' midpoint along the y-axis by more than 2 dva, the trial was repeated. Since we expected deviations along the x-axis to be induced by our stimulus, trials were not restarted when the gaze deviated along the x-axis. We ran a total of 200 trials per participant, counterbalanced and pseudo-randomized for the initial motion direction

(tilted left / right). The experiment was self-paced and took around 15 to 20 minutes to complete.

There was a small limitation with regard to the calibration of the eye tracking equipment, which could not be done by the experimenter due to restrictions related to the COVID-19 pandemic. Instead, participants had to calibrate the eye tracker and other equipment using oral instructions from the experimenter, who was monitoring the experiment from an adjacent room. This might have had a small impact on the overall accuracy of the eye tracking, but since all comparisons were done within each trial, the resulting noise in the data was well-controlled.

Apparatus

Participants viewed the stimuli on an AM250 OLED monitor (Flanders Scientific Inc, GA, USA; Cooper et al., 2013). The screen was set to 1920 x 1080 resolution. We used the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) for creation and display of the stimuli. Eye movements were monitored live and recorded using an EyeLink 1000 eye tracker (SR Research, Oakland ON, Canada, 1000 Hz sampling rate). Participants were stabilized with a chin rest and a forehead rest at a distance of 58 cm from the screen.

Results

As a first analysis step, the eye movement trace from each trial was split into two halves: the initial segment without internal drift of the Gabors and the second segment with internal drift. Next, we trimmed each half by 200 msec worth of gaze data at the beginning and end to exclude eye movement instabilities associated with starting, ending, and reversal, such as catch-up saccades. Figure 6 shows individual eye movement traces including saccades. We then used orthogonal regression (Deming, 1943) to fit two lines, one to each central portion of the downward and upward traces. The initial, downward segment without the illusory stimulus showed a good match between the path of the virtual midpoint and the eye movement path with an average deviation of 5.70° (95% confidence interval: -2.07° , 13.46°). The angle fit to the second segments varied widely across participants (ranging from -26.34° to -74.02°) but deviated consistently from the physical path in the same direction for all participants. The average deviation was -34.57° (95% confidence interval: -44.90° , -24.24°). We tested the difference between the angles of the first and second segments as a measure of the influence of the illusion on smooth pursuit. The average difference was 48.94° , significantly different from zero (95% confidence interval from 31.28° to 66.61° , $t(6)=6.78$, $p<0.001$, Cohen's $D=2.56$, 62.1% explained variance). When the analysis was repeated without excluding the initial start-up, reversal and end segments (black portions of average gaze trajectory on Figure 6), the results were essentially the same.

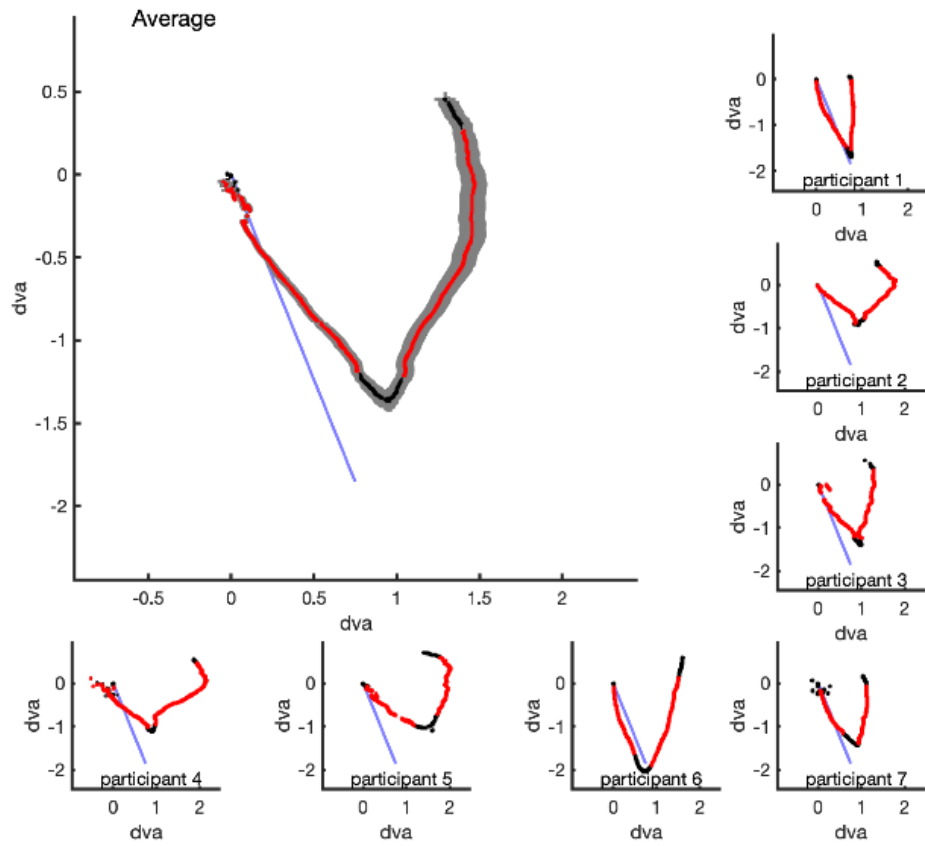


Figure 6. Average individual and group eye movement traces. Blue lines represent the path of the midpoint between the two Gabors on the screen (i.e., where participants were told to look). The black and red traces correspond to the mean trajectory of gaze starting at the origin (i.e., where they actually looked). The shaded area in the group average represents the standard error of the mean. Red segments of the traces were used in the analysis of the angles. Areas where the line appears broken coincide with time points when many saccades caused outliers in recorded gaze positions (i.e., beginning, end, and reversal). The black 200 ms segments of the traces were excluded from the data analysis.

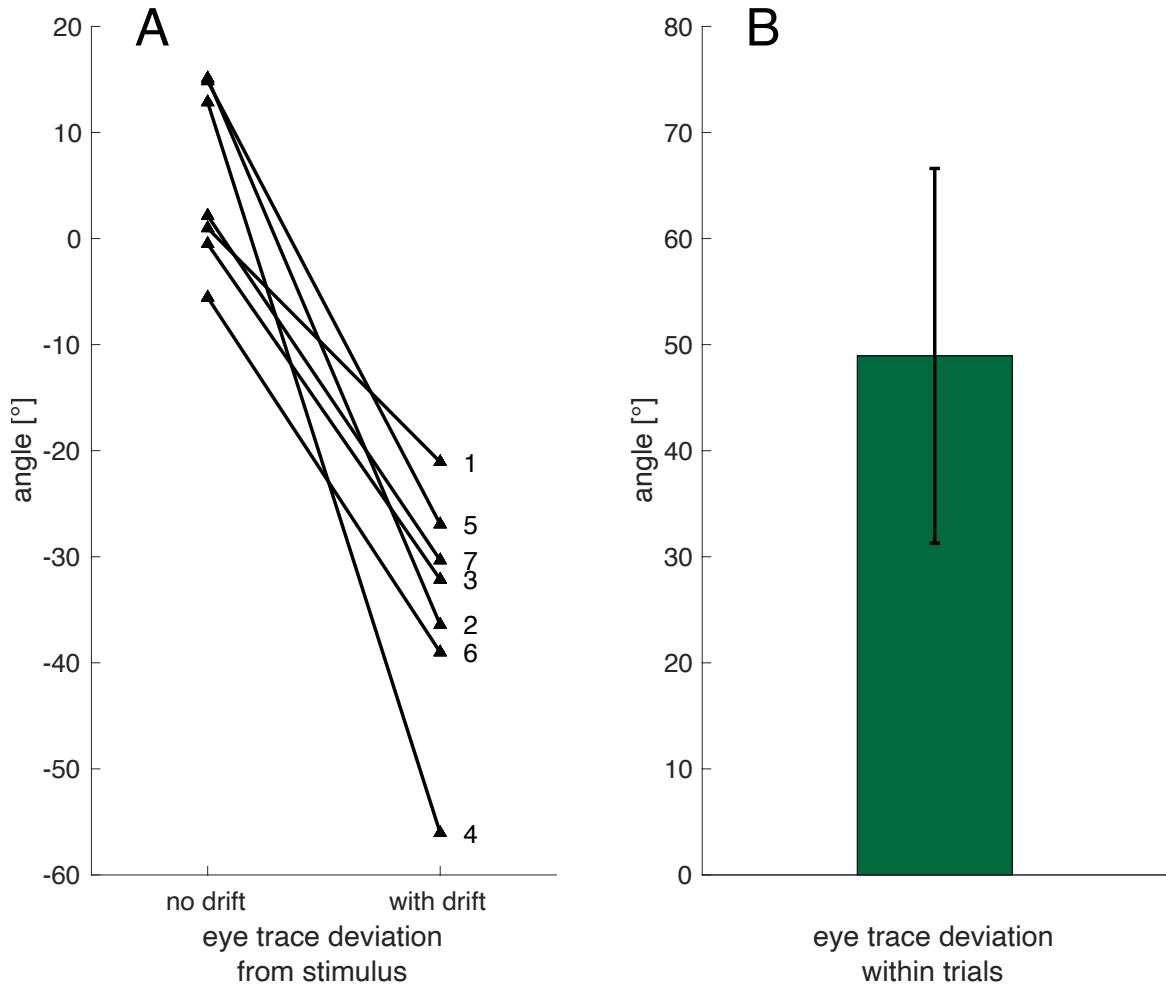


Figure 7. Differences in direction of eye traces with and without internal drift. A) Deviations in gaze tilt from the actual midpoint of the two Gabors. B) Average of within trial difference between the two traces and 95%-confidence interval. Numbers correspond to the individual participants as in Figure 6.

Our first analysis revealed that direction of the eye movements depended on perception, but it is possible that this effect was driven by saccades to the perceived midpoint, rather than by actual smooth pursuit of the perceived midpoint. To address this concern, we conducted multiple control analyses. First, with the goal of analyzing only trials without saccades, we used the algorithm of Engbert and Kliegl (2003) to detect

saccades and microsaccades. This allowed us to remove all trials from the analysis in which a saccade occurred during the double-drift part of the trial (66.9% of all trials). The downward segment of each trial still overlapped well with the physical midpoints (average angle between physical path and eye trace: 2.59° , 95% confidence interval from -5.81° to 11°), and in the double-drift condition the angle between eye trace and stimulus was still significantly larger (average 31.1° with a 95% confidence interval from 27.54° to 34.66°). The average within-trial difference in eye trace angle (33.69°) was also significantly higher than zero (95% confidence interval from 26.88° to 40.5° , $t(5)=12.72$, $p<.001$, Cohen's $d=5.19$, 87.1% variance explained), albeit smaller than when we included all trials.

We also calculated eye velocity gain for each half of each trial. Instead of removing entire trials with saccades, we de-saccaded the pursuit traces by removing segments that were identified as saccades. Then we calculated the average speed of the eye movements separately in each half of each trial. Since the illusion moves the perceived position of the Gabors away from the physical position along the x-axis, but not the y-axis, this analysis is matched for both conditions (with and without internal drift). The mean gain along the y-coordinate while the Gabors drifted downwards (without illusion / internal drift) was 0.6. Comparing this to the mean gain along the y-axis while the Gabors drifted upwards (with internal drift / illusion) showed no significant difference (the mean gain in this condition was 0.73, t-test for difference between the two: $t(6)= 1.167$, $p=.287$, Cohen's $d=0.48$, 5.4% variance explained).

Finally, we looked at the catch-up saccades themselves. Participants made between 0.2 and 1.2 saccades per transit (600 ms) within our analysis windows (200 ms

to 800 ms, and 1200 ms to 1800 ms), with an average of 0.6 saccades per transit for the control and 0.6 saccades per transit for the double-drift condition. At a saccade rate this low, it would not be possible to find the effect of the illusion from saccades alone.

Although the saccades were infrequent, we analyzed their directions to see if they were truly corrective — namely, to determine whether they aimed back to the midpoint of the physical or perceived paths. To this end, we plotted the saccade amplitude as a function of the distance of its starting point from the physical midpoint. We used data based only on the x-coordinates (see Figure 8) of the saccades occurring within the analysis windows. In other words, we analyzed horizontal saccadic amplitude as a function of horizontal distance from the physical midpoint. For the no-drift conditions, we expect corrective saccades to head to the right (i.e. have a positive amplitude value) if their start point was left of the physical midpoint (i.e. if the initial distance had a negative value), correcting the retinal offset of the fovea from the virtual target, and vice-versa. For the drift conditions, we expect corrective saccades to be to the right if they started to the left of the virtual midpoint of the perceived locations, and vice versa, independently of their location relative to the physical midpoint. The intercept with the x-axis of a regression line plotted through these data points reveals the average location of the saccade target, the location where the saccade vector switched from leftward to rightward. As can be seen in Figure 8, the inferred saccade target (x-intercept) for the no drift condition was not significantly offset from the physical midpoint (mean offset across participants: 0.31 dva, 95%-confidence interval from -0.12 dva to 0.74 dva, $t(6)=1.76$, $p=.129$, Cohen's $d=0.67$, 10.1% variance explained). In contrast, the x-intercept was significantly offset from 0 in the double-drift condition (mean offset across participants:

1.02°, 95%-confidence interval from 0.62 dva to 1.42 dva, $t(6)=6.23$, $p<.001$, Cohen's $d=2.35$, 58.1% variance explained) and also differed significantly from the x-intercept in the no-drift condition (mean difference across participants: 0.71 dva, $t(6)=2.66$, $p=.037$, Cohen's $d=1.01$, 20.2% variance explained). Additionally, we looked at the slopes of these regression lines, because negative slopes would indicate that saccades were corrective. Without drift, the average slope was -2.26 (95%-confidence interval from -3.77 to -0.76) while in the drift condition it was -1.29 (95%-confidence interval from -1.71 to -0.88). Note that this regression was computed and the x-intercept calculated individually for each participant. However, in Figure 8 we plot all saccades from all participants. The inferred saccade target location was consistent with the location of the midpoint between the perceived paths, which suggests that this location rather than the physical midpoint was the target of the catch-up saccades in the double-drift condition. This also explains why removing all trials with saccades from the analysis weakened the effect.

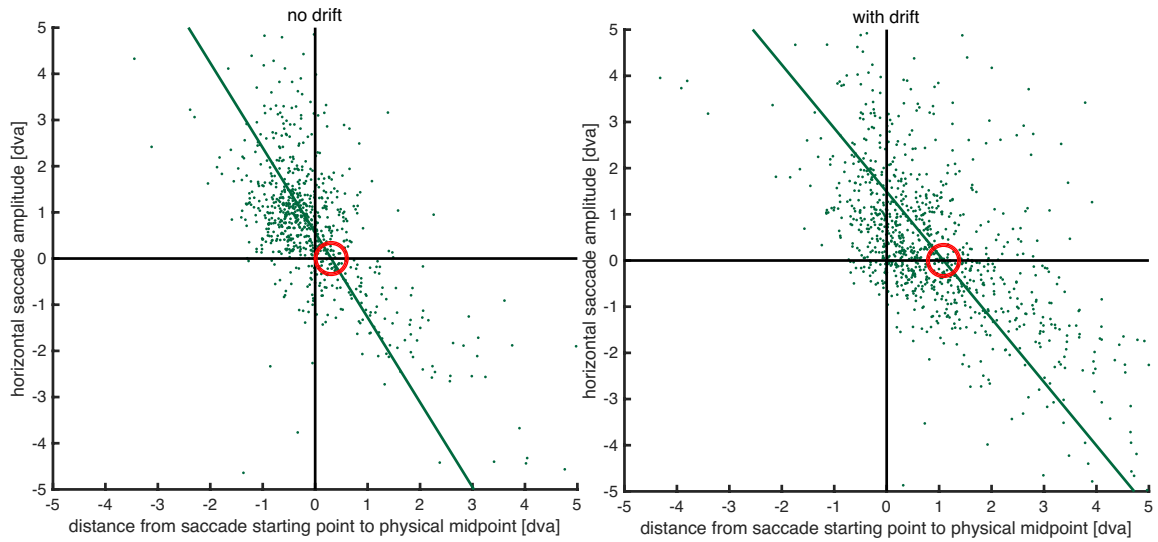


Figure 8. Scatter plot of the x-component of saccade vectors. The x-axis represents the distance between the saccade starting point and the physical midpoint between the Gabors (positive numbers indicate the direction of illusory motion). The y-axis represents the amplitude of the saccade, with positive numbers indicating that the saccade was made in the direction of illusory motion. The green line is the orthogonal regression (Deming, 1943) fit to all saccades across all participants. The x-intercept of the green line shows the average x-offset between physical midpoint and saccade targets (marked with a red circle), which was the physical midpoint in the control condition, but was offset in the direction of illusory motion in the double-drift condition.

Discussion

In this study we show that the target for smooth pursuit of the midpoint between two double-drift stimuli is derived from their perceived, not retinal positions. Corrective saccades likewise appear to be biased in the direction of the perceived rather than actual midpoint. These two findings are in direct contrast to the findings showing that saccades were made to the physical positions of this illusion, when targeting the Gabor directly (Lisi & Cavanagh, 2015; Massendari et al., 2018; Nakayama & Holcombe, 2020). We take these results to mean that the virtual target is calculated from perceived rather than retinal coordinates.

Whether smooth pursuit follows a stimulus or its percept is a question of long standing (Spering & Montagnini, 2011). Here, we report that for paired, double-drift stimuli, the smooth pursuit of their midpoint is driven by perception. First, we replicate earlier findings (Steinbach, 1976; Beutter & Stone, 2000; Hafed & Krauzlis 2008) showing that observers can pursue an inferred location. However, in these earlier studies, the inferred midpoints were based on stimulus components whose perceived locations were aligned with their retinal locations. In our study, the inferred targets are based on stimuli whose perceived locations differ dramatically from their physical locations. In this case, the smooth pursuit clearly followed the midpoint of perceived not physical locations. Thus, while visually guided saccades were found to be driven more by the physical position of the double-drift stimulus (Lisi & Cavanagh, 2015), the smooth pursuit system is instead driven by the perceptual representation of the double-drift stimulus even while it is present. This is also in line with findings that show smooth pursuit of a motion aftereffect seen on a static stimulus (Braun, Pracejus & Gegenfurtner,

2006; Matsumiya & Shioiri, 2015). Moreover, the direction of the internal motion of a Gabor (with or against the envelope motion) affected perception and smooth pursuit similarly (Hughes, 2018 – however, the perceptual error and smooth pursuit errors were not correlated in this study, suggesting some dissociation between pursuit and perception).

Catch-up saccades were directed to the inferred midpoint between the perceived coordinates of the double-drift rather than its physical coordinates. This is consistent with past findings that memory-guided saccades to a double-drift stimulus that has vanished also target its last perceived location (Massendari et al. 2018; Udeda, Abekawa, & Gomi, 2018). Thus far, only visually guided saccades to single, visible double-drift stimuli escape the illusion; both memory-guided saccades and saccades to inferred targets are instead largely biased by perceptual processing. This also ties in well with other similar studies on saccades targeting the Müller-Lyer illusion (Bruno, Knox & de Grave, 2010), as well as with the findings from Zivotofsky and colleagues (Zivotofsky, Rottach, Averbuch-Heller, Kori, Thomas, Dell’Osso & Leigh, 1996; Zivotofsky, Goldberg, & Powell, 2005) showing that targets derived from memory drive saccades to the perceived location of the Duncker illusion.

Although every participant in our study demonstrated the same effect, there were notable individual differences. As can be seen in Figure 6, some of our participants showed an effect akin to a saturation of the illusory effect near the end of the trial. After deviating from the physical direction of the Gabors, their gaze started to follow a line parallel to the physical midpoints but offset in the direction of internal motion by more than one degree. This might reflect saturation of the double-drift effect, which has been

reported previously (Kwon et al., 2015) and our data would show a variation in this saturation effect across participants with some showing it after 1.5 seconds and some not at all. This would suggest individual variation in the way observers weight position and motion information over time. Alternatively, this return of gaze towards the starting point might just reflect anticipatory return to the starting location of the next trial.

In addition to the studies using midpoint tracking (Steinbach, 1976; Beutter & Stone, 2000; Hafed & Krauzlis, 2008), several others have addressed whether smooth pursuit is driven by perception or by the retinal stimulus (for reviews see Kowler, 2011; Lisberger, 2010, Spering & Montagnini, 2011) but the results have been mixed. Mack, Fendrich, and Wong (1982) reported that when the retinal and perceived positions of a pursuit target were in conflict, smooth pursuit followed the retinal motion, when it was available. Spering and Gegenfurtner (2007) found a dissociation of pursuit speed and perceived speed of motion. In contrast, other authors have found that pursuit is influenced by perception. For instance, with the Duncker illusion, Wyatt and Pola (1979) reported that open-loop smooth pursuit (the first 100ms after the target's appearance) follows perceived direction even though closed-loop pursuit — after 100ms — does not (Zivotofsky et al., 1995; Zivotofsky, 2005). It should be noted that Zivotofsky's participants pursued the target foveally, which strongly impairs the Duncker illusion, so there might not actually have been a difference between the retinal and perceived directions.

More recently, Ma, Watamaniuk, and Heinen (2017) had participants smoothly pursue a set of four Gabors that had internal motion in addition to translation. They found clear evidence that the internal motion affected both perception and smooth pursuit. The

effects were significant but small, for example, about a 2° deviation in pursuit direction when the internal Gabor motion was orthogonal to the translation, compared to the 30° deviation we find in our experiment. This is likely a consequence of the small, 5° separation of their four Gabor patches in their experiment. In this case, gaze direction deviations would be limited to positions within the four Gabors. In addition, the Gabors in their display would fall quite close to the fovea, reducing the illusion. In our experiment, we use two Gabors separated by 20° and we are able to observe smooth pursuit unconstrained by the stimulus configuration. As a result, we advance these previous findings to show that that an illusion — the double-drift stimulus — influences both perception and closed-loop smooth pursuit equally. We keep the actual stimulus away from the fovea by tracking an inferred midpoint between two double-drift stimuli.

Other studies that have compared effects of perceived versus retinal motion on optokinetic nystagmus and ocular following have also found mixed results. Logothetis and Schall (1990) found that the slow phase of optokinetic nystagmus followed perception in a binocular rivalry study with monkeys. Spering and Carrasco (2011) using a similar paradigm found that involuntary eye movements in humans followed the stimulus motion and not perception. While these involuntary eye movements rely on many of the same cortical processing structures as smooth-pursuit eye movements (Ilg, 1997), our results with smooth pursuit appear more consistent with Wyatt and Pola (1979) showing that the open-loop portion of smooth pursuit was driven by perceived direction, as well as the findings by Ma, Watamaniuk and Heinen (2017) showing that closed-loop smooth pursuit is influenced as well.

Whereas Cavanagh and Tse (2019) stabilized a double-drift stimulus on the retina during smooth pursuit, here we do the opposite; we stabilize the Gabor in perception by having participants smoothly pursue the perceived midpoint between two doubly drifting Gabor patches. This did not stabilize the moving Gabor patches on the retina. When stabilized in perception, the double-drift Gabor patches must move across the retina orthogonally to the pursuit direction. The existence of the illusion here – now measured by the smooth pursuit itself — supports the conclusion that computations underlying the illusion must follow the recovery of the Gabor’s motion in world coordinates (i.e., after taking eye movements into account) as well as showing that tracking an inferred target involves using perceived, not retinal, coordinates as input.

There is substantial evidence suggesting that saccades and smooth pursuit are performed by largely overlapping systems (for reviews see Kowler, 2011; Krauzlis, 2004). Why then would smooth pursuit be subject to the double-drift illusion whereas saccades are not (Lisi & Cavanagh, 2015)? According to Krauzlis (2004) one of the few differences between the circuits underlying smooth pursuit and saccades is the involvement of MT/MST in smooth pursuit. This suggests that MT/MST might be the source of perceptual input to smooth pursuit that is absent in saccades, which is also supported by previous results from smoothly pursuing inferred locations (Beutter & Stone, 2000). The purpose of saccades is not tracking over time, but rapid foveation of peripheral stimuli. Perceived positions emerging in MT/MST or higher (Liu et al, 2019) undoubtedly take more time to compute, explaining one reason why visually guided saccades may be based on direct links from the retina to the superior colliculus, when available (Kato, Takaura, Ikeda, Yoshida, & Isa, 2011), and so avoid the perceptual

effects of the double-drift stimulus. We can only speculate whether inferred midpoints between moving stimuli are really computed in MT/MST. This will be a question for future research.

In conclusion, smooth pursuit of inferred targets uses location coordinates that are computed late in the visual position processing hierarchy, after the computation of the double-drift illusion. That is, the tracked target is stabilized in perception, not on the retina. This account of the smooth pursuit system reaffirms the notion that it is not merely controlled by the brainstem or the cerebellum and is instead driven largely by higher visual cortical areas (Krauzlis, 2004) that represent positions in the world as they are consciously experienced.

Supplementary Information

Saccade landings were additionally analyzed with respect to their origin. The area in which eye movements started and landed was first split up into a 7 by 7 grid. Saccades that started in each tile of this grid were averaged together to reveal the approximate targets of saccades relative to their approximate starting point (blue arrows in Figure 9). During the downward motion without illusory motion, participants' saccades targeted the physical midpoint of the stimulus. However, during the upward motion, when there was illusory motion, participants' saccades often targeted the perceived (illusory) midpoint. Compare Figure 9 for detailed results.

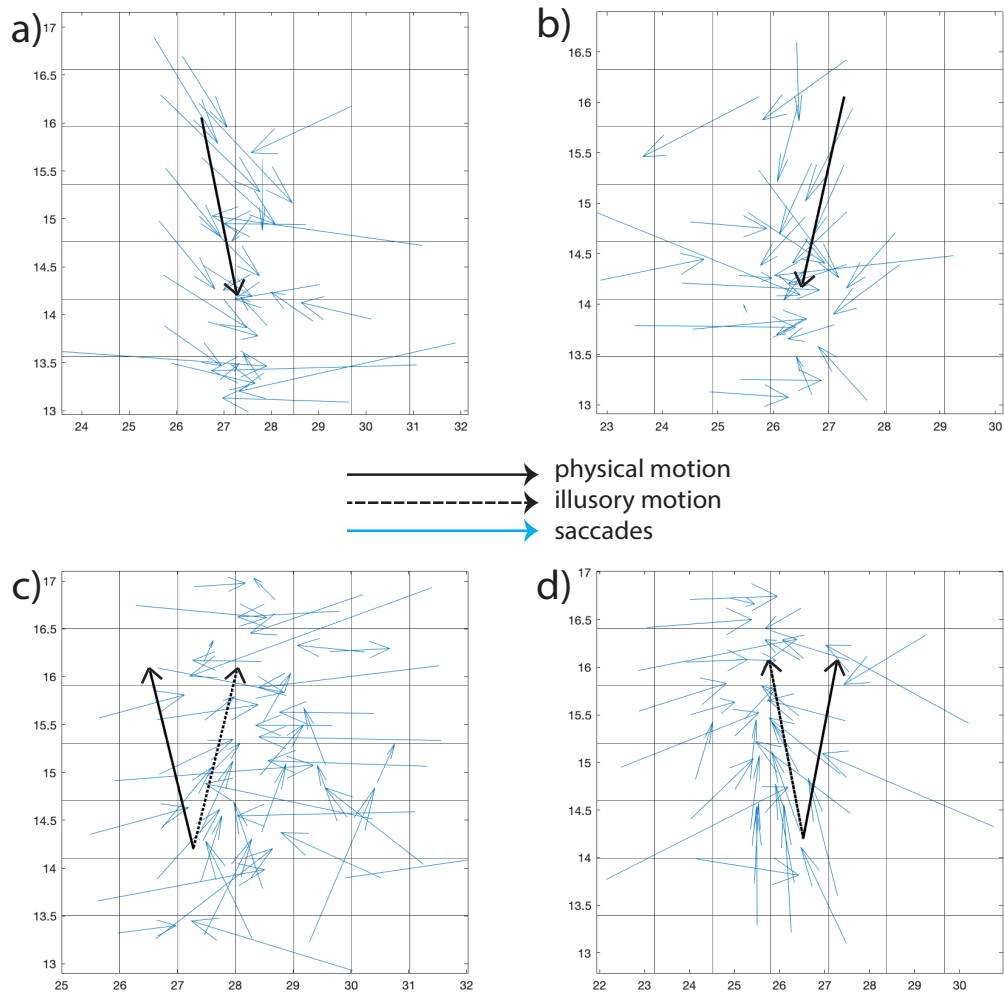


Figure 9. Visualization of catch-up saccades during smooth pursuit. Panels a) and b) show saccades made while the Gabors were moving downward to the right and left respectively. Panels c) and d) correspond to the upward moving Gabors with illusory motion. Corrective saccades were often made towards the perceived (illusory) midpoint between two double-drift Gabors.

Chapter 3: Hemifield Bias in Load Dependent Activity During Multiple Object Tracking

Authors: Marvin R. Maechler, Peter J. Kohler, Patrick Cavanagh & Peter U. Tse

Author Contributions: M.R.M designed the study, M.R.M. wrote the experiment code, and collected and analyzed the data. P.J.K, P.U.T. and P.C. oversaw analyses, M.R.M., P.J.K, P.C., and P.U.T wrote the paper. All authors provided critical feedback and helped shape the research, analysis, and manuscript.

Abstract

The difficulty of tracking multiple moving objects among identical distractors increases with the number of tracked targets. Previous research has shown that the number of targets tracked modulates activity in brain areas related to visuospatial attention, giving rise to so-called ‘attention response functions.’ While the hemifield/hemispheric effects of spatial attention (e.g., hemispatial neglect, hemifield capacity limits) are well described, whether these effects also impact attention response functions was previously unknown. Using functional magnetic resonance imaging, we show that the number of tracked objects modulates activity in a large network of areas bilaterally. Contralateral tracking load significantly covaried with activity throughout the visual system, while both contra- and ipsilateral load significantly influenced activity in the parietal and frontal lobes, specifically the dorsal attention network. Further, some areas were significantly more sensitive to contralateral than ipsilateral load. We replicate findings showing that a diverse set of brain areas contribute to tracking multiple targets and extend the canonical attention response functions to include hemifield bias. Given the hemifield-specific nature of speed and capacity limits to multiple object tracking, we suggest that those areas that show strong hemifield preference may be the source of overall tracking capacity and speed limits.

Introduction

Object tracking is a fundamental function of the visual system. Driving a car or participating in any sport would be unimaginable without the ability to know where an object came from and where it is going. Performance in object tracking is governed by several factors that severely constrain its effectiveness. For example, as objects get closer together or increase in speed, the ability to track them decreases (Intrilligator & Cavanagh, 2001; Maechler, Cavanagh & Tse, 2021; Cavanagh & Alvarez, 2005; Störmer, Alvarez & Cavanagh, 2014; Holcombe, 2023). Additionally, our capacity to track objects is limited, such that tracking performance drops as more targets are added (Franconeri, Lin, Enns, Pylyshyn & Fisher, 2008; Alvarez & Cavanagh 2005; Scholl, 2009).

How does the brain track multiple objects at once? Previous neuroimaging studies have identified several brain areas that may contribute to this task, which fit broadly into two categories. One set of areas show increasing blood-oxygen-level-dependent (BOLD) responses with the number of tracked objects (e. g., Culham, Cavanagh & Kanwisher, 2001). In contrast, a second set of areas responded during the task, but their activity was independent of the number of targets tracked. In both cases, activity is a function of how much attention is paid, with the former pattern of activity called load-dependent Attention Response Functions (ARFs) and the latter called task-dependent ARFs (Culham, Cavanagh & Kanwisher, 2001; Jovicich, Peters, Koch, Braun, Chang & Ernst, 2001; Shim, Alvarez, Vickery & Jiang, 2010; Jahn, Wendt, Lotze, Papenmeier & Huff, 2012; Alnæs, Sneve, Espeseth, Endestad, van de Pavert & Laeng, 2014; Nummenmaa, Oksama, Glerean & Hyönä, 2017). Load-dependent ARFs were found in the superior parietal lobule (SPL), the intraparietal sulcus (IPS) and the frontal eye fields (FEF) – all

areas that are canonically associated with visuo-spatial attention – but also in areas such as V5/hMT+ which are earlier in the visual processing stream.

While these prior studies have provided a great deal of insight into the neural substrate of multiple object tracking, the underlying mechanism is still debated. For example, although it is well known that some resources for attentional tracking are hemisphere specific, most of the fMRI studies of ARFs ignored whether the tracked targets were contra- or ipsilateral to the brain areas being imaged. This is important to account for, as both the number of trackable targets (Alvarez & Cavanagh, 2005) as well as their speed (Störmer, Alvarez & Cavanagh, 2014; Holcombe & Chen, 2012) can be increased much more if the targets are distributed across both visual hemifields rather than constrained to one. Additionally, tracking performance decreases when two targets cross the vertical meridian simultaneously (Strong & Alvarez, 2020). Thus, there is considerable behavioral evidence that tracking processes are constrained within a single hemifield, but we previously lacked a neural account of this effect.

Are load-dependent ARFs also hemifield specific? In one fMRI study (Shim et al., 2010), ARFs for different brain areas were assessed with respect to whether the targets were contra- or ipsilateral to the brain area. Their results suggested that both contra- as well as ipsilateral increases in tracking load increase activity in the parietal lobe, with ipsilateral targets having a slightly smaller effect. However, in this experiment the number of targets was only varied from one to two targets, while their speed was separately manipulated. Additionally, Shim and colleagues (2010) combined ROIs from the two hemispheres to increase their statistical power. In this way, they did not find a difference between hemispheres in their response to increasing attentional load.

There is extensive evidence for hemispheric differences in attention, with the right hemisphere being more important for spatial attention. For example, spatial attention to the left hemifield is more impacted by high concurrent working memory load than attention to the right hemifield (Naert, Bonato & Fias, 2018). In addition, hemispatial neglect, a neuropsychiatric disorder that renders some stroke survivors unable to attend to the contralesional side of space, occurs predominantly after injuries to the right hemisphere (Parton, Malhotra & Husain, 2003; Corbetta & Shulman, 2011). This right hemisphere dominance for neglect implies a corresponding asymmetry of the neural circuits underlying spatial attention in healthy brains. Older accounts of hemispatial neglect (Mesulam, 1981) claimed that the right hemisphere can deploy spatial attention to both hemifields, while the left hemisphere only attends to its contralateral hemifield. Based on this, one would predict ARFs in the left hemisphere to show contralateral bias, but not the right hemisphere. More modern explanations of neglect predict contralateral bias in both hemispheres (Corbetta & Shulman, 2011).

Here we examined the hemifield specificity of load-dependent ARFs. Specifically, we asked whether any brain areas with load-dependent ARFs showed contralateral bias. To do so, we limited each target's motion to either the left or right hemifield and varied the number of targets in each independently. We recorded fMRI data while participants tracked between zero and four targets, out of a total of eight objects evenly distributed across the hemifields. The number of targets in each hemifield varied independently between zero and two targets, allowing us to estimate the effect of increasing tracking load in each hemifield.

We found ARFs in a diverse network of brain areas, which is consistent with previous studies (Culham et al., 2001; Jovicich et al., 2001; Shim et al., 2010; Jahn et al., 2012; Alnæs et al., 2014). Crucially, we extended the canonical load-dependent ARFs to include the hemifield in which targets are tracked. While hemifield maps in the visual system showed strong hemifield preference, higher-order attention related areas were modulated by targets anywhere in the visual field. We found that several areas exhibited a contralateral bias in their ARF, such that the effect of tracking load was stronger for contralateral than ipsilateral targets. These areas were found bilaterally throughout the visual system, as well as in the IPS and SPL. Additionally, we found parts of the dorsal attention network in the parietal and frontal lobes, to respond to ipsilateral targets as well. These were a smaller subset of the wider network of areas that were modulated by the absolute number of tracked objects.

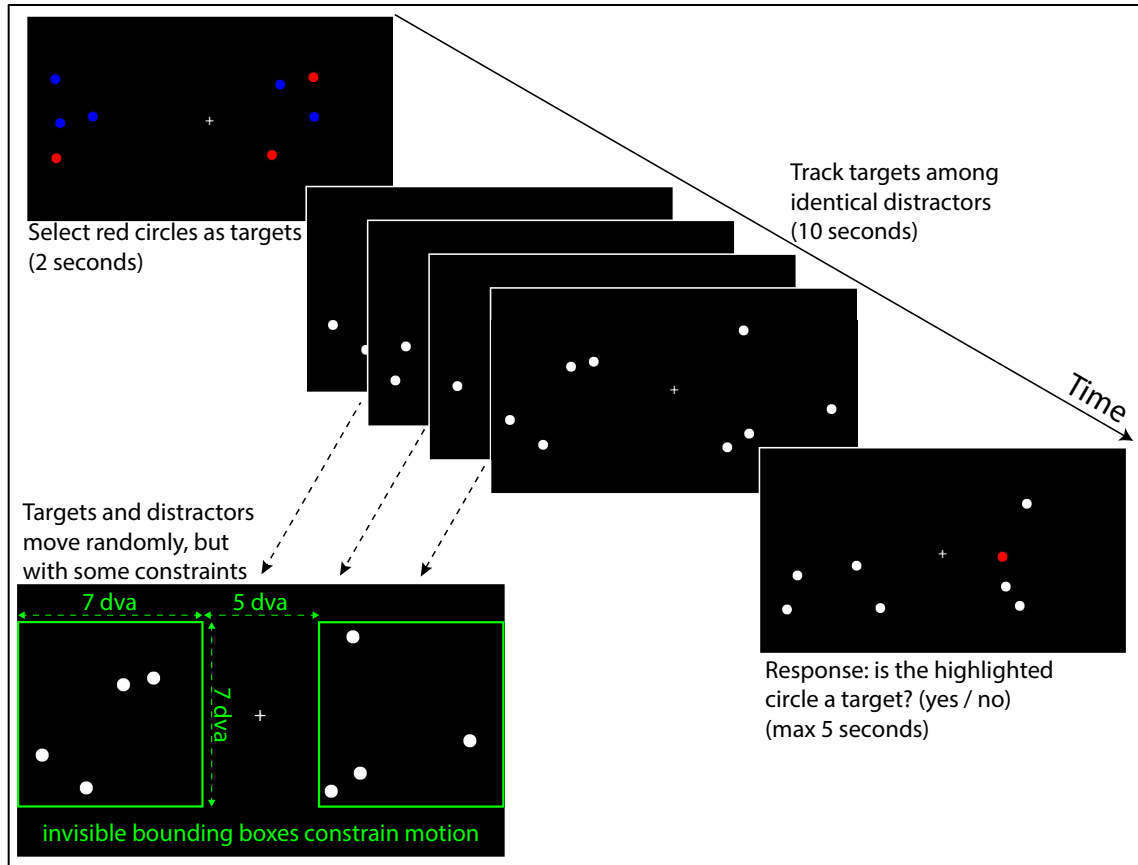


Figure 10. Task schematic showing a single trial. While fixating on the white cross in the center of the screen, participants first selected the highlighted (red) circles as targets. After 2 seconds, all circles turned white and moved randomly for 10 seconds. See bottom left for constraints on random movement. At the end, participants responded whether a randomly chosen circle was a target (two answer forced choice). There was a 16 second break between the end of one trial and the beginning of the next.

Method

Participants

Procedures were approved by Dartmouth's institutional review board. We recruited twenty volunteers to participate in the experiment. Their ages ranged from 20 to 60, with an average age of 28.8 years. The author of this dissertation was one of the participants. All participants gave informed consent and were reimbursed for their time with \$30. Participants' vision was normal or corrected-to-normal. Due to us incidentally finding a tumor in one participant's brain, their data became unusable, and they were excluded from all analyses.

Task

As in other multiple object tracking tasks, participants were told to maintain their attention on one or several objects among identical looking distractors as they moved over the screen. At the beginning of every trial eight circles with a diameter of 0.5 dva appeared on a black background, four of them in each hemifield. Some of those circles were highlighted in red to mark them as the tracking targets. The distractor circles were highlighted in blue. Then all circles' colors reverted to white, and they moved over the screen randomly. We instructed participants to covertly track the positions of all red highlighted target circles with their attention while fixating a white crosshair in the center of the screen. After the circles stopped moving, one circle was randomly highlighted, and participants responded whether it was one of the targets or one of the distractors in a two-answer forced choice format. Compare Figure 10 for a schematic.

The red and blue colors marking some circles as targets and others as distractors were displayed for two seconds. After the circles turned white, they moved over the screen for ten seconds. When the circles stopped moving, participants were given up to five seconds to respond whether the randomly highlighted circle was a target. There was a 16 second break in between trials to allow the BOLD signal to return to baseline.

All circles translated over the screen in random directions at a speed of 1.75 dva/s. Shim and colleagues (2010), as well as others (Franconeri, Lin, Enns, Pylyshyn & Fisher, 2008; Holcombe, 2023), have demonstrated that there is a strong interaction between load capacity and target speed, with slower object motion increasing the number of trackable targets. We kept object speed constant in all conditions and chose a speed that would allow most participants to track accurately at all tested tracking loads (Franconeri, Lin, Enns, Pylyshyn & Fisher, 2008).

Motion directions were random but constrained in several ways. The objects moved within squares of 7 by 7 dva that were offset to the left and right of fixation by 2.5 dva respectively, leading to a 5 dva minimum separation between the two areas of object motion. If their direction of motion would place them outside of the boundaries of these squares, their direction was changed randomly. Objects also maintained a minimum separation of 0.5 dva from each other and changed direction randomly when they were about to collide.

The number of targets varied from zero to two targets in each hemifield, with all possible combinations for the hemifields. Thus, the total tracking load ranged from zero to four targets. In this way, we were able to manipulate tracking load independently in each hemifield, leading to nine experimental conditions. In other words, every possible

number of targets (zero to two) in one hemifield was paired with every possible number of targets in the other hemifield in random order. A trial of each condition was included in every fMRI imaging run and the runs were repeated ten times per participant.

Apparatus

Stimuli were created using Matlab (Mathworks, Natick, MA), specifically Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). They were then projected onto a screen (18 dva by 7 dva projection area) inside the MRI scanner bore. Participants responded via button presses and were instructed to use the same finger to press the buttons throughout the experiment. Eye movements were recorded using an Eyelink 1000+ (SR Research, Ottawa, Canada) that was placed inside the scanner bore underneath the screen.

MRI acquisition

Images were acquired at the Dartmouth Brain Imaging Center on a 3T Siemens PRISMA scanner. After a brief localizer scan, three-dimensional fieldmaps were collected (anterior to posterior and posterior to anterior), which were used in the image preprocessing to aid with distortion correction. After the first five functional imaging runs (multiband T2* echo-planar imaging: TR = 1000 ms, TE = 30 ms, resolution = 2 mm³ isotropic, flip angle = 60°, multiband factor = 4, Grappa = 2), we acquired a high-resolution T1 image (MPRAGE, resolution = 0.94 mm³ isotropic, flip angle = 8°, TR = 2.3 s, TE = 2.3 ms). This was then followed by the remaining 5 functional runs.

MRI Preprocessing

Results included in this manuscript come from preprocessing performed using *fMRIPrep* 22.0.1 (Esteban, et al., 2019), which is based on *Nipype* 1.8.4 (Gorgolewski et al. 2011). A *fieldmap* was estimated based on two echo-planar imaging (EPI) references with *topup* (Andersson, Skare, and Ashburner (2003)). The T1-weighted (T1w) image was corrected for intensity non-uniformity (INU) with *N4BiasFieldCorrection* (Tustison et al. 2010), distributed with ANTs 2.3.3 (Avants et al. 2008), and used as T1w-reference throughout the workflow. The T1w-reference was then skull-stripped with a *Nipype* implementation of the *antsBrainExtraction.sh* workflow (from ANTs), using OASIS30ANTs as target template. Brain tissue segmentation of cerebrospinal fluid, white-matter, and gray-matter was performed on the brain-extracted T1w using *fast* (FSL 6.0.5.1:57b01774, Zhang, Brady, & Smith 2001). Brain surfaces were reconstructed using *recon-all* (FreeSurfer 7.2.0, Dale, Fischl, & Sereno 1999), and the brain mask estimated previously was refined with a custom variation of the method to reconcile ANTs-derived and FreeSurfer-derived segmentations of the cortical gray-matter of *Mindboggle* (Klein et al. 2017). Volume-based spatial normalization to one standard space (MNI152NLin2009cAsym) was performed through nonlinear registration with *antsRegistration* (ANTs 2.3.3), using brain-extracted versions of both T1w reference and the T1w template.

For each of the 10 BOLD runs per subject, the following preprocessing was performed. First, a reference volume and its skull-stripped version were generated using a custom methodology of *fMRIPrep*. Head-motion parameters with respect to the BOLD reference (transformation matrices, and six corresponding rotation and translation

parameters) are estimated before any spatiotemporal filtering using *mcflirt* (FSL 6.0.5.1:57b01774, Jenkinson et al. 2002). The estimated *fieldmap* was then aligned with rigid-registration to the target EPI (echo-planar imaging) reference run. The field coefficients were mapped on to the reference EPI using the transform. BOLD runs were slice-time corrected to 0.456s (0.5 of slice acquisition range 0s-0.912s) using *3dTshift* from AFNI (Cox & Hyde 1997). The BOLD reference was then co-registered to the T1w reference using *bbregister* (FreeSurfer) which implements boundary-based registration (Greve & Fischl 2009). Co-registration was configured with six degrees of freedom. The BOLD time-series were resampled into standard space, generating a *preprocessed BOLD run in MNI152NLin2009cAsym space*. A reference volume and its skull-stripped version were generated using a custom methodology of *fMRIPrep*. The BOLD time-series were resampled onto the following surfaces (FreeSurfer reconstruction nomenclature): *fsnative*, *fsaverage*. Surface resamplings were performed using *mri_vol2surf* (FreeSurfer).

Statistical Analyses of fMRI Data

We followed the same approach as previous studies investigating ARFs (e.g., Culham et al., 2001; Jovicich et al., 2001; Alnæs et al., 2014) and fit a Generalized Linear Model (GLM) to each vertex of each run to find parametric modulation of brain activity with tracking load. Besides the nuisance regressors for motion (6 rigid motion directions estimated during preprocessing) and for scanner drift (first- and second-order polynomials) we used several regressors of interest. One regressor captured passive viewing of the stimuli, another captured the average task response by modeling all

tracking trials equally, and crucially, a parametric modulation regressor was fit together with the task regressor to capture increases and decreases of task-related activity with changing tracking load. This analysis was repeated with the parametric modulating regressor split up into two regressors, one for each hemifield. It was not possible to estimate modulation with total load across the two hemifields and the hemifield-specific modulation in one GLM, because the regressors would have been correlated with each other ($r = 0.44$). A third GLM was fit with nine regressors of interest estimating brain responses to each experimental condition separately and used for ROI analyses. While the first two GLMs were fit to data in fsaverage space, the third GLM was fit to data in fsnative space, to extract atlas-based ROIs (Wang, Mruczek, Arcaro & Kastner, 2015).

Contrasts of interest were generated from the first-level GLM for each subject and each run. Group analysis was implemented at the second level based on the contrasts generated from the first-level GLMs. The contrast images were smoothed (FWHM of 4 vertices) to improve the signal-to-noise ratio and to account for individual differences in brain anatomy. At the second level, we used a linear mixed model to account for both the fixed experimental effects and the random subject effects. All statistical analyses were conducted in surface space.

Whole-brain analyses involve statistical tests at each vertex and therefore require correction for multiple comparisons (Bennet, Baird, Miller & Wolford, 2009). To this end, we used Random Field Theory, which estimates the probability of finding a cluster of vertices with a certain height empirically from the smoothness of the image by accounting for the spatial autocorrelation of fMRI data (Worsley, 2001; Chung, Worsley, Nacewicz, Dalton & Davidson, 2010). Group-level effects were estimated by fitting a

linear mixed effects model with fixed effects for modulation and run as well as a random effect for subjects to the data at each vertex. The data were then cluster corrected using random field theory with an initial cluster forming threshold of $p < .0001$. These analyses were implemented using the python packages *Nilearn* (version 0.10.0) and *BrainStat* (0.4.2). We used *Nilearn* for visualizations.

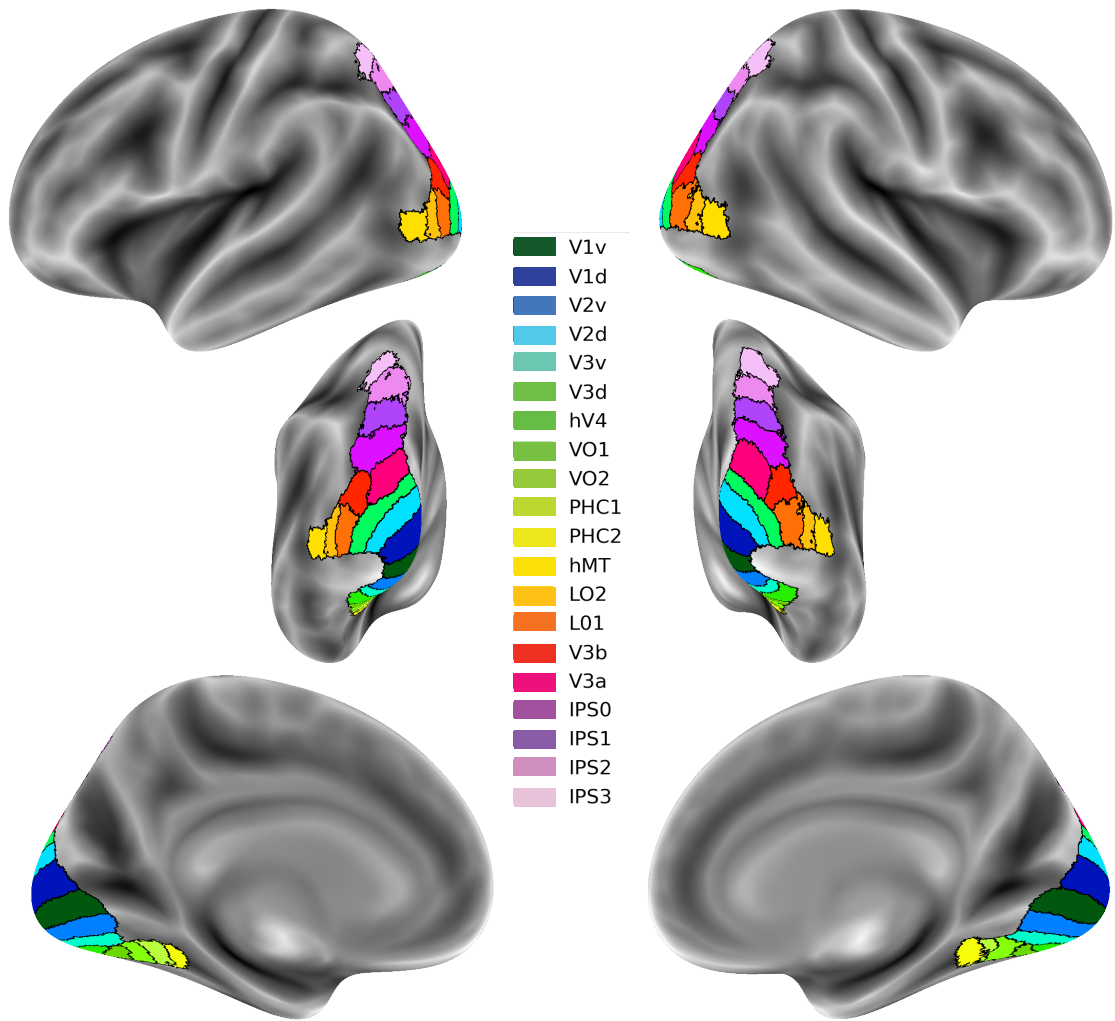


Figure 11. ROIs projected onto the inflated fsaverage brain. We used the outlined regions to perform ROI analyses. ROIs were based on the maximum probability map from Wang and colleagues (Wang, Mruczek, Arcaro & Kastner, 2015).

Results

Tracking Accuracy

Participants performed near ceiling in all conditions. Performance decreased slightly with increases in target load in both hemifields. This decrease was statistically assessed using linear mixed-effects models. A model with fixed effects for the number of targets on the left and on the right and a random effect for participants was compared to two nested alternative models using likelihood ratio tests (Luke, 2016). The alternative models were missing either the predictor for left or right load and were compared to the full model with respect to the likelihood estimates. Likelihood ratio tests revealed that load in both hemifields had a significant effect on accuracy during multiple object tracking (left: $X^2(1) = 23.85, p < .001$; right: $X^2(1) = 7.76, p = .005$). Participants' accuracy decreased on average by 3.7% with each target added in the left hemifield and by 1.5% with each target added in the right hemifield. There was no significant interaction between load on the left and right hemifields ($X^2(1) = 1.49, p = .226$).

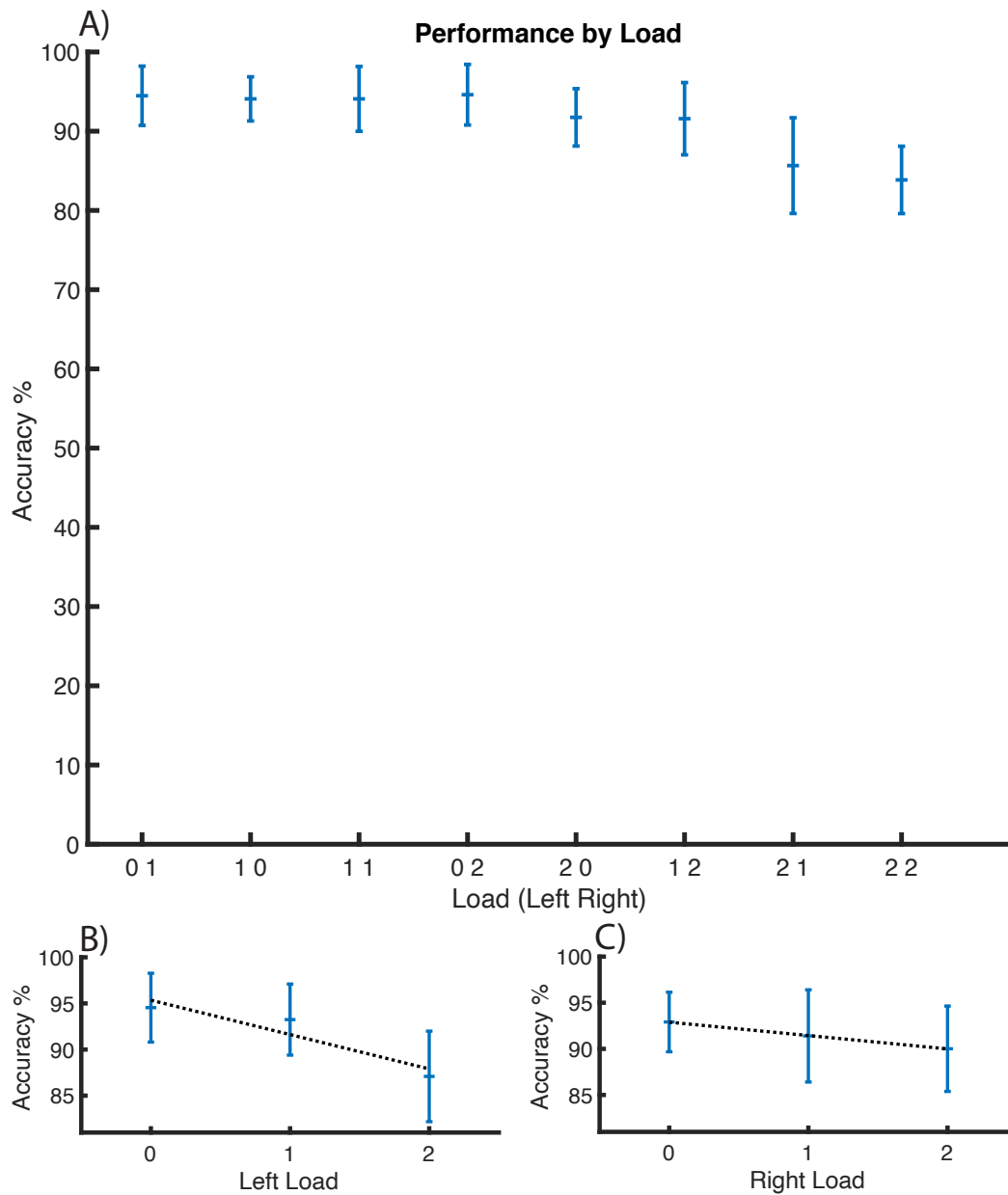


Figure 12. Performance in the tracking task by condition. A) Performance in each condition. X-axis labels indicate the number of targets in each hemifield with each digit corresponding to the number of targets in the left and right hemifield respectively. B) Performance averaged by number of targets in the left hemifield. The slope of the black dashed regression line is the estimated cost of an added target in the left hemifield. C) Same as B) for targets in the right hemifield. All error bars are 95% confidence intervals.

Eye Movements

Due to technical difficulties with the eye tracking equipment, only 17 participants' data was included in the analysis of eye movements. To verify that participants were fixating properly, their eye movements were recorded while they were performing the task inside the scanner. Fixations were usually close to the fixation cross on the center of the screen and participants rarely (~1% of the time) looked at the targets directly. Figure 13 shows an aggregate of the gaze recordings.

Spatial attention can bias fixational eye movements (Engbert & Kliegl, 2003). To characterize the degree to which the spatial distribution of targets influenced fixations, we fit gaussians to the mean heatmap of each participant and condition and compared their means. Trials with high load in one hemifield and no targets in the other hemifield should show the strongest biases in fixational eye movements if this bias was caused by imbalances in load. During trials with two targets on the left and zero targets on the right side of the screen participants fixated significantly more towards the left compared to trials with two targets on the right and none on the left ($t(16)=2.54$, $p=.022$, mean difference of 0.168 dva, 95% confidence interval from 0.03 dva to 0.3 dva).

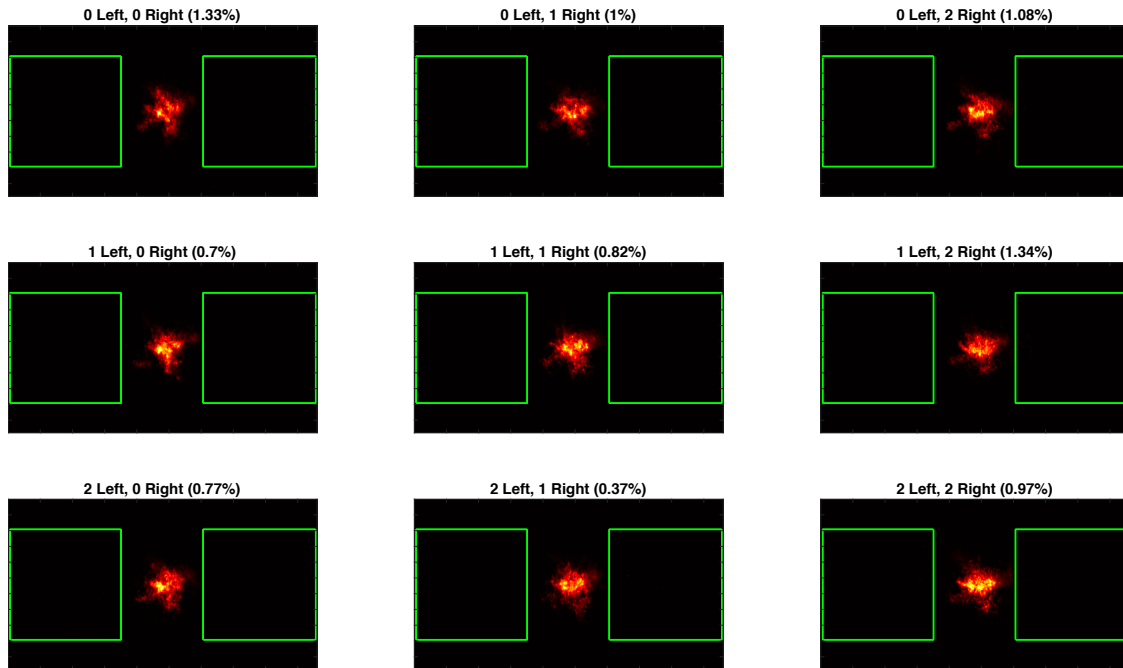


Figure 13. Average fixation heatmaps for each condition. The title of each subplot indicates the number of targets in each hemifield, as well as the percentage of time spent looking anywhere in the green boxes, where the targets and distractors were (compare bottom left panel of Figure 10).

Tracking Related Activity

In line with previous studies (Culham et al., 2001; Alnæs et al., 2014), we first calculated a linear contrast comparing tracking any number of targets to passive viewing of the same stimuli (i.e., no targets in either hemifield) revealing the “task areas”. This contrast showed large clusters (see Figure 14), where attentional tracking causes stronger activations than passive viewing. Parallel to what has been shown previously (Culham et al., 2001; Jovicich et al., 2001; Shim et al., 2010; Jahn et al., 2012; Alnæs et al., 2014; Nummenmaa, et al., 2017), we found strong and widespread activations in the occipital, parietal and frontal lobes.

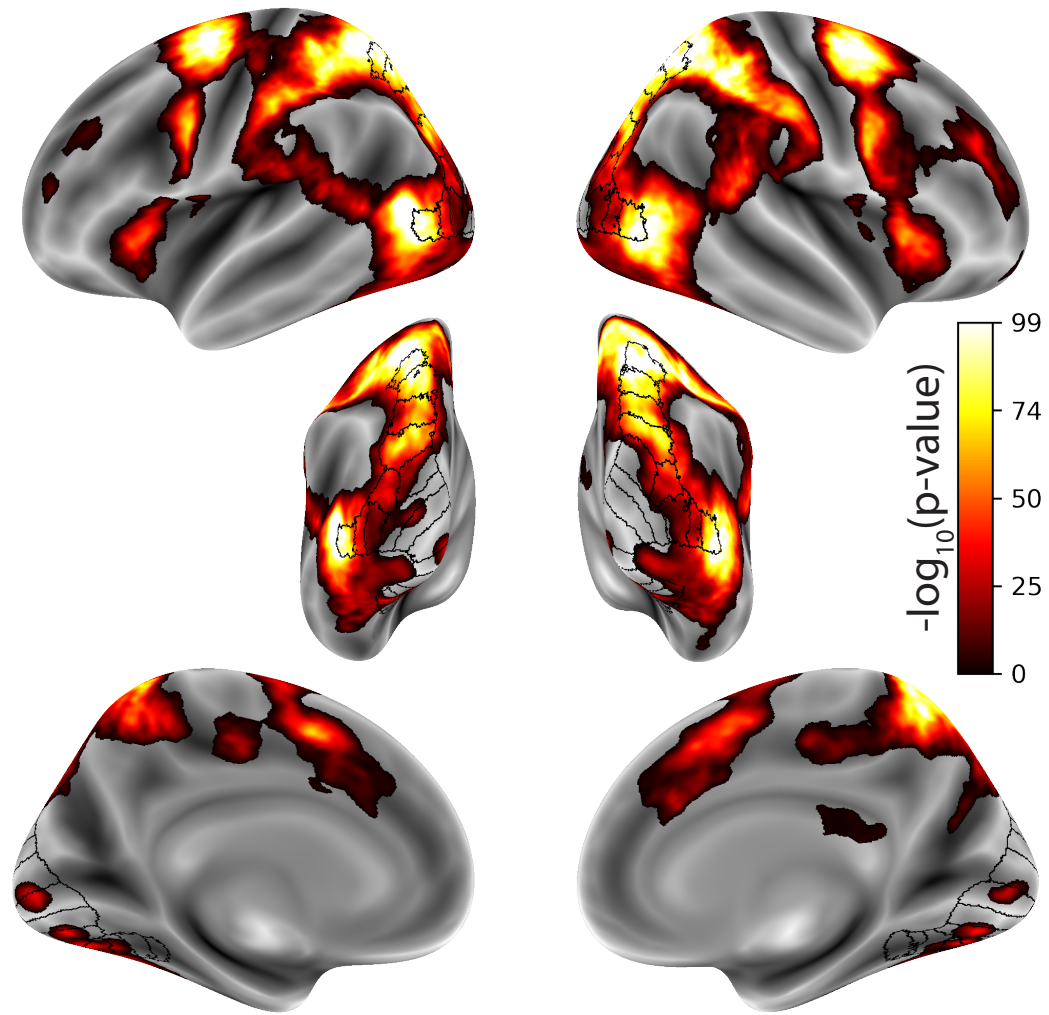


Figure 14. Statistical map comparing tracking to passive viewing. Clusters are shown where BOLD activity was significantly greater during tracking than during passive viewing.

Modulation by Total Attentional Load

Before assessing contralateral bias in ARFs it is necessary to show that we find the canonical ARFs (Culham et al., 2001; Jovicich et al., 2001; Shim et al., 2010; Jahn et al., 2012; Alnæs et al., 2014; Nummenmaa, et al., 2017). A wide array of brain areas

showed parametric modulation of activity depending on the total number of tracked targets (Figure 15). Significant clusters of activations were found in the dorsal visual stream (including but not limited to V3a, V3b, hMT+), as well as the dorsal and ventral attention network, IPS, SPL, FEF, and the anterior insula. The activations were essentially bilateral with slightly larger clusters in the right hemisphere. Additionally, areas, where activity was modulated by the number of tracked targets, were a subset of the areas that were significantly activated by tracking compared to passive viewing.

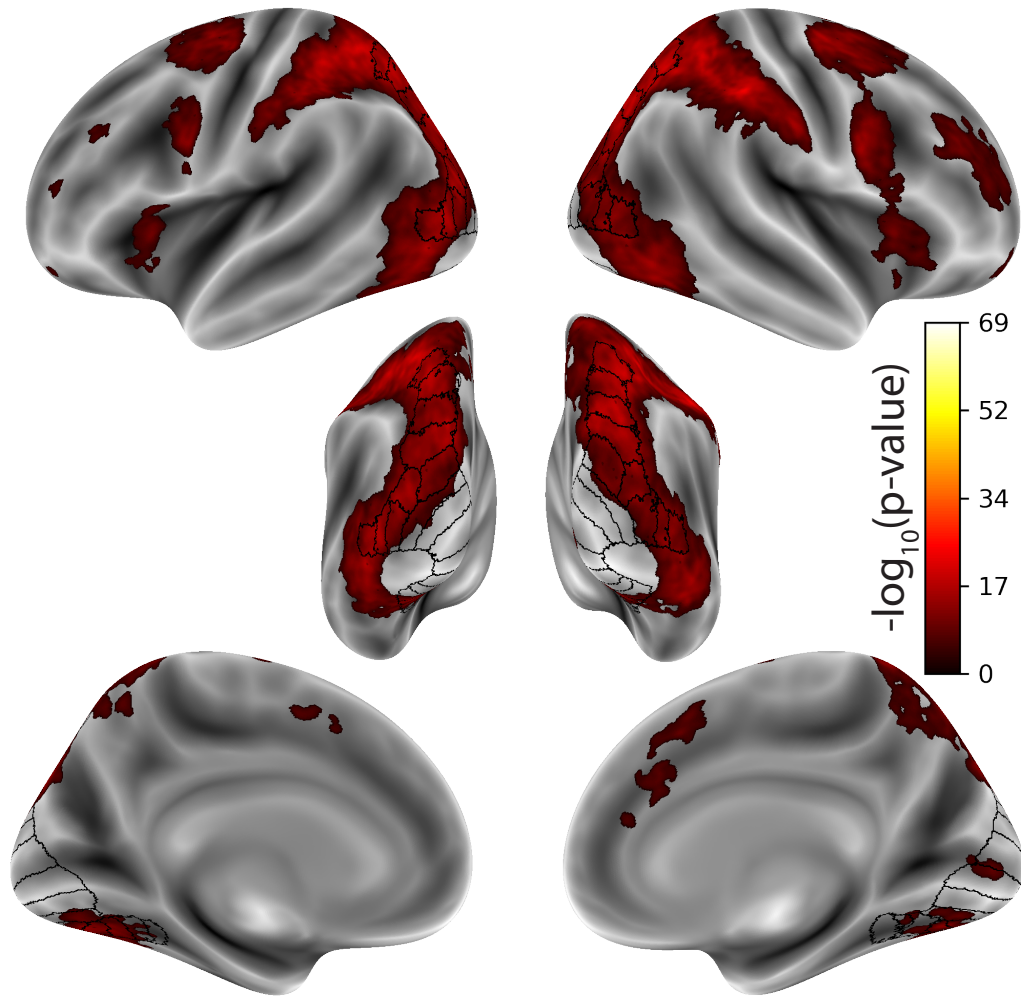


Figure 15. Statistical map of the canonical load-dependent ARF. Significant clusters where BOLD increased with increases in total attentional load, ignoring the exact locations of the targets.

Modulation by Contralateral and Ipsilateral Load

After establishing that our data replicate the canonical ARF findings, we sought to investigate whether this effect is driven by targets bilaterally, or whether ARFs show hemifield specific effects like other aspects of multiple object tracking (Alvarez &

Cavanagh, 2005; Störmer, Alvarez & Cavanagh, 2014; Holcombe & Chen, 2012). To this end, beta maps derived from the hemifield specific parametric modulation regressors in the first level analyses were compared to the baseline. We found contralateral load to drive brain activity throughout the dorsal visual stream, as well as the dorsal attention network. Activation maps for parametric modulation with total load and with contralateral load largely overlapped. Ipsilateral load on the other hand, also showed clusters where BOLD was modulated significantly, but these clusters emerged much later in the visual processing stream and were largely in attention related frontal and parietal areas.

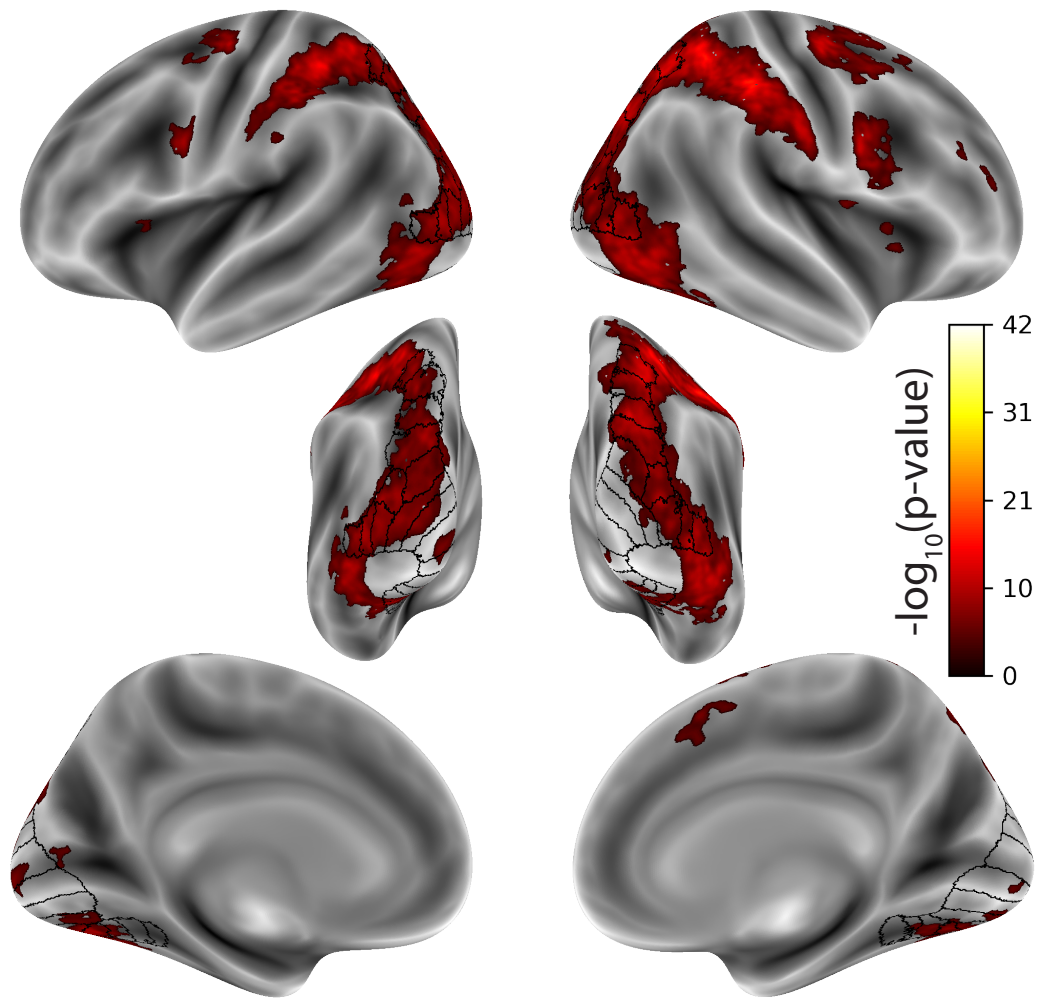


Figure 16. Statistical map of parametric modulation with contralateral load. Significant clusters show a positive linear relationship between contralateral load and brain activity.

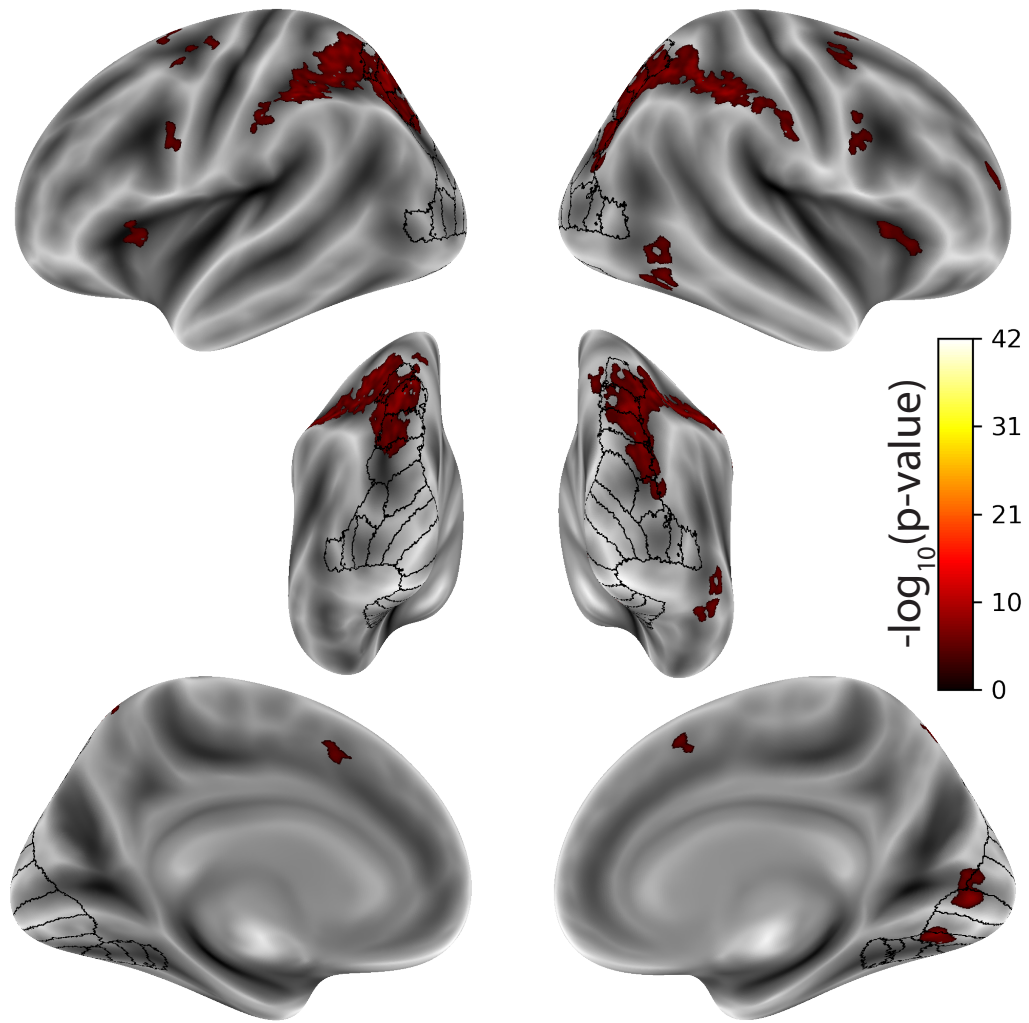


Figure 17. Statistical map of parametric modulation with ipsilateral load. Significant clusters show a positive linear relationship between ipsilateral load and brain activity.

Contralateral Bias

Areas where BOLD was more strongly modulated with contralateral than ipsilateral targets were revealed by directly contrasting the beta maps corresponding to modulation with targets on the left and right with each other. In both hemispheres clusters in the dorsal visual stream (e.g., V3, hMT+) showed stronger modulation with

contralateral than with ipsilateral load. Additionally, there were clusters in the SPL in both hemispheres, with larger clusters in the right hemisphere.

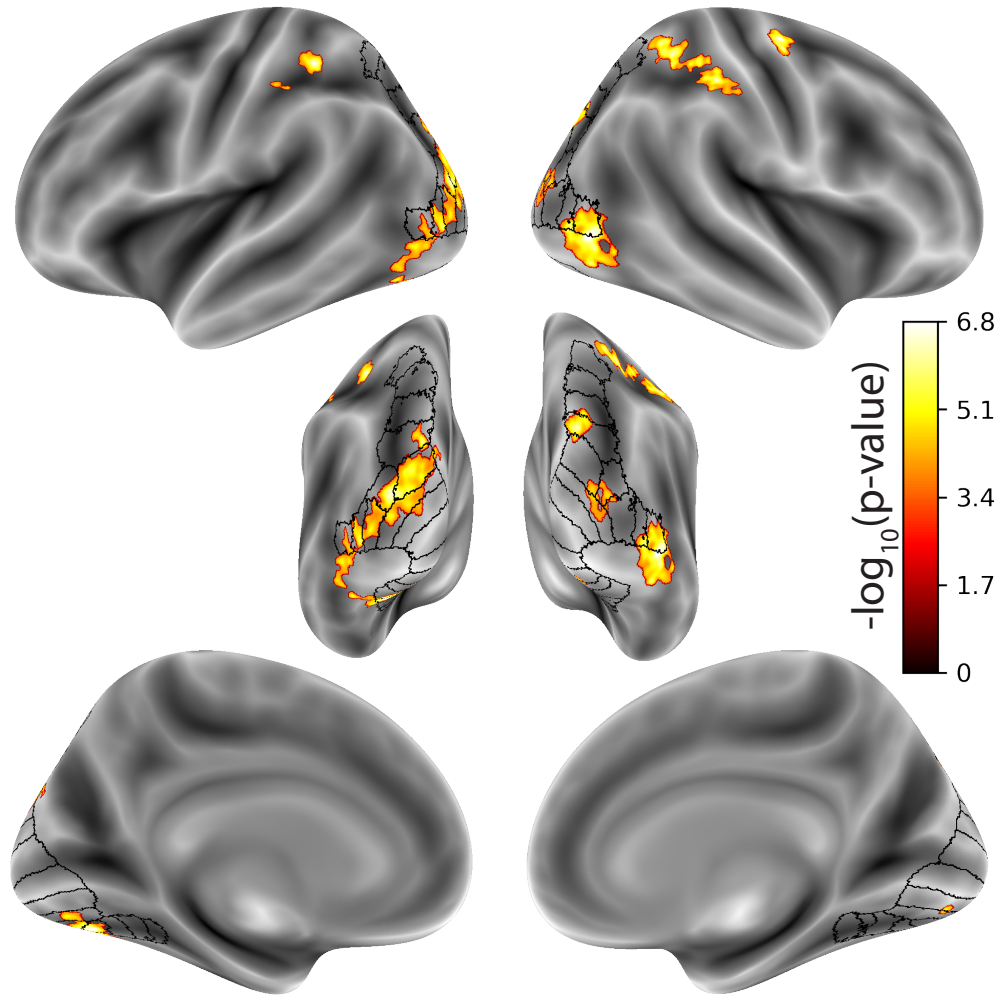


Figure 18. Statistical map comparing modulation with contralateral load to ipsilateral load. Significant clusters are shown where contralateral targets modulate activity more strongly than ipsilateral targets.

Regions of Interest

An analysis using pre-defined parcellations of the brain can be used as an alternative to whole-brain group analyses. Such regions of interest (ROI) analyses can reveal effects that would otherwise get drowned out on the vertex level. Additionally, comparing the brain hemispheres to each other at the vertex level should be avoided because the hemispheres are not fully symmetric. However, by using ROIs equivalent areas can be compared more easily. In addition, ROIs offer a more nuanced perspective on the clusters found in the whole-brain analyses. For example, the contrast comparing ipsilateral and contralateral modulation effects revealed significant clusters in visual cortex, but the exact extent of these clusters depends among other things on choices like the cluster forming threshold. An ROI analysis sacrifices some spatial resolution but increases power and can reveal whether contralateral bias also exists outside the clusters identified in the whole brain analysis. We used a probabilistic atlas of the visual cortex (Wang, Mruczek, Arcaro & Kastner, 2015) which assigns each vertex to the retinotopic brain area it is most likely a part of. While atlas-based ROI analyses are not as accurate as functionally defining a retinotopic map in each participant, they are more economical as they require significantly less data collection. Overlap of the probabilistic atlas we used with functionally defined ROIs from held-out subjects is high (around 40% - 60%) in the ROIs we used.

By fitting a linear model to the BOLD signal from each experimental condition, we can estimate the relationship between the BOLD signal and the number of targets in each hemifield. This allows us to determine the extent to which BOLD increases with each additional target, separately for each hemifield and hemisphere. Averaging the

conditions based on the number of targets in one hemifield provides an average level of brain activity for each load and hemifield. The slope of the line fitted to the mean activity for each number of targets per hemifield indicates the magnitude of brain activity increase with each added target (see Figures 19 and 20). There were no significant differences between hemispheres in their response to contralateral targets after correcting for multiple comparisons (p-values reported in Figure 20 are uncorrected). However, unexpectedly there was a significantly more negative response to ipsilateral targets in right VO1, VO2, V3 and V4. The ROI analysis additionally confirmed the findings from the whole-brain analyses and revealed that contralateral bias exists throughout the dorsal visual stream and even in some ventral visual areas. We also found ipsilateral modulation with load, but only in higher-level parietal areas, which are later in the processing stream again confirming the results of the whole-brain analysis. There was no modulation with load found in V1 and V2.

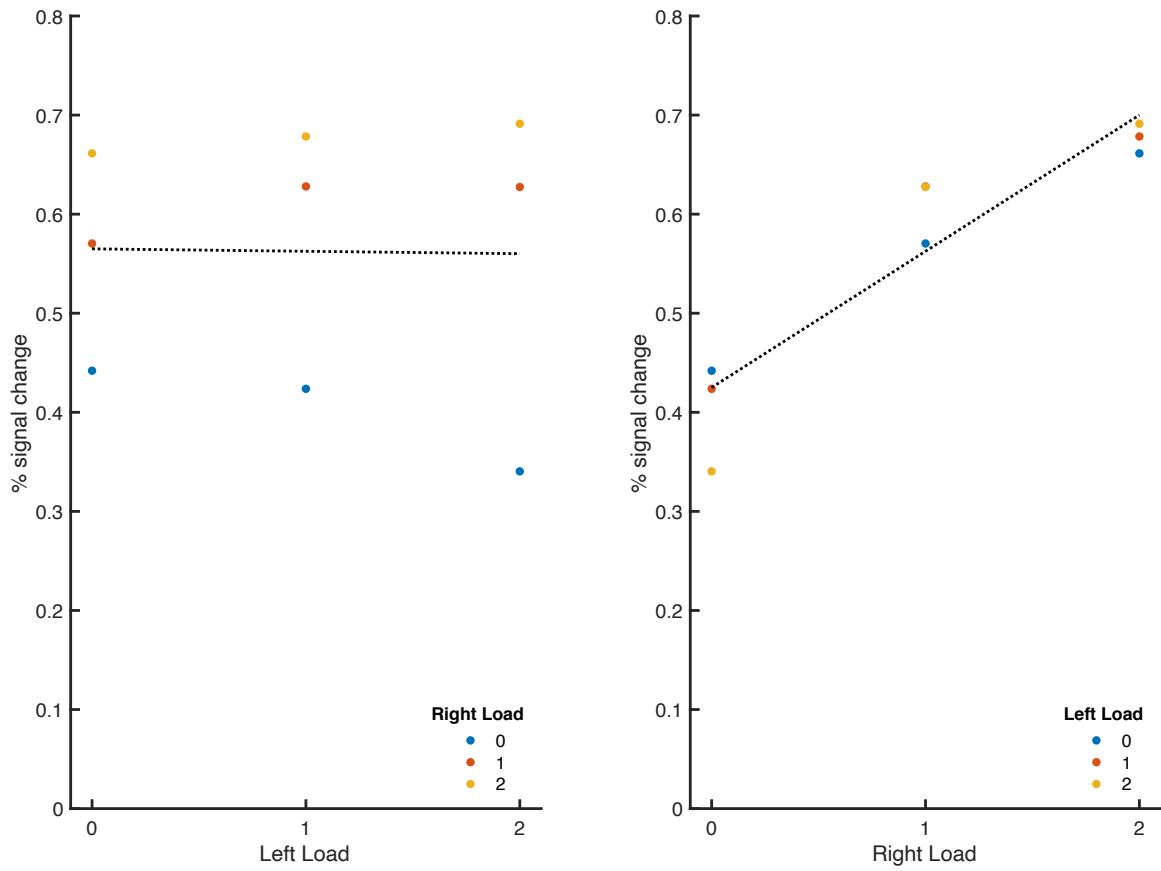


Figure 19. Determining the effect of adding targets separately to each hemifield. Both panels show the same data but with different arrangement. Each data point corresponds to one experimental condition. The slope of the linear fit (black dashed line) is the estimated effect of one added target. Here left hMT was used as an example ROI, where the left panel shows the (lack of a) response to changes in ipsilateral load, while the right panel shows a strong response to changes in contralateral load. See figure 20 for a summary of all ROIs under investigation.

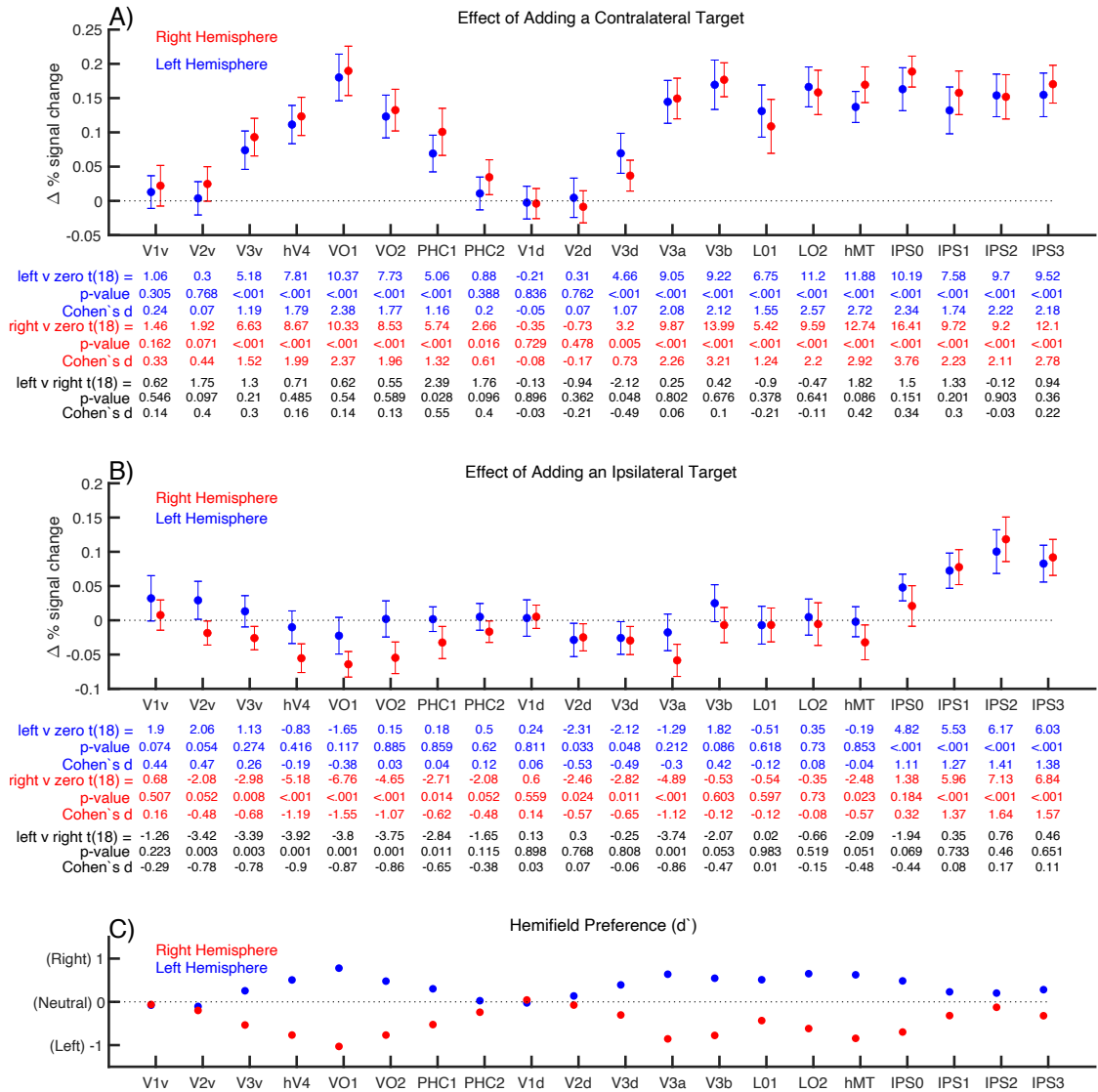


Figure 20. Change in BOLD signal caused by increasing load by one target. These changes were estimated by the slopes fit to data from each ROI as in Figure 19. Error bars show the 95% confidence intervals. A) Changes in brain activity caused by adding one contralateral target. Right hemisphere ROIs in red, left in blue. Statistical values below are comparing the change in BOLD to zero (red and blue) and to each other (black). B) Same as A) caused by adding an ipsilateral target. C) Hemifield preference assessed by comparing within each ROI the response to the left and right hemifields and calculating d-prime. Essentially a comparison of data from A) and B) for each ROI.

Discussion

During multiple object tracking brain activity in several visual, parietal, and frontal areas was strongly modulated by the number of tracking targets. In addition to replicating previous studies showing this phenomenon (Culham et al., 2001; Jovicich et al., 2001; Shim et al., 2010; Jahn et al., 2012; Alnæs et al., 2014), we demonstrate where in the brain this effect depends on the hemifield in which the targets are presented. While contralateral load had a modulating influence on activity in earlier visual areas, both contralateral and ipsilateral load modulated activity in attention-related parietal and frontal areas. Electrophysiological recordings during multiple object tracking have shown a consistent amplification of contralateral responses to tracked targets compared with distractors in early visual cortex (Störmer, Alvarez & Cavanagh, 2014; Störmer, Winther, Li & Andersen 2013; Adamian & Andersen, 2022). While this aligns with our findings, our data additionally revealed some higher-level areas in IPS and SPL involved in contralateral processing. Our results also extend the findings from Shim and colleagues (2010) who showed that contra- and ipsilateral targets influence BOLD activity in parietal lobe ROIs. Additionally, we found a significant contralateral bias in a subset of these load-sensitive areas, specifically in the dorsal visual stream, as well as some parietal and frontal regions.

Our findings validate previous investigations of ARFs (Culham et al., 2001; Jovicich et al., 2001; Shim et al., 2010; Jahn et al., 2012; Alnæs et al., 2014) and extend our understanding of ARFs to include hemifield and hemisphere specificity. Previous descriptions of ARFs, especially the ARFs of brain areas in the occipital lobe, did not incorporate target locations and ignored the hemifield specific processing in those areas.

Our findings also bridge electrophysiological findings of load dependent enhancement in occipital areas (e.g., Störmer, Alvarez & Cavanagh, 2014; Störmer, Winther, Li & Andersen 2013; Adamian & Andersen, 2022) with previous fMRI experiments showing load dependent activity mostly in parietal and frontal areas (Culham et al., 2001; Jovicich et al., 2001; Shim et al., 2010; Jahn et al., 2012; Alnæs et al., 2014).

The capacity for multiple object tracking as well as the maximum speed of trackable objects nearly double when targets are spread across hemifields (Alvarez & Cavanagh, 2005; Störmer, Alvarez & Cavanagh, Holcombe & Chen, 2012; Holcombe, 2023). Capacity and speed limits would not be hemifield specific if the tracking limit was imposed in areas that respond to targets bilaterally. Therefore, there does not appear to be any additional global limit on capacity that is more restrictive than the hemifield limit, otherwise, the hemifield independence would not be seen. In line with previous findings (Störmer, Alvarez & Cavanagh, 2014; Störmer, Winther, Li & Andersen 2013; Adamian & Andersen, 2022), our results imply that the processing bottleneck that limits multiple object tracking is found in early dorsal stream areas such as V3 or hMT+.

The same is not true for multiple identity tracking, which is similar to multiple object tracking with the added requirement to also keep track of which target is which. There, the hemifield-specific effects are less pronounced and the capacity limit dramatically lower (Hudson, Howe & Little, 2012). The capacity for multiple object tracking nearly doubles when targets are spread across hemifields, but during multiple identity tracking the proportional increase is much smaller (Hudson, Howe & Little, 2012; Holcombe, 2023). This suggests that the additional bottlenecks to multiple identity tracking, possibly related to working memory, are outside of the areas where multiple

object tracking activity shows contralateral bias. In comparison to multiple object tracking, multiple identity tracking also causes significantly more activation in the frontal lobe (Nummenmaa, et al., 2017).

A common description of spatial attention involves “gain” or “saliency” maps that send feedback to sensory maps to enhance or suppress processing at relevant locations (Itti & Koch, 2001; Cavanagh, Caplovitz, Lytchenko, Maechler, Tse & Sheinberg, 2023). Load-sensitive areas may reflect the activity in gain maps that represent the tracked targets as well as the consequences for sensory maps that receive feedback from gain maps. Given that target selection likely occurs late (Holcombe, 2023; Maechler, Cavanagh & Tse, 2021; Scholl, 2009) the load dependent ARFs in those areas are more likely consequences of selection (i.e., enhancement) rather than activity related to selection itself. In either case, the areas that were shown to be more sensitive to contralateral load than to ipsilateral load might be the candidate areas for hemifield-specific bottlenecks such as speed and capacity limits (Alvarez & Cavanagh, 2005; Holcombe & Chen, 2012; Hudson, Howe & Little, 2012; Störmer, Alvarez & Cavanagh, 2014).

We propose that the capacity bottleneck for multiple object tracking might arise in areas such as V3 and hMT+ where the modulation of activity with contralateral load was significantly stronger compared to ipsilateral load. However, we deliberately chose stimulus parameters that would allow participants to track four targets with high accuracy (Franconeri, Lin, Enns, Pylyshyn & Fisher, 2008). Due to this, accuracy was near ceiling, and we observed no notable hemifield-specific capacity. With higher difficulty, we might have replicated hemifield-independent resources (Alvarez & Cavanagh, 2014; Holcombe,

2023). Nevertheless, such increased task demands would likely lead to participants losing track of targets. To make reliable claims about the number of targets being tracked, it is crucial to ensure that participants are genuinely capable of tracking that specific number of targets. This shortcoming of our experiment can easily be overcome with follow up experiments. A future experiment investigating this further should include conditions with higher difficulties to find hemifield-specific capacity limits and link them to brain activity. If our prediction holds, activity in the dorsal visual stream, where contralateral targets show significantly stronger modulatory effects, should follow the hemifield capacity ceiling. Specifically, we would expect that load dependent ARFs show an increase in BOLD with the number of targets up to the capacity limit and asymptote with more targets.

Our study is the first to investigate hemifield- and hemisphere-specific effects during multiple object tracking with fMRI. There is a large literature demonstrating hemifield and hemisphere-specific effects in spatial attention, using many paradigms in addition to attentional tracking. For instance, in a spatial orienting task, where spatial attention is deployed according to a pre-stimulus cue, BOLD activity in the IPS and the FEF increases much more contralaterally to the attended side than ipsilaterally (Mayrhofer, Duecker, van de Ven, Jacobs & Sack, 2019). Contralateral bias in visual cortex has also been investigated in the context of visuospatial working memory, which likely shares mechanisms with spatial attention. While some experiments find contralateral bias during lateralized working memory tasks in both hemispheres (Killebrew, Mruczek & Berryhill, 2015) others find contralateral bias only in the left but not the right hemisphere (Sheremata, Bettencourt & Somers, 2010). Our findings expand

this literature to include contralateral bias in the visual system during multiple object tracking. In addition, we show a transition from earlier in the visual stream, where there is strong contralateral bias, to later, where ipsilateral targets also influence activity. This is not surprising, since these earlier visual areas have hemifield maps with no representations of ipsilateral targets, while the later areas have larger receptive fields that likely extend to both hemifields.

We observed a bilateral parametric modulation with target load, with a slightly stronger effect in the right hemisphere. This finding is consistent with the right hemisphere dominance that has been reported in previous attention paradigms. Posner and Petersen (Posner & Petersen, 1990; Petersen & Posner, 2012) proposed a taxonomy for spatial attention that is segregated into orienting, alertness, and executive control. In their review paper (Petersen & Posner, 2012) they note that many studies have found aspects of spatial attention to be right-lateralized in the brain. This is not surprising considering the vast literature on the lateralization of hemispatial neglect (Parton, Malhotra & Husain, 2003; Corbetta & Shulman, 2011; Mesulam, 1981; Corbetta & Shulman, 2011; Duecker & Sack, 2015; Heilman, & Van Den Abell, 1980; Kinsbourne, 1987; Duecker & Sack, 2015). Several theoretical accounts of neglect propose different degrees of contralateral bias of the hemispheres for spatial attention. One theory proposes that the left hemisphere controls attention shifts towards the right hemifield, while the right hemisphere controls attention shifts to both hemifields (Mesulam, 1981; Corbetta & Shulman, 2011; Duecker & Sack, 2015; Heilman, & Van Den Abell, 1980), while the other suggests that each hemisphere has a contralateral bias for attention shifting but also represents the ipsilateral hemifield. Our data showing contralateral bias in both

hemispheres, as well as responses to ipsilateral targets in both hemispheres, are more in line with the latter theory. A recent review shows that this theory is also superior at explaining data from neglect patients (Esposito, Shekhtman & Chen, 2021). Note, however, that multiple object tracking might not capture all aspects of spatial attention that are impaired by hemispatial neglect.

Conclusions

In conclusion, our study confirmed and extended previous research on the neural correlates of multiple object tracking. We found that brain activity in visual, parietal, and frontal cortex was strongly modulated by the number of tracking targets and that many load-sensitive areas respond to targets in both hemifields. Furthermore, we identified clusters showing a contralateral bias in a subset of brain areas implicated in multiple object tracking. Our findings suggest that these areas may be candidate regions for hemifield-specific limitations in attentional tracking. Crucially, our study contributes to the understanding of hemifield and hemisphere-specific effects in multiple object tracking. Overall, our study sheds light on the neural mechanisms underlying multiple object tracking and provides insights into the hemifield and hemisphere-specific effects of spatial attention.

General Discussion

The studies described in this dissertation answered several questions surrounding attentional tracking of moving objects. In the first chapter, we showed that covert attentional tracking selects object locations that are in perceived coordinates. This suggests that the gain maps that are used for attentional tracking are likely also in perceptual coordinates. One of the earliest questions surrounding attentional tracking was whether target selection for tracking occurs early or late in the visual hierarchy (Pylyshyn & Storm, 1988). Theories of multiple object tracking make varied predictions regarding the influence of higher-level cognitive properties, with Pylyshyn (1989) originally proposing an early, parallel, and automatic process, while more recent theories favor late selection (Oksama & Hyönä, 2004; Cavanagh & Alvarez, 2005). We argued in a recent review that objects are fully processed, up to the level of identity, before attention can select them, at least for object-based attention (Cavanagh, Caplovitz, Lytchenko, Maechler, Tse & Sheinberg, 2023). In Chapter One we showed that attention selects targets from perceived, illusory positions, when those differ from physical positions, suggesting that selection occurs late, at least after illusory positions come into play. Our finding fits with similar lines of investigation. Scholl and colleagues (Scholl, Pylyshyn & Feldman, 2001) showed that attention fully selects the whole of a tracked object, rather than just parts of it. In their study, participants were able to successfully track about four square targets among distractors but struggled to track the same locations when they were part of joined bar-shaped objects. Others (Howe, Drew, Pinto & Horowitz, 2011) have shown that attentional tracking remaps tracked positions to spatiotopic locations rather than relying on retinotopic coordinates. Together, these findings show that rather than an

early and automatic process, tracking selects targets very late in the visual processing stream, after illusory positions, spatiotopic coordinates, and object identities have been established.

In the second chapter, we linked this to smooth pursuit. We found that the oculomotor system uses perceptual coordinates to select targets for parafoveal smooth pursuit. This is in contrast with saccades made toward the same illusion (Lisi & Cavanagh, 2015; Nakayama & Holcombe, 2020), but it is in line with saccades that are aimed at the remembered location of the illusion after having disappeared from the screen (Massendari, Lisi, Collins & Cavanagh, 2018). We replicated earlier studies (Beutter & Stone, 2000; Hafed & Krauzlis, 2008; Steinbach, 1976) demonstrating that observers can pursue an inferred location. However, unlike previous research where the inferred midpoints were based on stimulus components that were aligned with retinal locations, our study utilizes stimuli with perceived locations significantly deviating from their physical locations. This aligns with similar findings showing smooth pursuit of a motion aftereffect on a static stimulus (Braun, Pracejus, & Gegenfurtner, 2006; Matsumiya & Shioiri, 2015), as well as studies demonstrating that the internal motion of a Gabor affects perception and smooth pursuit similarly (Hughes, 2018). We show that overt and covert tracking, just like pointing, utilize the perceived positions of targets. Perception and action are often linked (Goodale & Milner, 1992) and in contrast to previous findings of saccades targeting physical locations (Lisi & Cavanagh, 2015) we show that this link exists for tracking.

Given the deep link between saccadic eye movements and ballistic attention shifts (e.g., Awh, Armstrong & Moore, 2006; Moore & Zirnsak, 2017; Rolfs, Jonikaitis, Deubel

& Cavanagh, 2011) it would not be surprising to find a similar link between smooth pursuit and continuous attention shifts as in attentional tracking. Indeed, there are some studies attempting to link the two. For example, saccades are not disruptive to attentional tracking as long as spatiotopic coordinates are preserved (Howe, Drew, Pinto & Horowitz, 2011) while smooth pursuit concurrently with attentional tracking leads to decreases in performance (Howe, Pinto & Horowitz, 2010), suggesting that they use shared resources. Our results showing that smooth pursuit and attentional tracking both rely on perceived coordinates to select targets also suggest that they share a mechanism for selection, but further experiments will be needed to prove this. For example, if they do share neural circuitry, it should be possible to crossdecode the directions of attentional tracking based on classifiers trained on directions of smooth pursuit and vice versa.

The third chapter investigated brain activity during multiple object tracking. Not all visual maps represent the entire visual field. Some maps, especially earlier in the visual processing hierarchy, represent only the contralateral hemifield, meaning they selectively process information from either the left or right side of the visual field. Hemifield-specific effects have been observed in tracking (Alvarez & Cavanagh, 2005; Holcombe & Chen, 2012; Hudson, Howe & Little, 2012; Störmer, Alvarez & Cavanagh, 2014), which has implications for the maps from which tracking might select its targets. Since tracking resources are largely independent for each hemifield and hemisphere, it is plausible that the resource limitations imposed on attentional tracking lie early in the visual processing hierarchy. The capacity bottleneck that imposes the maximum number of trackable targets has been suggested to lie in early visual cortex based on electrophysiological recordings during tracking (Störmer, Alvarez & Cavanagh, 2014;

Adamian & Andersen, 2022). Our finding that activity is significantly more modulated by contra- than ipsilateral tracking load in early visual areas that to a large extent process the contralateral hemifield, aligns well with this idea.

Theoretical models of the parietal lobe explaining hemispatial neglect propose varying levels of contralateral bias in the hemispheres regarding spatial attention. An earlier theory postulated that the left hemisphere primarily directs attentional shifts toward the right hemifield, whereas the right hemisphere facilitates attention shifts toward both hemifields (Mesulam, 1981). In contrast, an alternative perspective suggests that each hemisphere exhibits a contralateral bias for attentional shifts while also representing the ipsilateral hemifield (Corbetta & Shulman, 2011; Duecker & Sack, 2015; Heilman & Van Den Abell, 1980). Our finding that ipsilateral load modulates activity in the superior parietal lobule in both hemispheres is more in line with the latter theories.

Tracking more targets requires more resources and leads to more brain activity in certain brain areas. Beyond this obvious correlation, it is not entirely clear how tracking resources relate to attentional response functions. Lack of cortical real estate could be one way in which resources for attentional tracking might be limited (Intrilligator & Cavanagh, 2001; Franconeri, Alvarez & Cavanagh, 2013). In the gain map metaphor, this would mean that when the targets are closely spaced, the activation for each is suppressed by the inhibitory surrounds of the other nearby spotlights, reducing activation and degrading performance. Alternatively, there might be some limit on how much gain enhancement there can be overall.

Overall, the findings of the studies presented in this dissertation expand our knowledge regarding the selection and enhancement of targets during tracking. However,

future studies will need to address some open questions. It is possible that smooth pursuit is influenced differently by the illusion if the order of illusory and non-illusory motion was changed (i.e., down with illusion, up without illusion), as there are some biases in downward perception (e.g., Phan, Harris, Kim & Kingdom, 2022). In Chapter Three we suggest that areas with significant contralateral bias are the source of capacity bottlenecks during attentional tracking. While this aligns with other findings (e.g., Störmer, Alvarez & Cavanagh, 2014; Störmer, Winther, Li & Andersen 2013; Adamian & Andersen, 2022) it will be necessary to confirm our prediction with another fMRI experiment, where the hemifield specific capacity limits are reached by increasing the task difficulty. If these areas truly are the source of capacity and speed limits to attentional tracking, then activity there should only increase with load so long as participants can successfully track that load. In trials with more targets than the capacity limit, activity in bottleneck areas should not exceed the activity associated with trials where participants track the maximum number of trackable targets. In other words, one would expect the load dependent ARFs to asymptote above the limit of trackable targets. Additionally, the spatial specificity of this study could be improved by measuring the retinotopic maps from each participant instead of using atlas-based ROIs.

Conclusions

In this dissertation, evidence has been presented that tracking, be it overt or covert, selects target positions from highly processed visual representations likely late in the processing stream. Specifically, perceived (illusory) positions rather than retinal locations determine whether targets get too close to distractors, and their perceived positions determine where attention is deployed. Target distractor spacing is only one kind of bottleneck for attentional tracking and it appears to emerge comparatively late in the process of constructing visual experiences. In contrast to this, the hemifield-specific capacity bottleneck for attentional tracking that limits the total number of trackable targets might appear rather early in the visual processing hierarchy. Since the capacity limit is hemifield specific, we would expect it to emerge in hemifield maps in the dorsal visual stream, where we found attentional modulation of processing to be hemifield specific.

References

- 't Hart, B. M., Henriques, D. Y., & Cavanagh, P. (2019). Manual tracking of the double-drift illusion. *Journal of Vision*, *19*(10), 286b-286b. doi: <https://doi.org/10.1167/19.10.286b>
- Adamian, N., & Andersen, S. K. (2022). Attentional enhancement of tracked stimuli in early visual cortex has limited capacity. *Journal of Neuroscience*, *42*(46), 8709-8715.
- Alvarez, G. A., & Cavanagh, P. (2005). Independent resources for attentional tracking in the left and right visual hemifields. *Psychological science*, *16*(8), 637-643.
- Andersson, J. L., Skare, S., & Ashburner, J. (2003). How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. *Neuroimage*, *20*(2), 870-888.
- Avants, B. B., Epstein, C. L., Grossman, M., & Gee, J. C. (2008). Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Medical image analysis*, *12*(1), 26-41.

Awh, E., Armstrong, K. M., & Moore, T. (2006). Visual and oculomotor selection: links, causes and implications for spatial attention. *Trends in cognitive sciences*, 10(3), 124-130.

Bennett, C. M., Miller, M. B., & Wolford, G. L. (2009). Neural correlates of interspecies perspective taking in the post-mortem Atlantic Salmon: An argument for multiple comparisons correction. *Neuroimage*, 47(Suppl 1), S125.

Beutter, B. R., & Stone, L. S. (2000). Motion coherence affects human perception and pursuit similarly. *Visual neuroscience*, 17(1), 139-153.
<https://doi.org/10.1017/s0952523800171147> [PubMed]

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial vision*, 10, 433-436.

Braun, D. I., Pracejus, L., & Gegenfurtner, K. R. (2006). Motion aftereffect elicits smooth pursuit eye movements. *Journal of Vision*, 6(7), 1-1. doi: <https://doi.org/10.1167/6.7.1> [PubMed]

Bruno, N. (2001). When does action resist visual illusions?. *Trends in cognitive sciences*, 5(9), 379-382. doi: [https://doi.org/10.1016/s1364-6613\(00\)01725-3](https://doi.org/10.1016/s1364-6613(00)01725-3) [PubMed]

- Bruno, N., Knox, P. C., & de Grave, D. D. (2010). A metanalysis of the effect of the Müller-Lyer illusion on saccadic eye movements: no general support for a dissociation of perception and oculomotor action. *Vision Research*, 50(24), 2671-2682. doi: <https://doi.org/10.1016/j.visres.2010.09.016> [PubMed]
- Cavanagh, P., & Alvarez, G. A. (2005). Tracking multiple targets with multifocal attention. *Trends in cognitive sciences*, 9(7), 349-354.
- Cavanagh, P., & Anstis, S. (2013). The flash grab effect. *Vision Research*, 91, 8-20.
- Cavanagh, P., Caplovitz, G. P., Lytchenko, T. K., Maechler, M. R., Tse, P. U., & Sheinberg, D. (2023). The architecture of Object-Based Attention. *Psychonomic Bulletin and Review*. DOI: 10.3758/s13423-023-02281-7
- Cavanagh, P., & Tse, P. U. (2019). The vector combination underlying the double-drift illusion is based on motion in world coordinates: Evidence from smooth pursuit. *Journal of Vision*, 19(14), 2-2.
- Chen, J., Valsecchi, M., & Gegenfurtner, K. R. (2017). Attention is allocated closely ahead of the target during smooth pursuit eye movements: Evidence from EEG frequency tagging. *Neuropsychologia*, 102, 206-216.

- Chung, M. K., Worsley, K. J., Nacewicz, B. M., Dalton, K. M., & Davidson, R. J. (2010). General multivariate linear modeling of surface shapes using SurfStat. *Neuroimage*, 53(2), 491-505.
- Cooper, E. A., Jiang, H., Vildavski, V., Farrell, J. E., & Norcia, A. M. (2013). Assessment of OLED displays for vision research. *Journal of Vision*, 13(12), 16-16. doi: <https://doi.org/10.1167/13.12.16> [[PubMed](#)]
- Cox, R. W., & Hyde, J. S. (1997). Software tools for analysis and visualization of fMRI data. *NMR in Biomedicine: An International Journal Devoted to the Development and Application of Magnetic Resonance In Vivo*, 10(4-5), 171-178.
- Dakin, S. C., Greenwood, J. A., Carlson, T. A., & Bex, P. J. (2011). Crowding is tuned for perceived (not physical) location. *Journal of Vision*, 11(9), 2-2.
- Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis: I. Segmentation and surface reconstruction. *Neuroimage*, 9(2), 179-194.
- Deming, W. E. (1943). Statistical adjustment of data. Wiley, NY (Dover Publications edition, 1985).
- De Valois, R. L., & De Valois, K. K. (1991). Vernier acuity with stationary moving Gabors. *Vision research*, 31(9), 1619-1626.

- Dosenbach, N.U., Visscher, K.M., Palmer, E.D., Miezin, F.M., Wenger, K.K., Kang, H.C., Burgund, E.D., Grimes, A.L., Schlaggar, B.L., & Petersen, S.E., (2006). A core system for the implementation of task sets. *Neuron*, 50(5), 799-812.
- Duncker, K. (1929). Über induzierte Bewegung. *Psychologische Forschung*, 12(1), 180-259. doi: <https://doi.org/10.1007/BF02409210>
- Eagleman, D. M., & Sejnowski, T. J. (2007). Motion signals bias localization judgments: A unified explanation for the flash-lag, flash-drag, flash-jump, and Frohlich illusions. *Journal of vision*, 7(4), 3-3. doi: <https://doi.org/10.1167/7.4.3> [PubMed]
- Esposito, E., Shekhtman, G., & Chen, P. (2021). Prevalence of spatial neglect post-stroke: a systematic review. *Annals of Physical and Rehabilitation Medicine*, 64(5), 101459.
- Esteban, O., Markiewicz, C. J., Blair, R. W., Moodie, C. A., Isik, A. I., Erramuzpe, A., ... & Gorgolewski, K. J. (2019). fMRIPrep: a robust preprocessing pipeline for functional MRI. *Nature methods*, 16(1), 111-116.
- Franconeri, S. L., Alvarez, G. A., & Cavanagh, P. (2013). Flexible cognitive resources: competitive content maps for attention and memory. *Trends in cognitive sciences*, 17(3), 134-141.

- Franconeri, S. L., Lin, J. Y., Enns, J. T., Pylyshyn, Z. W., & Fisher, B. (2008). Evidence against a speed limit in multiple-object tracking. *Psychonomic bulletin & review*, *15*, 802-808.
- Garcia-Pérez, M. A. (1998). Forced-choice staircases with fixed step sizes: asymptotic and small-sample properties. *Vision research*, *38*(12), 1861-1881.
- Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in neurosciences*, *15*(1), 20-25. doi: [https://doi.org/10.1016/0166-2236\(92\)90344-8](https://doi.org/10.1016/0166-2236(92)90344-8) [[PubMed](#)]
- Gorgolewski, K., Burns, C. D., Madison, C., Clark, D., Halchenko, Y. O., Waskom, M. L., & Ghosh, S. S. (2011). Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in python. *Frontiers in neuroinformatics*, *13*.
- Greene, D. J., Barnea, A., Herzberg, K., Rassis, A., Neta, M., Raz, A., & Zaidel, E. (2008). Measuring attention in the hemispheres: The lateralized attention network test (LANT). *Brain and cognition*, *66*(1), 21-31.
- Greve, D. N., & Fischl, B. (2009). Accurate and robust brain image alignment using boundary-based registration. *Neuroimage*, *48*(1), 63-72.

- Gurnsey, R., & Biard, M. (2012). Eccentricity dependence of the curveball illusion. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 66(2), 144.
- Hafed, Z. M., & Krauzlis, R. J. (2008). Goal representations dominate superior colliculus activity during extrafoveal tracking. *Journal of Neuroscience*, 28(38), 9426-9439.
doi: <https://doi.org/10.1523/jneurosci.1313-08.2008> [[PubMed](#)]
- He, S., Cavanagh, P., & Intriligator, J. (1996). Attentional resolution and the locus of visual awareness. *Nature*, 383(6598), 334.
- He, Z. J., & Nakayama, K. (1992). Surfaces versus features in visual search. *Nature*, 359(6392), 231.
- Heilman, K. M., & Van Den Abell, T. (1980). Right hemisphere dominance for attention: the mechanism underlying hemispheric asymmetries of inattention (neglect). *Neurology*, 30(3), 327-327.
- Hochstein, S., & Ahissar, M. (2002). View from the top: Hierarchies and reverse hierarchies in the visual system. *Neuron*, 36(5), 791-804.

- Holcombe, A. O., & Chen, W. Y. (2012). Exhausting attentional tracking resources with a single fast-moving object. *Cognition*, 123(2), 218-228.
- Holcombe, A. O., Chen, W. Y., & Howe, P. D. (2014). Object tracking: Absence of long-range spatial interference supports resource theories. *Journal of vision*, 14(6), 1-1.
- Horowitz, T. S., Holcombe, A. O., Wolfe, J. M., Arsenio, H. C., & DiMase, J. S. (2004). Attentional pursuit is faster than attentional saccade. *Journal of vision*, 4(7), 6-6.
- Howe, P. D., Drew, T., Pinto, Y., & Horowitz, T. S. (2011). Remapping attention in multiple object tracking. *Vision research*, 51(5), 489-495.
- Howe, P. D., Pinto, Y., & Horowitz, T. S. (2010). The coordinate systems used in visual tracking. *Vision research*, 50(23), 2375-2380.
- Hughes, A. E. (2018). Dissociation between perception and smooth pursuit eye movements in speed judgments of moving Gabor targets. *Journal of vision*, 18(4), 4-4. doi: <https://doi.org/10.1167/18.4.4> [PubMed]
- Hui, J., Wang, Y., Zhang, P., Tse, P. U., & Cavanagh, P. (2020). Apparent motion is computed in perceptual coordinates. *i-Perception*, 11(4), doi: <https://doi.org/10.1177/2041669520933309> [PubMed]

- Hutton, S. B., & Tegally, D. (2005). The effects of dividing attention on smooth pursuit eye tracking. *Experimental Brain Research*, 163, 306-313.
- Jahn, G., Wendt, J., Lotze, M., Papenmeier, F., & Huff, M. (2012). Brain activation during spatial updating and attentive tracking of moving targets. *Brain and cognition*, 78(2), 105-113.
- Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage*, 17(2), 825-841.
- Kaernbach, C. (1991). Simple adaptive testing with the weighted up-down method. *Perception & psychophysics*, 49(3), 227-229.
- Kato, R., Takaura, K., Ikeda, T., Yoshida, M., & Isa, T. (2011). Contribution of the retino - tectal pathway to visually guided saccades after lesion of the primary visual cortex in monkeys. *European Journal of Neuroscience*, 33(11), 1952-1960.
doi: <https://doi.org/10.1111/j.1460-9568.2011.07729.x> [PubMed]
- Klein, A., Ghosh, S. S., Bao, F. S., Giard, J., Häme, Y., Stavsky, E., ... & Keshavan, A. (2017). Mindboggling morphometry of human brains. *PLoS computational biology*, 13(2), e1005350.

Kosovicheva, A. A., Wolfe, B. A., & Whitney, D. (2014). Visual motion shifts saccade targets. *Attention, Perception, & Psychophysics*, 76(6), 1778-1788.

Kowler, E. (2011). Eye movements: The past 25 years. *Vision research*, 51(13), 1457-1483. doi: <https://doi.org/10.1016/j.visres.2010.12.014> [[PubMed](#)]

Krauzlis, R. J. (2004). Recasting the smooth pursuit eye movement system. *Journal of neurophysiology*, 91(2), 591-603. doi: <https://doi.org/10.1152/jn.00801.2003> [[PubMed](#)]

Kwon, O. S., Tadin, D., & Knill, D. C. (2015). Unifying account of visual motion and position perception. *Proceedings of the National Academy of Sciences*, 112(26), 8142-8147. doi: <https://doi.org/10.1073/pnas.1500361112> [[PubMed](#)]

Lisberger, S. G. (2010). Visual guidance of smooth-pursuit eye movements: sensation, action, and what happens in between. *Neuron*, 66(4), 477-491. doi: <https://doi.org/10.1016/j.neuron.2010.03.027> [[PubMed](#)]

Lisi, M., & Cavanagh, P. (2015). Dissociation between the perceptual and saccadic localization of moving objects. *Current Biology*, 25(19), 2535-2540. doi: <https://doi.org/10.1016/j.cub.2015.08.021> [[PubMed](#)]

Lisi, M., & Cavanagh, P. (2017). Different spatial representations guide eye and hand movements. *Journal of Vision*, 17(2), 12-12. doi: <https://doi.org/10.1167/17.2.12> [[PubMed](#)]

Liu, S., Tse, P. U., & Cavanagh, P. (2018). Meridian interference reveals neural locus of motion-induced position shifts. *Journal of neurophysiology*, 119(6), 2091-2099.

Liu, S., Yu, Q., Tse, P. U., & Cavanagh, P. (2019). Neural Correlates of the Conscious Perception of Visual Location Lie Outside Visual Cortex. *Current Biology*.

Logothetis, N. K., & Schall, J. D. (1990). Binocular motion rivalry in macaque monkeys: eye dominance and tracking eye movements. *Vision research*, 30(10), 1409-1419. doi: [https://doi.org/10.1016/0042-6989\(90\)90022-d](https://doi.org/10.1016/0042-6989(90)90022-d) [[PubMed](#)]

Lovejoy, L. P., Fowler, G. A., & Krauzlis, R. J. (2009). Spatial allocation of attention during smooth pursuit eye movements. *Vision Research*, 49(10), 1275-1285.

Luke, S. G. (2017). Evaluating significance in linear mixed-effects models in R. *Behavior research methods*, 49, 1494-1502.

Ma, Z., Watamaniuk, S. N., & Heinen, S. J. (2017). Illusory motion reveals velocity matching, not foveation, drives smooth pursuit of large objects. *Journal of vision*, 17(12), 20-20. doi: <https://doi.org/10.1167/17.12.20> [[PubMed](#)]

Mack, A., Fendrich, R., & Wong, E. (1982). Is perceived motion a stimulus for smooth pursuit. *Vision Research*, 22(1), 77-88. doi: [https://doi.org/10.1016/0042-6989\(82\)90169-9](https://doi.org/10.1016/0042-6989(82)90169-9) [[PubMed](#)]

Maechler, M. R., Cavanagh, P., & Tse, P. U. (2021). Attentional tracking takes place over perceived rather than veridical positions. *Attention, Perception, & Psychophysics*, 83, 1455-1462.

Maechler, M. R., Heller, N. H., Lisi, M., Cavanagh, P., & Peter, U. T. (2021). Smooth pursuit operates over perceived not physical positions of the double-drift stimulus. *Journal of Vision*, 21(11), 6.

Massendari, D., Lisi, M., Collins, T., & Cavanagh, P. (2017). Memory-guided saccades show effect of a perceptual illusion whereas visually guided saccades do not. *Journal of neurophysiology*, 119(1), 62-72.

Matsumiya, K., & Shioiri, S. (2015). Smooth pursuit eye movements and motion perception share motion signals in slow and fast motion mechanisms. *Journal of vision*, 15(11), 12-12. doi: <https://doi.org/10.1167/15.11.12> [[PubMed](#)]

Maus, G. W., Fischer, J., & Whitney, D. (2011). Perceived positions determine crowding. *PLoS One*, 6(5).

- Moore, T., & Zirnsak, M. (2017). Neural mechanisms of selective visual attention. *Annual review of psychology*, 68, 47-72.
- Nummenmaa, L., Oksama, L., Glerean, E., & Hyönä, J. (2017). Cortical circuit for binding object identity and location during multiple-object tracking. *Cerebral Cortex*, 27(1), 162-172.
- Oksama, L., & Hyönä, J. (2004). Is multiple object tracking carried out automatically by an early vision mechanism independent of higher-order cognition? An individual difference approach. *Visual Cognition*, 11(5), 631-671.
- Özkan, M., Tse, P. U. & Cavanagh, P. (2020). Pop-out for illusory rather than veridical trajectories with double-drift stimuli. in press in *Attention, Perception, & Psychophysics*.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial vision*, 10, 437-442.
- Pelli, D. G. (2008). Crowding: A cortical constraint on object recognition. *Current opinion in neurobiology*, 18(4), 445-451.

- Phan, M. H., Harris, L., & Kim, J. (2022). Downwards versus upwards motion-more Aristotelian than we think. *Journal of Vision*, 22(14), 4274-4274.
- Pylyshyn, Z. (1989). The role of location indexes in spatial perception: A sketch of the FINST spatial-index model. *Cognition*, 32(1), 65-97.
- Ramachandran, V. S., & Anstis, S. M. (1990). Illusory displacement of equiluminous kinetic edges. *Perception*, 19(5), 611-616.
- Rolfs, M., Jonikaitis, D., Deubel, H., & Cavanagh, P. (2011). Predictive remapping of attention across eye movements. *Nature neuroscience*, 14(2), 252-256.
- Schneider, M., Marquardt, I., Sengupta, S., de Martino, F., & Goebel, R. (2019). Linking apparent position to population receptive field estimates using a visual field projection model. Conference on Cognitive Computational Neuroscience, September, Berlin, Germany. doi: 10.32470/CCN.2019.1185-0
- Scholl, B. J., Pylyshyn, Z. W., & Feldman, J. (2001). What is a visual object? Evidence from target merging in multiple object tracking. *Cognition*, 80(1/2), 159–177.
- Shapiro, A., Lu, Z. L., Huang, C. B., Knight, E., & Ennis, R. (2010). Transitions between central and peripheral vision create spatial/temporal distortions: A hypothesis concerning the perceived break of the curveball. *PLoS One*, 5(10), e13296.

- Spering, M., & Carrasco, M. (2015). Acting without seeing: eye movements reveal visual processing without awareness. *Trends in neurosciences*, 38(4), 247-258. doi: <https://doi.org/10.1016/j.tins.2015.02.002> [PubMed]
- Spering, M., & Gegenfurtner, K. R. (2008). Contextual effects on motion perception and smooth pursuit eye movements. *Brain research*, 1225, 76-85. doi: <https://doi.org/10.1016/j.brainres.2008.04.061> [PubMed]
- Spering, M., & Montagnini, A. (2011). Do we track what we see? Common versus independent processing for motion perception and smooth pursuit eye movements: A review. *Vision research*, 51(8), 836-852. <https://doi.org/10.1016/j.visres.2010.10.017> [PubMed]
- Steinbach, M. J. (1976). Pursuing the perceptual rather than the retinal stimulus. *Vision research*, 16(12), 1371-1376. doi: [https://doi.org/10.1016/0042-6989\(76\)90154-1](https://doi.org/10.1016/0042-6989(76)90154-1). [PubMed]
- Störmer, V. S., & Alvarez, G. A. (2014). Feature-based attention elicits surround suppression in feature space. *Current Biology*, 24(17), 1985-1988.

Störmer, V. S., Alvarez, G. A., & Cavanagh, P. (2014). Within-hemifield competition in early visual areas limits the ability to track multiple objects with attention. *Journal of Neuroscience*, 34(35), 11526-11533.

Störmer, V. S., Winther, G. N., Li, S. C., & Andersen, S. K. (2013). Sustained multifocal attentional enhancement of stimulus processing in early visual areas predicts tracking performance. *Journal of Neuroscience*, 33(12), 5346-5351.

Suzuki, S., & Cavanagh, P. (1995). Facial organization blocks access to low-level features: An object inferiority effect. *Journal of Experimental Psychology: Human Perception and Performance*, 21(4), 901.

Tse, P. U., & Hsieh, P. J. (2006). The infinite regress illusion reveals faulty integration of local and global motion signals. *Vision research*, 46(22), 3881-3885.

Tullo, D., Faubert, J., & Bertone, A. (2018). The characterization of attention resource capacity and its relationship with fluid reasoning intelligence: A multiple object tracking study. *Intelligence*, 69, 158-168.

Ueda, H., Abekawa, N., & Gomi, H. (2018). The faster you decide, the more accurate localization is possible: Position representation of “curveball illusion” in perception and eye movements. *PloS one*, 13(8), e0201610.

- Wallach, H., Bacon, J., & Schulman, P. (1978). Adaptation in motion perception: Alteration of induced motion. *Perception & Psychophysics*, 24(6), 509-514. doi: <https://doi.org/10.3758/bf03198776> [[PubMed](#)]
- Whitney, D. (2002). The influence of visual motion on perceived position. *Trends in cognitive sciences*, 6(5), 211-216. doi: [https://doi.org/10.1016/s1364-6613\(02\)01887-9](https://doi.org/10.1016/s1364-6613(02)01887-9) [[PubMed](#)]
- Whitney, D., Goltz, H. C., Thomas, C. G., Gati, J. S., Menon, R. S., & Goodale, M. A. (2003). Flexible retinotopy: motion-dependent position coding in the visual cortex. *Science*, 302(5646), 878-881.
- Worsley, K. J. (2001). Statistical analysis of activation images. *Functional MRI: An introduction to methods*, 14(1), 251-270.
- Wyatt, H. J., & Pola, J. (1979). The role of perceived motion in smooth pursuit eye movements. *Vision Research*, 19(6), 613–618. [https://doi.org/10.1016/0042-6989\(79\)90237-2](https://doi.org/10.1016/0042-6989(79)90237-2) [[PubMed](#)]
- Zhang, Y., Brady, M., & Smith, S. (2001). Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE transactions on medical imaging*, 20(1), 45-57.

- Zhang, J., Yeh, S. L., & De Valois, K. K. (1993). Motion contrast and motion integration. *Vision Research*, 33 (18), 2721–2732. doi: [https://doi.org/10.1016/0042-6989\(93\)90231-k](https://doi.org/10.1016/0042-6989(93)90231-k) [PubMed]
- Zivotofsky, A. Z. (2005). A dissociation between perception and action in open-loop smooth-pursuit ocular tracking of the Duncker illusion. *Neuroscience letters*, 376(2), 81-86. doi: <https://doi.org/10.1016/j.neulet.2004.11.031> [PubMed]
- Zivotofsky, A. Z., Averbuch-Heller, L., Thomas, C. W., Das, V. E., Discenna, A. O., & Leigh, R. J. (1995). Tracking of illusory target motion: differences between gaze and head responses. *Vision Research*, 35(21), 3029-3035. doi: [https://doi.org/10.1016/0042-6989\(95\)00067-a](https://doi.org/10.1016/0042-6989(95)00067-a) [PubMed]
- Zivotofsky, A. Z., Goldberg, M. E., & Powell, K. D. (2005). Rhesus monkeys behave as if they perceive the Duncker Illusion. *Journal of cognitive neuroscience*, 17(7), 1011-1017. doi: <https://doi.org/10.1162/0898929054475235> [PubMed]
- Zivotofsky, A. Z. (1996). Rottach KG, Averbuch-Heller L, Kori AA, Thomas CW, Dell'Osso LF, and Leigh RJ. Saccades to remembered targets: the effects of smooth pursuit and illusory stimulus motion. *J Neurophysiol*, 76, 3617-3632. doi: <https://doi.org/10.1152/jn.1996.76.6.3617> [PubMed]