INHIBITION OF RETURN IN 3D SPACE

INHIBITION OF RETURN IS DEPTH-SPECIFIC, OBJECT-BASED, AND RELIES ON A WORLD-CENTERED FRAME OF REFERENCE IN 3D SPACE

BY HANNA HAPONENKO, BHSc

A thesis submitted to the School of Graduate Studies in Partial Fulfillment of the Requirements for the degree of Doctor of Philosophy (Psychology, Neuroscience, and Behaviour)

DOCTOR OF PHILOSOPHY (2022) McMaster University (Psychology)

TITLE: Inhibition of Return is Depth-Specific, Object-Based, and Relies on a World-Centered Frame of Reference in 3D Space

AUTHOR: Hanna Haponenko, BHSc (McMaster University)

SUPERVISOR: Dr. Hong-jin Sun

NUMBER OF PAGES: CLXXVII, 177

Lay abstract

Human attention can be distributed over space and affected by external events. Prior research using 2D environments has shown that some time after the first stimulus (a cue), the reaction time to a subsequent stimulus (a target) appearing in the same location is typically slower compared to when this target appears elsewhere. Thus, attention likely moves away from a previously observed to more novel location of interest. I examined, in a 3D environment, whether this “location” of reduced attention resides in the same 3D location or retinal location as that of the cue. I also assessed the impact on reaction time for when the cue and target belong to the same or different object and when their locations differ in reference to the observer or world environment. My research suggests that humans maintain a higher level of attention for nearer space when the cue previously appears at a farther location.

Abstract

The distribution of human attention in space can be modulated by spatial and temporal factors. This dissertation studied inhibition of return (IOR), a robust behavioural effect obtained through a spatial cueing paradigm where observers exhibit slower detection times to a target appearing over 300 ms after a cue in a previously cued location. Most research has studied the IOR effect in two-dimensional space; thus, it remains unclear whether, in three-dimensional space (3D) space, slower reaction times occur due to a target appearing in the same world location (defined in 3D coordinates) or in the same retinal location as the cue (i.e., anywhere along an observer's line of sight to the cue). My thesis examines IOR in a computer-simulated 3D environment, with the location of the cue and target residing in the same versus different depth/distance position either within the same or different object and either relative to the observer or to the world environment. Following a general literature review (Chapter 1), the first empirical chapter (Chapter 2) demonstrates that IOR is depth-specific when the direction of depth switch between cue to target occurs from far-to-near space, suggesting a behavioural advantage for near space in the human attention system. Chapter 3 shows that this depth-specificity and depth-asymmetry of IOR is maintained only when cues and targets are not part of the same object; object membership can therefore override the depth-specific property of IOR in 3D scenes. Chapter 4 introduces motion of the viewpoint, showing that IOR is depth-specific when the cue and target appear in different depth locations in the world environment even when located at the same relative distance from the observer's viewpoint. Thus, IOR could be the result of an inhibitory tag placed at a location relative to the environment rather than at a location relative to the viewpoint.

Preface

This thesis encompasses five chapters. Chapter 1 details the background literature for how the inhibition of return effect manifests in 3D environments. Additional summary is provided about how positioning cues and targets within the boundary of the same objects affects the spread of IOR when compared to when attention is cued in empty space. Finally, the literature review also provides a background for whether the IOR effect is affected by a viewer-centered or world-centered frame of reference. Chapters 2-4 are empirical chapters. Chapter 2 observes a depth-specific IOR effect in a 3D composed of pictorial depth cues. Chapter 3 suggests that this depth-specificity can only occur when cues and targets are positioned in different objects rather than when positioned within a single object. Chapter 4 investigates how the distances between viewer and cue, viewer and target, and cue and target affect the magnitude of IOR, suggesting that the world-centered reference frame influences IOR. Chapter 5 serves as a general discussion and conclusion chapter, discussing the findings and implications of each empirical chapter.

The research in this thesis was financially supported by the Natural Sciences and Engineering Research Council of Canada Graduate Scholarship awarded to Hanna Haponenko. It was also supported through an annual graduate stipend from 2017 to 2022 by McMaster University.

Declaration of Academic Achievement

The majority of Chapter 2 (apart from Figures 4 and 7, along with corresponding descriptions specific to the figures, as well as Equation 2 in long-form), have been submitted to *Attention, Perception, and Psychophysics* (Manuscript ID: PP-ORIG-22-148). Hanna Haponenko was involved in all aspects of experimental research, including the experimental design, virtual reality programming, data collection and analysis, writing, and editing. Dr. Hong-jin Sun made major contributions to the experimental design and manuscript revision. Dr. Brett Cochrane and Noah Britt made major contributions to manuscript revision.

Chapter 3 is being prepared for submission to *Attention, Perception, and Psychophysics*. Hanna Haponenko was involved in all aspects of experimental research, including the experimental design, virtual reality programming, data collection and analysis, writing, and editing. Dr. Hong-jin Sun made major contributions to the experimental design and manuscript revision.

Chapter 4 serves as part of a manuscript being prepared for submission to *Attention, Perception, and Psychophysics*. Hanna Haponenko was involved in all aspects of experimental research, including the experimental design, virtual reality programming, data collection and analysis, writing, and editing. Noah Britt was involved in data collection and analysis, as well as writing and editing. Dr. Hong-jin Sun made major contributions to the experimental design and manuscript revision.

Acknowledgements

Standing on the shoulders of giants is not only helpful, but downright necessary when completing a Ph.D. But growing and transforming into a more mature, humble, and passionate person was only made possible by the following groups of people.
Hong-jin Sun—For your wisdom, guidance, curiosity, and valuable constructive feedback. You have made me stronger as a researcher, a thinker, and a person. Thank you.

My committee—Pat Bennett, Jim Lyons, and Bruce Milliken—you have given me the unwavering support, structure, and diversity in thought that I always craved and sought throughout my journey. Thank you.

My lab mates and colleagues—Jiali Song, Chao Wang, Nadia Wong, Michael Wan, Laura Jin, Ben Sclodnick, Noah Britt, Shree Venkateshan, Darren Kenney, Ky Gauder, Joseph Voronov, and Andrew King—I'm honoured to have worked with you or otherwise alongside you. Thank you for thinking with me, laughing with me, venting with me, and simply talking with me. I am so utterly excited to see where we will all be in twenty more years. I appreciate you all for providing me with the opportunity to offer advice of my own whenever you asked for it (or didn't and I offered it anyway). Thank you.

To my friends—thank you for believing in me when I may not have believed in myself. Thank you for your patience, kindness, and warmth. Thank you.

My parents, Iryna and Valentin, and my sister, Mariya—I am left speechless when I think of you. My thanks comes from the overwhelming ache of my love for your absolute faith of and patience in me. Я не думаю, що є люди, які розуміють мене краще, ніж ви самі. Я люблю вас з глибини душі. Дякую. Thank you.

Finally, to the person who has taught me that there is more to life than these pages in front of me. To the person who sometimes knows me more than I know myself. To the person that expects nothing but gives everything. I love you, Tay.

Table of Contents

List of Figures

List of Terminologies

IOR	Inhibition of Return
RT	Reaction Time
SE	Standard Error
3D	Three Dimensional
2D	Two Dimensional
VsWs	Viewer-Same World-Same
VdWs	Viewer-Different World-Same
VsWd	Viewer-Same World-Different

Chapter 1: Introduction

1.1 Allocation of Visual Spatial Attention

We are often faced with a highly complex and dynamic environment in which we must shift our attention between different parts of our visual field. Research has explored how visuospatial attention is allocated in two-dimensional (2D) visual spaces (e.g., R. Klein, 1980; Posner, 1980; Posner & Marin, 2016; J. E. T. Taylor et al., 2015; Weaver et al., 1998). The widely accepted gradient metaphor of attention suggests that perceptual processing peaks at the centre of focus and decreases towards the periphery in 2D settings (Bao et al., 2013; Downing & Pinker, 1985; LaBerge & Brown, 1989).

In the natural world, observers attend to locations and objects that have an extent in depth. Depth perception is made possible with a variety of binocular and pictorial depth cues (e.g., stereopsis, linear perspective, relative size, ground intercept information) that aid in the visual segmentation of areas of interest (Atchley & Kramer, 1998; Maringelli et al., 2001; Royden et al., 2016; Warren & Rushton, 2009). To attend to one spatial location or object in a 3D environment, observers might partially or completely inhibit their attention elsewhere to meet the demands of their limited cognitive processing capacity (Knowles, 1963; Redden et al., 2021). However, relatively few behavioural studies have examined how visual spatial attention is distributed and inhibited along the depth axis in 3D space (Bourke et al., 2006; Casagrande et al., 2012; Theeuwes & Pratt, 2003; A. Wang et al., 2015, 2016).

Visual spatial attention is especially important when exploring a complex environment, as observers must constantly attend to locations and objects that vary in depth. A related consideration is that visual attention operates from viewer-centered and world-centered reference frames under different environmental constraints (Klinghammer et al., 2015; Matsumiya & Ando, 2009; Neggers et al., 2005; Parks & Corballis, 2006). When an observer attends to their environment from a viewer-centered frame of reference, they encode spatial coordinates of locations and objects relative to their own viewpoint. A world-centered frame of reference is used when observers encode the spatial coordinates of locations and objects based within a viewpoint–agnostic global environment. Thus, the reference frames employed in one's environment can heavily influence how visuospatial attention is distributed in 3D space—attention to certain areas of interest and even within entire objects can be facilitated or inhibited depending on the changing spatial parameters between these nodes of interest and the observer's viewpoint (Andersen & Kramer, 1993; Lv & Hu, 2020; Reppa et al., 2010).

This dissertation explores the spatial and temporal dynamics of the inhibition of covert attention allocation in 3D space. Specifically, this dissertation aims to provide a new perspective on how the inhibition of covert visuospatial attention acts on spatial locations and objects from the viewer-centered and world-centered reference frames within a virtual 3D setting defined entirely by pictorial depth cues. The overarching goal of the research is to guide and hopefully inspire the reader to conceptualize how this inhibition can be affected without the use of binocular vision.

1.2. Visuospatial Attention in 2D Space

1.2.1. Early Research on Spatial Cueing and Attention Allocation

The speed of covert target detection (e.g., manually detecting a target while eye gaze is maintained at central fixation) is influenced by the relation between where a target appears and where attention has been previously cued and oriented. The time elapsed between cue onset and target onset is also a crucial factor in determining the reaction time to a visual target. For example, reaction times to targets that appear at exogenously cued locations are typically shorter than those for targets that appear at uncued locations when the cue-target onset asynchrony (CTOA) is less than 300 ms. However, reaction times to targets appearing at a cued location are typically greater than those for targets appearing at an uncued location if the CTOA is greater than 300 ms (Maylor & Hockey, 1985; Posner & Cohen, 1984; Z. Wang & Klein, 2010a). This attentional cost, coined as inhibition of return (IOR), was first formally reported by Posner et al. (1984, 1985).

When spatial orienting research was in its infancy, a great deal of emphasis was placed on determining whether an observer's attention to cued or uncued targets was influenced by the movement of an observer's eyes (i.e., whether eye saccades were made). Klein (1980) and Posner (1980) were among the first researchers to study this issue. Within their research, they sought to determine whether saccades influenced attention and whether, inversely, attention influenced saccades. This bidirectional communication, referred to as the oculomotor readiness hypothesis (OMRH), was originally introduced by Wurtz and Mohler (1976) when studying the spatial attention of monkeys. Other researchers at that time suggested that attention could indeed be influenced without saccades (Klein, 1980; Shaw, 1978; Van Voorhis & Hillyard, 1977). However, the structural and functional relationship between saccades and distribution of attention had not yet been thoroughly supported.

The OMRH proposed that covert shifts of attention to a spatial location were made possible only by preparing overt saccades to that specific location. The OMRH also proposed that saccades could only be made with prior orienting of attention. Klein (1980) introduced the notion that attention was able to shift with eyes fixed, and controversially challenged the idea that saccades and covert shifts of attention were produced by one and the same physiological system. In Klein's (1980) experiment, participants were shown three horizontally arranged dots and were told to fixate on the central dot. During manual detection trials, participants produced manual responses after localizing whether the left or right dot brightened without making any saccades. On saccade generation trials, participants moved their eyes to a peripheral asterisk, the location of which was verbally prespecified to appear at the left or right side of fixation. Participants made saccades on most trials and made manual responses only on a minority of trials. Klein reasoned that if subjects were quicker to manually detect a change in luminance at the location to which they prepared to move their eyes, then there would be support for the OMRH. However, because there were no manual detection benefits or costs for targets located at the side to which the average participant's eyes were prepared to move, Klein rejected the OMRH.

In further exploration of the OMRH in the same study, Klein (1980) made manual detection the primary task and saccade generation the secondary task. An informative central cue correctly or incorrectly informed participants of an upcoming luminance change at the left or right side of the display. On most trials, participants were required to manually detect this luminance change without making saccades. On a minority of trials, a saccade was instead required in the attended direction. If the OMRH was true, Klein reasoned that saccades should be faster when directed toward the attended location. Klein's data showed that manual detection times to endogenously cued locations were faster than those at uncued locations. However, saccade reaction times were the same for both cued and uncued locations. Klein concluded that, contrary to the predictions of

OMRH, the preparation of saccades was not necessary for covertly shifting attention and a covert shift of attention was not necessary for the initiation of saccades.

The idea that attention could be directed without saccade generation prevailed (Posner, 1980; Remington, 1978). In Posner's (1980) study, participants whose covert attention and overt saccades were cued by a peripheral box that occurred to the left or right of fixation generated manual detection times that occurred 100 ms earlier than the average saccade generation time of 230 ms (Purves et al., 2001). Clearly, covert attention could be physiologically separated from generated saccades. Nevertheless, it was still unknown whether covert attention could be maintained at one spatial location while generating a peripheral saccade to an opposing location; this would answer the question of whether the two components could be decoupled.

Could covert and overt orienting be decoupled? To answer this question, Posner (1980) reasoned that if a covert shift of attention could directionally oppose an initiated saccade, covert and overt orienting may represent a decoupled system. First, participants were instructed to fixate at the left edge of a display. At time 0 ms, a peripheral cue appeared to the right of fixation. Then, 400 ms later, a second peripheral cue appeared to the right of the first cue. Targets occurred with 80% probability at the initial fixation, and with 20% probability at the first cue location. Participants were instructed to move their eyes from the initial fixation to the first cue and then to the second cue, regardless of target location. If participants were able to covertly return their attention back to fixation upon the presentation of the second cue, the OMRH in a strictly coupled physiological sense would be disproven—planning saccades would show to successfully occur in the opposing direction to a covert shift in attentional trajectory. This is indeed what happened. Manual detection times to the target at fixation were significantly faster than to the first cued location at which the participant was previously fixated but from which participants prepared to move their eyes to the second cue. Researchers now began to look at covert and overt orienting as being functionally, rather than anatomically, coupled. The results of Klein (1980) and Posner (1980) showed that covert orienting could occur efficiently without planned saccades.

Many early experiments combined endogenous and exogenous attentional orienting. Endogenous orienting can be influenced by probabilities of cue or target presentation, or with symbolic cues that indicate to which location a participant should allocate attention. Exogenous orienting is typically accomplished with changes in the luminance, often through abrupt onsets of cues and targets. Exogenous orienting can reflexively capture a person's visuospatial attention even in the presence of previously presented endogenous cues (Müller & Rabbitt, 1989; Santangelo & Spence, 2008). Perhaps overt and covert attention systems would be more closely coupled if both were cued endogenously without distracting exogenously presented targets. Shepherd et al. (1986) attempted to study the relation between overt and covert attention when either system was cued endogenously in the absence of exogenous targets that can automatically capture an observer's covert attention. Participants were shown a central arrow that pointed to the side at which they had to make a manual detection response. For a minority of trials, participants were also required to make overt saccades to the target location. The purpose of this experiment was to investigate whether overt saccades and covert detection were controlled by one system or by separate but interdependent systems when both were endogenously cued. By the conventional OMRH view, if manual detection times were no different from saccade times, then the two systems could be considered separate and likely independent. On the other hand, if manual detection times became faster with compatible saccades, then the two systems could be considered interdependent.

Shepherd et al. (1986) found that the cueing effect, represented by the subtraction of manual detection times of same from oppositely cued target locations, was greater when the endogenous cues were compatible with the target location. In addition, saccades were faster when covert attention was endogenously cued to the same location. Thus, Shephard et al. suggested a reciprocal relation between covert and overt orienting in the presence of endogenous cueing. On trials where participants had to fixate at centre, covert attention was successfully directed without saccades, represented by a positive cueing effect for informative endogenous cues. However, on trials where participants also made saccades to the target location, saccade latencies were reduced when covert attention was allocated to a compatible location. In addition, Shepard et al. discovered an asymmetric relation between covert and overt orienting. When saccades were cued in opposition to the manually detected target location, the magnitude of the covert cueing effect was largest. Shepherd et al. challenged the OMRH by proposing that covert attention operates effectively only during the inhibition of overt saccades to the same location. Supporting evidence by Klein and Pontefract (1994) corroborated Shepherd et al.'s (1986) findings. When participants fixated at centre and manually detected the side at which an endogenously cued target appeared, manual detection reaction times were facilitated. For this facilitated covert orienting to take place, saccade latencies, as part of the secondary task, increased in the covertly attended location (i.e., saccade reaction times were inhibited). This reciprocal and asymmetrical relation between saccades and covert orienting is also made prominent during the inhibition of attention. In Posner's (1985) study, participants were required to make a saccade in any direction they deemed most comfortable or natural after being exogenously cued to a peripheral location. Results revealed that participants were biased in making saccades to the uncued rather than cued side, even though the proportion of manual detection responses were uniformly distributed. Thus, the inhibition of manual detection responses may occur due to an inhibition placed on saccades to a previously cued location.

The covert-overt orienting system has since been deemed a tightly coupled yet disembodied system—covert orienting may exist functionally independently from overt orienting, yet may still depend on the inhibition of eye movements (Klein, 2020). Covert orienting precedes and does not require overt orienting in the same direction, and its operation may depend on inhibiting saccades to the same attended location (Klein, 2020). The rejection of the OMRH, at least in its strict anatomically coupled form, set the stage for studies validly testing covert shifts of spatial attention. One of the predominant themes that emerged as a candidate for studying spatial covert attention in the absence of overt orienting was the inhibition of manual detection responses.

1.2.2. Inhibition of Return: The Paradigm

The IOR paradigm introduced by Posner and Cohen (1984) featured a spatially uninformative peripheral cue appearing at the left or right side of fixation, followed by a peripheral target appearing at either the cued or an uncued location. When the CTOA was greater than 300 ms, significantly slower reaction times were observed for targets that appeared in the cued location than for targets that appeared in the uncued location. The slower response to targets at the cued location was explained to occur due to a process that inhibited overt orienting toward previously attended locations in the visual field (Posner et al., 1985).

Many researchers have used the term IOR to describe any effect in which there is an attentional cost associated with cue-target repetition, even when the study does not follow Posner's IOR method using non-predictive peripheral cues and targets (Dukewich & Klein, 2015). This dissertation only includes experiments that use non-predictive spatial cues, closely mimicking Posner and Cohen's (1984) covert inhibition experiments. Nonetheless, the data from this dissertation will also be compared to those of more recent studies, particularly since the dissertation topic deals with IOR in 3D space. To better understand how IOR works in a covert and exogenous spatial cueing environment in 3D space, the reader should first become familiarized with the effect's spatial distribution along the x and y axes in 2D space.

1.2.3. Spatial Distribution of Inhibition in 2D Space

The seminal attempt that defined the spatial distribution of IOR in 2D space was made by Maylor and Hockey (1985). The researchers applied a covert spatial cueing design to coarsely map the distribution of reaction times to 14 different target locations (i.e., seven locations located to either side of the central fixation point). While fixating, participants viewed a cue that appeared directly to the left or right of the fixation point. Following a CTOA of 700, 900, or 1300 ms, the luminance of one of the 14 targets changed, at which time the participants had to make a simple manual detection response. Reaction times were slowest for targets that appeared in the same location as the cue, with reaction times linearly decreasing for targets more vertically displaced from the cued location. Reaction times exhibited the greatest decrease when targets appeared at a location that was vertically displaced on the side opposite from the cued location. In other words, the inhibitory effect fell as a function of the target's Euclidean distance from the cue.

More recent studies have provided a higher resolution map of the IOR effect (Bennett & Pratt, 2001; J. E. T. Taylor et al., 2015). While fixating on a central point, participants viewed a 21°x21° grid and perceived one of four cues, each located in one of four quadrants of the grid. After a CTOA of 800 ms, a target then appeared in one of 441 locations. Participants were required to detect the target onset. Results were consistent with Maylor and Hockey's (1985) study—reaction times were slowest at locations centered around each cue and were fastest at the edges of the quadrants around the central fixation point. The inhibition gradually dissipated as a function of the target's distance from the cue. Bennett and Pratt (2001) concluded that a model encompassing the sum of independent excitatory and inhibitory effects best fit the data. That is, the inhibitory effect of a cue decreased while the excitatory effect increased as a Gaussian function based on the Euclidean distance between target and cue in the same quadrant.

Bennett and Pratt (2001) explained their data with the attentional momentum (AM) hypothesis. The AM hypothesis is buttressed on the assumption that participants make three movements of attention during a spatial cueing trial: from fixation to cue, from cue back to fixation, then from fixation to the target. If attention moves in a straight line, preserving its momentum, a participant's orienting speed to an uncued target speeds up—inhibition decreases and facilitation increases as attention moves from cue, to fixation, and to the uncued target along one trajectory. But, when attention moves from cue, to fixation, and back to the cued target (i.e., back to the original cue location), then attention loses its momentum. As a result, orienting to the cued target slows down, inhibition peaks, and facilitation declines.

Several studies have used paradigms involving multiple simultaneous cues to investigate the spatial distribution of IOR (Christie et al., 2013; R. M. Klein et al., 2005). Within these experiments, a net vector is calculated from the midpoint of the individual cue locations (i.e., from their centre of gravity). The studies showed that there is no significant difference in detection reaction times to cued and uncued targets when the centre of gravity of multiple simultaneous cues is at fixation (i.e., when the net vector is zero). This phenomenon has been suggested to be made possible via population coding, where the neural action system biases attentional orienting towards a spatial area's centre of gravity (average centre location) (McGowan et al., 1998; Melcher & Kowler, 1999; Tipper et al., 1998). Population coding involves global attentional averaging. In the presence of one cue, the level of inhibition peaks when a target appears at the cue location and declines as a function of cue-to-target distance. In the presence of multiple cues, the level of inhibition peaks at the centre of gravity of the cue cluster and declines with increasing distance between the target and the cluster's centre of gravity. So, if the cue cluster's centre of gravity lies at fixation, and that is where the target appears, there is no substantial IOR produced for cued minus uncued targets, largely because the cue-to-target net vector distance is nullified. Thus, the net magnitude of inhibition likely depends on the distance between the centre of gravity of the cue cluster and the single target, rather than on the distance between a single (and perhaps relatively closest) cue of a larger cluster and a single target. After population coding was conceptualized, research underlying the spatial distribution of IOR in 2D space started to take on a more refined shape.

As part of this refinement in research design, the spatial distribution of IOR in 2D space has also been analyzed across a wider range of CTOAs. In a recent study by Wang et al. (2018), participants responded to targets that appeared on one of two monitors located to the left and right of the fixation point on a central monitor. Cues appeared in any of 21 locations dispersed among two diagonals joined to form an "X" on each side monitor. Participants demonstrated a greater IOR for earlier CTOAs (i.e., 400 ms) compared to later CTOAs (i.e., 1200, 2400, and 3600 ms). Then at about 1200 ms, this inhibitory signal began to dissipate uniformly for all cue-target distances greater than zero until finally subsiding at the originally cue location (Samuel & Kat, 2003; B. Wang et al., 2018).

A possible limitation of B. Wang et al.'s (2018) study is that there were only two target locations, each located on their own separate monitor. Unfortunately, this setup prevented any analysis of IOR for cue-to-target distances residing within a single monitor (i.e., within a single object). IOR for cue-to-target distances greater than zero therefore may have been combined with object effects on IOR, which have been previously reported (Bourke et al., 2006; Jordan & Tipper, 1999; Tipper et al., 1999; Weaver et al., 1998). An earlier study by Samuel and Kat (2003) avoided this potential problem by presenting the cue and target on a single monitor. Even so, the spatiotemporal markers of IOR from their study, as well as results of a meta-analysis of 166 additional studies, were consistent with the results reported by B. Wang et al. (2018). Both Samuel and Kat's (2003) and B. Wang et al.'s (2018) data support the notion that IOR likely peaks at 300-600 ms, and shows a widening spatial spread up to CTOAs of 1200-1600 ms until beginning to retreat to the originally cued location and then dissipating completely at 3 s.

1.2.4. Objects Versus Spatial Locations on IOR in 2D Space

Studies of the spatial distribution of IOR raised a new question. In the real world, people attend to various objects that often occupy a grouped area of Euclidean space. These objects also often move through space. Is the inhibitory tag that produces IOR applied to objects, the space they occupy, or both? Object-based IOR, defined as the inhibitory tagging of cues structurally associated with objects rather than with spatial locations, emerged as the prominent theme of IOR research in the early 1990's.

Tipper et al. (1991, 1994) were the first researchers to explore object-based IOR. They sought to determine whether participants exhibited inhibited detection times to a previously cued moving object. The researchers showed participants two equidistant squares moving in a clockwise pattern around a central box. One of the squares was cued at the beginning of a trial and was later probed once the squares rotated 90° or 180° from the initial starting position. So, if the initially leftward square was cued, and the squares rotated clockwise 90°, the cued square now resided 90° above fixation and the uncued square resided 90° below fixation. If participants were to produce larger reaction times for the cued square, even if both squares were displaced an equal distance from the originally cued spatial location, then an object-based component of inhibition would be implicated. Similarly, if a cued square moved to the 180° location, then an object-based IOR would predict greater reaction times for targets in that cued square but uncued spatial location than for targets appearing in an uncued square but cued spatial location. However, if the IOR effect was based on a fixed location, reaction times would be faster for targets presented farthest away from the originally cued spatial location than for targets presented at or near the originally cued location. Alternatively, an object-based inhibition and a location-based facilitation could theoretically oppose one another, leading to a null effect for the 180° rotation condition.

Results showed that IOR was observed when the cued square appeared at its new 90° spatial location; that is, responses were slower for targets appearing in the cued square (now rotated 90° away from its original spatial location) than for targets appearing in an uncued square directly opposite the cued square, with both squares at an equal distance from the originally cued spatial location. Then, in the 180° condition, when the cued square began to rotate farther away from its original cued location to its spatially uncued 180° position, reaction times decreased compared to the 90° condition. When an uncued box appeared at the previously cued spatial location in the 180° condition, the reaction times for this uncued box decreased less compared to when the uncued box appeared in a previously uncued location in the 90° condition. This effect on reaction time was due to a residual location-based inhibition at that originally cued spatial location. As a result, there was no reaction time difference between cued and uncued square conditions (i.e., IOR = 0) at the maximum spatial cue-target distance of 180°. There likely existed a competition between object-based inhibition and a location-based inhibition gradient.

Furthermore, it was found that the IOR in the control static condition more than doubled the IOR observed in the moving condition. Participants showed a greater inhibitory effect for a target that was within the same object located at the same spatial coordinates as the cue. In the moving display, however, the displacement of the object from the spatial location separated the object based component from the spatial based component, leading to an overall reduced IOR in the moving condition (Tipper et al., 1991, 1994; Weaver et al., 1998). These results support the idea that location-based and object-based influences on attention are two separate components of IOR.

These results prompted further study of the nature of inhibition within a single object: Was inhibition limited to the part of the object that was cued, or did it spread evenly throughout the bounded object? The idea was introduced that perhaps attentional selection was even able to parse entire objects before mechanisms of search and detection acted within the spatial coordinate system of these objects (Egly et al., 1994). A behavioural mechanism of this type would lend support to a more efficient mode of attending to larger groupings of space (object-based), rather than attending to representations of space in a granular fashion (location-based).

It was yet to be determined whether inhibition could spread throughout an object via exogenous cueing. Jordan and Tipper (1999) extended previous research by Egly et al. (1994), who had previously reported results that pointed to an object-based facilitation effect. Jordan and Tipper (1999) followed up on this earlier work by examining whether object-based and location-based IOR could be separately observed in an exogenous, non-predictive cueing context, but within a single object. This study examined whether attention would orient to an entire object if only one part of the object was cued. Each participant was shown a peripheral cue that appeared in one of four ends of two rectangles, after which a target appeared in either end of the two rectangles. Performance differences were assessed by comparing reaction times to targets appearing in the exact same object and spatial location as the cue (location-based + object-based) versus reaction times to targets appearing in the same object but different spatial location as the cue (object-based). The researchers hypothesized that cueing an object in addition to a specific spatial location would increase IOR compared to just cueing an object. If the resulting IOR within an object was greater than zero regardless of spatial location, then inhibition could be said to spread throughout the entire rectangle. As predicted, the IOR for a target appearing within the same spatial location and object as the cue was tripled when compared to the IOR for a target appearing within a different spatial location but same object as the cue. However, the mere existence of an IOR for the object-based condition showed that IOR did indeed spread throughout the object, even though its magnitude was lessened with increased spatial distance between cue and target.

Jordan and Tipper (1999) also suggested that the salience of an object's boundaries influences the level of attention to an object. Their study featured a condition in which the rectangle shapes were made apparent—the objects appeared like rectangles via a Kanizsa illusion but lacked a distinct border. When reaction times to a target cued in the same object but different spatial location (object-based IOR) were compared with reaction times to a target that appeared in a different object and spatial location (uncued condition), there emerged no significant difference (i.e., IOR = 0). Inhibition was likely not able to be contained within the apparent object because of attention freely spreading from one apparent rectangle to another due to a lack of object boundaries. Compared to this apparent object condition, participants showed a greater IOR for rectangles whose boundaries were more salient and well-discriminated. This notion was further supported by a follow-up study that demonstrated the object-based component of IOR is further affected when cues and targets are separated by an additional local boundary within a global object (Leek et al., 2003). Therefore, when designing experiments measuring any form of object-based IOR, it is crucial to design objects that are easily recognizable with obvious physical boundaries and that are recognized as a single object of homogenous luminance, colour, and texture (Watson & Kramer, 1999).

To provide neuropsychological evidence for object-based IOR, Tipper et al., (1997) had split-brain and healthy control participants repeat Tipper et al.'s (1994) experiment. For the split-brain participants, it was predicted that inhibition should stay tagged with the cued object if the object remained in the non-affected visual hemifield. Along the same vein, the predicted effect of this inhibitory tag was that the reaction times would be similar for cued or uncued objects that were probed in the opposite hemifield (i.e., IOR would equal zero). Results showed that object-based IOR was robust for objects probed in the same hemifield in which they were cued. However, when an object moved into the opposite hemifield, split-brain participants showed facilitation for the probed cued object. These results contrasted with findings for the control participants, who produced an object-based IOR effect for both the within and between hemifield conditions. In the split-brain participants, object-based inhibition may have been overshadowed with a location-based facilitation the farther the object moved from the cued spatial location. Object-based IOR is suggested to involve callosal transfer, while the spatial facilitation effect may be caused by a separate, possibly subcortical, region of the brain.

While research analyzing location- and object-based IOR focused initially on 2D contexts, newer technology later made it possible to study both components of IOR in a more realistic 3D context. The patterns of IOR were at last ready to be studied with ecologically valid experiments. The following few chapters introduce the depth cues needed to reliably perceive a 3D scene, before discussing how location- and object-based components of IOR space behave differently in a 3D context.

1.3. Perceiving Depth

The visual system uses several depth cues to segment locations and objects in three-dimensional (3D) space. One particularly strong depth cue is stereopsis, induced by binocular disparity. In humans, one eye receives a slightly shifted image compared to the other eye due to the eyes being separated by an average 6 cm distance. Stereopsis, produced with a reliable disparity signal, has been shown to help observers perform tasks such as illumination discounting in 3D space, depth ordering, and flow parsing (Kitazaki et al., 2008; Warren & Rushton, 2009; Wilcox et al., 2005).

Without disparity signals, however, pictorial depth cues, which only necessitate monocular vision, can be used by the observer to make depth ordering judgements (Wilcox et al., 2005). Stereoscopic displays provide depth information only over short spatial ranges; binocular disparity is most useful for depth discrimination at distances of a few metres from the observer's viewpoint (McCann et al., 2018). Pictorial depth cues become more useful than binocular stereo cues as the spatial range increases past a few metres.

Linear perspective is one pictorial depth cue that allows for the segmentation of near relative to far space. With respect to the central line of sight, in the retina, nearer spatial locations reside horizontally away and vertically lower, while farther spatial locations reside horizontally closer and vertically higher. If two objects are equally sized in world coordinates, the near object will appear larger on the retina with the farther object appearing smaller, a phenomenon known as relative size. Linear perspective and relative size can enhance the impression of depth. These pictorial depth cues allow for an observer to perceive object location and motion in depth more quickly and accurately than in a control condition without perspective or relative size cues (Liao & Johnson, 2016; Royden et al., 2016). Out of these two pictorial depth cues, linear perspective has been shown to have a slightly greater benefit than relative size for parsing out objects from near relative to far space (Warren & Rushton, 2009) and for increasing search efficiency of targets from distractors (Wolfe, 1998).

An especially strong pictorial depth cue used in this dissertation's experiments is optic flow, applied in the context of forward self-motion. The relation between retinal motion, self-motion, and an object's location is given by Equation 1.

$$u = \frac{WX}{Z^2}, \qquad \text{Eq. 1}$$

where *u* is horizontal retinal velocity, *W* is forward or backward self-motion velocity, *X* is the horizontal location of the optic flow vector relative to the observer's viewpoint, and *Z* denotes the depth of the optic flow vector relative to the observer's viewpoint. Thus, the horizontal velocity of an object across the observer's retina is inversely proportional to the square of the distance to the attended object (Peltier et al., 2020). As a result, far objects move more slowly and cover less retinal distance than near objects, assuming both far and near objects move at the same world-relative velocity. This phenomenon is known as optic flow (Koenderink, 1986; Royden & Holloway, 2014; Warren & Rushton, 2009).

Equation 1 indicates that the retinal motion of an object reflects the motion of the object *and* the observer. A proposed mechanism for separating the components of object and self motion is so-called flow-parsing: the optic flow field due to self-motion is globally subtracted from the retinal flow field, which leaves a retinal flow field caused by object motion alone. An observer moving forward produces a radially expanding retinal flow field. Subtracting this radial motion leaves a flow field that is more closely tied to the motion of objects in the visual field, but this procedure works more effectively in the presence of othe rich pictorial depth cues such as texture gradients, linear perspective, and relative size (Royden et al., 2016; Simpson, 1993; Warren & Rushton, 2009). Peltier et al. (2020) studied whether monkeys experiencing simulated forward self-motion exhibited the expected localization bias for farther objects versus nearer objects in the absence of pictorial depth cues (i.e., the retinal size, velocity, and eccentricity were all kept constant). Results showed that when the monkeys moved forward, depth differences between objects, created via binocular disparity, did not differentially affect the perception of a farther or nearer object's optic flow relative to the world. Thus, the flow parsing hypothesis was not supported for forward observer motion when stereopsis was the only available depth cue. In the absence of rich pictorial depth cues, the visual system may approximate flow parsing by only considering the purely 2D optical flow field, thereby completely disregarding changing world coordinates of objects located at different depths relative to the observer. Pictorial depth cues may be essential in segmenting objects in 3D space when an observer is experiencing simulated forward motion.

Certain pictorial depth cues may also best allow an observer experiencing simulated forward motion to parse out a moving target among a set of non-moving distractors located at various simulated depths. In Royden et al.'s (2016) experiment, participants were required to identify the target whose image speed differed from the distractors. In the disconnected condition, the objects were presented at the horizon—this condition only offered relative size as a pictorial depth cue. The connected condition was similar except that a vertical line now connected the objects to the ground plane, which added elements of linear perspective and ground-intercept. The tombstone condition, with the richest pictorial depth cues, consisted of textured rectangular objects, called tombstones, that matched the texture of the ground plane. This condition enabled the motion parallax of objects to be more pronounced due to the added shear between the textured rectangles and ground plane. Participants in the tombstone condition were most accurate at identifying the moving target. Even when modifying the distance between the objects to account for changing retinal size during self-motion, the results indicated that richer pictorial depth cues like linear perspective and motion parallax led to enhanced depth segmentation.

Pictorial depth cues often are crucial in allowing an observer to segment spatial locations and objects at different depths. This line of thinking inspired the current experimental series to study IOR within the context of rich pictorial depth cues.

1.4. Visuospatial Attention in 3D Space

1.4.1. Spatial Distribution of Attention in 3D Space

Attentional orienting has been researched in three-dimensional (3D) settings mainly in the context of attentional facilitation across depths (Andersen, 1990; Andersen & Kramer, 1993; Arnott & Shedden, 2000; de Gonzaga Gawryszewski et al., 1987a; Downing & Pinker, 1985; Han et al., 2005; Miura et al., 2002; Nakayama & Silverman, 1986; Robertson & Kim, 1999; Song et al., 2021). Attentional facilitation has been shown to behave differently in 3D space, with facilitation being larger for attentional trajectories made from far to near space than vice versa.

Attention along the depth axis resembles an asymmetrical gradient. One of the ways in which this gradient can be measured is by comparing reaction times between conditions in which distractors are, by some visual feature, either compatible or incompatible with a central target. This difference score can be referred to as the interference effect: Interference is defined as the difference in reaction times between incompatible and compatible distractor-target trials. Greater interference means that performance suffers more when distractors are visually incompatible with targets (e.g., in shape, size, colour) (Nakayama & Silverman, 1986). When visually incompatible distractors are located past the point of fixation (i.e., at farther, uncrossed disparities), interference is reduced (Andersen & Kramer, 1993; Posner & Marin, 2016). Thus, attention is degraded beyond the point of fixation, the focus of attention, in depth. Consistent with this idea, Andersen and Kramer (1993) found greater interference effects at crossed disparities (i.e., nearer than the point of fixation) and a steeper gradient of attention that decreased drastically at locations beyond the point of fixation at uncrossed disparities.

The idea that the observer's attention is more sensitive to distractors that are closer to the participant than the target could mean that the near depth plane carries more behavioural relevance to an observer's peripersonal region of space. In an attention switching paradigm, observers were required to make same/different judgements about two targets presented at different stereoscopic depth planes (Arnott & Shedden, 2000). Results showed that participants were faster and more accurate in making these judgements when switching attention from far-to-near space compared to near-to-far space. Not only does there seem to exist a degradation in attention beyond the point of fixation, but there also appears to exist a propensity for attention to preferentially attend to an object nearer to an observer's peripersonal space.

Attentional orienting within a 3D setting has been suggested to contain elements that depend on the observer's egocentric position. Miura et al. (2002) found that when an observer moved forward along the depth dimension and responded to targets cued at various depths, there was a significant advantage for targets appearing nearer to the observer. That is, observers were fastest to react to a target that appeared in front of the fixation nearest to them. Targets farther than fixation did not exhibit this advantage even though they were located at the same world distance relative to fixation as nearer targets. When observers were stationary and resided in a single egocentric position, this asymmetry in advantage was not exhibited, a result possibly caused by the absence of rich pictorial depth cues (i.e., optic flow) that promote the segmentation of stimuli across depth planes (Peltier et al., 2020).

The attentional preference or facilitation for locations and objects nearer to the observer has also been shown to exist in a virtual environment employing pictorial depth cues such as linear perspective and optic flow (Song et al., 2021). Participants experiencing visually simulated forward motion were required to localize peripheral targets that appeared at various horizontal eccentricities and depths relative to the observer. Targets appeared on a checkerboard wall on either side of central fixation. Results showed that detection times were faster and more accurate for nearer targets than for farther targets, particularly those located at greater eccentricities. This near advantage, however, was partly attributed to the relative size of checks on the checkerboard walls. The observer's gradient of attention, in a scene with helpful depth information, is likely facilitated when a location or object is nearer to the observer, with inhibition prevailing for areas farther from the observer.

1.4.2. Spatial Distribution of Inhibition in 3D Space

The inhibition of return (IOR) effect has also been observed across depth planes, with researchers most commonly measuring IOR using stereoscopic displays (Bourke et al., 2006; Casagrande et al., 2012; Theeuwes & Pratt, 2003; A. Wang et al., 2016). Research has shown that IOR decreases as a function of the distance between cue and target in two-dimensional (2D) space. If IOR is solely based on retinal or 2D world coordinates, it would mean only the 2D coordinates of cue and target matter, which would render IOR depth-blind in a 3D context. If IOR is based on world coordinates, it would mean that spatial locations can be inhibited along the depth axis as well. In other words, when the cue and target are presented at different depth planes, IOR would be different from when the cue and target are presented at same depth planes.

The first research article measuring IOR in 3D space deemed the effect insensitive to depth differences between cue and target (Theeuwes & Pratt, 2003). The experiment induced stereopsis of rectangles nearer and farther away from the central fixation point. Two rectangles resided on the left side and two on the right side of the central fixation point. During a trial, one of these four rectangles brightened, acting as the cue. One of the four rectangles was replaced by a 2D letter (e.g., an H or an S), acting as the target. The target appeared at the same or different side as the cue, and in the same or different depth plane as the cue. Theeuwes and Pratt (2003) also tested a 2D condition where the rectangles appeared at the same depth plane as the central fixation point (i.e., in the absence of stereopsis). If the pattern of IOR changed depending on the attentional trajectory made in the 3D condition, then IOR was depth-specific. Otherwise, if IOR was the same regardless of whether participants made an depth switch between depths in the 3D condition, then IOR was depth-blind. In addition, if the patterns of IOR were comparable to those seen in the 2D condition, then it could be said that IOR depended on retinal, not world, coordinate differences between cue and target. Results showed that the pattern of IOR did not vary when comparing whether targets were in the same or different depth planes as the cue. The pattern of IOR was similar for 3D and 2D conditions. Therefore, a depth-specific effect on IOR was not found.

Bourke et al. (2006) suggested that Theeuwes and Pratt (2003) may have failed to find a depth effect because the rectangles in their stimulus occluded one another, making it difficult for the observer to perceive world depth differences between cue and target. Bourke et al. therefore used a modified stimulus which eliminated occlusion by slightly shifting the rectangles up and away from one another. It was predicted that IOR would be less for cues and targets appearing between different depth planes than within the same depth plane. This is exactly what they found: IOR nearly doubled when cues and targets appeared within the same depth. This demonstrated that IOR can show some form of depth-specificity in 3D space and was not solely reliant on x-y coordinates in 2D space.

Subsequent studies supported the notion that IOR is sensitive to depth differences within a 3D scene, independent of retinal coordinate differences along the x,y axes. Furthermore, this depth-specificity has also been demonstrated to be asymmetric: performing an depth switch from far to near space results in an IOR magnitude different from that when changing attention from near to far space. Most recently, A. Wang et al. (2016) conducted a study that manipulated depth while controlling for retinal location and size. Participants were shown a stereoscopic display with three rows of three white squares on a black screen. During fixation, attention was oriented to a target that appeared in the same or different depth plane and at the same or different side as the initial cue. A. Wang et al. (2016) found a difference in IOR when comparing targets cued within versus between depths, but only when the target appeared at a near depth relative to the observer. This depth asymmetry has previously been shown for the facilitation of detection responses for far-to-near switches (Andersen & Kramer, 1993; Arnott & Shedden, 2000; Reppa et al., 2010).

Research into how IOR behaves in 3D space is past its infancy, but it remains unclear whether IOR is affected by a more diverse set of pictorial depth cues, or how inhibitory tags operate on objects that span multiple depth planes. The following chapter will review the sparse literature detailing the separation of location- and object-based components of IOR in 3D space.

1.4.3. Objects Versus Spatial Locations on IOR in 3D Space

Early experiments proposed that IOR was tied to spatial locations (Maylor & Hockey, 1985; Posner & Cohen, 1984). None of those experiments, however, controlled for objects that shared the spatial locations of cues and targets. For example, in Maylor and Hockey's (1985) experiments, the cue and target always appeared in the same or different location of an LED bulb, an object. Therefore, IOR could have been influenced by spatial location, the object's structure, or both.

The first experiment that attempted to isolate the effects of spatial coordinates from object coordinates on IOR in 3D space was conducted by Gibson and Egeth (1994). The cues and targets were placed at various locations on a simulated brick whose orientation was transformed in space (e.g., see their Figure 3). The six surfaces of this brick were considered to belong to a single object. Cues and targets, in the form of dots, appeared in either the same spatial location along the x-y axes, on the same object surface rotated in depth, or both. It was hypothesized that the magnitude of IOR could be separately affected by the spatial locations or the object surfaces on which the cue and targets appeared. Results showed that two components were at play: i) object-based IOR, where IOR spread across one of the object's planes even though the cue and target appeared at different spatial locations; and ii) location-based IOR, where IOR appeared when the cue and target location shared the same spatial location but not the same object surface. However, one major issue with Gibson and Egeth's (1994) experimental design was that the surfaces of the object were never separated from the brick's global structure during its rotation, possibly influencing the observer to treat the individual object planes as part of a larger unitary object (i.e., the brick itself), thereby conflating the contributions brought forth by the location and object-based components of attentional orienting.

Bourke et al. (2006) set the groundwork for comparing the contribution of location- and object-based components of IOR in a stereoscopic setting. The researchers analyzed how IOR behaved between and within separate and conjoined objects that extended in depth. In their first experiment, participants were required to detect a target that appeared in the same or different depth planes with the cue and target appearing in separate placeholders (see the top of their Figure 1). IOR was smaller when the cue and target appeared in different depth planes (13 ms) than when they appeared in the same depth plane (23 ms). In their second experiment, when the placeholder was a single object extended across different depth planes (see the bottom of their Figure 1), IOR was not different when the cue and target appeared in the same depth or different depths (25 ms in both conditions). The researchers concluded that only a location-based component of IOR acted for the condition with two separate objects at different depth planes, and that this location-based component of IOR likely operated in a depth-blind fashion tied only to 2D x-y coordinates. The authors suggested that the object-based component of IOR was responsible for IOR's depth-specificity when the discrete object appeared in different depths. However, Bourke et al. (2006) did not measure IOR in an object-free condition where the cues and targets were unattached to objects, and therefore could only infer that only object-based IOR operated in 3D coordinates.

Casagrande et al. (2012) criticized Bourke et al.'s (2006) conclusion by correctly stating that their design lacked an object-free condition that permitted them to state location-based IOR operated in only a 2D depth-blind fashion. Casagrande et al. (2012) modified Bourke et al.'s (2006) experiment by presenting cues and targets in a virtual space devoid of objects (see their Figure 2). Participants fixated on a point that was centered on a gray rectangle representing the back wall. Four patterned parallelograms represented the top, bottom, left, and right-side walls, all slightly slanted and appropriately shaded to render the scene compelling in depth. A trial began with the participants fixating on the central fixation point. The cue (a 3D square frame) appeared before the target (a black sphere), which appeared in one of four peripheral locations. If the target was at a nearer depth plane relative to the observer, then it appeared at a larger size and horizontal eccentricity. If the target was at a depth plane farther from the observer, then it appeared at a smaller size and horizontal eccentricity. These pictorial depth cues were programmed within a stereoscopic setting. The results from this 3D condition were compared to a condition that only used pictorial depth cues free of stereopsis.

Casagrande et al. (2012) found a significantly greater IOR within the same depth plane than between depth planes in the stereoscopic condition. Therefore, the pattern of IOR changed across depth conditions. This depth-specificity was not observed for the non-stereoscopic pictorial depth cue condition. The changes in IOR along the depth axis in the stereoscopic condition appeared to be completely location-based: the targets and cues were not contextually attached to any objects, such as rectangles or other shapes, yet the depth-specificity of IOR prevailed. Pictorial depth cues such as relative eccentricity and size likely interacted with a strong depth context introduced by stereopsis, allowing for differential IOR along the depth axis. Casagrande et al. concluded that the depth-specificity of IOR could be influenced purely by spatial location differences between cue and target, but only in a stereoscopic scene and likely not in a scene with pictorial depth cues. Otherwise, in an environment where a single object extends over multiple depth planes, the object-based component of IOR likely dominates and overrides the need for observers to attend to differences in coarse spatial coordinates within the object.

1.5. Egocentric Versus Allocentric Reference Frames of IOR

Spatial locations can be defined by different spatial frames of reference. When thinking about the manifestation of IOR, one must consider the observer's reference frame in relation to the spatial locations and objects in the larger global environment. Is it only the world-coordinate differences formed between spatial locations and objects that feed into the IOR effect? Or does the viewer's egocentric position contribute in some way to the visual detection of targets in a spatial cueing paradigm?

In most IOR studies, spatial locations can be specified by a viewer-centered frame of reference or a world-centered spatial frame of reference. The two spatial frames of reference tend to specify the same and stable location if both the scene and viewpoint/viewing direction remain static between the presentation of the cue and target. For example, if a stationary participant fixates on a display's central fixation point, both cue and target location can be specified in reference to either the observer's retina or the world (e.g., on the left side of the fixation marker). If, after cue onset there is an eye movement before target onset for a static scene, the relative retinal location between cue and target will be different from that specified by world coordinates.

For attentional orienting to favour novel locations in a dynamic and complex environment, particularly with an observer constantly making saccades or moving their body, IOR could operate from a viewer-centered or world-centered frame of reference. If IOR was solely coded in viewer-centered retinal coordinates, the inhibitory tagging of locations and objects would be greatly affected by saccades being made between cue and target onset. In an experiment that separated the effects of retinal and world coordinates on IOR, participants were required to make a saccade following the offset of a peripheral cue (Maylor & Hockey, 1985). The cue and target either both appeared in the same world coordinates relative to the centre of the display or appeared on the same location on the observer's retina (see their Figure 4). The magnitude of IOR was greater when cue and target shared world coordinates than when cue and target shared retinal coordinates. Similarly, recent studies have found that IOR exists from both a world-centered and viewer-centered perspective, albeit with the increasing distance between cue and target in world coordinates producing a consistently greater IOR effect (M. D. Hilchey et al., 2012; Krüger & Hunt, 2013; Pertzov et al., 2010; Satel et al., 2012). Other spatial cueing experiments, including those measuring IOR in 3D space, have found an asymmetric cueing effect based on an observer orienting attention using a viewer-centered frame of reference (Reppa et al., 2010; A. Wang et al., 2016). Relative to a stationary observer, the detection of a nearer target cued from farther away has been shown to be facilitated even if the world coordinate differences between the two remain identical.

Nonetheless, the viewer-centered and world-centered reference frames have been conflated. In the aforementioned studies, any time a target, for example, appeared in the far depth plane, it also appeared farther from the observer (and vice versa for the nearer depth plane). The analysis of IOR in 3D space has not yet been thoroughly studied in paradigms that separate the effects of viewer-centered and world-centered reference frames on attention distribution. The empirical data in Chapter 4 of this dissertation aims to fill gaps within this domain of research.

1.6. Additional Considerations: Input-Based Versus Output-Based IOR

IOR is accepted as a mechanism that occurs because of attention being inhibited to a previously attended location (Posner et al., 1985). Early research showed that IOR can be produced covertly, with eyes fixed at centre while making manual detection responses, usually in response to peripherally presented targets either exogenously or endogenously cued. IOR can also be produced overtly (i.e., with the requirement of saccadic eye movements towards the target) (Posner et al., 1985). But does the IOR effect occur due to a deficit in perceptual or cognitive analysis at the input-based moment of attentional capture or at the level of a delay in the output-based motor response (Chica et al., 2006; Lupiáñez et al., 2007; T. L. Taylor & Klein, 1998)?

Converging evidence suggests that IOR bares input- and output-based components, with the strength of each component's inhibition depending on whether trajectorial responses such as eye movements are part of the task set (R. Klein, 2000; T. L. Taylor & Klein, 2000). In a systematic investigation of IOR, Taylor and Klein (2000) assessed the relationship between covert and overt attention using central or peripheral signals with manual detection or saccadic reaction time as dependent variables. The idea was to counterbalance the response required for the cue (either ignoring the cue, manually detecting it, or making a saccade in the direction it specified) and the target (manually detecting or making a saccade towards it). The cue and target could either be a central arrow or peripheral stimulus. This central versus peripheral manipulation made it possible to analyze the contribution of perceptual IOR since central and peripheral locations were never superimposed. Therefore, if IOR was, for example, found for an ignore-manual trial where a central arrow endogenously cued a peripheral target, then a perceptual/sensory cause of IOR could be ruled out since the central cue and peripheral target would have never appeared in the same physical location. Results indeed showed that when saccades were inhibited during a trial, IOR was only observed for peripheral targets, even when cued centrally. Since IOR was still generated for peripheral targets in the absence of saccades, a higher-order input mechanism representing spatial location, either cued centrally or peripherally, was a possibility. On the other hand, when saccades were required, IOR manifested for both central and peripheral targets. In this regard, IOR likely functions due to output processes sharing the same trajectory from central fixation to cue and from central fixation to target (Chang & Ro, 2005; M. D. Hilchey et al., 2012; T. L. Taylor & Klein, 2000; Tremblay et al., 2005). Additionally, manual IOR was always less when saccades were made to the cue beforehand compared to when the eyes remained fixated. Thus, the oculomotor component of IOR likely interferes when the more input-based mode of IOR is required during response, a notion corroborated by other studies (M. Hilchey et al., 2014; Satel et al., 2013; T. L. Taylor & Klein, 2000).

However, many IOR experiments, including the ones in this dissertation, cannot resolve input versus output streams of processing and their relative contributions to IOR. The dependent variable in IOR studies always involves a motor output (e.g., saccade, manual detection) to a target location. Thus, the IOR effect can occur because of inhibited visual input of a spatial location just as much as it could occur because of a delayed motor output to that same location. This conflation can be partly resolved by using a non-spatial discrimination task in assessing IOR.

When participants are told to stare at fixation, covertly attend to a peripheral cue onset, and then manually discriminate a non-spatial feature of a peripheral target (e.g., its colour), IOR emerges for the non-spatial target discrimination task without the presence of a higher accuracy trade-off (Chica et al., 2010). But, when saccades are made to the cue and back to fixation before the colour discrimination task, IOR also increases but only amid an accuracy increase (i.e., a speed-accuracy trade-off). The oculomotor component of IOR likely delays the speed or responding without affecting the accumulation of information about the target; importantly, the delayed response is made at a time when there is more information. Thus, the former motoric IOR component likely involves a delayed output-based response whereas the latter input-based IOR involves a failure to quickly and accurately visually process a conceptualized area of cued space.

Neuropsychological evidence suggests that input- and output-based components of IOR exist. For example, saccadic IOR has shown to interact with the fixation being removed during a trial, which is related to the response of the superior colliculus (SC), while manual IOR interacts with perceptual manipulation of target onset (luminance) when saccades are not made (Hunt & Kingstone, 2003). Event-related potential research reports that when IOR is generated by S-cone cues (which are not visible to the SC but are visible cortically), IOR only occurs for manual detection responses, rather than saccades (Sumner et al., 2004). This indicates that when the cortical system is cued, only manual responses are inhibited. Similarly, when eyes are fixated, only the amplitude of early sensory rapid temporal processing components are found, as opposed to changes in response-associated lateralized readiness potentials, which are more closely related to planning of motor processes (Prime & Ward, 2004). The subcortical system is associated with oculomotor activation and therefore the IOR could have been created because of inhibited motor response and inhibited perceptual analysis in conjunction. Note, however, that both cortical (Bourgeois et al., 2012, 2013; Posner et al., 1985) and subcortical (Gabay et al., 2010; Smith et al., 2004) neural regions involved with saccade preparation and output have been shown to influence IOR.

The IOR effect likely occurs due to a motor component that acts in conjunction with the perceptual component of attentional inhibition. This notion is supported by studies that show that an increase in stimulus-response probability increases the magnitude of IOR. For example, one-choice detection key presses lead to a greater IOR than two- or four-choice discrimination responses (Adam et al., 2005). In this dissertation’s studies, IOR was measured via a two-choice localization task via the detection of exogenously presented targets. While this dissertation does not attempt to make a distinction between the perceptual-cognitive (input-based) and motor (output-based) components of IOR, it is important to note that IOR has shown to increase when different effectors are used between the first and second target in a target-target paradigm. This suggests that the motor pairings associated with stimulus presentation may influence the planning or execution of motor movements, which can result in different manifestations of the IOR effect (Howard et al., 1999; Ivanoff & Klein, 2001; Lyons et al., 2006; Tipper et al., 1998; Tremblay et al., 2005). For example, one of the most relevant studies to our experiment required participants to either make a manual pointing response that directly mapped to the location of the target (leftward point to a leftward target) or indirectly mapped to the location of the target (upward point for a leftward target) (Khatoon et al., 2002). Results showed that IOR emerges more quickly (at shorter CTOAs) when there is a direct mapping involved compared to an indirect mapping, in which IOR manifests at later CTOAs (above 900 ms).

IOR can be influenced by the responses that observers pair with perceived stimuli. Spatial attention is rarely a unitary event—actions are frequently paired with spatial locations or objects that are selected for action (Allport, 1987; Tipper et al., 1998). Clearly, input-based IOR is not the only factor at play; there exists a motor bias that should not be disregarded as a potential mechanism in how IOR is generated.

1.7. Overview of Empirical Chapters

In Chapter 2, two peripheral cued target detection experiments were conducted, which used a modified IOR paradigm in a 3D setting. Peripheral cues and targets appeared upon placeholder boxes placed on the surface of a textured ground plane with linear perspective and relative size as the main pictorial depth cues. Both experiments found that in an environment without stereopsis but with rich pictorial depth cues, IOR showed a depth-specific effect. The depth-specific effect found in our 3D condition (factoring both 3D and 2D results) was not entirely explained by 2D parameters. The IOR difference between same versus different depths (for the far cue and near target condition) was much greater than any other conditions. A large source of this difference was because far-to-near depth switch trajectories in the 3D setting resulted in a reduced IOR magnitude than near-to-far depth switches. The near depth plane of space likely has a higher baseline of attentional focus, favouring facilitation rather than inhibition.

Chapter 3 clarified whether the depth-specific and depth-asymmetric IOR effect observed in Chapter 2 was mostly due to location- or object-based components. A cued target detection experiment was conducted, which again employed a modified IOR paradigm in a 3D scene. Cues and targets appeared within placeholder boxes situated on the surface of a textured ground plane presenting linear perspective, texture gradient, and relative size as the main pictorial depth cues. When the cue appeared in a farther location than that of the subsequent target, the IOR effect was depth-specific. Thus, IOR was smaller when the cue and target appeared in different depth planes than when cued and target appeared in the same depth plane (when compared with the 2D control). This depth effect was only present when the cue and target appeared in two placeholder boxes separated in depth, but not when cue and target appeared in one placeholder box while all other spatial properties remained the same. Our results suggest that IOR can be depth-specific, but the depth effect is limited to the specific direction of the depth switch across depth and object-based inhibition can override this depth-specific effect. The main issue with the experiments conducted in Chapters 2 and 3 was that when the target was farther than the cue, it was also farther from the observer (and vice versa when the target was nearer). Therefore, world-centered IOR effects were never isolated from viewer-centered IOR effects. This contrast was investigated in Chapter 4.

A viewer-centered frame of reference typically goes hand in hand with variance in the structure of a scene—as an observer experiences a change in their environment, either through simulated locomotion or a spatial change in luminance, the input of a scene changes. However, it is when the viewer experiences change and the structure of a scene remains the same, that invariant features of a scene may also arise. Thus, with Chapter 4, the variant viewer-centered frame of reference was controlled to make it possible to study the invariant world-centered frame of reference. Chapter 4 isolated the relative contributions of viewer-centered IOR and world-centered IOR by controlling for the viewer-to-cue and viewer-to-target distance versus the cue-to-target distance along the depth axis. Previous studies in Chapters 2 and 3 found evidence for a depth-specific IOR—IOR was reduced most drastically for far-to-near depth switches. However, it is not clear what frame of reference is involved in this depth-specific IOR since viewer-centered and world-centered frame of references commonly specify the same depth relation.

In Chapter 4, the spatial frame of reference for the depth-specific effect was investigated by creating a discrepancy between depth information specified by viewer-centered and world-centered coordinate systems. Target detection was tested in a spatial cueing paradigm in a simulated 3D space illustrated through pictorial depth cues (such as linear perspective and optic flow). In the control condition, the cue and target were always presented in the same depth plane relative to a stationary viewpoint. In the experimental conditions, the cue and target appeared at different viewer-centered distances (Experiment 1 and 2) or different world-centered depth planes (Experiment 2), with the competing frame of reference controlled between cue and target onset. Results showed that in Experiment 1, the IOR only changed when targets located very near to the viewer's apparent peripersonal space were cued from farther away relative to the viewer. The near depth plane of space likely has a higher baseline of attentional focus, promoting a sense of behavioural urgency when the viewer approaches the task-relevant placeholder wall. In Experiment 2, IOR was drastically reduced only when cues and targets were located on placeholder walls separated in depth, indicating that the world-centered reference frame strongly contributes to how IOR is generated.

1.8. Summary

The study of inhibition of return in manual detection responses originated in studies of covert shifts of spatial attention. When a spatial cueing task strictly requires covert orienting, reaction times for targets occurring in the same spatial location as the cue are oftentimes significantly slower than those for targets occurring in a different spatial location as the cue (Posner et al., 1985; Posner & Cohen, 1984). This inhibitory cueing effect for peripheral targets typically occurs at CTOAs greater than 300 ms. Some researchers posit that the IOR effect is related to endogenous suppression of the oculomotor system, a notion that fits conceptually with covert attention having been shown to act independently of overt saccades (M. Hilchey et al., 2014; Satel et al., 2013; T. L. Taylor & Klein, 2000).

As research on IOR evolved, researchers began to study the spatial distribution of this effect in the 2D x-y coordinates. It was found that the magnitude of IOR decreased monotonically the farther a target was from the originally cued spatial location. Around the same time, it was shown that IOR also depended on whether the cue and target were associated with the same object. That is, the IOR effect was significantly larger when cue and target not only appeared at the same spatial location, but also both appeared within the same object (Bennett & Pratt, 2001; J. E. T. Taylor et al., 2015; Tipper et al., 1991, 1994).

With the study of more technologically and ecologically valid settings, recent research has begun exploring the spatial distribution of IOR in 3D scenes. These 3D scenes, mostly composed of stereopsis depth cues, have been leveraged to showcase that IOR unveils a depth specificity in depth space—observers exhibit a significantly larger IOR for targets cued within the same depth. This depth-specificity is asymmetrical depending on whether targets appear farther or nearer than the cue. In addition, this depth-specificity has been shown to be manifested by spatial locations as well as object associations. The observer's egocentric or the world-coordinate differences between cue and target may drive the depth asymmetry of attention in 3D space, but no directly relevant research to date has investigated this issue.

This dissertation analyzes whether IOR exhibits depth-specificity in a 3D scene composed only of pictorial depth cues and whether this depth-specificity depends on objects demarcating depth planes at which cues and targets appear. Furthermore, the viewer-centered and world-centered reference frames are each isolated to study their respective contributions to the generation of the IOR effect.

1.9. References

Adam, J., O'Donnell, C., & Pratt, J. (2005). Response selection influences inhibition of return. *European Journal of Cognitive Psychology*, *17*(3), 319–328. https://doi.org/10.1080/09541440440000069

Allport, D. A. (1987). Selection for action: Some behavioral and neurophysiological nonsiderations of attention and action. In H. Heuer & H. F. Sanders (Eds.), *Perspectives on Perception and Action* (pp. 395–419). Lawerence Erlbaum.

Andersen, G. J. (1990). Focused attention in three-dimensional space. *Perception & Psychophysics, 47*(2), 112–120. https://doi.org/10.3758/BF03205975

Andersen, G. J., & Kramer, A. F. (1993). Limits of focused attention in three-dimensional space. *Perception & Psychophysics*, *53*(6), 658–667. https://doi.org/10.3758/BF03211742

Arnott, S. R., & Shedden, J. M. (2000). Attention switching in depth using random-dot autostereograms: Attention gradient asymmetries. *Perception & Psychophysics, 62*(7), 1459–1473. https://doi.org/10.3758/BF03212146

Atchley, P., & Kramer, A. F. (1998). Spatial cuing in a stereoscopic display: Attention remains "depth-aware" with age. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *53B*(5), 318–323. https://doi.org/10.1093/geronb/53B.5.P318

Bao, Y., Lei, Q., Fang, Y., Tong, Y., Schill, K., Pöppel, E., & Strasburger, H. (2013). Inhibition of return in the visual field. *Experimental Psychology*, *60*(6), 425–431. https://doi.org/10.1027/1618-3169/a000215

Bennett, P. J., & Pratt, J. (2001). The spatial distribution of inhibition of return. *Psychological Science*, *12*(1), 76–80. https://doi.org/10.1111/1467-9280.00313

Bourgeois, A., Chica, A. B., Migliaccio, R., de Schotten, M. T., & Bartolomeo, P. (2012). Cortical control of inhibition of return: Evidence from patients with inferior parietal damage and visual neglect. *Neuropsychologia*, *50*(5), 800–809. https://doi.org/10.1016/j.neuropsychologia.2012.01.014

Bourgeois, A., Chica, A. B., Valero-Cabré, A., & Bartolomeo, P. (2013). Cortical control of inhibition of return: Causal evidence for task-dependent modulations by dorsal and ventral parietal regions. *Cortex*, *49*(8), 2229–2238. https://doi.org/10.1016/j.cortex.2012.10.017

Bourke, P. A., Partridge, H., & Pollux, P. M. J. (2006). Additive effects of inhibiting attention to objects and locations in three-dimensional displays. *Visual Cognition*, *13*(5), 643–654. https://doi.org/10.1080/13506280544000309

Casagrande, M., Barbato, M., Mereu, S., Martella, D., Marotta, A., Theeuwes, J., & Collinson, S. (2012). Inhibition of return: A “depth-blind” mechanism? *Acta Psychologica*, *140*, 75–80. https://doi.org/10.1016/j.actpsy.2012.02.011

Chang, E., & Ro, T. (2005). Inhibition of return in perception and action. *Visual Cognition*, *12*, 443–472. https://doi.org/10.1080/13506280444000391

Chica, A. B., Lupiáñez, J., & Bartolomeo, P. (2006). Dissociating inhibition of return from endogenous orienting of spatial attention: Evidence from detection and discrimination tasks. *Cognitive Neuropsychology*, *23*(7), 1015–1034. https://doi.org/10.1080/02643290600588277

Chica, A. B., Taylor, T. L., Lupiáñez, J., & Klein, R. M. (2010). Two mechanisms underlying inhibition of return. *Experimental Brain Research*, *201*(1). https://doi.org/10.1007/s00221-009-2004-1

Christie, J., Hilchey, M. D., & Klein, R. M. (2013). Inhibition of return is at the midpoint of simultaneous cues. *Attention, Perception, & Psychophysics*, *75*(8), 1610–1618. https://doi.org/10.3758/s13414-013-0510-5

de Gonzaga Gawryszewski, L., Riggio, L., Rizzolatti, G., & Umiltá, C. (1987). Movements of attention in the three spatial dimensions and the meaning of “neutral” cues. *Neuropsychologia*, *25*(1A), 19–29.

Downing, C., & Pinker, S. (1985). The spatial structure of visual attention. In *Attention and performance: Vol. XI* (pp. 171–187). Erlbaum.

Dukewich, K., & Klein, R. (2015). Inhibition of return: A phenomenon in search of a definition and a theoretical framework. *Attention, Perception & Psychophysics*, *77*, 1647–1658. https://doi.org/10.3758/s13414-015-0835-3

Egly, R., Driver, J., & Rafal, R. D. (1994). Shifting visual attention between objects and locations: Evidence from normal and parietal lesion subjects. *Journal of*

Experimental Psychology. General, *123*(2), 161–177. https://doi.org/10.1037//0096-3445.123.2.161

Gabay, S., Henik, A., & Gradstein, L. (2010). Ocular motor ability and covert attention in patients with Duane Retraction Syndrome. *Neuropsychologia*, *48*(10), 3102–3109. https://doi.org/10.1016/j.neuropsychologia.2010.06.022

Gibson, B. S., & Egeth, H. (1994). Inhibition of return to object-based and environment-based locations. *Perception & Psychophysics*, *55*(3), 323–339. https://doi.org/10.3758/bf03207603

Han, S., Wan, X., & Humphreys, G. W. (2005). Shifts of spatial attention in perceived 3-D space. *The Quarterly Journal of Experimental Psychology Section A*, *58*(4), 753–764. https://doi.org/10.1080/02724980443000548

Hilchey, M. D., Klein, R. M., Satel, J., & Wang, Z. (2012). Oculomotor inhibition of return: How soon is it “recoded” into spatiotopic coordinates? *Attention, Perception, & Psychophysics*, *74*(6), 1145–1153. https://doi.org/10.3758/s13414-012-0312-1

Hilchey, M., Hashish, M., MacLean, G., Satel, J., Ivanoff, J., & Klein, R. (2014). On the role of eye movement monitoring and discouragement on inhibition of return in a go/no-go task. *Vision Research*, *96*, 133–139. https://doi.org/10.1016/j.visres.2013.11.008

Howard, L. A., Lupiáñez, J., & Tipper, S. P. (1999). Inhibition of return in a selective reaching task: An investigation of reference frames. *The Journal of General Psychology*, *126*(4), 421–442. https://doi.org/10.1080/00221309909595374

Hunt, A. R., & Kingstone, A. (2003). Inhibition of return: Dissociating attentional and oculomotor components. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(5), 1068–1074. https://doi.org/10.1037/0096-1523.29.5.1068

Ivanoff, J., & Klein, R. M. (2001). The presence of a nonresponding effector increases inhibition of return. *Psychonomic Bulletin & Review*, *8*(2), 307–314. https://doi.org/10.3758/bf03196166

Jordan, H., & Tipper, S. P. (1999). Spread of inhibition across an object's surface. *British Journal of Psychology*, *90*(4), 495–507. https://doi.org/10.1348/000712699161576

Khatoon, S., Briand, K. A., & Sereno, A. B. (2002). The role of response in spatial attention: Direct versus indirect stimulus–response mappings. *Vision Research*, *42*(24), 2693–2708. https://doi.org/10.1016/S0042-6989(02)00327-9

Kitazaki, M., Kobiki, H., & Maloney, L. T. (2008). Effect of pictorial depth cues, binocular disparity cues and motion parallax depth cues on lightness perception in three-dimensional virtual scenes. *PLOS ONE*, *3*(9), 1–9. https://doi.org/10.1371/journal.pone.0003177

Klein, R. (1980). Does oculomotor readiness mediate cognitive control of visual attention? In R. *Attention and Performance VIII*, *8*. https://doi.org/10.1037/e665402011-424

Klein, R. (2000). Inhibition of return. *Trends in Cognitive Sciences*, *4*(4), 138–147.

Klein, R. M. (2020). Is covert spatial orienting embodied or disembodied cognition? A historical review. *Quarterly Journal of Experimental Psychology*, *73*(1), 20–28. https://doi.org/10.1177/1747021819889497

Klein, R. M., Christie, J., & Morris, E. P. (2005). Vector averaging of inhibition of return. *Psychonomic Bulletin & Review*, *12*(2), 295–300. https://doi.org/10.3758/BF03196375

Klein, R. M., & Pontefract, A. (1994). Does oculomotor readiness mediate cognitive control of visual attention? Revisited! In *Attention and performance 15: Conscious and nonconscious information processing* (pp. 333–350). The MIT Press.

Klinghammer, M., Blohm, G., & Fiehler, K. (2015). Contextual factors determine the use of allocentric information for reaching in a naturalistic scene. *Journal of Vision, 15*(13), 24. https://doi.org/10.1167/15.13.24

Knowles, W. B. (1963). Operator loading tasks. *Human Factors*, *5*, 155–161. https://doi.org/10.1177/001872086300500206

Koenderink, J. J. (1986). Optic flow. *Vision Research*, *26*(1), 161–179. https://doi.org/10.1016/0042-6989(86)90078-7

Krüger, H. M., & Hunt, A. R. (2013). Inhibition of return across eye and object movements: The role of prediction. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(3), 735–744. https://doi.org/10.1037/a0030092

LaBerge, D., & Brown, V. (1989). Theory of attentional operations in shape identification. *Psychological Review*, *96*(1), 101–124. https://doi.org/10.1037/0033-295X.96.1.101

Leek, E. C., Reppa, L., & Tipper, S. P. (2003). Inhibition of return for objects and locations in static displays. *Perception & Psychophysics*, *65*(3), 388–395. https://doi.org/10.3758/bf03194570

Liao, M.-J., & Johnson, W. (2016). The Effects of Monocular Cues on a 3-D Target Acquisition Task: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. https://doi.org/10.1177/154193120204601729

Lupiáñez, J., Ruz, M., Funes, M. J., & Milliken, B. (2007). The manifestation of attentional capture: Facilitation or IOR depending on task demands. *Psychological Research*, *71*(1), 77–91. https://doi.org/10.1007/s00426-005-0037-z

Lv, M., & Hu, S. (2020). Asymmetrical switch costs in spatial reference frames switching. *Perception*, *49*(3), 268–280. https://doi.org/10.1177/0301006620906087

Lyons, J., Glazebrook, C. M., Keetch, K. M., Dhillon, V. P., & Elliott, D. (2006). Influence of endogenous and exogenous orientations of attention on inhibition of return in a cross-modal target-target aiming task. *Journal of Motor Behavior*, *38*(3), 219–228. https://doi.org/10.3200/JMBR.38.3.219-228

Maringelli, F., McCarthy, J., Steed, A., Slater, M., & Umiltà, C. (2001). Shifting visuo-spatial attention in a virtual three-dimensional space. *Cognitive Brain Research*, *10*(3), 317–322. https://doi.org/10.1016/S0926-6410(00)00039-2

Matsumiya, K., & Ando, H. (2009). World-centered perception of 3D object motion during visually guided self-motion. *Journal of Vision*, *9*(1), 15–15. https://doi.org/10.1167/9.1.15

Maylor, E. A., & Hockey, R. (1985). Inhibitory component of externally controlled covert orienting in visual space. *Journal of Experimental Psychology: Human Perception and Performance*, *11*(6), 777–787. https://doi.org/10.1037/0096-1523.11.6.777

McCann, B. C., Hayhoe, M. M., & Geisler, W. S. (2018). Contributions of monocular and binocular cues to distance discrimination in natural scenes. *Journal of Vision*, *18*(4). https://doi.org/10.1167/18.4.12

McGowan, J. W., Kowler, E., Sharma, A., & Chubb, C. (1998). Saccadic localization of random dot targets. *Vision Research*, *38*(6), 895–909. https://doi.org/10.1016/s0042-6989(97)00232-0

Melcher, D., & Kowler, E. (1999). Shapes, surfaces and saccades. *Vision Research*, *39*(17), 2929–2946. https://doi.org/10.1016/S0042-6989(99)00029-2

Miura, T., Shinohara, K., & Kanda, K. (2002). Shift of attention in depth in a semi-realistic setting. *Japanese Psychological Research*, *44*(3), 124–133. https://doi.org/10.1111/1468-5884.00015

Müller, H. J., & Rabbitt, P. M. (1989). Reflexive and voluntary orienting of visual attention: Time course of activation and resistance to interruption. *Journal of Experimental Psychology. Human Perception and Performance*, *15*(2), 315–330. https://doi.org/10.1037//0096-1523.15.2.315

Nakayama, K., & Silverman, G. H. (1986). Serial and parallel processing of visual feature conjunctions. *Nature*, *320*(6059), 264–265. https://doi.org/10.1038/320264a0

Neggers, S. F. W., Schölvinck, M. L., van der Lubbe, R. H. J., & Postma, A. (2005). Quantifying the interactions between allo- and egocentric representations of space. *Acta Psychologica*, *118*(1–2), 25–45. https://doi.org/10.1016/j.actpsy.2004.10.002

Parks, N. A., & Corballis, P. M. (2006). Attending to depth: Electrophysiological evidence for a viewer-centered asymmetry. *NeuroReport*, *17*(6), 643.

Peltier, N. E., Angelaki, D. E., & DeAngelis, G. C. (2020). Optic flow parsing in the macaque monkey. *Journal of Vision*, *20*(10), 8. https://doi.org/10.1167/jov.20.10.8

Pertzov, Y., Zohary, E., & Avidan, G. (2010). Rapid formation of spatiotopic representations as revealed by inhibition of return. *Journal of Neuroscience*, *30*(26), 8882–8887. https://doi.org/10.1523/JNEUROSCI.3986-09.2010

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*(1), 3–25. https://doi.org/10.1080/00335558008248231

Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. *Attention and Performance X*, *32*, 531–556.

Posner, M. I., & Marin, O. S. M. (2016). *Attention and Performance XI*. Routledge.

Posner, M. I., Rafal, R. D., Choate, L. S., & Vaughan, J. (1985). Inhibition of return: Neural basis and function. *Cognitive Neuropsychology*, *2*(3), 211–228. https://doi.org/10.1080/02643298508252866

Prime, D. J., & Ward, L. M. (2004). Inhibition of return from stimulus to response. *Psychological Science*, *15*(4), 272–276. https://doi.org/10.1111/j.0956-7976.2004.00665.x

Purves, D., Augustine, G. J., Fitzpatrick, D., Katz, L. C., LaMantia, A.-S., McNamara, J. O., & Williams, S. M. (2001). Types of eye movements and their functions. *Neuroscience. 2nd Edition*. https://www.ncbi.nlm.nih.gov/books/NBK10991/

Redden, R. S., MacInnes, W. J., & Klein, R. M. (2021). Inhibition of return: An information processing theory of its natures and significance. *Cortex*, *135*, 30–48. https://doi.org/10.1016/j.cortex.2020.11.009

Remington, R. W. (1978). *Visual attention, detection and the control of saccadic eye movements*. https://www.proquest.com/openview/4b8cfcc862f44c17985d174f686645f3/1.pdf?pq-origsite=gscholar&cbl=18750&diss=y

Reppa, I., Fougnie, D., & Schmidt, W. C. (2010). How does attention spread across objects oriented in depth? *Attention, Perception, & Psychophysics*, *72*(4), 912–925. https://doi.org/10.3758/APP.72.4.912

Robertson, L. C., & Kim, M.-S. (1999). Effects of perceived space on spatial attention. *Psychological Science*, *10*(1), 76–79. https://doi.org/10.1111/1467-9280.00110

Royden, C. S., & Holloway, M. A. (2014). Detecting moving objects in an optic flow field using direction- and speed-tuned operators. *Vision Research*, *98*, 14–25. https://doi.org/10.1016/j.visres.2014.02.009

Royden, C. S., Parsons, D., & Travatello, J. (2016). The effect of monocular depth cues on the detection of moving objects by moving observers. *Vision Research*, *124*, 7–14. https://doi.org/10.1016/j.visres.2016.05.002

Samuel, A. G., & Kat, D. (2003). Inhibition of return: A graphical meta-analysis of its time course and an empirical test of its temporal and spatial properties. *Psychonomic Bulletin & Review*, *10*(4), 897–906. https://doi.org/10.3758/BF03196550

Santangelo, V., & Spence, C. (2008). Is exogenous orienting of spatial attention truly automatic? Evidence from unimodal and multisensory studies. *Consciousness and Cognition*, *17*, 989–1015. https://doi.org/10.1016/j.concog.2008.02.006

Satel, J., Hilchey, M. D., Wang, Z., Story, R., & Klein, R. M. (2013). The effects of ignored versus foveated cues upon inhibition of return: An event-related potential study. *Attention, Perception, & Psychophysics*, *75*(1), 29–40. https://doi.org/10.3758/s13414-012-0381-1

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

Shaw, M. L. (1978). A capacity allocation model for reaction time. *Journal of Experimental Psychology: Human Perception and Performance*, *4*(4), 586–598. https://doi.org/10.1037/0096-1523.4.4.586

Shepherd, M., Findlay, J. M., & Hockey, R. J. (1986). The relationship between eye movements and spatial attention. *The Quarterly Journal of Experimental Psychology Section A*, *38*(3), 475–491. https://doi.org/10.1080/14640748608401609

Simpson, W. A. (1993). Optic flow and depth perception. *Spatial Vision*, *7*(1), 35–75. https://doi.org/10.1163/156856893X00036

Smith, D. T., Rorden, C., & Jackson, S. R. (2004). Exogenous orienting of attention depends upon the ability to execute eye movements. *Current Biology*, *14*(9), 792–795. https://doi.org/10.1016/j.cub.2004.04.035

Song, J., Bennett, P. J., Sekuler, A. B., & Sun, H.-J. (2021). The effect of apparent distance on peripheral target detection. *Journal of Vision*, *21*(10), 1–8. https://doi.org/10.1167/jov.21.10.8

Sumner, P., Nachev, P., Vora, N., Husain, M., & Kennard, C. (2004). Distinct cortical and collicular mechanisms of inhibition of return revealed with S cone stimuli. *Current Biology*, *14*(24), 2259–2263. https://doi.org/10.1016/j.cub.2004.12.021

Taylor, J. E. T., Chan, D., Bennett, P. J., & Pratt, J. (2015). Attentional cartography: Mapping the distribution of attention across time and space. *Attention, Perception, & Psychophysics*, *77*(7), 2240–2246. https://doi.org/10.3758/s13414-015-0943-0

Taylor, T. L., & Klein, R. M. (1998). On the causes and effects of inhibition of return. *Psychonomic Bulletin & Review*, *5*(4), 625–643. https://doi.org/10.3758/BF03208839

Taylor, T. L., & Klein, R. M. (2000). Visual and motor effects in inhibition of return. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(5), 1639–1656. https://doi.org/10.1037/0096-1523.26.5.1639

Theeuwes, J., & Pratt, J. (2003). Inhibition of return spreads across 3-D space. *Psychonomic Bulletin & Review*, *10*(3), 616–620. https://doi.org/10.3758/BF03196523

Tipper, S. P., Driver, J., & Weaver, B. (1991). Object-centred inhibition of return of visual attention. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, *43*(2), 289–298. https://doi.org/10.1080/14640749108400971

Tipper, S. P., Howard, L. A., & Houghton, G. (1998). Action-based mechanisms of attention. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *353*(1373), 1385–1393.

Tipper, S. P., Jordan, H., & Weaver, B. (1999). Scene-based and object-centered inhibition of return: Evidence for dual orienting mechanisms. *Perception & Psychophysics*, *61*(1), 50–60. https://doi.org/10.3758/BF03211948

Tipper, S. P., Rafal, R., Reuter-Lorenz, P. A., Starrveldt, Y., Ro, T., Egly, R., Danzinger, S., & Weaver, B. (1997). Object-based facilitation and inhibition from visual orienting in the human split-brain. *Journal of Experimental Psychology. Human Perception and Performance*, *23*(5), 1522–1532. https://doi.org/10.1037//0096-1523.23.5.1522

Tipper, S. P., Weaver, B., Jerreat, L. M., & Burak, A. L. (1994). Object-based and environment-based inhibition of return of visual attention. *Journal of Experimental Psychology. Human Perception and Performance*, *20*(3), 478–499. https://doi.org/10.1037/0096-1523.20.3.478

Tremblay, L., Welsh, T. N., & Elliott, D. (2005). Between-trial inhibition and facilitation in goal-directed aiming: Manual and spatial asymmetries. *Experimental Brain Research*, *160*(1), 79–88. https://doi.org/10.1007/s00221-004-1987-x

Van Voorhis, S., & Hillyard, S. A. (1977). Visual evoked potentials and selective attention to points in space. *Perception & Psychophysics*, *22*(1), 54–62. https://doi.org/10.3758/BF03206080

Wang, A., Liu, X., Chen, Q., & Zhang, M. (2016). Effect of different directions of attentional shift on inhibition of return in three-dimensional space. *Attention, Perception, & Psychophysics*, *78*(3), 838–847. https://doi.org/10.3758/s13414-015-1055-6

Wang, A., Yue, Z., Zhang, M., & Chen, Q. (2015). Interaction between spatial inhibition of return (IOR) and executive control in three-dimensional space. *Experimental Brain Research*, *233*(11), 3059–3071. https://doi.org/10.1007/s00221-015-4374-x

Wang, B., Yan, C., Klein, R. M., & Wang, Z. (2018). Inhibition of return revisited: Localized inhibition on top of a pervasive bias. *Psychonomic Bulletin & Review*, *25*(5), 1861–1867. https://doi.org/10.3758/s13423-017-1410-9

Wang, Z., & Klein, R. M. (2010). Searching for inhibition of return in visual search: A review. *Vision Research*, *50*(2), 220–228. https://doi.org/10.1016/j.visres.2009.11.013

Warren, P. A., & Rushton, S. K. (2009). Perception of scene-relative object movement: Optic flow parsing and the contribution of monocular depth cues. *Vision Research*, *49*(11), 1406–1419. https://doi.org/10.1016/j.visres.2009.01.016

Watson, S. E., & Kramer, A. F. (1999). Object-based visual selective attention and perceptual organization. *Perception & Psychophysics*, *61*(1), 31–49. https://doi.org/10.3758/bf03211947

Weaver, B., Lupiáñez, J., & Watson, F. L. (1998). The effects of practice on object-based, location-based, and static-display inhibition of return. *Perception & Psychophysics*, *60*(6), 993–1003. https://doi.org/10.3758/BF03211934

Wilcox, L. M., Wildes, R. P., Lakra, D., & Spengler, R. A. (2005). The contribution of binocular and monocular texture elements to depth ordering. *Journal of Vision*, *5*(8), 772–772. https://doi.org/10.1167/5.8.772

Wolfe, J. M. (1998). Visual search. In *Attention*. University College London Press.

Wurtz, R. H., & Mohler, C. W. (1976). Organization of monkey superior colliculus: Enhanced visual response of superficial layer cells. *Journal of Neurophysiology*, *39*(4), 745–765. https://doi.org/10.1152/jn.1976.39.4.745

Chapter 2: Inhibition of Return in a 3D Scene Depends on the Direction of Depth Switch

Abstract

Inhibition of Return (IOR) is a phenomenon that reflects slower target detection when the target appears at a previously cued rather than uncued location. In the present study, we investigated the extent at which IOR occurs in 3D environments with pictorial depth cues. Peripheral cues and targets appeared on top of 3D rectangular boxes placed on the surface of a textured ground plane. It was revealed that when the target appeared at a farther location than the cue, the magnitude of the lateral IOR effect remained similar regardless of whether the cue and target appeared at different depths (i.e., IOR was depth-blind). When the target appeared at a nearer location than the cue, the magnitude of the lateral IOR effect was significantly attenuated (i.e., IOR was depth-specific). The present findings address inconsistencies in the literature on the effect of depth on IOR and support the notion that visuospatial attention exhibits a near space advantage.

Introduction

The speed of target detection can be affected by spatial cues appearing prior to target onset. Reaction times to targets at exogenously cued locations are typically shorter than those for uncued target locations when the cue-target onset asynchrony (CTOA) is less than 300 ms. However, if the CTOA is greater than 300 ms, reaction times are typically longer when the target appears at the cued rather than uncued location (Maylor & Hockey, 1985; Posner & Cohen, 1984; Z. Wang & Klein, 2010). This effect is known as Inhibition of return (IOR; Posner et al., 1985). IOR is commonly regarded by researchers as an effect that results from an orienting mechanism that biases observers away from previously attended locations (Bennett & Pratt, 2001; Bourke et al., 2006; Gibson &

Egeth, 1994; Klein, 2000; Maylor & Hockey, 1985; Posner & Cohen, 1984; Pratt & Abrams, 1995; Soto & Blanco, 2004; Tipper et al., 1991, 1994, 1999).

The IOR effect is typically revealed in spatial cueing paradigms where cues and targets are presented at various positions limited by the horizontal (x) and vertical (y) axes of 2D space. However, human observers usually interact with stimuli in 3D space, and therefore, it is important to examine how IOR is affected by depth. Besides the advantage of studying IOR in environments of greater ecological validity, research on the allocation of attention in 3D space can also provide important theoretical insight. In a 2D setting, spatial locations can be specified in either the retinal coordinate or world-centered coordinate system. That is, the retinal coordinate system indicates when the location of the cue and target stimuli are positioned at the same location on the retina, whereas the world-centered coordinate system indicates that the location of the cue and target stimuli are at the same location in the environment. From studies on IOR in 2D settings, the question of which of the two coordinate systems the brain relies on more in representing spatial information cannot be answered unless a dynamic scenario (with either object or eye movement) is introduced during the time interval between cue and target presentation. However, in a static scene, more than one 3D spatial location expressed in world-centered coordinates can match the same location in retinal coordinates. In other words, two different locations in the world that lie along the same line of sight are situated on the same retinal location.

In almost all previous studies investigating IOR in 3D space, depth information has been provided through stereoscopic displays (Bourke et al., 2006; Casagrande et al., 2012; Theeuwes & Pratt, 2003; A. Wang et al., 2015, 2016). In one of the first studies to specifically research IOR in a 3D setting, depth information of the cue and target was not found to modulate the magnitude of IOR (Theeuwes & Pratt, 2003). Using stereoscopic displays, two rectangles were located nearer to and two rectangles farther from the observer and to the left or right side of the central fixation point. On each side of central fixation, the centres of the two rectangles were nearly matched in retinal coordinates and only varied in depth. During a trial, the cue appeared as one of the four rectangles brightened. After the cue offset, one of the four rectangles was replaced by an ‘H’ or an ‘S’ acting as the target. Only a main effect of side validity was found; the magnitude of IOR was similar regardless of whether the cue and target appeared in the same or different depth plane. Theeuwes and Pratt concluded that retinal coordinates, rather than depth, affected IOR.

Challenging the conclusions of Theeuwes and Pratt (2003)’s failure to find a depth effect, Bourke et al. (2006) suggested that, in Theeuwes and Pratt’s study, the two rectangles located at different depths but same side overlapped one another along the line of sight such that differences in depth were not perceivable. Accordingly, Bourke et al. (2006) modified Theeuwes and Pratt’s layout by slightly shifting the rectangles located on the same side away from one another along the horizontal and vertical axes to address this line of sight issue. This time, when a target appeared at the same depth plane as the cue, the IOR effect was larger than when a target and cue appeared at different depth planes. Given that the magnitude of IOR decreased when cues and targets appeared at different depth planes, Bourke et al. concluded that IOR was depth-specific. Other studies have supported this conclusion by showing significantly reduced IOR effects when orienting attention between depth planes when compared to within a single depth plane (Casagrande et al., 2012; A. Wang et al., 2015, 2016). Essentially, if one seeks a single value to describe the depth specificity of IOR, it can be calculated through Equation 1.

$$Depth\ Specificity = \left(RT_{same\ side} - RT_{diff\ side}\right)_{same\ depth} - \left(RT_{same\ side} - RT_{diff\ side}\right)_{diff\ depth} \quad \text{Eq. 1}$$

Given the conclusion that IOR can be depth-specific, when a cue and target appear sequentially in different depth planes, the direction of depth switch is an additional variable to consider. In attentional facilitation paradigms (CTOA < 300 ms), the literature suggests that participants commonly need more time to respond to a far stimulus when originally cued to a near stimulus, and less time to respond to a near stimulus when cued from farther away (de Gonzaga Gawryszewski et al., 1987; Downing & Pinker, 1985). This viewer-dependent depth-asymmetry has been explained by a behavioural preparedness that may exist for unexpected targets approaching a viewer's peripersonal space. However, the conclusion from these studies should be interpreted with more caution. For example, in the study by Downing & Pinker (1985), targets of the same physical sizes (LED lights) appeared at different distances. The lowered attentional cost for near targets could have been contributed by their larger retinal size.

Importantly, most studies investigating IOR in 3D environments have utilized stereoscopic displays—that is, visual depth was produced by inducing a disparity across the images presented to each eye. Unfortunately, in many of these studies, only the difference in IOR values for when cues and targets appear within the same depth versus different depths was analyzed. Thus, the effect of depth switch direction (far-cue to near-target versus near-target to far-cue) was never thoroughly examined. Wang et al. (2015) was the only study, to our knowledge, that examined the effect of depth switch direction on IOR in the traditional single cue to target IOR paradigm. Wang et al. (2015) reported that far-to-near orienting resulted in a reduced IOR than depth switches made in the other direction. However, the design in Wang et al. (2015) only examined the conditions when the cue and target appeared in different depth planes. The conditions in which the cue and target appeared in the same depth planes were not tested; thus, the effect of depth switch direction could have been influenced by the effect of the target's depth plane.

However, visual depth perception is also achieved by means of monocular-based pictorial depth cues. Further, stereoscopic displays that leverage binocular disparity only provide depth information over short spatial ranges, as binocular disparity is most useful for depth discrimination at distances of a few metres from the observer (McCann et al., 2018). Pictorial depth information becomes more useful than binocular stereo cues as the spatial range increases. Human observers are frequently required to process stimuli at different depth planes beyond peripersonal space when standing, walking, or participating in more critical behaviours like driving a vehicle. Pictorial depth cues might not offer the best precision in offering information about absolute depth, but they are often powerful enough in specifying relative depth information (e.g., McCann et al., 2018; Warren & Rushton, 2009).

One previous experiment, to our knowledge, has specifically investigated the depth-specificity of IOR using pictorial depth cues (Casagrande et al., 2012). This study presented cues and targets floating in a virtual empty room with the environment's depth simulated by linear perspective. To create the impression of 3D space, targets appearing nearer the observer were of larger retinal size and horizontal eccentricity while targets appearing farther were of smaller retinal size and horizontal eccentricity. Casagrande et al. found that IOR in the pictorial depth display remained the same regardless of whether targets were cued within the same or different depth plane. However, one might argue that it is possible that the pictorial depth cues used in this experiment were not strong enough to elicit the intended perception of depth. For example, the depth of the cue and target was only signaled through relative size and eccentricity. Size information is a useful indicator for distance judgements, but only for familiar objects, which were not present in this study (Hochberg & Hochberg, 1952; Maltz & Culham, 2020). Objects suspended in space, without intercepting the ground plane, may also have introduced ambiguity in discerning relative depth planes, since objects creating a larger retinal image can be either closer to the eye or larger in physical size. Overall, it remains unclear whether IOR exists when only pictorial depth information is available.

Accordingly, the purpose of the present study was twofold: (i) to address whether depth-specific IOR effects can be observed with pictorial depth cues, and (ii) to address whether the IOR effect varied based on near-far directionality in depth. In two experiments, we adapted Posner's exogenous cueing paradigm (Posner & Cohen, 1984) while simulating visual depth with pictorial cues. The source of depth information was provided by the texture gradient of the ground plane from an elevated perspective, and the relative size and linear perspective of the 3D rectangular boxes presented in the display.

In the virtual environment, on the surface of the ground plane, two columns of horizontal rectangular boxes (one on the left and one on the right side of the central fixation) were arranged along the z-axis (i.e., across depth). Each individual box was oriented horizontally (i.e., along the x-axis). These individual boxes were of the same size expressed in world coordinates. In the rendering of the 3D environment, the boxes located farther from the observer were made to be smaller in retinal size compared to the nearer boxes. The cue and target appeared on top one of these boxes. Although the size of the boxes varied in accordance with the geometric rules for 3D presentation (i.e., the box size followed an inverse relation between on-screen size and distance-to-viewer), we intentionally maintained a constant on-screen size and horizontal eccentricity for both cues and targets to equate their visual saliency at different depths.

The observer's virtual viewpoint was above the array of rectangular boxes, so stimuli at the same horizontal location but different depths did not fall on the same retinal location. Figure 1 illustrates a side view of the projection setup and depicts a vertical plane parallel with the observer's line of sight and orthogonal to the ground surface. From the observer's perspective, when the location of point N and point F differed in depth, their retinal on-screen locations (N' and F') differed as well. A nearer location in 3D space (e.g., N) appeared on-screen in a lower position (N') than that of a farther location (e.g., F, which projects to the F' on-screen position).

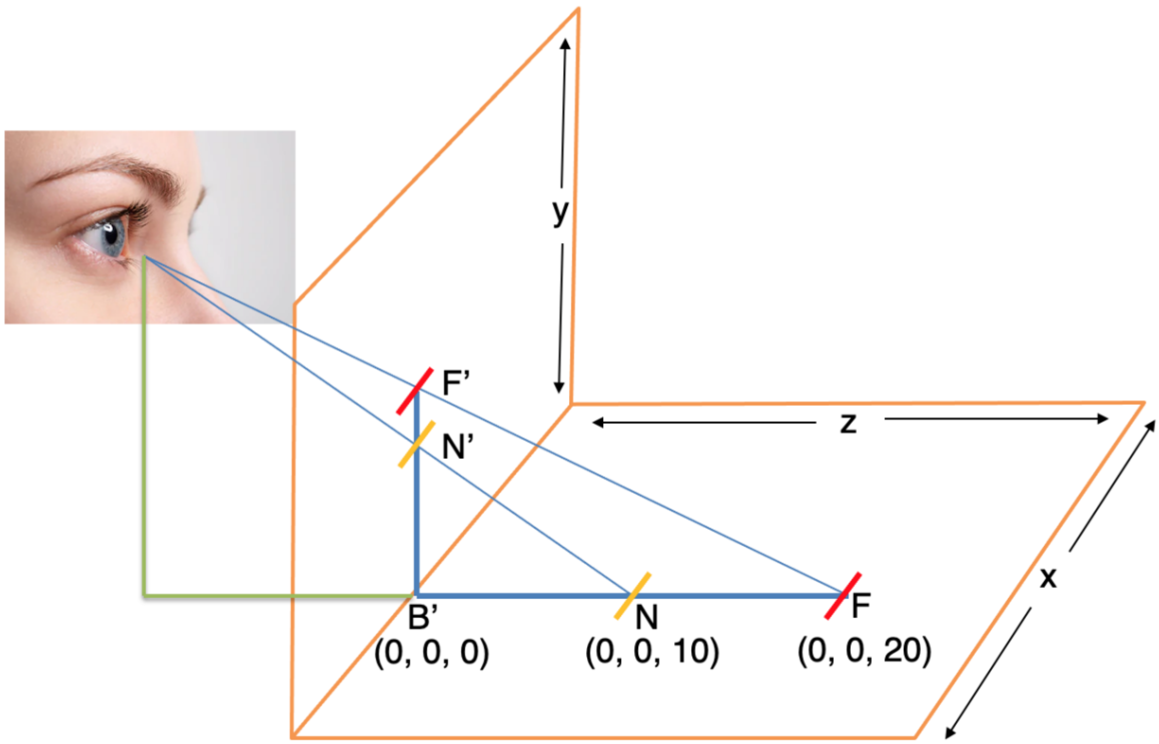


Figure 1. Side view of the projection setup for the 3D environment. Note that the z-coordinate difference between the near depth plane position and far depth plane position will always be of greater magnitude than the difference between corresponding on-screen positions (bottom elevation and top elevation) along the y-axis. Figure is not to scale.

In IOR studies involving the display of spatial locations on a 2D surface, it was found that when the cue and target appeared on the same side and in the exact same (x,y) location, IOR was maximal, and decreased monotonically with separation between cue and target on the same side of space (Bennett & Pratt, 2001; Maylor & Hockey, 1985). Based on this finding, when the cue and target are located at different depths in our monocular illustration of 3D space, IOR could decrease due to the vertical offset of cue and target location on the retina rather than in world space.

To discount the contribution of this vertical offset, we introduced a 2D condition where only the visual properties of the cue and target in the 3D condition were preserved but the contextual information for depth (ground plane and rectangular boxes) was removed. In effect, by analyzing the resulting IOR difference between 3D and 2D conditions, we were able to remove the confound caused by the on-screen displacement along the y-axis for cues and targets appearing at different depths. If IOR is sensitive to depth separation, as opposed to the on-screen 2D separation between cue and target, we should see a greater decrease in IOR in the 3D scene compared to the 2D scene when cue and target appeared at different depths.

We conducted two experiments that manipulated the cue and target position to be at either the same or different depth. In Experiment 1, two columns of three horizontal rectangular boxes were presented throughout the experiment. Cues appeared on either the left or right side and on top of any one of the three placeholders in each array, either at a near, middle, or far depth relative to the observer. The target only appeared at the middle depth. In Experiment 2, two columns of two horizontal rectangular boxes were presented throughout the experiment. Both cues and targets appeared on top of the left or right box at the near or far depth plane.

In both experiments, for the different-depth condition, we were able to compare the effect of different directions of depth switch between cue and target. In Experiment 1, depth switches were made either from the near to middle boxes or from the far to middle boxes. Given the target depth was constant, the comparison of two different depth switch directions could not have been affected by absolute target depth. However, the magnitude of depth switch (between far-to-middle and near-to-middle) was determined arbitrarily as 20 virtual metres (vm) for the far-to-middle depth switch and 10 vm for the near-to-middle depth switch. In Experiment 2, the attentional switches across depths were from the near to far boxes or from the far to near boxes. This time, both a constant magnitude of depth switch and direction switch in elevation was preserved, although the effect of depth switch on IOR could have been influenced by the absolute target depth.

Through these two experiments, we compared the differences in IOR when the cue and target appeared at the same verse different depth planes in our 3D scenes against the differences in IOR for the same versus different on-screen elevations in the corresponding 2D scenes. We also assessed whether the magnitude of the IOR effect varied based on the directionality of depth switch between cue and target onset (i.e., near-to-far and far-to-near) in the 3D scene relative to the corresponding locations in the 2D scene.

Experiment 1

As a first pass, we implemented the exogenous cueing procedure noted above where three pairs of rectangular boxes were on the left and right of central fixation and were arranged along the z-axis (i.e., across depth) and aligned horizontally (i.e., along the x-axis). The cue locations were aligned vertically along the y-axis and were presented on top of the rectangular boxes. We compared the differences in IOR when the cue appeared across these three possible depth planes (near, middle, or far depth), with the target only ever appearing in a single depth plane (middle depth).

Methods

Participants

A total of 54 adult participants (18-22 years old, n = 34 female) were randomly assigned to either a 3D or 2D perspective condition (with an equal number of participants in each perspective condition). All participants were granted a research participation credit or were compensated $10 CAD for their participation. An *a priori* minimum sample size of n = 54 was determined to achieve 95% power, as calculated by GLIMMPSE for the interaction between depth validity x perspective for farther versus nearer targets (Kreidler et al., 2013).

Stimuli and Apparatus

The simulation was created with Vizard 4.0 virtual reality software and back-projected onto a screen located in a dark tent free of external visual cues. Participants sat 150 cm in front of the screen, which measured 107.4 cm tall x 144.8 cm wide, with a steering wheel response device positioned comfortably in front of them.

In both 2D and 3D conditions, a dual-tinted blue cuboid appeared as the cue (diameter of 1.15° x elevation of 1.33°) followed by a dual-tinted red cylinder target (diameter of 1.15° x elevation of 1.33°). A short beep sound was made every time the cue or target appeared. A short quack-like sound was made every time the participant made an error responding to the side of the target. During a trial, participants were instructed to press the left flap on the backside of a steering wheel when the target appeared on the left side and the right flap on the backside of a steering wheel when the target appeared on the right side.

In the 3D condition, a red vehicle (1.8° x 2.4°) was displayed centrally on screen. The target and cue were positioned atop the rectangles in the visual display. The nearest pair of rectangular boxes were positioned 10 vm away from the virtual origin (9.05° vertically from the bottom of the screen and 6.6° horizontally from the centre of the lead car with a retinal size of 1.5° x 20.3°). The second pair of boxes was positioned at the same x-axis as that of the car's mid frontal plane 20 vm away (13.0° vertically from the bottom of the screen and 4.4° horizontally from the centre of the lead car with a retinal size of 0.95° x 22.4°). The farthest pair of boxes was positioned 40 vm away (15.2° vertically from the bottom of the screen and 3.1° horizontally from the centre of the lead car with a retinal size of 0.5° x 23.5°). The viewpoint was positioned 20 vm behind and 4 vm above the lead red car, with the lead red car appearing 12.7° from the bottom of the screen (see Figure 2). The ground plane was made by placing a grey tiled texture on an array of the planes spread out along the x and z axes, with the sky remaining visible above the horizon.

The on-screen size of the three boxes (and horizontal separation between each pair of left and right boxes) varied in accordance with the geometric rules for 3D presentation. However, the on-screen size and horizontal eccentricity (24°) were held constant for both cues and targets (positioned atop the boxes) across depth space. In addition, a red vehicle (1.8° x 2.4°) was displayed centrally on screen, which served as a point of fixation.

The 2D condition was identical to the 3D condition except that the cues and targets were presented on a blue background. Further, instead of a red car, a white cross (1.33° x 1.33°) was the central fixation point (as shown in the bottom portion of Figure 2).

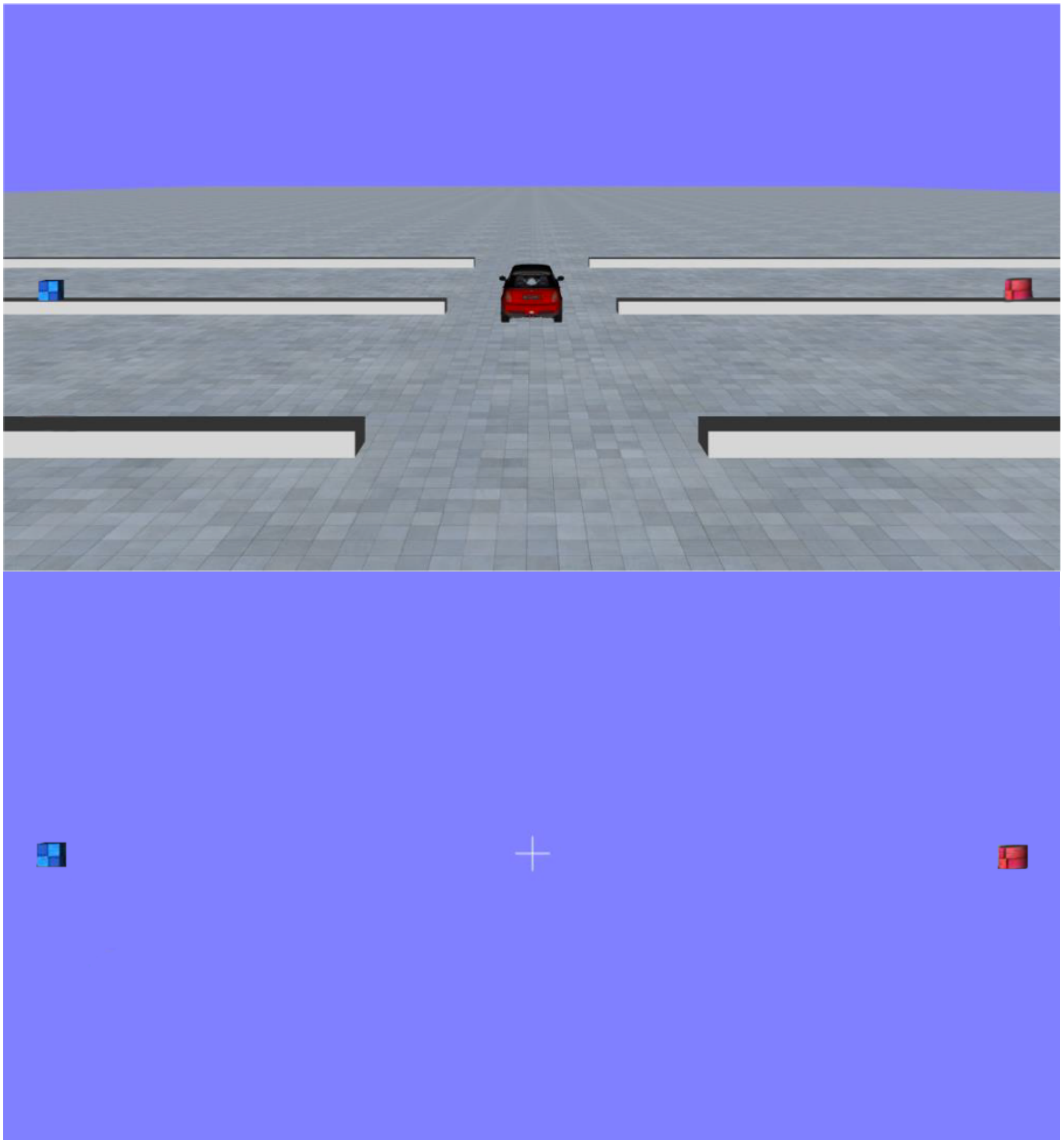


Figure 2. Top: An example of the cue, target, and 3D display of Experiment 1. Bottom: An example of the cue, target, and 2D display of Experiment 1.

Procedure

Participants were instructed that throughout the experiment, they should fixate on the central red car. Each trial started with a 1000 ms delay, then a cue appeared for 50 ms at one of the six possible locations. Following a CTOA of 1150 ms, the target appeared for 200 ms. During a trial, participants were instructed to press the left flap on the backside of a steering wheel when the target appeared on the left side and the right flap on the backside of a steering wheel when the target appeared on the right side. Participants were instructed not to make eye movements during a trial. The next trial began 1 s after the participant's response.

Participants performed 24 practice trials, which was followed by 936 experimental trials. The experimental trials were organized into 39 blocks of 24-trials, and participants were given the opportunity to take a break at the end of each block. Each block had an equal number of trials combining the following factors, which were randomized and repeated twice within a block: cue side (left versus right), target side (left versus right), and cue position (far, mid, or near depth plane in the 3D condition or top, mid, or bottom elevation in the 2D condition).

Design

This experiment used a between-subjects mixed factorial design. The within-subject variables were cue side (left/right), target side (left/right), and cue position (far, mid, or near depth in the 3D condition or top, mid, or bottom elevation in the 2D condition). The between-subjects variable was perspective (2D or 3D). For brevity, both 3D depth (far/mid/near) and 2D elevation (top/mid/bottom) will be referred to in terms of depth position.

Results

Reaction times (RTs) longer than 1500 ms and shorter than 150 ms, as well as incorrect localization responses were excluded from the analyses (1.38% of trials). Further outliers (1.9% of trials) were identified using a threshold criterion of three times the mean absolute deviation of the median of raw reaction time data for each participant and were excluded from the analyses (Leys et al., 2013). Altogether, approximately 3.28% of the experimental trials was removed from further analysis. The dependent variable was IOR magnitude, which was computed by subtracting RTs when the cue and target appeared at the same minus different side along horizontal axis. See Table 1 for the summary data of Experiment 1.

Table 1. The summary data of Experiment 1.

Exp 1 Independent Variables			Dependent Variables			
Target Depth	Perspective	Depth or Elevation	Side (RT in ms); [SE in ms]	IOR (ms); SE [ms]	IOR difference between same vs different depth/elevation (ms); SE [ms]	Depth Effect: IOR difference with 2D effect removed (ms); SE [ms]
NEARER (OR LOWER) THAN CUE	3D	SAME	SAME (402); [16.2]	17.4; [2.2]	9.0; [2.0]	7.1; [2.5]
			DIFF (384.6); [16.7]			
		DIFF	SAME (394.67); [16.5]	8.3; [1.7]		
			DIFF (386.33); [16.8]			
	2D	SAME	SAME (372.7); [11.0]	19.5; [2.5]	1.9; [1.8]	
			DIFF (353.2); [10.7]			
		DIFF	SAME (370.5); [11.3]	17.6; [2.1]		
			DIFF (352.9); [10.6]			
FARTHER (OR HIGHER) THAN CUE	3D	SAME	SAME (402); [16.2]	17.4; [2.2]	10.4; [2.6]	1.5; [3.0]
			DIFF (384.6); [16.7]			
		DIFF	SAME (394.22); [16.3]	7.1; [1.8]		
			DIFF (387.15); [16.5]			
	2D	SAME	SAME (372.7); [11.0]	19.5; [2.5]	8.9; [2.5]	
			DIFF (353.2); [10.7]			
		DIFF	SAME (365.8); [10.9]	10.6; [2.1]		
			DIFF (355.2); [10.6]			

We conducted separate analyses that assessed when the cue was presented nearer or farther than the target. That is, mean IOR values for each participant were entered into two separate two-way analyses of variance (ANOVAs), where one ANOVA assessed when the cue position was nearer than the target, and the other ANOVA assessed when the cue position was farther than the target. Each mixed factor ANOVA shared independent variables that treated perspective (3D /2D) as a between-subjects factor and depth (same/different) as a within-subject factor. In the case of the same depth situation, only observations when the cue always appeared in the middle depth plane along with the target were included.

For the ANOVA when the target was farther than the cue, we observed a main effect of depth validity, $F(1, 52) = 25.7$, $p < .001$, $\eta_p^2 = .33$, indicating a larger IOR effect when the target and cue were presented at the same depth (18.45 ms) than different depths (8.85 ms). The main effect of perspective was not significant, $F(1, 52) = 1.8$, $p = .186$, $\eta_p^2 = .03$. The interaction between perspective x depth validity was not significant, $F(1, 52) = 0.12$, $p = .73$, $\eta_p^2 = .002$ (see the left portion of Figure 3).

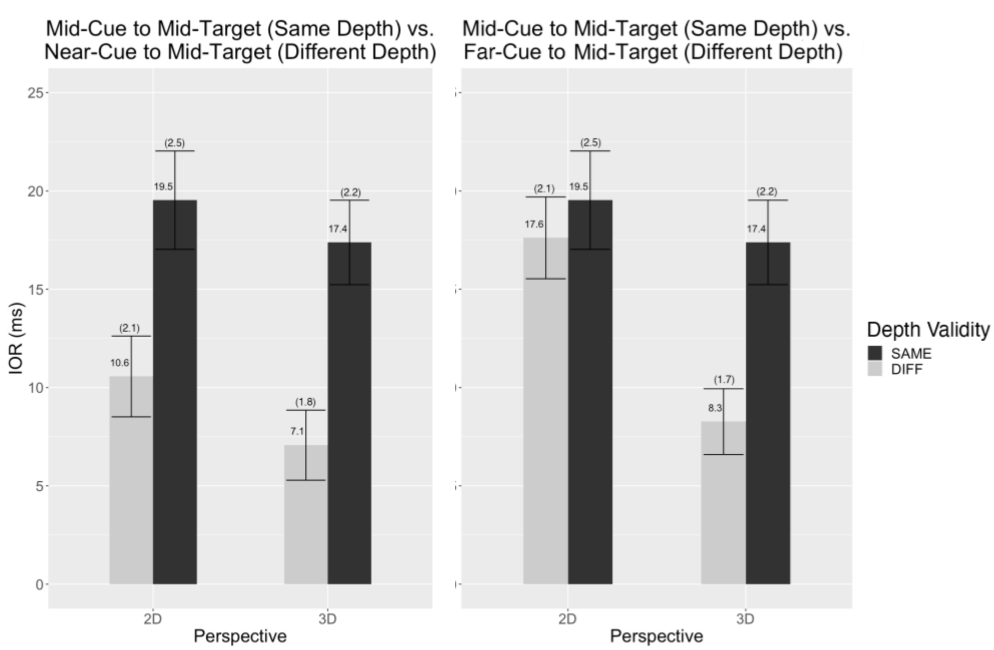


Figure 3. Left portion: The IOR effects when the target was farther than the cue in depth (different condition) and when the cue and target appeared in the middle depth (same condition) of Experiment 1. Right portion: The IOR effects when the target was nearer

than the cue in depth (different condition) and when the cue and target appeared in the middle depth (same condition) of Experiment 1.

For the ANOVA assessing when the target was nearer than the cue, we observed a main effect of perspective, $F(1, 52) = 5.8$, $p < .05$, $\eta_p 2 = .10$, reflecting a larger IOR effect for the 2D (18.55 ms) than 3D (12.85 ms) condition. There was a main effect of depth validity, $F(1, 52) = 15.5$, $p < .001$, $\eta_p 2 = .23$, reflecting a larger IOR effect for same depth (18.45 ms) than different depths (12.95 ms). Most importantly, there was also an interaction between perspective x depth validity, $F(1, 52) = 8.1$, $p < .01$, $\eta_p 2 = .13$. This interaction was represented by the IOR difference for same minus different depths being larger for the 3D condition (9.0 ms) compared to the IOR difference for same minus different elevations for the 2D condition (1.9 ms; see the right portion of Figure 3). To examine the source of the above two-way interaction, two planned comparisons were performed. The IOR was significantly greater for the same depth than different depths for the 3D condition, $F(1, 26) = 19.9$, $p < .001$, $\eta_p 2 = 0.43$, but not for the 2D condition, $F(1, 19) = 0.72$, $p = 0.41$, $\eta_p 2 = 0.03$.

The significant two-way interaction between perspective x depth validity for targets nearer than the cue (and the analogous non-significant interaction for targets farther than the cue), can be conceptualized by calculating the difference in the difference of IOR scores, as represented by what we refer to as the depth effect (DE), shown in Equation 2.

$$Depth\ effect = \left[IOR_{3D(Same\ Depth)} - IOR_{3D(Different\ Depth)}\right] - \left[IOR_{2D(Same\ Elevation)} - IOR_{2D(Different\ Elevation)}\right],$$

where $$IOR_{3D(Same\ Depth)} = (RT_{same\ side\ \&\ depth} - RT_{diff\ side\ \&\ same\ depth});$$

$$IOR_{3D(Different\ Depth)} = (RT_{same\ side\ \&\ different\ depth} - RT_{diff\ side\ \&\ diff\ depth});$$

$$IOR_{2D(Same\ Elevation)} = (RT_{same\ side\ \&\ same\ elevation} - RT_{diff\ side\ \&\ same\ elevation});$$

$$IOR_{2D(Different\ Elevation)} = (RT_{same\ side\ \&\ diff\ elevation} - RT_{diff\ side\ \&\ diff\ elevation}) \quad \text{Eq. 2}$$

We input reaction times for targets presented in either the same or different side or same or different depth plane/elevation as the cue. When DE > 0, IOR was judged to be depth-specific and when DE = 0, IOR was judged to be depth-blind. According to our results, the DE was 7.1 ms for far-to-near depth switches (i.e., depth-specific IOR), but 1.5 ms for near-to-far depth switches (i.e., depth-blind IOR) in Experiment 1 (see the left portion of Figure 4).

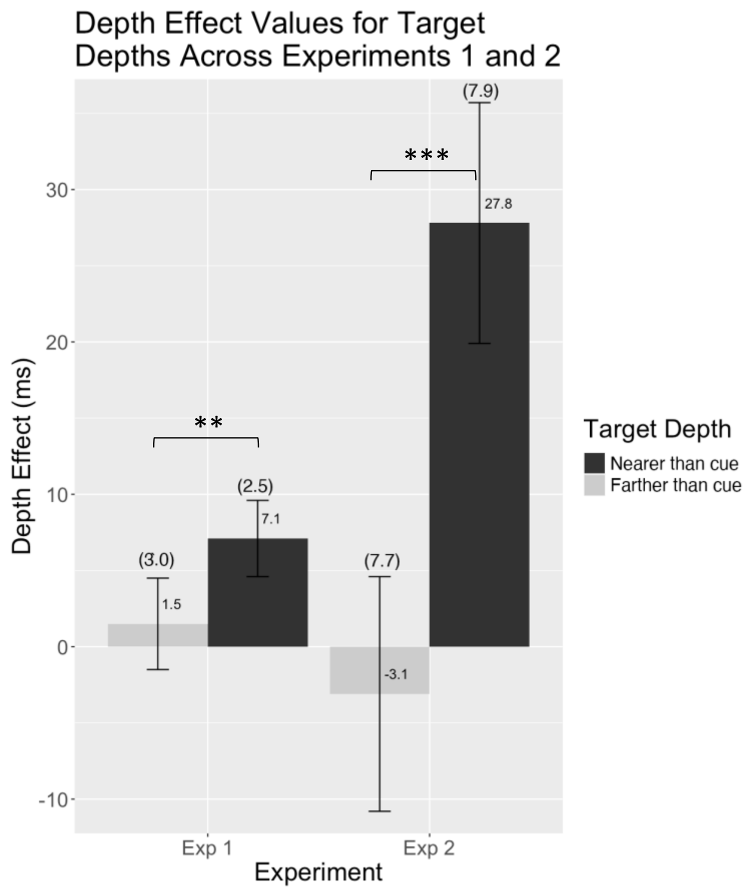


Figure 4. Summary of depth effect results. The magnitude of reduction in IOR when cues and targets appeared at different versus same depth planes in a 3D scene compared to a 2D scene according to Equation 2. Standard errors are in parentheses.

Discussion

The results indicate that the magnitudes of IOR were smaller for the two "different" conditions (different-depth in 3D or different-elevation in 2D) compared to the two "same" conditions (same-depth in 3D or same-elevation in 2D). When this difference in depth was caused by the cue appearing farther than the target, an interaction was found. Specifically, the mean IOR in the "different" condition was much smaller than the IOR in the "same" condition in the 3D condition, but not so in the 2D condition. However, when the difference in depth was caused by the cue appearing nearer than the target, this interaction did not occur. Conceptually, this was represented by a greater depth effect for targets appearing nearer than the cue (7.1 ms) compared to targets appearing farther than the cue (1.5 ms). This depth-specific IOR for targets nearer than the cue suggests that the brain might be capable of modulating attention based on the interpretation of monocular-ascribed depth information. A depth-specific IOR in the context of pictorial depth cues complements prior stereoscopic studies that have considered binocular disparity as a major contributor to a depth-specific effect on IOR (Casagrande et al., 2012; A. Wang et al., 2015, 2016).

Experiment 2

The results of Experiment 1 suggest that the IOR effect is attenuated for a target appearing nearer than the cue, likely because of an observer's bias for responding preferentially to spatial locations within their virtual peripersonal space. However, we should also consider other factors that could potentially confound this depth-specific effect. The depth effect could also have been affected by the magnitude of the depth difference between cue and target. In Experiment 1, the horizontal rectangular boxes on which the cue and target were positioned were located 10, 20, and 40 vm along the z-axis in virtual space.

We intentionally used a larger virtual spatial gap between middle and far rectangular boxes to increase the illusion of linear perspective. We predicted that this increased linear perspective would influence observers to place, relative to the larger on-screen vertical location between near and middle rectangular boxes, greater weight on the larger depth distance between middle and far rectangular boxes. The depth-specific IOR found for a middle target preceded by a far cue could have been influenced by the long (20 vm) z-axis distance between far and middle rectangular boxes. Thus, the unequal z-axis separation in the two depth differences (10 vm versus 20 vm) might have helped create the depth-blind and depth-specific IOR effects, respectively. Furthermore, because of the geometry involved in realistic 3D rendering, the on-screen retinal distance between the far and middle rectangular boxes was less than half the retinal distance between the near and middle rectangular boxes. Thus, the compression of space for more distal locations in the 3D display could have enlarged the differential IOR results in 3D and 2D conditions. To remove the confound of the magnitude of depth difference, we held the depth difference between cue and target constant in Experiment 2.

Moreover, in Experiment 1, the depth information in the 2D display was ambiguous due to lack of context. Thus, in Experiment 2, the cue and target were made to appear on the surface of a vertical wall erected in the 3D environment, which provided clear information that the cue and target were on the same depth plane. This would also make the appearance of 2D and 3D scenes more comparable. In addition, we increased the number of possible depth planes at which the target could appear, thus making target detection less predictable.

Methods

Participants

Data from 20 young adults (18-22 years old, n=14 female) were analyzed. All participants were either compensated with a course credit or $10 CAD for their participation. The minimum sample size of n = 12 was determined by GLIMMPSE (Kreidler et al., 2013) as the appropriate sample size based on the perspective x depth validity for nearer targets in Experiment 1. All variables were within-subject variables now able to be placed in a general omnibus ANOVA analyzing the three-way interaction between perspective x depth validity x target depth due to targets no longer just appearing in a single depth plane.

Stimuli and Apparatus

The stimuli and apparatus of Experiment 1 were identical to Experiment 2 with the following exceptions. A dual-tinted yellow cuboid appeared as the cue followed by a green spherical target (radius = 0.67°). The cues and targets appeared in one of four positions: on the left or right side of the central fixation in the far or near depth plane in the 3D condition or in the top or bottom elevation in the 2D condition.

In the 3D condition, there were only two pairs of rectangular boxes. To enhance the perception of depth, we stretched the rectangular boxes 0.66 vm along the z-axis. The central red car (1.4° x 1.9°) was positioned equidistantly along the depth-axis between the near and far pair of rectangular boxes at 15 vm (10.2° from the bottom of the screen) in front of the viewer. The near rectangles were positioned 10 vm away from the viewer (5.9° vertically from the bottom of the screen and 5.2° horizontally from the centre of the lead car with a retinal size of 1.5° x 23.5°). The far rectangles were 20 vm away from the viewer (12.9° vertically from the bottom of the screen and 2.8° horizontally from the centre of the lead car with a retinal size of 0.7° x 20.3°; see top portion of Figure 5).

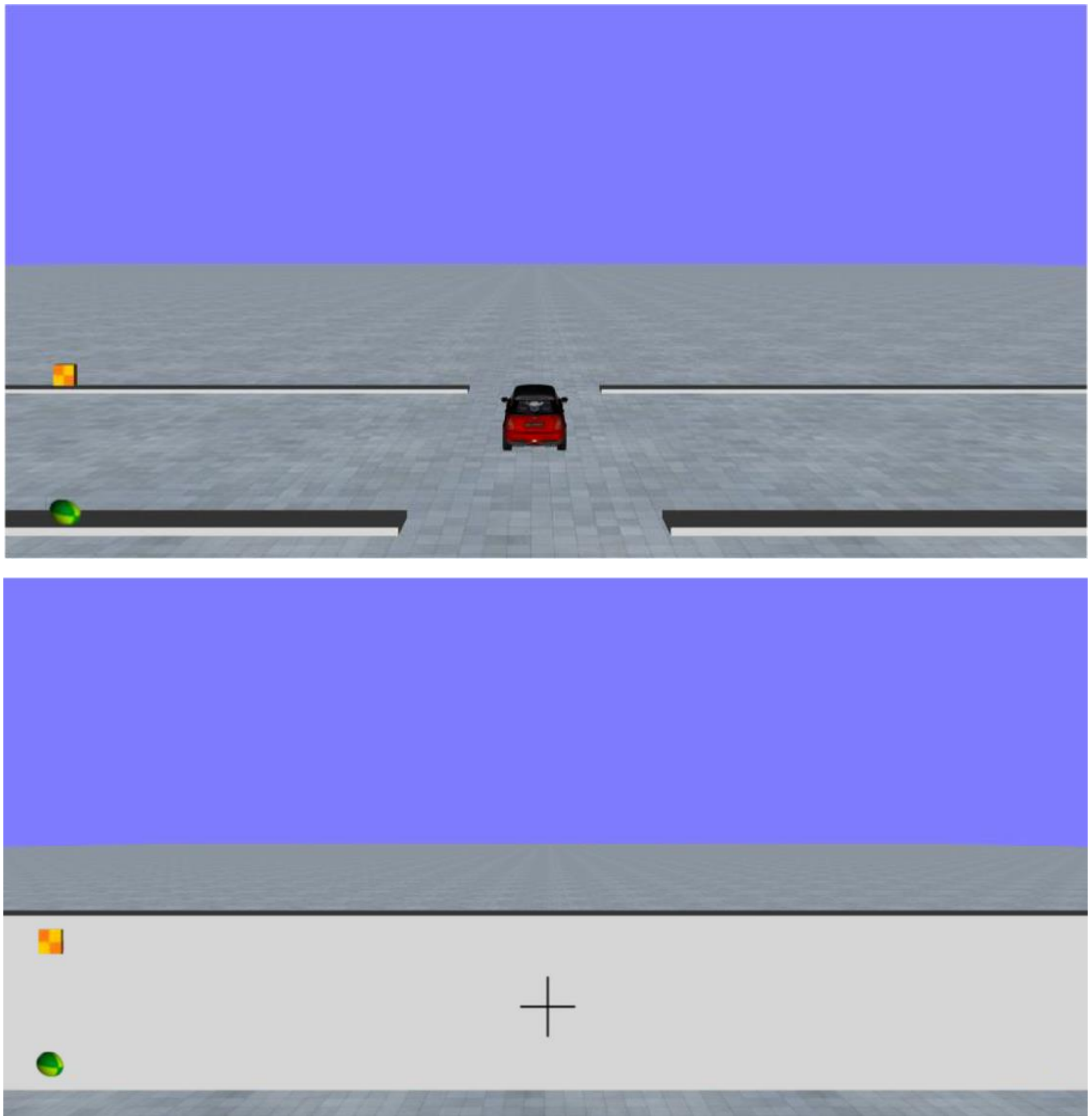


Figure 5. Top: An example of the cue, target, and 3D display of Experiment 2. Bottom: An example of the cue, target, and 2D display of Experiment 2.

In the 2D condition, the stimuli were presented on a beige panel located 10 vm away from the viewer (5.9° vertically from the bottom of the screen with a retinal size of 10.1° x 44.0°). A black-lined cross (1.8° x 1.8°) was the central fixation point. The cues and targets appeared in front of this placeholder wall. A quack-like sound was made every time the participant made an anticipatory error, took longer than 2 seconds to respond, or made a localization error (see bottom portion of Figure 5).

Procedure

The experimental procedure was identical to Experiment 1 with the following exceptions. 1500 ms after the offset of the fixation cross, the cue appeared for 50 ms. Following a CTOA that randomly ranged from 750 to 1050 ms, the target appeared for 100 ms. The next trial began 1500 ms after a participant's response.

Participants now performed 864 experimental trials. The experimental trials were organized into 24 blocks of 36 trials each, which included 4 catch trials in each block. Each block had an equal number of trials combining the following factors: cue side (left versus right), target side (left versus right), target depth (near versus far), cue depth (near versus far). Perspective was counterbalanced (3D block followed by a 2D block and vice-versa).

Design

The design was identical to Experiment 1 except that there were now only near and far levels of the depth factor, with targets being able to appear in both depth planes. Perspective (3D versus 2D) was now a within-subject variable.

Results

RTs longer than 1000 ms and shorter than 150 ms, as well as incorrect localization responses and anticipatory errors, were excluded from the analyses (2.84% of experimental trials). Observations constituted as outliers (2.0% of experimental trials) based on a threshold criterion of three times the mean absolute deviation of the median of raw reaction time data for each participant were also excluded from the analyses (Leys et al., 2013). Altogether, approximately 4.84% of the experimental trials was removed from further analysis. The dependent variable was IOR magnitude which was computed by subtracting RTs from when the cue and target appeared at the same than at a different location.

See Table 2 for the summary data of Experiment 2. Overall, for Experiment 2, the effect of depth specificity was 27.8 ms for the far-to-near depth switch (i.e., depth-specific IOR), but -3.1 ms for the near-to-far depth switch (i.e., depth-blind IOR) with both depth effect values seen in the last column in Table 2. The conclusions of "depth-specific" versus "depth-blind" effect on IOR are supported by the statistical analysis described below.

Table 2. The summary data of Experiment 2.

Exp 2	Independent Variables		Dependent Variables			
Target Depth	Perspective	Depth or Elevation	Side (RT in ms); [SE in ms]	IOR (ms); SE [ms]	IOR difference between same vs different depth/elevation (ms); SE [ms]	Depth Effect: IOR difference with 2D effect removed (ms); SE [ms]
NEAR	3D	SAME	SAME (486.9); [28.1]	36.0; [4.1]	31.9; [5.5]	27.8; [7.9]
			DIFF (450.9); [25.7]			
		DIFF	SAME (464.55); [26.6]	4.1; [3.6]		
			DIFF (460.49); [28.1]			
	2D	SAME	SAME (463.6); [26.1]	18.6; [4.7]	4.1; [5.2]	
			DIFF (445.0); [27.2]			
		DIFF	SAME (453.2); [26.6]	14.5; [4.0]		
			DIFF (438.7); [25.6]			
FAR	3D	SAME	SAME (488.76); [27.1]	26.4; [4.6]	9.26; [5.9]	-3.1; [7.7]
			DIFF (462.32); [28.2]			
		DIFF	SAME (469.3); [27.9]	17.2; [3.9]		
			DIFF (452.1); [27.7]			
	2D	SAME	SAME (471.54); [26.4]	23.1; [4.6]	12.35; [4.6]	
			DIFF (448.49); [27.0]			
		DIFF	SAME (457.26); [26.9]	10.7; [5.0]		
			DIFF (446.54); [25.9]			

Mean IOR values were first entered into a factorial ANOVA that treated target depth (near/far), depth validity (same/different), and perspective (2D/3D) as within-subject factors. The main effect of depth validity was significant, $F(1, 19) = 28.8$, $p < .001$, η_p^2 =

0.6, reflecting that the IOR effect was larger when the cue and target appeared at the same than different depths. The main effect of perspective was not significant, $F(1, 19) = 0.64$, $p = .44$, $\eta_p^2 = .03$. The main effect of target depth was not significant, $F(1, 19) = 0.01$, $p = .94$, $\eta_p^2 < .001$. The two-way interaction of depth validity x perspective was significant, $F(1, 19) = 7.0$, $p < .05$, $\eta_p^2 = .30$. The two-way interaction between depth validity x target depth was not significant, $F(1, 19) = 1.7$, $p = .21$, $\eta_p^2 = .08$. The two-way interaction between perspective x depth validity was not significant, $F(1, 19) = 0.02$, $p = .88$, $\eta_p^2 = .001$. The three-way interaction of depth validity x perspective x depth directionality was significant, $F(1, 19) = 5.6$, $p < .05$, $\eta_p^2 = .20$. The IOR effects were assessed further by separate ANOVAs that treated depth validity (same/different) and perspective (2D/3D) as factors for each level of target depth.

The ANOVA when the target was presented farther than the cue revealed a significant main effect of depth validity, $F(1, 19) = 8.91$, $p < .01$, $\eta_p^2 = .07$, reflecting a larger IOR effect when the cue and target were presented at the same depth (24.75 ms) than at different (13.95 ms) depths. The main effect of perspective was not significant, $F(1, 19) = 0.52$, $p = .48$, $\eta_p^2 = .03$. The interaction between depth validity x perspective was not significant, $F(1, 19) = 0.22$, $p = .64$, $\eta_p^2 = .01$ (see the left portion of Figure 6).

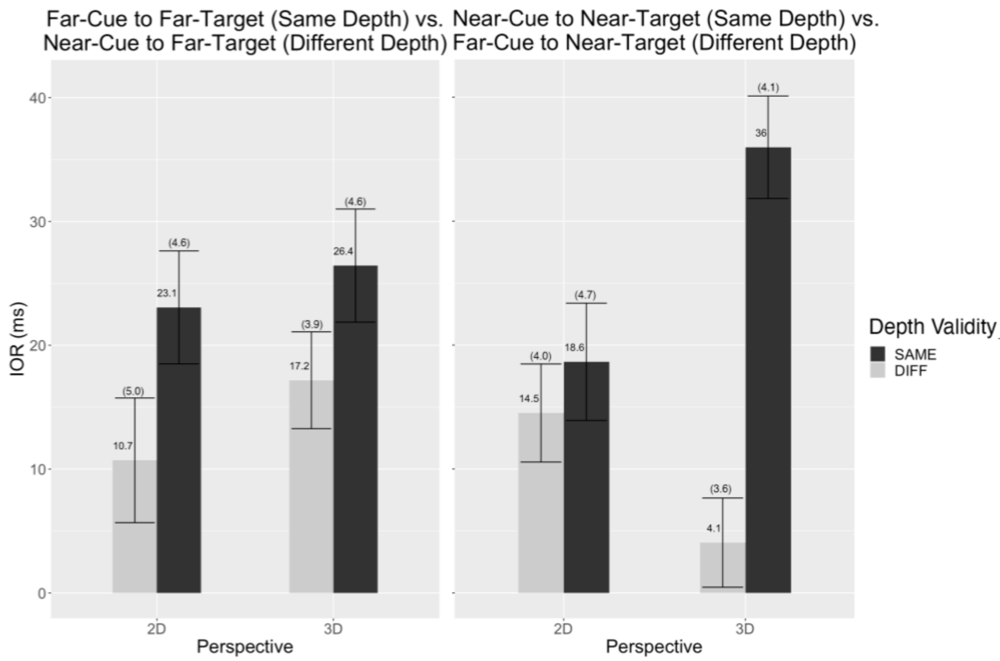


Figure 6. Left portion: The IOR effects when the target was farther than the cue in depth and when the cue and target appeared at the same depth of Experiment 2. Right portion: The IOR effects when the target was nearer than the cue in depth and when the cue and target appeared at the same depth of Experiment 2.

The ANOVA for when the target was presented nearer than the cue revealed a main effect of depth validity, $F(1, 19) = 21.5$, $p < .001$, $\eta_p^2 = .53$, reflecting a larger IOR effect when the cue and target were presented at the same depth (27.3 ms) than at different (7.4 ms) depths. The main effect of perspective was not significant, $F(1, 19) = 0.30$, $p = .59$, $\eta_p^2 = .02$. There was a significant interaction of depth validity x perspective, $F(1, 19) = 11.5$, $p < .001$, $\eta_p^2 = .40$. This interaction was represented by the IOR difference for same minus different depths being larger for the 3D condition (31.9 ms) compared to the IOR difference for same minus different elevations for the 2D condition (4.1 ms) (see the right portion of Figure 6). To examine the source of the two-way interaction, two planned comparisons were performed. The IOR was significantly greater for same depths versus different depths for the 3D condition, $F(1, 19) = 30.4$, $p < .001$, $\eta_p^2 = .62$, but not for the 2D condition, $F(1, 19) = 0.45$, $p = .51$, $\eta_p^2 = 0.02$. According to our results, the DE was

27.8 ms for far-to-near depth switches (i.e., depth-specific IOR), but -3.1 ms for near-to-far depth switches (i.e., depth-blind IOR) in Experiment 2 (see the right portion of Figure 4).

Discussion

Like in Experiment 1, when targets were presented nearer than the cue, the effect of depth/elevation on IOR differed between the 3D and 2D conditions. Specifically, the magnitude of the IOR effect was greater in the same than different depth condition in the 3D scene but not in the 2D scene. However, when targets appeared at the far locations, the effect of depth validity did not differ significantly between the 3D and 2D conditions. In Experiment 2, because the same magnitude of depth difference was used for both depth switch directions, the findings cannot be explained by the differences in depth separation across the near and far locations.

General Discussion

Overall, both experiments found that the IOR effect was smaller when the cue and target were presented at the different than same depths in the 3D displays, and when the cue and target were presented at the different elevations in the 2D displays. Most importantly, the reduction in IOR in 3D displays was greater than that in the corresponding 2D displays, but only when the difference in depth was caused by a target appearing nearer than the cue. In other words, there was a depth-specific IOR effect when the target was presented nearer than the cue and the IOR effect was depth insensitive when the target was presented farther from the cue.

In the current study, the perceived depth differences may have been amplified by the rectangular boxes on which the cues and targets appeared. The depth locations of cues and targets were clearly demarcated by rectangular objects. We introduced these placeholders to reduce the ambiguity in the depth information. Without them, the available spatiotemporal information provided by the cues and targets alone (as in Casagrande et al., 2012) may have been unreliable in producing a depth-specific IOR. Indeed, the IOR effect has not previously exhibited depth specificity in 3D scenes composed of pictorial depth cues because the depth locations of the stimuli were not clearly demarcated by intercepting with the ground plane. Instead, previous designs established depth locations only by changing retinal size and horizontal/vertical eccentricities of cues and targets (Casagrande et al., 2012).

In both Experiments 1 and 2 of our study, the IOR difference between the same and different elevations in the 2D conditions was also affected by the direction switch in elevation. For Experiment 2, for instance, the IOR value for cues and targets at the same bottom elevation (18.6 ms) was comparable to the IOR for when the observer made a downward switch from a higher cue to lower target (14.5 ms). However, the IOR value for cues and targets at the same top elevation (23.1 ms) was much greater than IOR for when the observer shifted their attention upward from a lower cue to a higher target (10.7 ms). This directional difference in the 2D conditions may have contributed to the difference in the depth effect found between the depth switch directions in 3D versus 2D space. However, the overall depth effect (factoring both 3D and 2D results as noted by Equation 2) cannot be entirely explained by 2D results—note that the IOR difference (36.0 ms) between same versus different depths (for the far cue and near target condition in the 3D condition) was much greater than any other conditions in Experiment 2.

The finding that IOR magnitude decreases for far-to-near depth switches but not for near-to-far depth switches has not been reported in most of the studies using stereopsis depth. In fact, most of these studies did not report the results for the variable of depth switch direction. Instead, they combined the data from the two depth switch directions and only focused on the comparison of IOR between cue and target appearing at the same versus different depth (Bourke et al., 2006; Casagrande et al., 2012; Theeuwes & Pratt, 2003). The finding that IOR magnitude decreases for far-to-near depth switches but not for near-to-far depth switches has, to our knowledge, only been suggested in two prior studies (A. Wang et al., 2015, 2016), albeit with stimulus configurations and calculations not directly comparable with our study and the rest of the literature. First of all, in Wang et al (2015, 2016), the physical size of the targets was made to be the same across depth, which consequently made the retinal size of the near targets to be greater than that of the far targets. Therefore, the effect of depth switch direction could have been confounded by differences in low level visual properties across depth. Second, their designs were not the same as ours. In the Wang et al. (2015)'s detection condition for Experiment 2, participants were told to fixate centrally and respond to peripheral targets. The cues and targets appeared at either both the same side and same depth or both different side and different depth. Thus, when cues and targets appeared in the same depth plane, they always appeared at the same side. When cues and targets appeared in different depth planes, they always appeared at different sides of central fixation. This side-biased setup was not counterbalanced between participants. Thus, IOR in their study was calculated as shown in Equation 3.

$$IOR = (RT_{same\ side})_{same\ depth} - (RT_{diff\ side})_{diff\ depth} \qquad \text{Eq. 3}$$

In other words, IOR was never calculated by encompassing RTs to stimuli presented at different x-locations within a single depth plane or at the same x-location at different depth planes. Furthermore, the only depth switches were made from the left near-depth plane to right far-depth plane (or from the right far-depth plane to the left near-depth plane), but never across the orthogonal diagonal trajectory in depth space (i.e., right near-depth plane to or from the left far-depth plane). Therefore, the change in the spatial extent of IOR was not evaluated comprehensively for all these trajectories as typically evaluated in the literature (see Equation 1).

The reduced IOR found for far-to-near depth switches in our study appears to be consistent with findings of a near advantage in the spatial cueing literature examining the facilitation effect (for smaller CTOAs). Studies of the facilitatory effect of attentional orienting in 3D space have also found that a smaller attentional cost exists for far-to-near switches, with observers tending to react more quickly to targets unexpectedly appearing closer to them and in the unattended hemispace (Chen et al., 2012; de Gonzaga Gawryszewski et al., 1987a; Downing & Pinker, 1985; Maringelli et al., 2001).

The depth-specific IOR effect in spatial cueing paradigms is consistent with results from studies examining the default distribution of attention in 3D space (albeit without spatial cueing). Overall, a near advantage in attention allocation is typically shown in these studies. For example, Li, Watter, and Sun (2011) demonstrated that peripheral target detection performance was more accurate for near targets in real space. Song, Bennett, Sekuler and Sun (2021) also found an advantage for near targets that involved observers experiencing visually simulated forward motion in a virtual environment.

Figure 7 provides an illustration of how attention might vary across different depths following cue and target presentations in this current study. For each depth switch direction (i.e., a far cue followed by a near target or a near cue followed by a far target), there are two possible outcomes: a depth-blind or depth-specific IOR. Our results are aligned with the two respective outcomes underlined in Figure 7 (a depth-specific IOR for far-to-near depth switches and a depth-blind outcome for near-to-far depth switches).

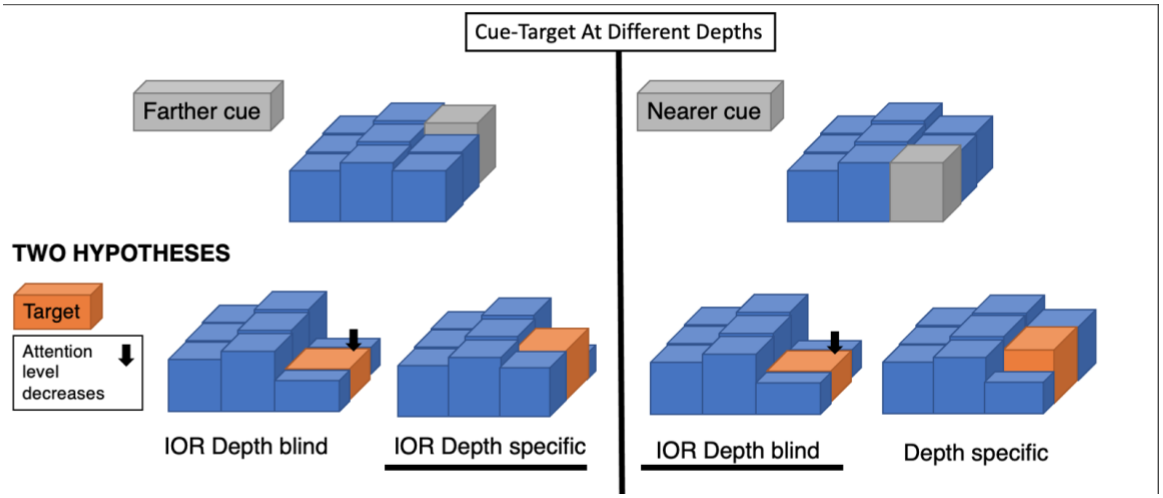


Figure 7. Conceptual model of attention distribution in 3D space: Our results align with the two respective outcomes underlined in Figure 7 (a depth-specific IOR for far-to-near depth switches and a depth-blind outcome for near-to-far depth switches). A depth-specific IOR occurs when significant inhibition is absent for targets cued from a different depth plane. If IOR is coded in world coordinates, and if the cue appears at the farther depth plane, attention would not be affected for a target appearing in the nearer depth plane. Attention would still be held at a high level due to a lack of inhibition to that uncued depth plane. This would mean that the 3D separation between cue and target is perceived, and it is done so in world coordinates. A depth-blind IOR occurs when inhibition is still present for targets cued from a different depth plane. If IOR in coded in retinal coordinates, there would be no difference in IOR values for a target appearing at a different depth from the cue. Attention would be equally inhibited for all depth planes that shared the retinal coordinates of cue onset.

In the current study, the reduction in IOR was greatest for switches made from far-to-near space. Given that near space likely affords a higher level of baseline attention than that of far space, this depth-specific effect could have been contributed by a greater facilitatory effect in near space. Moreover, the default level of attention for far space is likely low, with inhibition for a cued location in far space less effectively spreading to near space. A similar reason could explain why IOR reduction was not seen when a depth switch was made from near to far space. The baseline attention level in far space is smaller, thus making the IOR effect more evident. Moreover, given that attention to near space is high by default, the inhibition for a cued location may be more effective in spreading from near to far space.

This asymmetry in switching between depths may have evolved to deal with the inherent urgency typically required in responding to a near target. If a cue appears farther from the observer, and it is followed by a target appearing in a nearer space, then the urgency and preparedness placed upon this unexpected target is facilitated (or least not inhibited) for both same and different side conditions nearly equally, thus masking and lessening the IOR magnitude.

Conclusion

We conducted two peripheral cued target localization experiments that used a modified IOR paradigm in a 3D setting. Both experiments found that in an environment without stereopsis but with rich pictorial depth cues, IOR showed a depth-specific effect. The depth-specific effect found in our 3D condition (factoring both 3D and 2D results) was not entirely explained by 2D parameters. The IOR difference between same versus different depths (for the far cue and near target condition) was much greater than any other conditions. A large source of this difference was because far-to-near switch trajectories in the 3D setting resulted in a reduced IOR magnitude than near-to-far switches. The near depth plane of space likely has a higher baseline of attentional focus, favouring facilitation rather than inhibition of attention. We realize that the depth-specific differences in IOR found in our study could have been influenced by object-based in addition to world-coordinate differences. We also acknowledge that object-based IOR is only found under limited conditions (List & Robertson, 2007; Pilz et al., 2012). Regardless, future studies could examine whether the depth-asymmetric IOR effect found in the current study is mostly due to location- or object-based components.

References

Arnott, S. R., & Shedden, J. M. (2000). Attention switching in depth using random-dot autostereograms: Attention gradient asymmetries. *Perception & Psychophysics*, *62*(7), 1459–1473. https://doi.org/10.3758/BF03212146

Bennett, P. J., & Pratt, J. (2001). The spatial distribution of inhibition of return. *Psychological Science*, *12*(1), 76–80. https://doi.org/10.1111/1467-9280.00313

Bourke, P. A., Partridge, H., & Pollux, P. M. J. (2006). Additive effects of inhibiting attention to objects and locations in three-dimensional displays. *Visual Cognition*, *13*(5), 643–654. https://doi.org/10.1080/13506280544000309

Casagrande, M., Barbato, M., Mereu, S., Martella, D., Marotta, A., Theeuwes, J., & Collinson, S. (2012). Inhibition of return: A “depth-blind” mechanism? *Acta Psychologica*, *140*, 75–80. https://doi.org/10.1016/j.actpsy.2012.02.011

Chen, Q., Weidner, R., Vossel, S., Weiss, P. H., & Fink, G. R. (2012). Neural mechanisms of attentional reorienting in three-dimensional space. *The Journal of Neuroscience*, *32*(39), 13352–13362. https://doi.org/10.1523/JNEUROSCI.1772-12.2012

de Gonzaga Gawryszewski, L., Riggio, L., Rizzolatti, G., & Umiltá, C. (1987). Movements of attention in the three spatial dimensions and the meaning of “neutral” cues. *Neuropsychologia*, *25*(1A), 19–29.

Downing, C., & Pinker, S. (1985). The spatial structure of visual attention. In *Attention and performance: Vol. XI* (pp. 171–187). Erlbaum.

Gibson, B. S., & Egeth, H. (1994). Inhibition of return to object-based and environment-based locations. *Perception & Psychophysics*, *55*(3), 323–339. https://doi.org/10.3758/bf03207603

Han, S., Wan, X., & Humphreys, G. W. (2005). Shifts of spatial attention in perceived 3-D space. *The Quarterly Journal of Experimental Psychology Section A*, *58*(4), 753–764. https://doi.org/10.1080/02724980443000548

Hochberg, C. B., & Hochberg, J. E. (1952). Familiar size and the perception of depth. *The Journal of Psychology*, *34*(1), 107–114. https://doi.org/10.1080/00223980.1952.9916110

Klein, R. (2000). Inhibition of return. *Trends in Cognitive Sciences*, *4*(4), 138–147.

Kreidler, S. M., Muller, K. E., Grunwald, G. K., Ringham, B. M., Coker-Dukowitz, Z. T., Sakhadeo, U. R., Barón, A. E., & Glueck, D. H. (2013). GLIMMPSE: Online power computation for linear models with and without a baseline covariate. *Journal of Statistical Software*, *54*(10), i10.

Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, *49*(4), 764–766. https://doi.org/10.1016/j.jesp.2013.03.013

Li, T., Watter, S., & Sun, H. J. (2011). Differential visual processing for equivalent retinal information from near versus far space. *Neuropsychologia*, *49*(14), 3863–3869. https://doi.org/10.1016/j.neuropsychologia.2011.10.002

Maltz, M., & Culham, J. (2020). Familiar size affects size and distance perception for real objects, even in the presence of oculomotor cues. *Journal of Vision*, *20*(11), 1568–1568. https://doi.org/10.1167/jov.20.11.1568

Maringelli, F., McCarthy, J., Steed, A., Slater, M., & Umiltà, C. (2001). Shifting visuo-spatial attention in a virtual three-dimensional space. *Cognitive Brain Research*, *10*(3), 317–322. https://doi.org/10.1016/S0926-6410(00)00039-2

Maylor, E. A., & Hockey, R. (1985). Inhibitory component of externally controlled covert orienting in visual space. *Journal of Experimental Psychology: Human Perception and Performance*, *11*(6), 777–787. https://doi.org/10.1037/0096-1523.11.6.777

McCann, B. C., Hayhoe, M. M., & Geisler, W. S. (2018). Contributions of monocular and binocular cues to distance discrimination in natural scenes. *Journal of Vision*, *18*(4). https://doi.org/10.1167/18.4.12

Miura, T., Shinohara, K., & Kanda, K. (2002). Shift of attention in depth in a semi-realistic setting. *Japanese Psychological Research*, *44*(3), 124–133. https://doi.org/10.1111/1468-5884.00015

Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. *Attention and Performance X*, *32*, 531–556.

Pratt, J., & Abrams, R. A. (1995). Inhibition of return to successively cued spatial locations. *Journal of Experimental Psychology. Human Perception and Performance*, *21*(6), 1343–1353. https://doi.org/10.1037//0096-1523.21.6.1343

Song, J., Bennett, P. J., Sekuler, A. B., & Sun, H.-J. (2021). The effect of apparent distance on peripheral target detection. *Journal of Vision*, *21*(10), 1–8. https://doi.org/10.1167/jov.21.10.8

Soto, D., & Blanco, M. J. (2004). Spatial attention and object-based attention: A comparison within a single task. *Vision Research*, *44*(1), 69–81. https://doi.org/10.1016/j.visres.2003.08.013

Theeuwes, J., & Pratt, J. (2003). Inhibition of return spreads across 3-D space. *Psychonomic Bulletin & Review*, *10*(3), 616–620. https://doi.org/10.3758/BF03196523

Tipper, S. P., Driver, J., & Weaver, B. (1991). Object-centred inhibition of return of visual attention. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, *43*(2), 289–298. https://doi.org/10.1080/14640749108400971

Tipper, S. P., Jordan, H., & Weaver, B. (1999). Scene-based and object-centered inhibition of return: Evidence for dual orienting mechanisms. *Perception & Psychophysics*, *61*(1), 50–60. https://doi.org/10.3758/BF03211948

Tipper, S. P., Weaver, B., Jerreat, L. M., & Burak, A. L. (1994). Object-based and environment-based inhibition of return of visual attention. *Journal of*

Experimental Psychology. Human Perception and Performance, *20*(3), 478–499. https://doi.org/10.1037/0096-1523.20.3.478

Wang, A., Liu, X., Chen, Q., & Zhang, M. (2016). Effect of different directions of attentional shift on inhibition of return in three-dimensional space. *Attention, Perception, & Psychophysics*, *78*(3), 838–847. https://doi.org/10.3758/s13414-015-1055-6

Wang, A., Yue, Z., Zhang, M., & Chen, Q. (2015). Interaction between spatial inhibition of return (IOR) and executive control in three-dimensional space. *Experimental Brain Research*, *233*(11), 3059–3071. https://doi.org/10.1007/s00221-015-4374-x

Wang, Z., & Klein, R. M. (2010). Searching for inhibition of return in visual search: A review. *Vision Research*, *50*(2), 220–228. https://doi.org/10.1016/j.visres.2009.11.013

Chapter 3: Object-Based IOR Is Depth-Blind

Abstract

A cued target detection experiment measured inhibition of return (IOR) in a computer-simulated three-dimensional (3D) setting. Cues and targets appeared within garden beds situated on the surface of a textured ground plane. The 3D scene featured linear perspective, texture gradient, and relative size as the main pictorial depth cues. When the cue appeared in a farther location than that of the subsequent target, the IOR effect was depth-specific. IOR was smaller when the cue and target appeared in different depth planes than in the same depth plane. While all other spatial properties remained the same, this depth effect was only present when the cue and target appeared in two garden beds separated in depth, but not when they appeared in within a single garden bed stretched along the depth axis. These results suggest that IOR can be depth-specific when making a far-to-near depth switch, a finding nullified by object-based inhibition.

Introduction

Visual attention can be allocated flexibly to locations in space (Bennett & Pratt, 2001; Casagrande et al., 2012; de Gonzaga Gawryszewski et al., 1987; Maylor & Hockey, 1985; Posner, 1980) and also to objects in space (Bourke et al., 2006; Leek et al., 2003; Tipper et al., 1991, 1994, 1997, 1999; Watson & Kramer, 1999). Attention has been consistently shown to be more strongly influenced by location-based effects rather than object-based effects (B. S. Gibson & Egeth, 1994; Krüger & Hunt, 2013; Pilz et al., 2012). This result likely occurs because spatial attention can exist independent of object-based properties in our visual field (Bennett & Pratt, 2001; Casagrande et al., 2012; Posner, 1980; Tipper et al., 1997), while object-based attention cannot exist without spatial boundaries demarcating the outlines of relevant objects (Leek et al., 2003; Tipper et al., 1999; Watson & Kramer, 1999). However, attention cued to a single object remains preferentially relegated within that object, making it difficult to orient to a novel object in space (Atchley & Kramer, 2001; Jordan & Tipper, 1999; Tipper et al., 1994, 1999).

Reaction times to targets at exogenously cued locations are typically shorter than those for uncued target locations when the cue-target onset asynchrony (CTOA) is less than 300 ms. However, if the CTOA is greater than 300 ms, reaction times to a target appearing at the cued location are typically greater than those for an uncued location (Maylor & Hockey, 1985; Posner & Cohen, 1984; Z. Wang & Klein, 2010b). This reaction time delay is known as the inhibition of return (IOR). IOR is typically revealed in a spatial cueing paradigm where cues and targets are presented at various positions demarcated by a horizontal axis (x-axis) and a vertical axis (y-axis) running along two-dimensions (i.e., in 2D space). The target may appear at either the same or different location as a previous cue. IOR is then calculated by subtracting the average reaction time for the uncued target location from the cued target location. Researchers commonly regard the orienting mechanism underlying IOR as essential in enhancing the detection efficiency of stimuli, biasing observers away from previously attended locations to novel locations (Bennett & Pratt, 2001; Bourke et al., 2006; B. S. Gibson & Egeth, 1994; R. Klein, 2000; Maylor & Hockey, 1985; Posner & Cohen, 1984; Pratt & Abrams, 1995; Soto & Blanco, 2004; Tipper et al., 1991, 1994, 1999). IOR also has been shown in 3D space by comparing conditions where targets may appear in the same or different depth plane as the cue, eliciting observers to orient their attention along the z-axis. If IOR is significantly smaller when cues and targets appear at different depth planes than in the same depth plane, IOR is considered depth-specific (Bourke et al., 2006; Casagrande et al., 2012; A. Wang et al., 2016).

IOR can exhibit an object-based component as well. With the spatial distance between cue and target remaining equal, IOR has consistently been shown to be greater when both the cue and target are located within an object than when located between two different objects (Jordan & Tipper, 1999; Tipper et al., 1991, 1994, 1997, 1999). This object-based component of IOR also increases when the boundaries of an object are made more salient with continuous and homogenously brightened, coloured, or textured perimeters or areas (Jordan & Tipper, 1999; Leek et al., 2003; Watson & Kramer, 1999). Although many studies have found a noticeable object-based IOR, location-based IOR appears to dominate attentional orienting. For example, when the distance between a cue and target increases within an object, IOR decreases (Jordan & Tipper, 1999). Similarly, IOR decreases the farther a cued object moves from its originally cued spatial position (Tipper et al., 1994). These findings suggest that object-based properties may be limited in their influence on IOR.

Furthermore, object-based effects have also been shown to be small in certain task-specific settings. For example, unpredictable object movements (Krüger & Hunt, 2013) and shortened presentation time of objects or CTOAs (Avrahami, 1999; List & Robertson, 2007) have been shown to pre-empt the manifestation of object-based effects due to an observer's inability to parse the objects in enough time. Additionally, research has shown that while nearly all observers demonstrate significant location-based attention, very few exhibit object-based effects (Pilz et al., 2012). These findings suggest that significant object-based IOR may be more task-specific and may not occur in many naturalistic viewing conditions.

Three-dimensional environments presented through stereoscopic cues have demonstrated an object-based IOR effect independent of location-based properties (Bourke et al., 2006; Casagrande et al., 2012; B. S. Gibson & Egeth, 1994). However, even in 3D scenes, the object-based component of IOR can easily degrade because of viewers preferring to attend to the coarse 2D spatial coordinates of a scene if an object's features frequently change, such as when the object is in high-speed motion (B. S. Gibson & Egeth, 1994; Krüger & Hunt, 2013). When the boundaries of the object degrade and the cues and targets are located between depth planes in an empty 3D space, the object-based component of IOR is reduced, with the location-based component of IOR remaining (Bourke et al., 2006; Casagrande et al., 2012; B. S. Gibson & Egeth, 1994).

The first experiment that attempted to separate the effects of spatial coordinates from object coordinates on IOR in 3D space was conducted by Gibson and Egeth (1994). Cues and targets were placed at various locations on a simulated brick whose orientation was transformed in space (e.g., see their Figure 3). The six surfaces of this brick were considered to belong to a single object. Cues and targets, in the form of dots, appeared in either the same spatial location along the (x,y) axes, or on the same object surface rotated in depth, or both. It was hypothesized that the magnitude of IOR could be separately affected by the spatial locations or the object surfaces on which the cue and targets appeared. Gibson and Egeth (1994) found that IOR consisted of two components: i) an object-based component, where IOR spread across an object's surface even though the cue and target appeared at different spatial locations; and ii) a location-based component, where IOR appeared when the cue and target location shared the same spatial location but not the same object surface. They found the location-based component of IOR was significantly greater than the object-based component of IOR. However, one major issue with Gibson and Egeth's experimental design was that the surfaces of the object were never separated from the brick's global structure during its rotation, possibly preventing the observers from placing inhibitory tags on the individual planes. As a result, the contributions brought forth by the location and object-based components of attentional orienting may have been conflated.

Bourke et al. (2006) compared the contributions of location- and object-based components of IOR in a stereoscopic setting by analyzing how IOR behaved between and within objects that extended in depth. In their first experiment, cues and targets were presented in separate boxes (see the top of their Figure 1) and IOR was smaller when the cue and target appeared in different depth planes (13 ms) than when they appeared in the same depth plane (23 ms). In their second experiment, when the boxes were joined to form a single object that extended across two different depth planes (see the bottom of their Figure 1), IOR was the same (25 ms) regardless of whether the cue and target appeared at the same or different depth planes. Bourke et al. concluded that only a location-based component of IOR acted for the condition with two separate objects at different depth planes, and that this IOR likely operated in 2D (x,y) coordinates irrespective of depth differences. When an individual object extended across different depths, the researchers suggested that an object-based component of IOR, operating in 3D coordinates, was additive with a location-based component of IOR, which operated in "depth-blind" 2D coordinates. Bourke et al., however, did not measure IOR in an object-free condition where the cues and targets were unattached to objects, and therefore could only infer that the object-based component of IOR operated in 3D coordinates. In other words, their findings are insufficient to suggest the location-based component of IOR only operated in depth-blind 2D coordinates.

Casagrande et al. (2012) criticized Bourke et al.'s (2006) conclusion by correctly stating that their design lacked an object-free condition, which only permitted them to state that object-based IOR operated in 3D coordinates. Casagrande et al. modified Bourke et al.'s experiment by presenting cues and targets in a virtual space devoid of objects (see their Figure 2). Participants fixated on a point that was centered on a gray rectangle representing the back wall. Four patterned parallelograms represented the top, bottom, left, and right-side walls, all slightly slanted and appropriately shaded to render the scene compelling in depth. A trial began with the participants fixating on the central fixation point. The cue (a 3D square frame) appeared before the target (a black sphere), which appeared in one of four peripheral locations. If the target was at a nearer depth plane relative to the observer, then it appeared at a larger size and horizontal eccentricity. If the target was at a depth plane farther from the observer, then it appeared at a smaller size and horizontal eccentricity. These pictorial depth cues were presented within a stereoscopic setting. The results from this 3D condition were compared to a condition that only used pictorial depth cues free of stereopsis. Casagrande et al. found a significantly greater IOR within the same depth plane (28 ms) than between depth planes (18 ms) in the stereoscopic condition, but not in the non-stereoscopic pictorial depth cue condition. The changes in IOR along the depth axis in the stereoscopic condition appeared to be completely location-based: the targets and cues were not contextually attached to any objects, such as rectangles or other shapes, yet the depth-specificity of IOR prevailed. Pictorial depth cues such as relative eccentricity and retinal size likely interacted with a strong depth context introduced by stereopsis, allowing for the differential IOR pattern along the depth axis. Casagrande et al. concluded that the depth-specificity of IOR could be influenced purely by spatial location differences between the cue and target, but only in a stereoscopic scene and likely not in a scene with only pictorial depth cues.

In a 3D environment composed only of pictorial depth cues, the presence of objects demarcating various depth planes may be required to elicit a reliable location-based IOR. When a single object extends over multiple depth planes in the same environment, the object-based component of IOR may dominate and override the need for observers to attend to differences in coarse spatial coordinates within the object. To date, no studies have explored the presence of the object-based IOR effect in 3D scenes using only pictorial depth cues. Thus, the following study attempts to separate the location-based versus object-based components of IOR in a 3D scene composed of pictorial depth cues.

Current Study

The current study used a modified Posner cueing paradigm to investigate IOR in a 3D scene simulated through depth cues only interpretable from the pictorial layout of the scene. The study investigated whether IOR was affected by depth switches made between cues and targets located at different depth planes. We also examined the difference in

IOR when the cue and target appeared at two distinct depth planes occupied by a single object, or two different objects. As shown in Figure 1, the major source of depth information was provided by a ground surface beneath garden beds located at the left and right sides along the horizontal axis. These garden beds contained the potential cue and target onset locations.

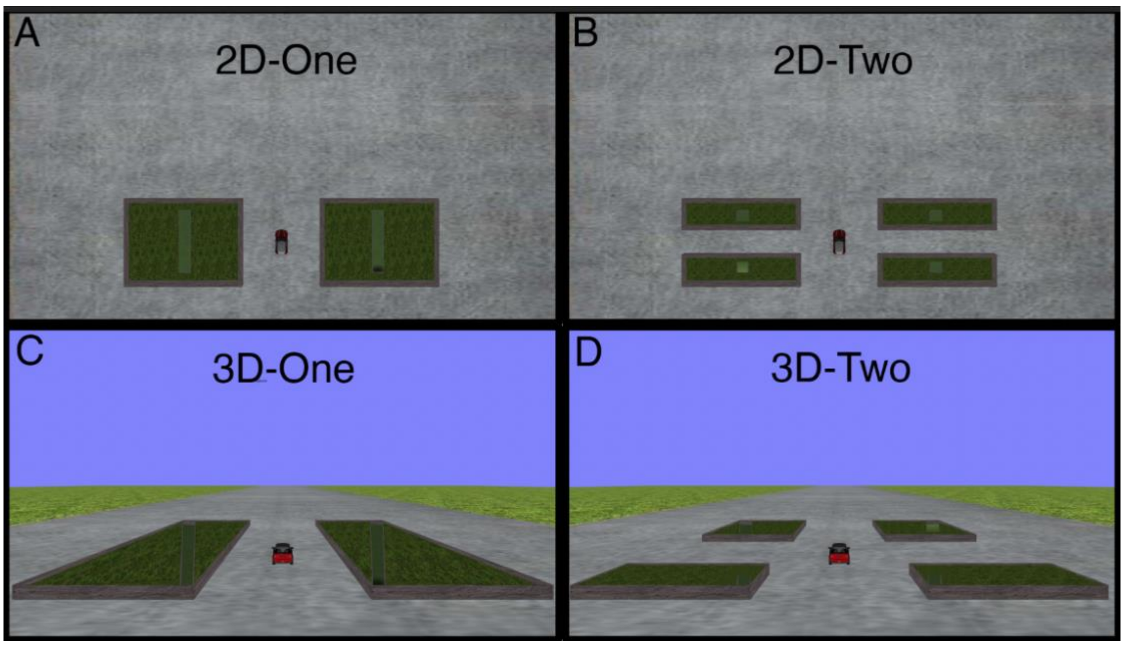


Figure 1. Stimuli in Experiment 1**.** In this experiment, there was a condition with one garden bed on either side of fixation (1A and 1C) or two garden beds on either side of fixation (1B and 1D). The red vehicle served as the point of fixation and was always 14 vm in front of the viewer in the 3D scene. The cue was a shadow and the target a patch of light. Figure 1A illustrates the 2D setting where a cue appears in the bottom elevation of the right garden bed. Figure 1B illustrates the 2D setting where a target appears in the bottom elevation of the left lower garden bed. Figure 1C illustrates the 3D setting where a cue appears in the near depth plane within the right garden bed. Figure 1D illustrates the 3D setting where a target appears in the far depth plane of the right farther garden bed. In all conditions in this experiment, both cue and target either appeared in the same depth plane or different depth planes.

Figure 2 illustrates the geometric relation between locations in simulated 3D space and corresponding 2D on-screen locations. The illustration shows a side view of the study's projection setup, which depicts a vertical plane (i.e., where x = 0 on the horizontal axis) parallel with the observer's line of sight and orthogonal to the ground surface of the 3D scene. The observer's viewpoint was located slightly above ground. From the observer's perspective, the locations N and F separated along the z-axis would be displayed in the locations of N' and F' separated in elevation on-screen. A nearer location in 3D space (e.g., N) would appear on-screen in a lower position (N') than that of a farther location (e.g., F, which would project to the F' on-screen position). Note that stimuli at the same horizontal but different depth locations did not fall on the same retinal location.

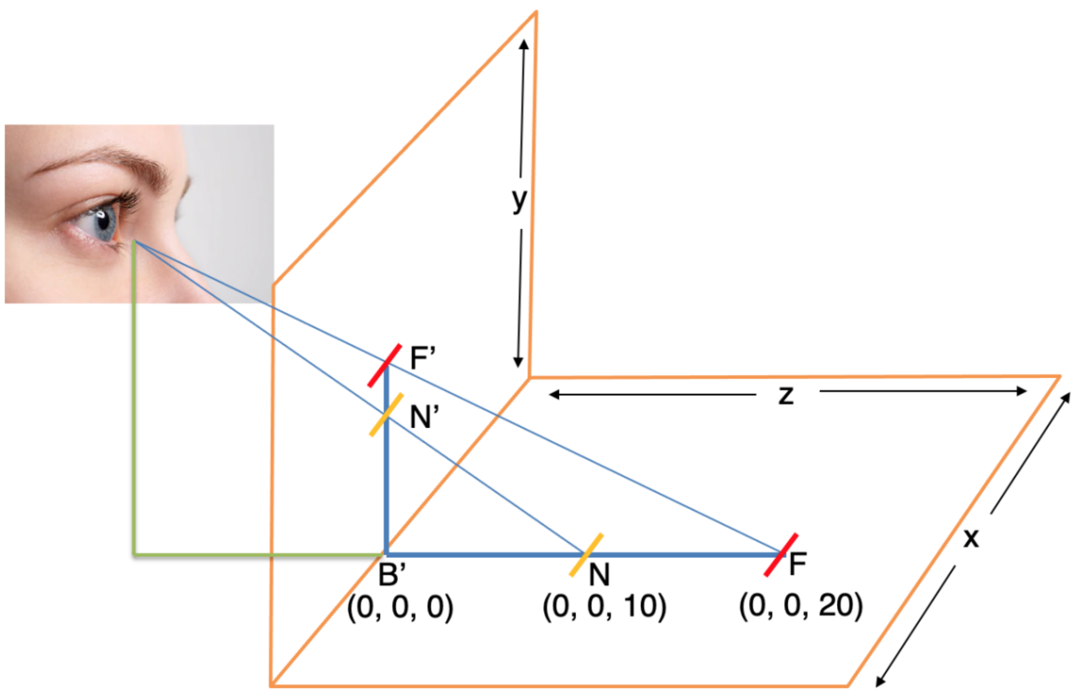


Figure 2. Side view of the projection setup for the 3D environment. Note that the z-coordinate difference between the near depth plane position and far depth plane position will always be of greater magnitude than the difference between corresponding on-screen positions (bottom elevation and top elevation) along the y-axis. Figure is not to scale.

The experiments in Chapter 2 showed that IOR was sensitive to depth separation (i.e., the effect is likely depth-specific). That is, when the cue and target appeared in different depth planes compared to when they appear in the same depth plane, IOR decreased in magnitude. The magnitude of this depth effect likely reflected the extent of separation along the z-axis in world space between near and far depth planes (N and F in Figure 2, respectively). In 2D space, however, it is commonly found that IOR is maximized when the cue and target appear at the same location and decreases monotonically with increasing separation between cue and target in (x,y) space (Bennett & Pratt, 2001; Maylor & Hockey, 1985). Based on this finding, when the cue and target in the current study are located at different depths within the pictorial illustration of 3D space, IOR could decrease simply due to the vertical offset of cue and target location on the retina. So, like in Chapter 2, the current study, with control 2D conditions, discounted the contribution of cue-target separation in elevation in the 2D scene from the analogous IOR established in the 3D scene. Given that the z-axis distance between N and F is greater than the corresponding on-screen separation between these two points (N' and F'), the extent of IOR reduction for the different-depth condition (compared to the same-depth condition) should also be greater in the 3D condition than that in the corresponding 2D condition within the current study. This would once again establish IOR as depth-specific.

Earlier work from Chapter 2 also showed that the depth-specific IOR is found for the far-to-near but not for the near-to-far depth switch. In other words, the IOR effect is significantly smaller when the cue and target appear in different depth planes, but only when the cue appears at farther depth and the target nearer. Although some previous research on the effect of depth (presented through stereoscopic information) on IOR has reported a depth-specific effect (Bourke et al., 2006; Casagrande et al., 2012; A. Wang et al., 2015, 2016), only two of these prior studies have reported an effect of depth switch direction across depth, albeit not according to the traditional calculation of lateral IOR within and across depth planes as noted in Equation 1.

$$IOR_{depth} = (RT_{same\ side} - RT_{diff\ side})_{same\ depth} - (RT_{same\ side} - RT_{diff\ side})_{diff\ depth} \quad \text{Eq. 1}$$

To investigate the context in which depth-specific IOR was generated in our 3D scenes, we examined how IOR (calculated according to Equation 1) changed when a cue and target appeared at different depth planes but within a single object. The phenomenon of object-based IOR was evaluated by comparing the IOR depth effect when the cue-target separation was within one object or between two objects. It is not clear whether the depth-specific effect for far-to-near depth switches would be found when the cue and target appear at different depths but still within the same object extended in depth. Given previous reports of inhibition spreading throughout a cued object (Bourke et al., 2006; Jordan & Tipper, 1999), it was predicted that direction-specific and depth-specific IOR would disappear when the depth switch between depths occurred within an object. In other words, finding that IOR does not decrease for a far-to-near depth switch in a single object would indicate that object-based IOR in that circumstance is similar in 3D and 2D viewing conditions (Jordan & Tipper, 1999; Leek et al., 2003; Tipper et al., 1994).

Methods

Participants

Data from 39 young adults (18-23 years old, n = 32 female) attending McMaster University were analyzed. All participants were granted a research participation credit for the first session and had the option of receiving a research participation credit or being compensated $10 CAD for the second session. The minimum sample size of n = 32, calculated by GLIMMPSE (Kreidler et al., 2013), gave the study a power of 95% for the effect of the interaction between depth validity x perspective x target depth x number of garden beds on IOR.

Stimuli and Apparatus

The simulation was created with Vizard 4.0 virtual reality software and back-projected onto a screen located in a dark tent. Participants sat 150 cm in front of the screen, which measured 107.4 cm tall x 144.8 cm wide, with a steering wheel response device positioned comfortably in front of them.

Figure 1 shows a sample environment presented to the participants. On the surface of the ground plane in the 3D condition, either one or two garden beds on either side of fixation were arranged along the z-axis (across the depth plane) with their sides and top views visible. In the 2D condition, the garden beds were arranged along the vertical axis, with the observer seeing a birds-eye view of the scene. These individual garden beds resembled rectangular boxes with equal physical lengths and equal widths. In the rendering of the 3D environment, however, the horizontal widths of the garden beds that were located farther from the observer were smaller in retinal size compared to the widths of the nearer boxes. Inside these garden beds appeared a greener and elevated texture, resembling shrub hedge, that stood within the confines of an outer concrete boundary.

In addition to this green texture, the middle of the garden beds housed a smaller, elongated rectangle, with the bottom floor of the box containing a uniform green colour, framed by inner concrete-textured separators. The elongated garden bed was described to participants as a hollowed-out well within the garden, which was being redeveloped. The well was oriented along the vertical on-screen axis in both 3D and 2D conditions. Consequently, in the 3D condition, the garden bed well appeared to be oriented along the depth direction inside the larger garden bed. In the 2D condition, this well appeared completely vertical on screen. In both 3D and 2D conditions, for the one versus two objects condition, the garden bed well's retinal dimensions were identical.

The garden beds appeared on either side of a red car (central fixation), which measured (width = 1.2° x elevation = 3.2°) in the 2D condition and 1.8° x 3.0° in the 3D condition. In the Full Condition, one garden appeared on either side of central fixation. In the Split Condition, two garden beds appeared on either side of central fixation.

In the 2D condition, the observer's perspective was oriented 90° towards the ground plane at 24.5 virtual metres (vm) above the ground. The result was that the observer perceived the top of the central red car, the placeholders, and the ground plane. The observer did not see the sky because the view was simulated as if the observer was lying in a prone position looking above the ground plane from an aerial viewpoint. The central fixation was positioned 12.7 vm north of the world origin (origin being x = 0, y = 0, z = 0), or 9° from the bottom edge of the projector screen.

In the 2D-One Condition (see Figure 1a), the top view of one garden bed appeared on either side of central fixation. The horizontal distance from central fixation to the middle of the garden bed's inner edge was 3.65 vm (3.9°). The garden bed's dimensions measured 10.5 vm (11.7°) horizontally and 7.5 vm (11.6°) vertically on the projection screen.

In the 2D-Two Condition (see Figure 1b), the top view of two garden beds appeared on either side of central fixation. The horizontal distances from central fixation to the middle of the garden beds' inner edges measured 3.65 vm (3.9°). Both garden beds measured 10.5 vm (11.7°) horizontally and 2.5 vm (4.3°) vertically. There was a ground space gap of 1.75 vm (3.0°) in between the top and bottom elevations.

In the 3D condition, the observer was situated 4 vm above origin and looked forward, perceiving the back of the central red car, the placeholders, ground plane and surrounding grass extending in depth, as well a clear blue sky located above the horizon. The central fixation was positioned along the depth-axis between the nearest and farthest set of blocks at 14 vm (9.5°) in front of the viewer.

In the 3D-One Condition (see Figure 1c), a garden bed extending along a simulated depth axis appeared on either side of central fixation. The horizontal distance from the central fixation to the middle of the garden bed's inner edge measured 3.05 vm (5.8°). The garden bed's dimensions measured 6 vm horizontally (11.7°, at the bed's z-axis midpoint) and 16 vm along the depth axis (10.6°, at the bed's x-axis midpoint).

In the 3D-Two Condition (see Figure 1d), two garden beds each extending along a simulated depth axis appeared on either side of central fixation. The horizontal distance from the central fixation to the middle of the farther garden bed's inner edge (i.e., the garden bed located higher on the projection screen) measured 3.1 vm (4.1°). This farther garden bed measured 6.1 vm (7.9°) horizontally and 7 vm (3°) along the depth axis but vertically on the projection screen. The horizontal distance from the central fixation to the middle of the nearer garden bed's inner edge (i.e., the garden bed located lower on the projection screen) measured 3.1 vm (7.7°). This nearer garden bed measured 6.1 vm (12.9°) horizontally and 2.5 vm (4.9°) along the depth axis but vertically on the projection screen. There was a ground space gap of 6.5 vm (2.7°) along the depth axis between the far and near garden bed on either side.

The cue and target appeared in either ends (far/near ends in 3D or top/bottom end in 2D) of these garden bed wells, with matching retinal properties in the 2D and 3D conditions. The appearances of the cue or target were signalled by the luminance change (an onset of a shadow or light patch, respectively) of either end of the elongated garden bed well. The luminance change formed a gradient of luminance difference with the greatest contrast close to the end of the bed, reducing gradually until diminishing at the middle of the garden bed well. This arrangement was made to maximize the impression that the elongated garden bed was one object, and the cue and target were parts of that object. Both cues and targets possessed a constant on-screen size and horizontal eccentricity to equate their visual saliency at different depths or different on-screen elevations.

The cue was a black shadow patch, followed by a target patch of light (both measuring 1.15° x 0.65°). In the 2D condition, the cues and targets appeared in the bottom or top elevation. The bottom elevation was 9 vm (7.2°) and the top elevation 16.4 vm (14.6°) from origin (or bottom of the projection screen). In the 3D condition, the cues and targets appeared either in the near or far depth plane. The near depth plane was 9.25 vm (7.2°) and the far depth plane 25.25 vm (14.6°) away from the origin point ($x = 0$, $y = 4$, $z = 0$). The cue and target always appeared at the same horizontal retinal eccentricity in both 2D and 3D conditions (9.4°). The vertical retinal distance between the central car and each depth plane/elevation was 2.8°.

A short beep sound was made every time the cue or target appeared. A quack-like sound was made every time the participant made an anticipatory error or made a mistake in detecting the side at which the target was located. During a trial, participants were instructed to press the left flap on the backside of a steering wheel when the target appeared on the left side and the right flap on the backside of a steering wheel when the target appeared on the right side.

Procedure

Participants began a trial by fixating on the central fixation point (i.e., the red car). After 1500 ms, a cue appeared for 50ms. Following a CTOA that randomly ranged from 750 – 1050 ms, the target appeared for 100 ms. During a trial, participants were instructed not to make any eye movements. The participants were told to press the left or right flap located on the backside of a steering wheel device for a target that appeared on the left or right side of the display, respectively. A new trial began 1500 ms after the participant's response. All trials except catch trials required a response and did not advance without a manual input from the participant. A trial was interrupted if an anticipatory or localization error was made, which was followed by the 1500 ms intertrial interval before a new trial began.

Design

This was a within-subjects design that had 864 experimental trials over two, 1-hour sessions. One session involved both the 2D and 3D conditions with one garden bed on either side of fixation whereas the other session featured two garden beds on either side. The order of sessions was counterbalanced across participants. Each participant received 12 practice trials. One block contained 36 experimental trials, of which 4 (11.1% of trials) were catch trials. Each block had an equal number of trials combining the following randomized within-block factors: cue side (left versus right), target side (left versus right), target depth (near versus far), cue depth (near versus far), and perspective (3D versus 2D). Trials were presented in a random order. An optional break was granted after every block.

In each perspective condition, the cues and targets appeared either in the same garden bed or in either of the two garden beds on either side of fixation. When the cue and target appeared at different depths in the 3D condition, the depth plane switches included: far-cue to near-target and near-cue to far-target. The design in the 2D condition was the same as that in the 3D condition, except that cues and targets now appeared within 2D elevations rather than 3D depth planes. In the 2D condition, the elevation switches included: top-cue to bottom-target and bottom-cue to top-target. For brevity, both 3D depth (far//near) and 2D elevation (top//bottom) will be referred to in terms of depth.

Results

Reaction times longer than 1000 msec and shorter than 200 msec, as well as incorrect localization responses and anticipatory responses, were counted as errors (1.03% of experimental trials). Outlier data (3% of experimental trials) were also removed on either side of the reaction time distribution using a highly conservative threshold criterion of three times the mean absolute deviation of the median of raw reaction time data for each participant (Leys et al., 2013). Altogether, approximately 4.03% of the experimental trials was removed from further analysis. See Tables 1a and 1b for the tables of reaction time values for conditions featuring one or two garden beds on either side of fixation, respectively. Also see Figures 3a and 3b for an overview of IOR values across target depth and number of garden bed conditions.

Table 1a. Reaction time, IOR, and depth effect values for one garden bed on either side of fixation

One garden bed on either side of fixation

Independent Variables — Dependent Variables

Target Depth	Perspective	Depth or Elevation	Side (RT in ms); [SE in ms]	IOR (ms); SE [ms]	IOR difference between same vs different depth/elevation (ms); SE [ms]	Depth Effect: IOR difference with 2D effect removed (ms); SE [ms]
NEAR	3D	SAME	SAME (488.9); [14.9]	12.3; [2.2]	0.9; [3.0]	-2.5; [3.7]
			DIFF (476.6); [15.6]			
		DIFF	SAME (483.2); [15.7]	11.4; [2.0]		
			DIFF (471.8); [15.2]			
	2D	SAME	SAME (490.8); [14.4]	16.7; [2.1]	3.4; [2.5]	
			DIFF (474.1); [15.2]			
		DIFF	SAME (481.5); [14.9]	13.3; [2.5]		
			DIFF (468.1); [15.2]			
FAR	3D	SAME	SAME (495.6); [15.6]	17.2; [2.4]	7.0; [3.0]	-0.6; [3.6]
			DIFF (478.3); [15.4]			
		DIFF	SAME (486.1); [15.6]	10.2; [2.4]		
			DIFF (476); [15.6]			
	2D	SAME	SAME (492.7); [14.8]	18.4; [2.4]	7.6; [2.4]	
			DIFF (474.2); [14.7]			
		DIFF	SAME (482.3); [14.9]	10.8; [1.8]		
			DIFF (471.5); [15.2]			

Table 1b. Reaction time, IOR, and depth effect values for two garden beds on either side of fixation

Two garden beds on either side of fixation

Independent Variables — Dependent Variables

Target Depth	Perspective	Depth or Elevation	Side (RT in ms); [SE in ms]	IOR (ms); SE [ms]	IOR difference between same vs different depth/elevation (ms); SE [ms]	Depth Effect: IOR difference with 2D effect removed (ms); SE [ms]
NEAR	3D	SAME	SAME (490.9); [15.1]	12.5; [2.4]	10.2; [2.8]	7.0; [3.1]
			DIFF (478.4); [15.6]			
		DIFF	SAME (476.2); [15.1]	2.3; [2.4]		
			DIFF (473.9); [15.6]			
	2D	SAME	SAME (495.7); [14.6]	15.9; [2.6]	3.2; [2.7]	
			DIFF (479.8); [15.7]			
		DIFF	SAME (484.3); [14.9]	12.7; [1.8]		
			DIFF (471.7); [15.2]			
FAR	3D	SAME	SAME (497.3); [15.6]	14.3; [2.4]	5.0; [2.8]	-4.0; [3.6]
			DIFF (483); [15.9]			
		DIFF	SAME (486.9); [15.7]	9.3; [1.5]		
			DIFF (477.6); [15.7]			
	2D	SAME	SAME (501); [14.8]	19.1; [2.6]	9.0; [3.2]	
			DIFF (481.9); [15.3]			
		DIFF	SAME (485.2); [15.2]	10.1; [2.4]		
			DIFF (475.1); [15.3]			

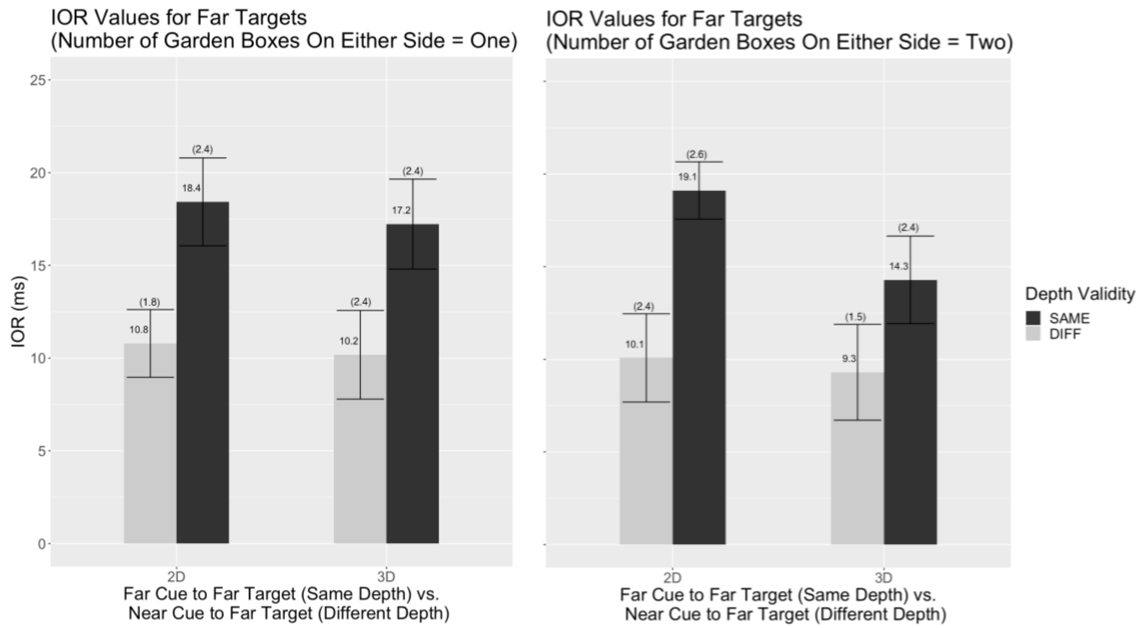


Figure 3a. IOR values for far targets (top elevation targets) for 3D and 2D conditions across depth validity conditions comparing settings with one garden bed versus two garden beds on either side of fixation.

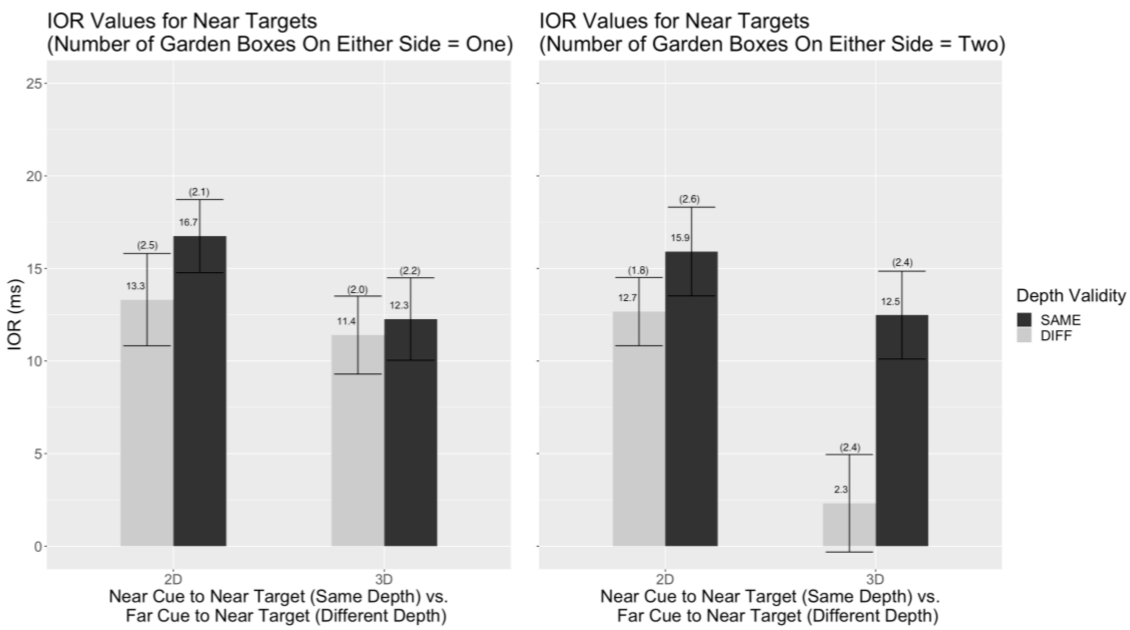


Figure 3b. IOR values for near targets (bottom elevation targets) for 3D and 2D conditions across depth validity conditions comparing settings with one garden bed versus two garden beds on either side of fixation.

The depth effect was defined by Equation 2 below. Mean depth effect values for each participant were entered into a 2 (target depth: near versus far) x 2 (number of garden beds on either side of fixation: one versus two) analysis of variance (ANOVA). The main effects of target depth, $F(1, 38) = 0.17$, $p = .69$, $\eta_p^2 = 4.34e\text{-}03$, and number of garden beds, $F(1, 38) = 1.20$, $p = .28$, $\eta_p^2 = .03$, were not significant. However, the interaction between target depth and number of garden beds was significant, $F(1, 38) = 5.5$, $p < .05$, $\eta_p^2 = .13$.

$$Depth\ effect = \left[IOR_{3D(Same\ Depth)} - IOR_{3D(Different\ Depth)}\right] - \left[IOR_{2D(Same\ Elevation)} - IOR_{2D(Different\ Elevation)}\right],$$

where $$IOR_{3D(Same\ Depth)} = (RT_{same\ side\ \&\ depth} - RT_{diff\ side\ \&\ same\ depth});$$

$$IOR_{3D(Different\ Depth)} = (RT_{same\ side\ \&\ different\ depth} - RT_{diff\ side\ \&\ diff\ depth});$$

$$IOR_{2D(Same\ Elevation)} = (RT_{same\ side\ \&\ same\ elevation} - RT_{diff\ side\ \&\ same\ elevation});$$

$$IOR_{2D(Different\ Elevation)} = (RT_{same\ side\ \&\ diff\ elevation} - RT_{diff\ side\ \&\ diff\ elevation})$$ Eq. 2

To examine the source of the above interaction, two final planned comparisons were performed. The depth effect values (see Figure 4) were compared between one versus two garden bed conditions in two separate one-way ANOVAs, one for near targets and one for far targets. The main effect of the number of garden beds was significant, $F(1, 38) = 6.3$, $p < .05$, $\eta_p^2 = .14$ for near targets, but not for far targets, $F(1, 38) = 0.95$, $p = .34$, $\eta_p^2 = .02$.

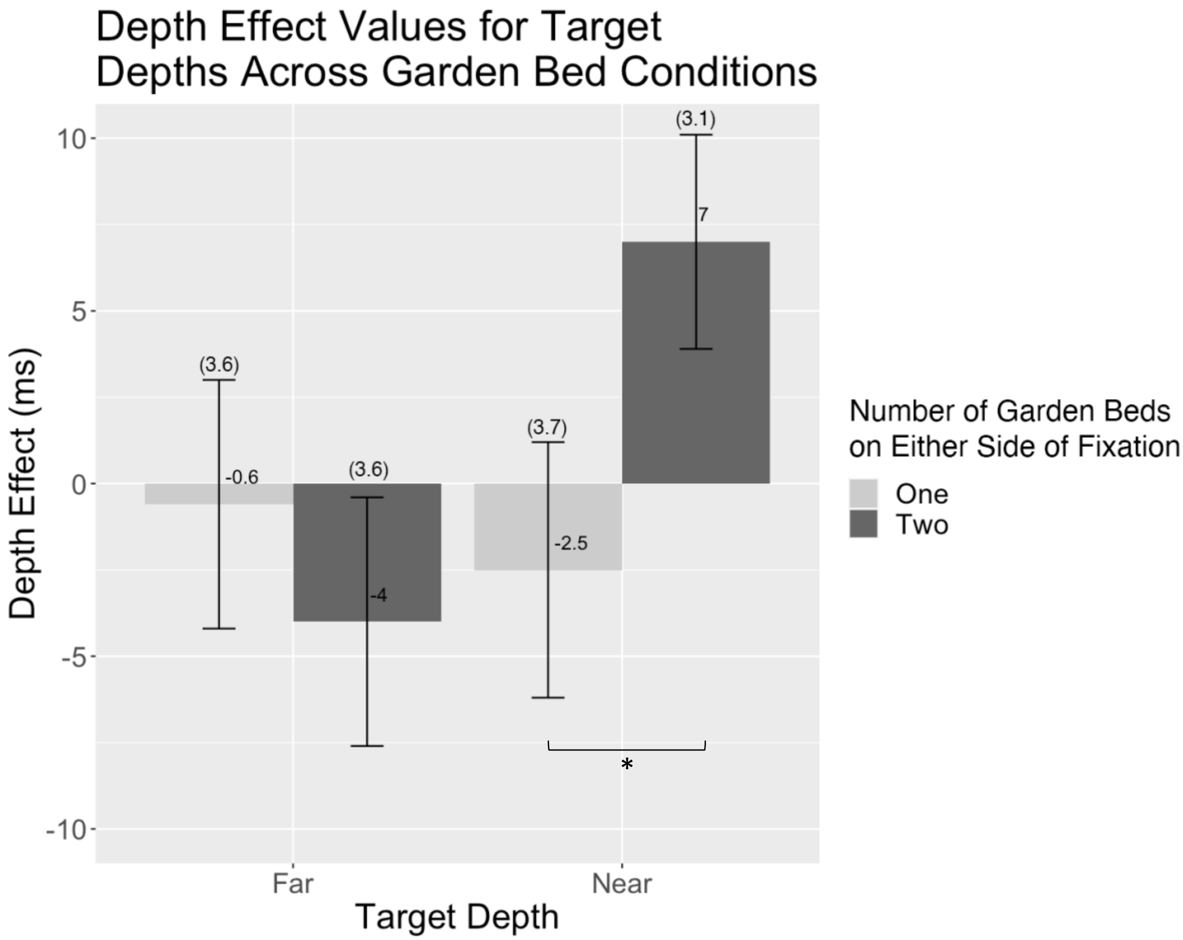


Figure 4. Depth effect values for far (left portion) and near (right portion) target depths for one and two garden beds on either side of fixation. Standard errors are in parentheses.

Discussion

The present study was conducted to answer a follow-up question concerning the depth-specificity of IOR described in Chapter 2, where the magnitude of IOR reduction when the cue and target appeared in different depths in the 3D displays was greater than in the corresponding 2D displays, but only when attention was oriented from far-to-near space. In Chapter 2, when the cue and target appeared in different depths, they also appeared within the confines of different placeholders. This design therefore made it impossible to differentiate the contribution of spatial location-based and object-based contributions to the IOR effect. To explore the object-based effects, the present study ensured that there was a condition in which cues and targets could appear at different depths within a single object. By comparing the data from this condition to one where the cues and targets appeared in different depth planes and different objects, the magnitude of object-based IOR could be isolated.

The results from the condition in which two garden beds lay on either side of fixation constitute a replication of the key result from Chapter 2. When there were two garden beds on either side of fixation in the 3D condition, the magnitude of IOR was the smallest for the far-to-near depth switch (2.3 ms). The resulting depth effect (difference between IOR for 3D and 2D conditions) was largest for this condition (7.0 ms). However, the corresponding depth effect when one garden bed was situated on either side of fixation was not significant (-2.5 ms). In the condition with one object box on either side of fixation, the inhibitory tag was likely placed on the entire object, so that spatial-based and object-based effects both contributed to IOR. The inhibition was present within the entire object, and the near advantage seen in Chapter 2 was no longer present. Graphs summarizing the depth effects for far and near targets with one or two garden beds on either side of fixation are summarized in the left and right portions of Figure 4, respectively.

Under certain conditions IOR may contain a location-based component and an object-based component. The location-based component of IOR is influenced by the Euclidian distances between (x, y, z) coordinates of cues and targets, with the resulting differences in the coarse spatial distances contributing to difference in the magnitude of IOR (Casagrande et al., 2012). The location-based component of IOR has been said to be more robust, being consistently present across different experimental design configurations (B. S. Gibson & Egeth, 1994; Krüger & Hunt, 2013; Pilz et al., 2012). This study did not feature a 3D scene without objects, so any claims about location-based IOR are not able to be made. However, when a cue and target were spatially separated and appeared on two separate objects, a depth-specific IOR effect was generated. This depth-specific IOR effect was not observed when cues and target appeared within a single object. Thus, it is possible that the depth-specific IOR observed in this current study was coded in coarse spatial coordinates. In previous literature, IOR has been observed to increase when cues and targets are located at different depths but still bounded by the perimeter of a single 3D object compared to when the cues and targets are located at different depths within two different 3D objects (Bourke et al., 2006; Jordan & Tipper, 1999; Leek et al., 2003; Tipper et al., 1999; Weaver et al., 1998). It could very well be that participants in the present study allocated their attention to the entire object on either side of fixation as opposed to the individual locations within the single object.

The contribution of object membership can also be examined in our 2D conditions. One would argue that if object-based attention played a role (on top of the location-based IOR), one should see greater IOR values in the one-object condition (i.e., Full condition) compared to the two-object condition (i.e., Split condition) (Jordan & Tipper, 1999; Leek et al., 2003). Interestingly, in the 2D scenes of this current study, the IOR values across both 2D conditions were quite comparable. When the cue appeared higher on the screen than the target, IOR was 13.3 ms and 12.7 ms for one and two objects, respectively. When the cue appeared lower on screen than the target, IOR was 10.8 ms and 10.1 ms for one and two objects, respectively. It is not surprising that object-based attention contributed little in the 2D situation, as object-based effects of very small magnitudes have been reported in previous 2D literature (Krüger & Hunt, 2013; Pilz et al., 2012). Yet, in the 3D condition, the object-based inhibition was strong enough to override the IOR reduction when orienting across depths. One speculation could be that in the 2D condition, the location-based contribution is increasingly dominant, but in the 3D condition, the attention spread in depth is very sensitive to object membership, with the IOR increasing with cues and targets occurring in the same object.

The rich depth information in this study's 3D condition may have strongly contributed to the presence of the depth-specific IOR effect in the condition with two garden beds on either side of fixation. No IOR research to date has suggested that pictorial depth features may contribute to depth-specific IOR. However, a handful of studies have suggested that providing participants with a stereoscopic view would enable an easier distinction between the near and far depth planes, contributing to a larger magnitude of IOR within as opposed to between depth planes (Bourke et al., 2006; Casagrande et al., 2012). Some claims have been made about this depth-specific IOR effect not existing in the presence of pictorial depth cues like relative size and eccentricity, in addition to the global linear perspective of a scene (Casagrande et al., 2012; A. Wang et al., 2016). No study to date, however, has reported results about the effect that changing more local objects associated with cues and targets may have on differences in IOR. In the 3D condition of this study, for instance, the garden beds contained rich pictorial depth cues including linear perspective, relative size, shading, and texture differences, which altogether seemed to aid in the distinct segmentation of far versus near depth space.

Conclusion

Object-based effects of IOR were shown to exist in a 3D setting composed only of pictorial depth cues. Future studies can investigate how this object-based IOR would be affected by increasing or decreasing the realism of depth features. Furthermore, one unanswered question in this study was whether it was the world-centered or viewer-centered coordinate differences between cue and target that produced the depth-specific IOR effect. The farther cue/target always appeared on the northernmost garden bed and the nearer cue/target on the southernmost garden bed. Understanding whether the depth-specific IOR occurs due to the world-centered separation between cue and target or the distance between these stimuli and the viewer is a separate topic to explore. Answering this question would provide insight into whether observers attend to cues and targets relative to their own position in space or relative to each other. This is a question that Chapter 4 attempts to answer.

References

Atchley, P., & Kramer, A. F. (2001). Object and space-based attentional selection in three-dimensional space. *Visual Cognition*, *8*(1), 1–32. https://doi.org/10.1080/13506280042000009

Avrahami, J. (1999). Objects of attention, objects of perception. *Perception & Psychophysics*, *61*(8), 1604–1612. https://doi.org/10.3758/BF03213121

Bennett, P. J., & Pratt, J. (2001). The spatial distribution of inhibition of return. *Psychological Science*, *12*(1), 76–80. https://doi.org/10.1111/1467-9280.00313

Bourke, P. A., Partridge, H., & Pollux, P. M. J. (2006). Additive effects of inhibiting attention to objects and locations in three-dimensional displays. *Visual Cognition*, *13*(5), 643–654. https://doi.org/10.1080/13506280544000309

Casagrande, M., Barbato, M., Mereu, S., Martella, D., Marotta, A., Theeuwes, J., & Collinson, S. (2012). Inhibition of return: A "depth-blind" mechanism? *Acta Psychologica*, *140*, 75–80. https://doi.org/10.1016/j.actpsy.2012.02.011

de Gonzaga Gawryszewski, L., Riggio, L., Rizzolatti, G., & Umiltá, C. (1987). Movements of attention in the three spatial dimensions and the meaning of "neutral" cues. *Neuropsychologia*, *25*(1A), 19–29.

Gibson, B. S., & Egeth, H. (1994). Inhibition of return to object-based and environment-based locations. *Perception & Psychophysics*, *55*(3), 323–339. https://doi.org/10.3758/bf03207603

Jordan, H., & Tipper, S. P. (1999). Spread of inhibition across an object's surface. *British Journal of Psychology*, *90*(4), 495–507. https://doi.org/10.1348/000712699161576

Klein, R. (2000). Inhibition of return. *Trends in Cognitive Sciences*, *4*(4), 138–147.

Kreidler, S. M., Muller, K. E., Grunwald, G. K., Ringham, B. M., Coker-Dukowitz, Z. T., Sakhadeo, U. R., Barón, A. E., & Glueck, D. H. (2013). GLIMMPSE: Online power computation for linear models with and without a baseline covariate. *Journal of Statistical Software*, *54*(10), i10.

Krüger, H. M., & Hunt, A. R. (2013). Inhibition of return across eye and object movements: The role of prediction. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(3), 735–744. https://doi.org/10.1037/a0030092

Leek, E. C., Reppa, L., & Tipper, S. P. (2003). Inhibition of return for objects and locations in static displays. *Perception & Psychophysics*, *65*(3), 388–395. https://doi.org/10.3758/bf03194570

Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, *49*(4), 764–766. https://doi.org/10.1016/j.jesp.2013.03.013

List, A., & Robertson, L. C. (2007). Inhibition of return and object-based attentional selection. *Journal of Experimental Psychology. Human Perception and Performance*, *33*(6), 1322–1334. https://doi.org/10.1037/0096-1523.33.6.1322

Maylor, E. A., & Hockey, R. (1985). Inhibitory component of externally controlled covert orienting in visual space. *Journal of Experimental Psychology: Human Perception and Performance*, *11*(6), 777–787. https://doi.org/10.1037/0096-1523.11.6.777

Pilz, K. S., Roggeveen, A. B., Creighton, S. E., Bennett, P. J., & Sekuler, A. B. (2012). How prevalent Is object-based attention? *PLoS ONE*, *7*(2). https://doi.org/10.1371/journal.pone.0030693

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*(1), 3–25. https://doi.org/10.1080/00335558008248231

Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. *Attention and Performance X*, *32*, 531–556.

Pratt, J., & Abrams, R. A. (1995). Inhibition of return to successively cued spatial locations. *Journal of Experimental Psychology. Human Perception and Performance*, *21*(6), 1343–1353. https://doi.org/10.1037//0096-1523.21.6.1343

Soto, D., & Blanco, M. J. (2004). Spatial attention and object-based attention: A comparison within a single task. *Vision Research*, *44*(1), 69–81. https://doi.org/10.1016/j.visres.2003.08.013

Tipper, S. P., Driver, J., & Weaver, B. (1991). Object-centred inhibition of return of visual attention. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, *43*(2), 289–298. https://doi.org/10.1080/14640749108400971

Tipper, S. P., Jordan, H., & Weaver, B. (1999). Scene-based and object-centered inhibition of return: Evidence for dual orienting mechanisms. *Perception & Psychophysics*, *61*(1), 50–60. https://doi.org/10.3758/BF03211948

Tipper, S. P., Rafal, R., Reuter-Lorenz, P. A., Starrveldt, Y., Ro, T., Egly, R., Danzinger, S., & Weaver, B. (1997). Object-based facilitation and inhibition from visual orienting in the human split-brain. *Journal of Experimental Psychology. Human Perception and Performance*, *23*(5), 1522–1532. https://doi.org/10.1037//0096-1523.23.5.1522

Tipper, S. P., Weaver, B., Jerreat, L. M., & Burak, A. L. (1994). Object-based and environment-based inhibition of return of visual attention. *Journal of Experimental Psychology. Human Perception and Performance*, *20*(3), 478–499. https://doi.org/10.1037/0096-1523.20.3.478

Wang, A., Liu, X., Chen, Q., & Zhang, M. (2016). Effect of different directions of attentional shift on inhibition of return in three-dimensional space. *Attention, Perception, & Psychophysics*, *78*(3), 838–847. https://doi.org/10.3758/s13414-015-1055-6

Wang, A., Yue, Z., Zhang, M., & Chen, Q. (2015). Interaction between spatial inhibition of return (IOR) and executive control in three-dimensional space. *Experimental Brain Research*, *233*(11), 3059–3071. https://doi.org/10.1007/s00221-015-4374-x

Wang, Z., & Klein, R. M. (2010). Searching for inhibition of return in visual search: A review. *Vision Research*, *50*(2), 220–228. https://doi.org/10.1016/j.visres.2009.11.013

Watson, S. E., & Kramer, A. F. (1999). Object-based visual selective attention and perceptual organization. *Perception & Psychophysics*, *61*(1), 31–49. https://doi.org/10.3758/bf03211947

Weaver, B., Lupiáñez, J., & Watson, F. L. (1998). The effects of practice on object-based, location-based, and static-display inhibition of return. *Perception & Psychophysics*, *60*(6), 993–1003. https://doi.org/10.3758/BF03211934

Chapter 4: Spatial Frame of Reference for Inhibition of Return in 3D Space

Abstract

The previous studies in Chapters 2 and 3 found evidence for a depth-specific inhibition of return (IOR): IOR was reduced when a cue appeared at a farther location in depth than the target compared to when a cue appeared in the same depth plane as the target. However, it is not clear what frame of reference was involved in this depth-specific IOR since viewer-centered and world-centered frame of references specified the same depth relation. In this current study, we investigated the spatial frame of reference for the depth-specific IOR effect by creating a discrepancy between depth information coded in viewer-centered and world-centered coordinates. We tested target detection in a spatial cueing paradigm in a simulated three-dimensional (3D) space illustrated through pictorial depth cues (such as linear perspective and optic flow). In the experimental conditions, the cue and target appeared at different viewer-centered distances (Experiment 1 and 2) or different world-centered depth planes (Experiment 2). This manipulation was achieved by simulating forward self-motion during a trial, and consequently changing the viewpoint in depth between cue onset and target onset.

In the first experimental condition, the cue and target appeared at different viewer-centered distances but in the same world-centered depth plane. The IOR in this condition was not less than the IOR in the control condition in which the cue and target appeared in the same world depth plane and at the same viewer-centered distance. In the second experimental condition, the cue and target appeared in different world-centered depth planes but at the same viewer-centered distance. IOR in this condition was much less than IOR in the conditions where the cue and target appeared within the same world depth plane (regardless of whether the viewer-centered distance was the same or different

between cue and target onset). Collectively, these results suggest that similar to what is found in 2D space, IOR in 3D space involves inhibitory tagging of areas mapped in world-centered coordinates rather viewer-centered coordinates.

Introduction

In studies of exogenous spatial cueing, an observer's reaction time to a target is typically faster if that target appears within 300 ms following a cue appearing in the same location. Once more than 300 ms passes between cue and target onset, a viewer's reaction becomes inhibited, or slowed down for cued locations compared to uncued locations. This reaction delay is pivotal in forming the inhibition of return (IOR), calculated as the reaction time difference between targets appearing in the same versus different location as the cue. IOR is thought to reflect a process that facilitates foraging, with neural mechanisms of attention favouring novel environmental locations over locations already observed (R. Klein, 2000; Posner & Cohen, 1984).

In 2D space, the gradient of IOR peaks at the cue location and monotonically degrades as a function of cue-to-target distance. IOR decreases as a function of cue-to-target distance because reaction times for targets appearing farther from the original cue are lowered drastically (Bennett & Pratt, 2001; J. E. T. Taylor et al., 2015). IOR is thought to involve the inhibitory tagging of areas mapped in world-centered coordinates, largely because it has been shown that IOR exists without needing to make a saccade between cue and target onset (M. D. Hilchey et al., 2012; Krüger & Hunt, 2013; Pertzov et al., 2010; Posner et al., 1985). However, this is not to say that viewer-centered retinotopic coordinates do not have some role to play in the manifestation of IOR.

The retinotopic coordinate system is likely used when the delay between cue and target (i.e., the cue-target onset asynchrony or CTOA) is approximately 200 ms (Golomb et al., 2008). In studies examining the spatial frame of reference of IOR involving eye movements, participants typically are told to fixate at centre, covertly attend to a peripheral cue, and then make a saccade before the onset of a target that could appear either at the same spatial (world-centered) location or same retinotopic (viewer-centered) location as the cue. If the CTOA is 100-200 ms, manual detection performance is typically facilitated when the target matches the cue location in retinal coordinates (Golomb et al., 2008; Posner & Cohen, 1984). If the CTOA is greater than 300 ms, manual detection or saccadic performance is typically inhibited when the target appears in the same spatiotopic (rather than retinotopic) coordinates as the cue (M. D. Hilchey et al., 2012; Maylor & Hockey, 1985; Pertzov et al., 2010; Satel et al., 2012). Furthermore, this spatiotopic IOR occurs either before or during the initiation of a saccade, suggesting that a world-centered reference frame is likely dissociated from a viewer-centered reference frame and used for attentional orienting at larger CTOAs (Abrams & Pratt, 2000; Astle, 2009). At larger CTOAs, inhibition has shown to increase while facilitation decreases when both cue and target are presented at greater peripheral retinal eccentricities, independent of cortical magnification (Bao et al., 2013; Feng & Spence, 2017). To say that IOR is purely world-centered or spatiotopic is misleading, however, as numerous studies have indicated that a viewer-centered or retinotopic IOR also exists, just not at the same magnitudes as spatiotopic IOR (Krüger & Hunt, 2013; Maylor & Hockey, 1985; Pertzov et al., 2010; Satel et al., 2012).

The gradient of visuospatial attention in depth may represent that of a horizontal ellipse stretched along its minor axis from the viewer's head: attention may be strongest within near space extending into the horizontal periphery, declining with increasing height and depth distance from the viewer (Andersen & Kramer, 1993; de Gonzaga Gawryszewski et al., 1987a). Visuospatial cueing experiments measuring target detection or localization along the depth axis have been shown that observers are quicker to respond to targets nearer to them, particularly if they are closer to the viewer than the cue (Miura et al., 2002; Reppa et al., 2010; A. Wang et al., 2016). Song et al. (2021) found the attentional preference or facilitation for locations and objects in a space nearer to the viewer in a virtual environment that employed pictorial depth cues such as linear perspective and optic flow. Participants experiencing visually simulated forward motion were required to localize peripheral targets that appeared at various horizontal eccentricities and depths. Results showed that detection times were faster and more accurate for nearer targets than farther targets at all eccentricities, but particularly so at greater eccentricities. This default level of attention for near space was partly attributed to the relative size of the squares on checkerboard walls. Thus, in a scene with helpful depth information, the observer's detection speed is likely facilitated when a location or object is nearer, with inhibition prevailing for areas farther from the observer.

Studies measuring the effect of depth on IOR have found that that IOR is affected by far-to-near depth switches, but not near-to-far switches, even when the retinal coordinates of the cue and target are held constant (A. Wang et al., 2016). As shown in Chapters 2 and 3 of this dissertation, IOR was reduced drastically for far-to-near depth switches when cues and targets appeared in different world-centered depth planes. This result is consistent with the notion that the level of attention in 3D space is greatest near the observer, with inhibition more easily spreading from near to far space than far to near space. Nevertheless, no research articles to date have measured whether IOR is biased towards the world-centered or viewer-centered reference frame in 3D space when retinal coordinates are controlled.

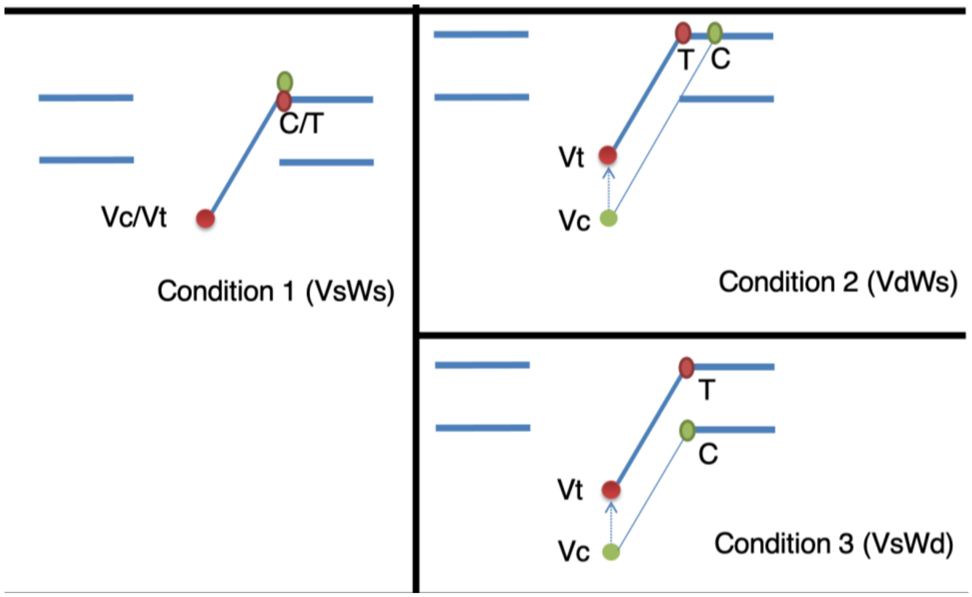
Is IOR most influenced by depth differences between cue and target (i.e., world-centered coordinates) or depth differences among cue, target, and viewer (i.e., viewer-centered coordinates)? The experiments of Chapters 2 and 3 do not address this question because world-centered and viewer-centered reference frames were not controlled. For example, when the cue appeared at the northernmost depth plane it also appeared at the most distant depth plane from the viewer. Also, when the cue appeared on the southernmost depth plane it also appeared at the nearest depth plane. In this case, is the world-centered depth between cue and target more task relevant or is there also an element of egocentrism that dictates how a viewer visuospatially responds to areas of interest located along the depth axis? Previous research has shown that the spatiotopic (world-centered depth) reference frame is dominant in guiding a viewer's spatial attention in 2D space (Astle, 2009; M. D. Hilchey et al., 2012; Maylor & Hockey, 1985; Pertzov et al., 2010). Our experiments aimed to examine whether a viewer-centered or world-centered reference frame might strongly contribute to IOR in 3D space.

Current Study

Studies of attention modulation in 2D space have been shown that IOR involves the inhibitory tagging of cued locations primarily mapped in world-centered coordinates, rather than viewer-centered retinal locations. To our knowledge, no studies have explored the spatial frame of reference of IOR in 3D space. Results from Chapters 2 and 3 showed that, for a stable viewpoint, IOR for peripheral targets was drastically reduced for far-cue to near-target depth switches. However, it is not clear whether the depth difference between cue and target that caused the reduction of IOR (i.e., the depth-specific IOR) was encoded in a viewer-centered or world-centered frame of reference. In fact, the expression of "far" versus "near" was based on a language assuming a viewer-centered frame of reference for the sake of simplicity. However, the switch of attention can also be defined in world-centered coordinates (i.e., from a north to south depth plane, assuming the viewpoint is at the southernmost location facing a northern direction). In other words, the reduction of IOR found in Chapters 2 and 3 could be caused by a far-cue to near-target viewer-centered depth switch or by a north-cue to south-target world-centered depth switch. To identify which of the two frames of references lead to the depth-specific IOR found in Chapters 2 and 3, the current study measured IOR in scenarios where the depth relation between cue and target differed in the two frames of references.

In the current study, observers viewed cues and targets presented in a 3D space composed of pictorial depth information (mostly linear perspective, ground texture, and optic flow). The cue and target appeared on one or two walls vertically stretched and oriented along the horizontal plane (i.e., the x-axis). However, the viewpoint of the observer could be either in the same or different depth position along the z-axis when the cue and target appeared. The observer's viewer-centered distance to the walls during the time interval between the appearance of cue and then appearance of target could be the same or different due to visually simulated self-motion of the observer's viewpoint.

In both experiments in the current study, the control condition (Condition 1, illustrated in Figure 1a as a bird's eye view of the design) involved a scenario where the cue and target appeared in the same world-centered depth plane and same distance to the viewer. The experimental conditions involved a scenario where the cue and target appeared either in different world-centered depth planes or at different viewer-centered distances. In Chapters 2 and 3, the cue and target appeared in different depth planes in terms of both world-centered coordinates and viewer-centered coordinates. In the current study, however, the cue and target appear in different depth planes only in one frame of reference. Specifically, in Condition 2 of Experiment 1 (Figure 1b), the cue and target appeared in the same depth plane defined by world coordinates, but the viewer position was different along the z-axis between cue and target onset. In Experiment 2, in addition to the replication of conditions 1 and 2 in Experiment 1, a scenario was tested where the cue and target appeared in different depth planes, with the observer's position in depth changing between cue and target onset. Consequently, the viewer-to-cue and viewer-to-target distance was matched in this condition (Condition 3, Figure 1c). In other words, the cue and target appeared in two different world-centered depth planes with the extent of depth separation matching the observer's distance of travel.



Condition	CF to cue	CF to target	Viewer to cue	Viewer to target
1, VsWs	10 vm	10 vm	32 vm	32 vm
2, VdWs	32 vm	10 vm	54 vm	32 vm
3, VsWd	10 vm	10 vm	32 vm	32 vm

Figure 1. Illustration of the spatial relations among locations of the cue, target and viewpoint depicting a birds-eye view of the current study's simulated 3D scenes. The four thick horizontal lines represent the walls which served as placeholders for the cue and target. The lines between the viewpoints and the cue/target were drawn for illustration purposes and were not shown in the actual stimulus. In Experiment 1, Condition 1, two levels of viewer-to-target distances were tested ("middle" and "near", but only one is illustrated in this figure). In Experiment 1, only Conditions 1 and 2 were tested. In Experiment 2, all three conditions were tested as illustrated. Labels in the figure: C: cue position; T: target position; Vc: Viewpoint at the moment of cue appearance; Vt: Viewpoint at the moment of target appearance; CF: central fixation (lead red vehicle).

Condition 1 (VsWs): cue and target both appear at the same distance relative to stationary viewer and same world depth plane
Condition 2 (VdWs): the viewer-to-cue and viewer-to-target distance is different, but the cue and target both appear within the same world depth plane
Condition 3 (VsWd): the viewer-to-cue and viewer-to-target distance is the same, but the cue and target appear at a different world depth plane

Experiment 1

Results from Chapters 2 and 3 showed that for a stable viewpoint, the IOR for peripheral targets was reduced for far-cue to near-target depth switches. In Experiment 1, we created the same depth switch direction by having a greater viewer-centered distance for cue onset compared to target onset as the viewpoint moved from "far" to "near" depth planes relative to the wall displaying both cue and target (see Figure 1C, 1D). If IOR is reduced for the condition where the cue appears farther than the target relative to the viewer, compared to the control condition where the cue and target appear at the same viewer-centered and world-centered depths, then it could be said depth-specific IOR depends on viewer-centered coordinates. However, if the results do not show such a reduction, this would ultimately suggest that viewer-centered distances to the cue and target have minimal to no involvement in the generation of IOR.

Experiment 1 presented cues and targets in a virtual 3D environment. On each trial, the participant viewed a cue that appeared on a wall at a particular viewer-centered spatial position. The observer then underwent simulated forward motion through the 3D scene, and finally responded to a target that appeared on the same wall and eccentricity as the cue. In all conditions, the cue and target always shared the same world depth plane: they appeared on the same placeholder wall located at a single depth plane and therefore the world-centered cue-to-target depth difference was always zero. However, because the viewpoint changed during the trial, the viewer-centered cue-to-target depth difference was always non-zero. Thus, the observer always saw the cue from a viewer-centered distance that was greater than the viewer-centered target distance. In the control (static) condition, the viewer did not move and therefore saw the cue and target from the same viewer-centered and world-centered depth planes. By comparing IOR in the control and experimental conditions, it was possible to examine whether the viewer-centered frame of reference affected IOR. If IOR is larger in the control condition, then IOR likely depends on the viewer-centered frame of reference. Otherwise, if IOR does not differ in the dynamic and static conditions, then IOR likely does not depend on the viewer-centered frame of reference.

Methods

Participants

In total, data from 34 adults (18 – 24 years old, n =17 female) were analyzed. All participants were credited 1.0 research participation credit for their participation. An *a priori* minimum sample size of n = 21 was determined because it gave the study 95% power, as calculated by GLIMMPSE for the interaction between viewer-centered depth validity x viewer-to-target distance (Kreidler et al., 2013).

Stimuli and Apparatus

The simulation was created with Vizard 4.0 virtual reality software and back-projected onto a screen located in a dark tent free of external visual cues. Participants sat 150 cm in front of the screen, which measured 107.4 cm tall x 144.8 cm wide, with a steering wheel response device positioned comfortably in front of them.

The simulation consisted of a central lead vehicle (width = 2.4° x elevation = 1.8°) surrounded by a checkerboard wall measuring 21 virtual metres (vm) horizontally on either side. The viewer was positioned 25 vm behind and 1.82 vm above the lead vehicle.

A purple-white checkerboard square appeared as the cue (1.13° x 1.13°), followed by a black-and-white checkerboard circle target (diameter = 1.13°). The cue and target appeared at 24° horizontal eccentricity.

The two major conditions, same and different viewer-centered distance condition, consisted of two levels each. In the same viewer-centered distance condition, the cue and target always appeared at the same distance relative to the viewer; the cue and target both appeared 25 vm (middle distance) or 10 vm (near distance) from the viewer (Figure 2a, 2b). In the different viewer-centered distance condition, the cue and target appeared at different distances relative to the viewer. There were two types of different distance trials (Figures 2c & 2d). In one trial type (Figure 2c), the cue appeared 40 vm from the viewer (i.e., "far" distance) and, as the viewer moved forward in space, the target appeared 25 vm from the viewer (i.e., "middle" distance). In the second trial type (Figure 2d), the cue appeared 25 vm from the viewer (i.e., "middle" distance) and, as the viewer moved forward in space, the target appeared 10 vm from the viewer (i.e., "near" distance). At the beginning of a different distance trial, the viewer passively moved forward 15 vm before cue onset. For example, if the trial consisted of a far cue appearing at 40 vm from the viewer, the viewer would start at 55 vm from the placeholder wall, experience simulated motion for 15 vm, and then see the cue.

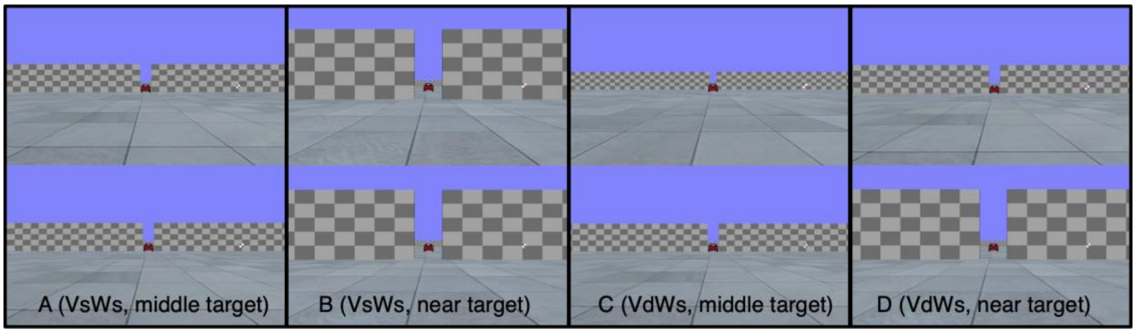


Figure 2. Stimuli in Experiment 1**.** The cue was a purple/white square and the target a black/white circle, both with checkerboard texture. The figure illustrates the conditions with the side validity being the same. In all conditions in this experiment, both cue and target appeared in the same depth plane demarcated by a placeholder wall. Two levels of viewer-to-target distance were varied: "middle" and "near". In the condition involving the cue and target appearing in the same viewer-centered and world-centered depths, the viewpoint was 25 vm from the cue and target in the middle condition (Condition 1, VsWs, **2A**) or 10 vm from the cue and target in the near condition (Condition 1, VsWs, **2B**). In the condition involving the cue and target appearing at different viewer-centered but same world-centered depths, the viewpoint was 40 vm from the cue and 25 vm from the target in the middle condition (Condition 2, VdWs, **2C**) or 25 vm from the cue and 10 vm from the target in the near condition (Condition 2, VdWs, **2D**).

For each of the four figure columns (A to D), the top and bottom images illustrate the view where the cue and target appeared, respectively. The red vehicle served as the point of fixation and was always 25 vm in front of the observer. Both the viewpoint and the vehicle remained stationary in Condition 1 and moved forward in Condition 2 during the time interval between cue onset and target onset.

When the viewer was 10 vm from the wall (near distance) the wall was located 16.3° from the bottom of the screen and measured 23.6° x 17.7°. When the viewer was 25 vm from the wall (middle distance), the wall was located 18.2° from the bottom of the screen and measured 24.9° x 7.1°. When the viewer was 40 vm from the wall (far distance), the wall was located 18.7° from the bottom of the screen and measured 25.2° x 4.5°.

The speed of the lead vehicle, linked to the viewer's speed and varied within each block, averaged 52.2 km/h. The lead vehicle's speed equalled the sum of three sine wave functions with frequencies of 0.033, 0.083, and 0.117 Hz, and amplitudes of 9.722, 3.889, and 2.778 km/h, respectively. The phase shifts of two sine components were generated randomly on each trial, and the lowest frequency sine component was chosen such that the sum of the phase shifts of all three sine components was zero. This sine-wave jitter was embedded to emulate a realistic driving environment since the speeds of vehicles in the real world never maintain a constant value.

A short beep sound was made every time the cue or target appeared. A quack-like sound was made every time the participant made a mistake in detecting the side of target onset. During a trial, participants were instructed to press the left flap on the backside of a steering wheel when the target appeared on the left side and the right flap on the backside of a steering wheel when the target appeared on the right side.

Procedure

Participants began a trial by fixating the central fixation point. After approximately 1000 ms, a cue appeared on the placeholder wall for 500 ms. Following a CTOA of 650 ms, the target appeared for 300 ms. Participants were instructed not to make eye movements during a trial, and to press a left flap on the backside of a steering wheel for a leftward target and a right flap on the backside of a steering wheel for a rightward target. After a participant's response was made, the display went blank until the next trial began after approximately 1000 ms.

Design

This was a within-subject design with 768 experimental trials. Each participant received 24 practice trials for each condition. An experimental run involved alternating between four blocks of 192 trials, with the order of same versus different viewer-centered distance conditions counterbalanced across participants. Trials were presented in a random order. An optional break was granted after every 24 trials. For each 24-trial experimental block, there was a randomized number of trials combining the following factors: cue horizontal location (left versus right), target horizontal location (left versus right), cue depth (far, middle), target depth (middle, near). In each block, there was an equal number of trials featuring the aforementioned factors, with viewer-centered distance validity (cue and target appear at same or different distances relative to viewer) counterbalanced between blocks.

Results

Reaction times longer than 1000 ms and shorter than 200 ms, as well as incorrect localization responses, were counted as errors (5.12% of experimental trials). Outlier data (3% of experimental trials) also were removed using a highly conservative threshold criterion from the mean absolute deviation of the median of raw reaction time data for each participant (Leys et al., 2013). Altogether, approximately 8% of experimental trials were removed from further analysis.

The average reaction time for a near target cued from a different side and different viewer-centered distance was the lowest (418.8 ms) compared to all other conditions. Reaction times were fastest for the near targets cued from a farther distance relative to the viewer's position. The reaction times for all conditions are plotted in Figure 3.

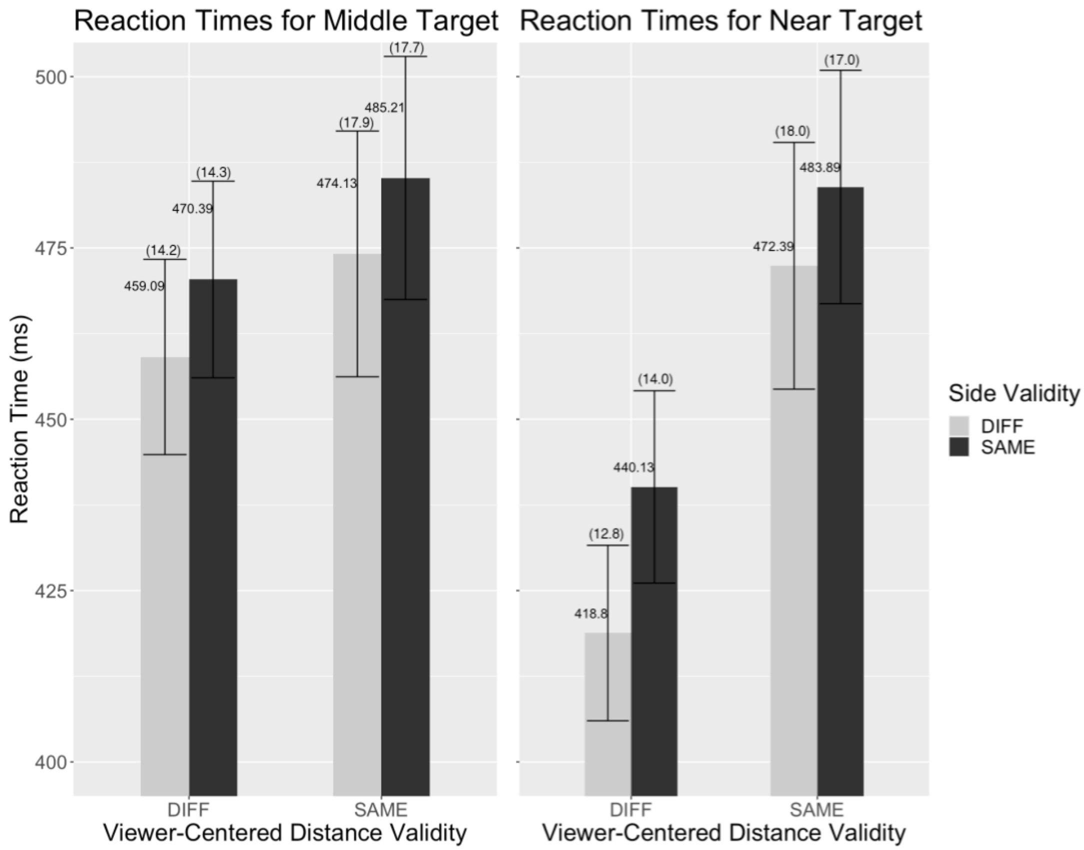


Figure 3. Reaction times in Experiment 1 as a function of side validity when cue and target appear at the same and different viewer-centered distances. Left graph: "middle" target; Right graph: "near" target.

IOR, calculated by subtracting RT on different side trials from RT on same side trials, is plotted in Figure 4. Mean IOR values for each participant were entered into a 2 (viewer-centered distance validity: same versus different) x 2 (viewer-to-target distance: middle versus near) repeated-measures analysis of variance model (ANOVA). The main effect of viewer-to-target distance was significant, $F(1, 33) = 8.0$, $p < .01$, $\eta_p^2 = .19$, represented by higher IOR values for the near distance (16.4 ms) rather than middle distance condition (11.55 ms). The main effect of viewer-centered distance validity was not, $F(1, 33) = 2.1$, $p = .16$, $\eta_p^2 = .06$. The interaction between viewer-centered distance validity x viewer-to-target distance was significant, $F(1, 33) = 5.6$, $p < .05$, $\eta_p^2 = .15$.

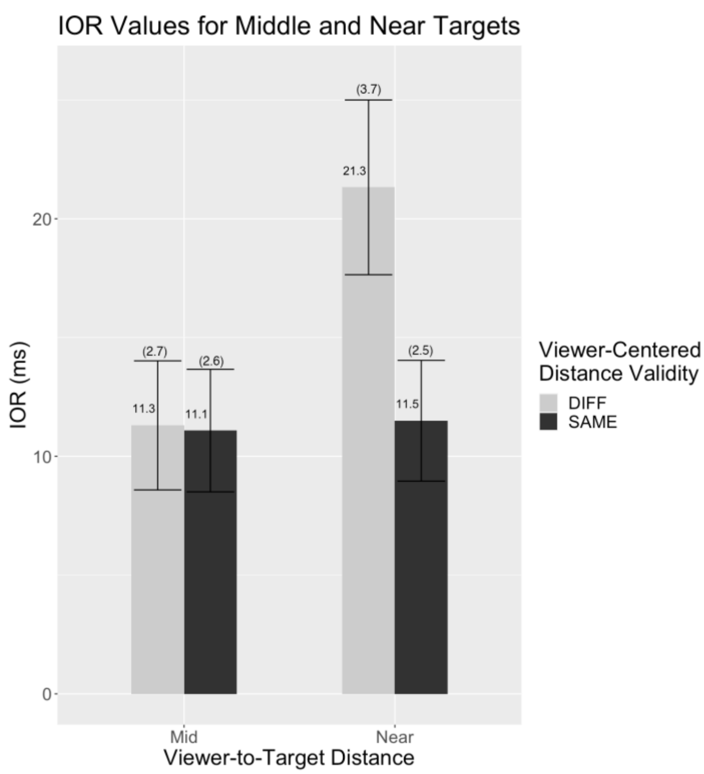


Figure 4. IOR results in Experiment 1 for same and different viewer-to-cue and viewer-to-target distances and for two levels of viewer-to-target distances tested: “middle” and “near.”

The interaction between viewer-centered distance validity x viewer-to-target distance was further analyzed by conducting two one-way ANOVAs, one for middle targets and one for near targets. For targets at the middle distance, the effect of viewer-centered distance validity was not significant, $F(1, 33) = 0.004$, $p = .90$, $\eta_p^2 = 1.29e-04$. For near targets, the effect of viewer-centered distance validity was significant, $F(1, 33) = 4.7$, $p < .05$, $\eta_p^2 = .12$. IOR for near targets was significantly larger (21.3 ms) when the cue appeared at a different viewer-centered distance compared to when the cue appeared at the same viewer-centered distance (11.5 ms). See Table 1 for an overview of all reaction times and IOR values for Experiment 1.

Table 1. Mean reaction times (RTs) and IOR values in ms for all conditions in Experiment 1.

Exp 1	Independent Variables		Dependent Variables	
	Viewer-to-Target Distance	Viewer-Centered Distance Validity	Side (RT in ms); [SE in ms]	IOR (ms); SE [ms]
	MIDDLE	SAME	SAME (485.21); [17.7]	11.1; [2.6]
			DIFF (474.13); [17.9]	
		DIFF	SAME (470.39); [14.3]	11.3; [2.7]
			DIFF (459.09); [14.2]	
	NEAR	SAME	SAME (483.89); [17.0]	11.5; [2.5]
			DIFF (472.39); [18.0]	
		DIFF	SAME (440.13); [14.0]	21.3; [3.7]
			DIFF (418.80); [12.8]	

IOR was calculated as the mean same side RT minus the mean different side RT for each condition. Standard error values are shown in square brackets.

Discussion

Experiment 1 investigated the effect of the viewer-centered reference frame on IOR by manipulating the viewer-to-cue and viewer-to-target distance difference. In the different viewer-centered distance condition, the viewer-to-cue distance was always greater than the viewer-to-target distance along the z-axis. In the same viewer-centered distance condition, the cue and target were located at the same distance relative to the viewer along the z-axis. In both conditions, the target either appeared 25 vm (middle target) or 10 vm (near target) from the viewer along the depth axis. The cue and target always appeared on the same placeholder wall, so the viewer did not need to adjust their attention across different world-centered depth planes. If the viewer-centered reference frame influenced IOR, a smaller IOR was expected to be seen for the different viewer-centered distance condition compared to the same viewer-centered distance condition. If the viewer-centered reference frame did not influence IOR, then IOR was expected to be equal in all conditions since the cue and target were always located at the same world depth plane but at different distances relative to the viewer.

We found no significant difference in IOR in the same versus different viewer-centered distance conditions when targets appeared at the middle distance. However, when analyzing the IOR for near targets, a larger IOR (21.3 ms) was produced when the viewer-to-cue depth distance was greater than the viewer-to-target distance along the depth axis. Previous literature has shown that a observers respond more quickly to information appearing closer to their peripersonal region of space (Andersen & Kramer, 1993; Arnott & Shedden, 2000). In an experiment that had observers follow a lead vehicle and detect a light change in an array of lights located above the lead vehicle at various distances from the observer (Andersen et al., 2011), the fastest detection times were observed for nearer light changes. Further, particularly in the case of an observer experiencing visually simulated forward motion, targets apparently nearer to them are perceived as having a greater shear in the periphery of the retina (Koenderink, 1986). A viewer-distance dependency, most likely caused by shear of nearer targets on the peripheral area of the retina, may have influenced faster detection times for targets appearing nearer to the observer. In this current study, observers experiencing visually simulated motion might have allocated more spatial attention resources to nearer targets.

It is important to note that the heightened attention for near targets (i.e., faster reaction times) was only shown when the cue appeared at a farther distance ("middle" distance in our terminology), while reaction times in the other three conditions were comparable. The greatest IOR was found for this middle-cue to near-target condition. This greater IOR may have been influenced by a heightened attention for the target located at the opposite side of the cue and/or a relatively greater inhibition for the target located at the same side of the cue. The current design does not offer the possibility to differentiate these two alternatives. However, based on the general findings of a near advantage in the literature, it is more likely that the heightened attention for targets located at the uncued side plays a major role in generating IOR.

To further investigate the role of the world-centered frame of reference on the IOR effect, several improvements to Experiment 1 were implemented in Experiment 2. For instance, Experiment 1 did not feature a condition where the cue and target appeared at different world depth planes. Experiment 2 included a condition where the viewer-centered distance difference was zero, but the world-centered depth difference was greater than zero. To implement this condition, the cue appeared on a south wall and target on a north wall during visually simulated forward motion. Inclusion of this condition in a within-group design would offer a stronger test of whether the viewer-centered or world-centered reference frame contributes to the generation of IOR.

Experiment 2

The data from Experiment 1 indicate that the difference in viewer-to-cue and viewer-to-target distances was not essential in generating IOR. However, the contribution of the world-centered reference frame to IOR was not directly tested. For Experiment 2, an additional condition was added to more directly test the influence of the world-centered depth reference frame on the generation of IOR. In this condition, the world-centered depth between cue and target was different —the cue now appeared on a placeholder wall closer to origin (i.e., the south wall) and the target appeared on a placeholder wall further from origin (i.e., the north wall; see Figure 1c). Even though the observer experiences visually simulated forward motion in 3D space, the viewer-centered distance was kept constant between viewer-to-cue and viewer-to-target position (i.e., same viewer-centered distance). The IOR from this Viewer-same World-different (VsWd, Condition 3) condition was compared with two other conditions (Conditions 1 and 2) where i) the viewer-centered distance and the world-centered depth was the same (VsWs), and ii) the viewer-centered distance was different, and the world-centered depth was the same (VdWs). It was reasoned that if the VsWd IOR is the same as the VsWs IOR, then world-centered depth differences have no effect on IOR. Otherwise, if the IOR is larger for the VsWs condition compared to the VsWd condition, the increased world-centered depth difference between cue and target in the VsWd condition decreases IOR (Bennett & Pratt, 2001; J. E. T. Taylor et al., 2015). In addition, if the VsWs IOR is the same as the VdWs IOR, then the viewer-centered distance differences have no effect on IOR. Otherwise, if the IOR is larger for the VsWs condition compared to the VdWs condition, the increased viewer-centered distance difference between cue and target in the VdWs condition decreases IOR.

Methods

Participants

A total of 49 McMaster students (18-22 years old; 38 female) took part in Experiment 2. All participants reported normal or corrected-to-normal vision, and normal colour vision. All participants received partial course credit or were compensated $15 CAD for their participation. Participants completed an informed consent form before starting the experiment. The study was approved by the McMaster University Research Ethics Board. The minimum sample size of n = 23 was determined because it gave the study 95% power, as calculated by GLIMMPSE for the main effect of viewing condition (Kreidler et al., 2013).

Stimuli and Apparatus

The stimuli and apparatus of Experiment 1 were identical to Experiment 2 with the following exceptions. Participants were asked to remain seated and instructed to maintain fixation on a central red vehicle (width = 2.18° x elevation = 1.94°). The participant's viewpoint was 22 vm behind the red fixation vehicle and 2 vm above ground on every trial.

Both south and north walls featured a checkerboard texture, which served as a pictorial depth cue. The distance between the south and north wall was held constant at 22 vm. Participants took part in three conditions: one control condition with a static viewpoint and two experimental conditions in which the viewpoint shifted between cue and target onset (i.e., visually simulated forward motion). In the condition involving simulated forward motion, when the viewpoint was 32 vm in front of the south wall, the south wall on either side of fixation measured 5.08° tall with a central gap of 18.27° and the north wall measured 3° tall with a central gap of 1.95°. When the viewpoint was 10 vm in front of the south wall, the south walls had receded into the periphery, and the north wall measured 5.16° tall with a central gap of 3°. In the static condition, the south walls were manually added to induce an impression of depth difference between south and north walls. When the viewpoint was 10 vm in front of the south wall, the south wall measured 16° tall with a central gap of 31.37° and the north wall measured 5.16° tall with a central gap of 3°.

Following the presentation of a peripheral blue-white checkered square cue (1.15° x 1.15°) participants were told to respond to a peripheral red-white checkered square target (1.15° x 1.15°) that appeared at either the same or different world-centered coordinates along the z-axis as the cue. In all conditions, the peripheral cue and target appeared at a 12.07° horizontal eccentricity.

The speed of the lead red vehicle was manipulated in the same way as in Experiment 1 except that its mean speed now ranged between 25-50 km/h per trial, which corresponded to the trial's CTOA (the faster the vehicle speed, the shorter the CTOA, and vice versa). This change in speed was included to reduce predictability of target onset. A short beep sound was made every time the cue or target appeared, as well as during a catch trial. An audio 'quack' feedback noise was provided for incorrect localization reactions or reactions made prior to the target onset (including catch trials). During a trial, participants were instructed to press the left flap on the backside of a steering wheel when the target appeared on the left side and the right flap on the backside a steering wheel when the target appeared on the right side.

Procedure

Before beginning the experimental trials, participants completed an 18-trial practice block containing all three conditions to become familiar with the environment and reaction conditions. For the experimental trials, participants began a trial by fixating on the central red vehicle. After 1000 ms, a cue appeared on either the left or right side of a north wall for 50 ms. In changing viewpoint trials, the cue offset occurred at motion onset.

Following a random CTOA between 650-1050ms (which coincided with the duration of perceived forward motion in dynamic trials), the target appeared on the north wall until a response was recorded.

Participants were instructed not to make eye movements and to record their reaction by pressing the left or right flap on the backside of a steering wheel corresponding to the side on which the target appeared. On catch trials, the cue would appear on the south or north wall, but no target would appear on the north wall. Participants were instructed not to respond during a catch trial; after 2000 ms, the trial would end automatically.

Design

The experiments used a within-subjects design where each participant took part in three conditions undergoing 480 experimental trials (160 per condition) and 60 catch trials (11.1% of trials). Every 18-trial experimental block (30 blocks) involved just one viewing condition, and the block presentation was counterbalanced across participants in one of three pre-randomized orders. Each block presented equal combinations of the cue side (left versus right) and target side (left versus right), repeated four times. The CTOA was randomly set between 650-1050 ms per trial.

The three conditions were designed to separate the effects of the viewer-centered versus world-centered spatial frame of reference. In all three conditions, participants viewed a peripheral cue and then a target. The z-axis distance between observer and target onset was always 32 vm. The distance between the south and north walls was 22 vm.

The control Viewer-same World-same (VsWs) condition is pictured in Figure 5a. The cue and target both appeared on the north wall, maintaining the same viewer-centered and world-centered depths. The viewer-to-cue distance and viewer-to-target distance was 32 vm.

The Viewer-different World-same (VdWs) condition is pictured in Figure 5b. The cue appeared at motion onset on the north wall. The cue and target were presented at the same world-centered depth (i.e., both on the north wall), but differed in viewer-centered distance. The viewer-to-cue distance was 54 vm. Following cue onset, the viewpoint translated forward 22 vm. The target then appeared at motion offset also on the north wall. The viewer-to-target distance was 32 vm.

The Viewer-same World-different (VsWd) condition is pictured in Figure 5c. The peripheral cue appeared on the south wall at motion onset. Then, after the viewpoint translated forward 22 vm, a subsequent target appeared on the north wall at motion offset. The viewer-to-cue distance and viewer-to-target distance were both 32 vm. Thus, the viewer-centered distance change was 22 vm, while the cue and target appeared at different world-centered depths 22 vm apart (i.e., cue on the south wall, target on the north wall).

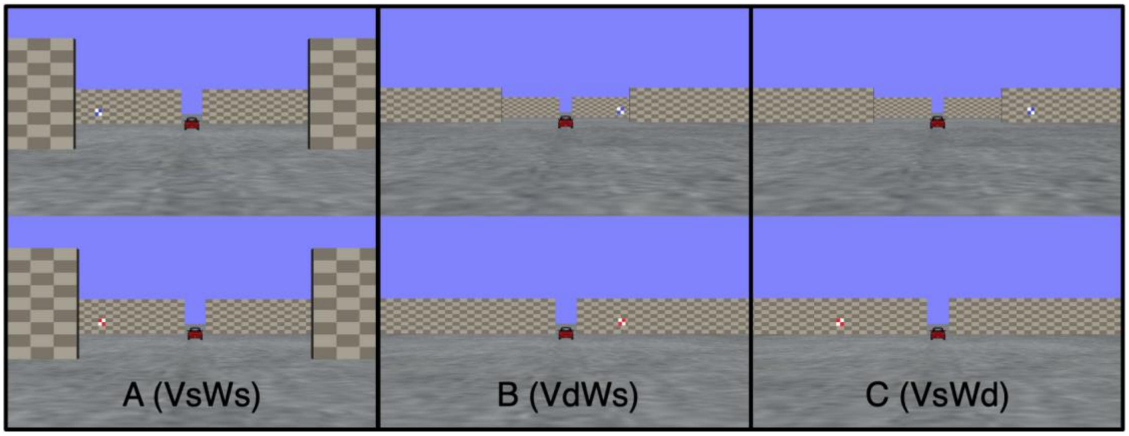


Figure 5. Stimuli in Experiment 2**.** The cue was a blue/white square and the target a red/white square, both with checkerboard texture. The red vehicle again served as the point of fixation and with the same spatial parameters as that in Experiment 1. For each of the three figure columns (A to C), the top and bottom images illustrate the view where the cue and target appeared, respectively.

A - Condition 1 (VsWs): Both the cue and target appeared on the north wall and viewer-to-cue and viewer-to-target distances were the same (22 vm).
B - Condition 2 (VdWs): Both the cue and target appeared on the north wall and the viewer-to-cue and viewer-to-target distances were 54 vm and 32 vm respectively, resulting in a viewer-centered change of 22 m.
C - Condition 3 (VsWd): The cue and target appeared on the south wall and north wall respectively with a wall-to-wall depth gap of 22 vm. Both the viewer-to-cue and viewer-to-target distance was 32 vm. The magnitude of viewpoint change was controlled so that depth difference between the two walls also matched the magnitude of distance travelled by the viewer (22 vm).

In the VsWs and VdWs condition, the cue appeared on the north wall. In the VsWs condition, the cue was 10 vm in front of the red vehicle, which acted as central fixation. In the VdWs condition, the cue was 32 vm in front of central fixation. In the VsWd, the cue appeared on the south wall, 10 vm in front of central fixation. In all conditions, the target always appeared on the left or right north walls 10 vm in front of central fixation. The viewpoint was always located 22 vm behind central fixation.

Results

Reaction times longer than 1000 msec and shorter than 200 msec, as well as incorrect localization reactions and anticipatory responses, were counted as errors (2.04% of experimental trials). Outlier data (2% of experimental trials) were also removed on either side of the reaction time distribution using a highly conservative threshold criterion of three times the mean absolute deviation of the median of raw reaction time data for each participant (Leys et al., 2013). Altogether, approximately 3.3% of the experimental trials was removed from further analysis.

Mean reaction time values were entered into a 2 (side validity: same versus different cue-target horizontal location) x 3 (condition: VsWs, VdWs, VsWd) univariate type III two-way repeated-measures analysis of variance (ANOVA) model. Deviations from sphericity were corrected for using a Huynh-Feldt correction. The ANOVA revealed a significant main effect of validity ($F(1,48) = 109.1$, $p < .001$, $\eta_p^2 = .69$), indicating a target appearing on the same side as a cue (M = 419.8 ms, se = 13.6 ms) was significantly slower than for a different side (M = 395.7 ms, se = 13.0 ms)—consistent with IOR.

The ANOVA also revealed a significant main effect of condition ($F(2,96) = 96.8$, $p < .001$, $\eta_p^2 = .67$). Using Tukey's method of honest significant differences to observe the pairwise comparisons, the VsWs condition (M = 430.3 ms, se =14 ms) was significantly slower ($p < .001$) than both the VsWd condition (M = 396 ms, se = 12.2 ms) and the VdWs condition (M = 397.3 ms, se = 13.5 ms). The VsWs and VdWs conditions were not significantly different ($p = .70$). This result suggests that the conditions involving perceived forward self-motion resulted in faster reaction times to targets than in the static VsWs condition.

Lastly, the ANOVA found a significant side validity x condition interaction ($F(2,96) = 55.4$, $p < .001$, $\eta_p^2 = .54$). To further examine the interaction, simple main effects for validity on reaction times were examined across all three conditions using three separate one-way within-subject ANOVAs. The simple main effect ANOVAs revealed a significant simple main effect of validity at all three conditions: VsWd ($F(1,48) = 13.7$, $p < .001$, $\eta_p^2 = .22$), VsWs ($F(1,48) = 97.3$, $p < .001$, $\eta_p^2 = 0.67$), and VdWs ($F(1,48) = 110$, $p < .001$, $\eta_p^2 = .70$). The IOR for the VsWs, VsWd, and VdWs conditions were 32.4 ms, 6.9 ms, and 33.4 ms, respectively. These findings are illustrated in Figure 6.

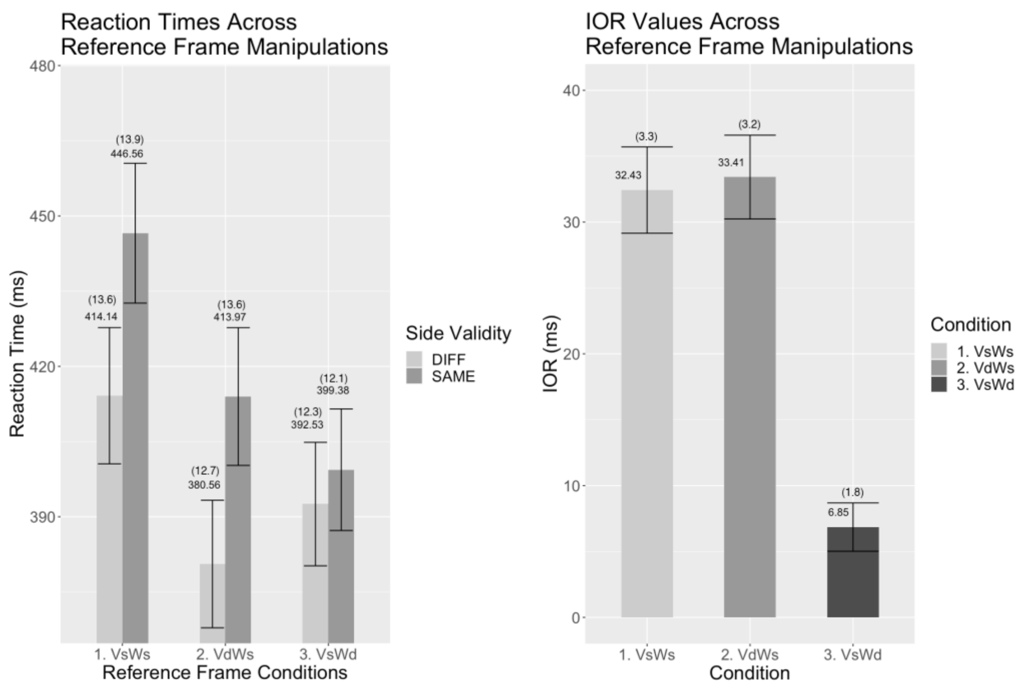


Figure 6. Reaction times (left portion) and IOR values (right portion) for the three conditions in Experiment 2.
Condition 1, VsWs: same viewer-centered distance and same world depth.
Condition 2, VdWs: different viewer-centered distance and same world depth.
Condition 3, VsWd: same viewer-centered distance and different world depth.

After calculating the IOR scores, paired sample t-tests were used to examine the differences in IOR. The results revealed a significant difference between the VsWs and VsWd conditions, ($t(48) = 6.8$, $p < .001$) and between the VdWs and VsWd conditions, (t(48) = 7.2 , $p < .001$). There was no significant difference between the VsWs and VdWs conditions, (t(48) = 0.23, $p = .82$). All mean values are displayed in Table 2.

Table 2. Mean reaction times (RTs) and IOR values in ms for all conditions in Experiment 2.

Exp 2	Independent Variables	Dependent Variables	
	Condition	Side (RT in ms); [SE in ms]	IOR (ms); SE [ms]
	1. VsWs	SAME (446.5); [13.9]	32.4; [3.3]
		DIFF (414.1); [13.6]	
	2. VdWs	SAME (414); [13.6]	33.4; [3.2]
		DIFF (380.6); [12.7]	
	3. VsWd	SAME (399.4); [12.1]	6.9; [1.8]
		DIFF (392.5); [12.3]	

IOR was calculated as the mean valid RT minus the mean invalid RT for each condition. Standard error values are shown in square brackets.

Discussion

Experiment 2 isolated the effects of viewer-centered and world-centered reference frames on IOR. The Viewer-same World-same (VsWs) condition involved the viewer remaining in a single viewer-centered position while responding to a target that appeared at the same world depth plane as the cue. The Viewer-same World-different (VsWd) condition involved the viewer experiencing visually simulated forward motion along a ground plane, seeing a cue appear on a south wall, and responding to a target appearing on a north wall relative to origin. However, the viewer-to-cue and viewer-to-target viewer-centered distances were kept the same. Finally, the Viewer-different World-same (VdWs) condition involved the viewer also experiencing visually simulated forward motion along a ground plane, but now seeing both the cue and target appear at the same world depth plane. If the viewer-centered reference frame influenced IOR, a reduced IOR was expected to be seen for the VdWs condition than for the VsWs condition. If the viewer-centered reference frame had no bearing on IOR, the IOR was expected to be the same for the VdWs and VsWs conditions. If the world-centered reference frame influenced IOR, a reduced IOR was expected to be seen for the VsWd condition than for the VsWs condition. Otherwise, the world-centered reference frame likely did not influence IOR.

Results indicated that the world-centered reference frame influenced IOR most greatly. The IOR was significantly smaller in the VsWd condition (6.9 ms) than in the VsWs condition (32.4 ms). However, the IOR in the VsWs condition was not significantly different from the IOR in the VdWs condition (33.4 ms), which indicates that differences in viewer-centered distance relative to the cue and target had no or very minimal influence on IOR.

It was the difference along the depth axis between cue and target that led to the drastic reduction in IOR for the VsWd condition. In the VsWd condition, the cue appeared on the south wall, which was located 22 vm south of the target that appeared on the north wall. In the VdWs and VsWs conditions, this depth difference was non-existent as the cue and target always appeared on the north wall within a single depth plane. Thus, observers considered the 22 vm difference between walls, leading to the reduction in IOR for the VsWd condition. The world-centered reference frame was used.

General Discussion

This study explored the contribution of the viewer-centered and world-centered reference frames on the generation of IOR. The viewer-centered reference frame was defined as the difference between viewer-to-cue and viewer-to-target distance. The world-centered reference frame was defined as the depth difference between cue and target. In both cases, this distance was considered solely along the depth axis. Results from Experiment 1 suggest that the difference in viewer-to-cue distance and viewer-to-target distance did not contribute greatly to IOR, with data from Experiment 2 confirming this notion by directly demonstrating that the world-centered reference frame is used when generating IOR. However, the results of Experiment 1 suggest that a small viewer-to-target distance may evoke a sense of behavioural urgency that causes observers to respond more quickly to close targets, especially when they appear opposite to the cue, inflating IOR.

The frame of reference that an observer uses to guide their attention may depend on the environmental features and structure of scene to which the observer is exposed. The conditions of Experiments 1 and 2 differed significantly in ways that may have affected the reference frame observers used to orient their attention within the simulated space. In Experiment 1, the only condition in which the IOR significantly increased was for the middle-cue to near-target dynamic condition. This was the only condition that involved an observer experiencing visually simulated forward motion while detecting a target on a wall located extremely near (10 vm) to their apparent region of peripersonal space. The reaction time for these near targets in Experiment 1 was reduced when they appeared at the opposite side from the cue, resulting in a larger IOR. In Experiment 2, observers were never at an apparent distance of less than 32 vm from any task-relevant depth plane, which likely dissuaded them from incorporating a viewer-centered frame of reference as that seen in the middle-cue to near-target condition in Experiment 1.

The attentional level for observers in similar dynamic spatial cueing situations have exhibited comparable effects of behavioural urgency. The viewer-centered reference frame has been suggested to bias behaviour in situations involving an moving toward a salient and task-relevant location (Andersen et al., 2011; Miura et al., 2002) or a task-relevant location or target approaching an observer (Reppa et al., 2010; A. Wang et al., 2016). In addition, an observer moving forward in space sees nearer objects move a greater distance, faster, and with a greater perceived shear on the peripheral region of the retina (Koenderink, 1986). This near advantage has been shown to be inflated when rich pictorial depth cues such as relative size are incorporated into the environmental design (Song et al., 2021). However, in Experiment, 2 the world-centered frame of reference likely dominated the generation of IOR because the north and south walls were both made task-relevant for spatial cueing.

The world-centered depth separation between the north and south walls probably caused the near elimination of IOR for the VsWd condition in Experiment 2. Observers likely encoded the depth separation between cue and target, while ignoring the viewer-centered distances to the cue and target. Similarly, when cues and targets both appeared on a single wall, the IOR increased because of cues and targets appearing within a single world-centered depth plane. The depth-specificity of IOR introduced in Chapters 2 and 3 may rely, therefore, on the differences in world-centered depth planes within which cues and targets appear, as opposed to viewer-centered distance differences.

Several limitations of the study pre-empted a closer analysis of how certain environmental features influenced the viewer-centered frame of reference's contribution to IOR. The first limitation was that, in Experiment 2, the target never appeared on a placeholder wall that was extremely near to the observer's viewpoint. In Experiment 1, the target appeared 10 vm from the observer's viewpoint. In Experiment 2, however, the observer's viewpoint was always 32 vm in front of a task-relevant placeholder wall on which a target appeared. In effect, there is a lack of supporting evidence against or in favour of the behavioural urgency hypothesis for targets located nearer to a viewer's peripersonal region of space. To remedy this limitation, future experiments could involve manipulating the effect that various viewer-to-target distances have on the distribution of IOR. A second limitation of both Experiments 1 and 2 was that the influence of viewer-centered depth switch trajectories on IOR were not compared in the dynamic condition. An additional condition, therefore, could involve alternating the world depth plane at which the cue and target appear while equating the viewer-to-cue and viewer-to-target distance. The walls on which the cue and target appear could move forward at the same speed as the viewer. The cue could appear on the north wall and the target on the south wall (and vice versa). With this revised design, different depth switch trajectories can be compared. The viewer-centered and world-centered depth difference could be held constant between cue and target onset, while examining the influence of depth switches in either direction (farther north wall to the nearer south wall and vice versa) on IOR.

Another limitation of the experiments in this chapter was that when cues and targets were located at different depth planes in the VsWd condition, they also appeared on different wall objects. Thus, there was no way of telling whether IOR was coded in world-centered coordinates based on coarse spatial locations or because of an inhibitory tag placed on objects residing at different depths (as was observed in Chapter 3). A way to delineate the two sources of influence (location-based versus object-based IOR) would be to extend the wall object along the depth axis and place the cue and target along different world depth planes but within the same object. Along the same vein, when both the cues and target are located within the same depth plane, they could also be located upon different wall objects along the horizontal axis to identify the influence of inhibitory tags placed on objects along the horizontal axis.

One might also think about the contribution of depth information in relation to central fixation. In the VdWs condition, the fixation-to-cue distance was 32 vm, which was different from the fixation-to-target distance of 10 vm. This 22 vm difference was greater compared to that in the other two conditions (0 vm difference). However, given that across three conditions the viewer-to-target distance was always 32 vm, and the viewer-to-fixation distance was always 22 vm, the fixation likely did not influence the difference in results between conditions. In the VdWs condition the cue may have been less effective in inhibiting attention because it was 22 vm farther from the central fixation compared to other conditions. However, the IOR for this condition is not smaller relative to the control condition, and the IOR values between the control condition VsWs and experimental condition VsWd differed even though the fixation-to-cue distance equalled the fixation-to-target distance. Therefore, the larger fixation-to-cue distance in the VdWs condition likely did not play a significant role in the generation of IOR.

Conclusion

In this study's IOR paradigm, cues and targets appeared within or between world depth planes while the viewer-to-cue and viewer-to-target distances remained equal or were changed. In Experiment 1, the IOR only changed when targets located very near to the viewer's apparent peripersonal space were cued from farther away. The near depth plane of space likely has a higher baseline of attentional focus, promoting a sense of behavioural urgency when the viewer approaches a task-relevant placeholder wall. In Experiment 2, IOR was drastically reduced only when cues and targets were located on walls separated in depth, indicating that the world-centered reference frame strongly contributes to how IOR is generated.

References

Abrams, R., & Pratt, J. (2000). Oculocentric coding of inhibited eye movements to recently attended locations. *Journal of Experimental Psychology. Human Perception and Performance*, *26*, 776–788. https://doi.org/10.1037/0096-1523.26.2.776

Andersen, G. J., & Kramer, A. F. (1993). Limits of focused attention in three-dimensional space. *Perception & Psychophysics*, *53*(6), 658–667. https://doi.org/10.3758/BF03211742

Andersen, G. J., Ni, R., Bian, Z., & Kang, J. (2011). Limits of spatial attention in three-dimensional space and dual-task driving performance. *Accident; Analysis and Prevention*, *43*(1), 381–390. https://doi.org/10.1016/j.aap.2010.09.007

Arnott, S. R., & Shedden, J. M. (2000). Attention switching in depth using random-dot autostereograms: Attention gradient asymmetries. *Perception & Psychophysics*, *62*(7), 1459–1473. https://doi.org/10.3758/BF03212146

Astle, D. E. (2009). Going from a retinotopic to a spatiotopic coordinate system for spatial attention. *The Journal of Neuroscience*, *29*(13), 3971–3973. https://doi.org/10.1523/JNEUROSCI.5929-08.2009

Bao, Y., Lei, Q., Fang, Y., Tong, Y., Schill, K., Pöppel, E., & Strasburger, H. (2013). Inhibition of return in the visual field. *Experimental Psychology*, *60*(6), 425–431. https://doi.org/10.1027/1618-3169/a000215

Bennett, P. J., & Pratt, J. (2001). The spatial distribution of inhibition of return. *Psychological Science*, *12*(1), 76–80. https://doi.org/10.1111/1467-9280.00313

de Gonzaga Gawryszewski, L., Riggio, L., Rizzolatti, G., & Umiltá, C. (1987). Movements of attention in the three spatial dimensions and the meaning of "neutral" cues. *Neuropsychologia*, *25*(1A), 19–29.

Feng, J., & Spence, I. (2017). The effects of spatial endogenous pre-cueing across eccentricities. *Frontiers in Psychology*, *8*(888), 1–16. https://doi.org/10.3389/fpsyg.2017.00888

Golomb, J. D., Chun, M. M., & Mazer, J. A. (2008). The native coordinate system of spatial attention Is retinotopic. *The Journal of Neuroscience*, *28*(42), 10654–10662. https://doi.org/10.1523/JNEUROSCI.2525-08.2008

Hilchey, M. D., Klein, R. M., Satel, J., & Wang, Z. (2012). Oculomotor inhibition of return: How soon is it "recoded" into spatiotopic coordinates? *Attention, Perception, & Psychophysics*, *74*(6), 1145–1153. https://doi.org/10.3758/s13414-012-0312-1

Klein, R. (2000). Inhibition of return. *Trends in Cognitive Sciences*, *4*(4), 138–147.

Koenderink, J. J. (1986). Optic flow. *Vision Research*, *26*(1), 161–179. https://doi.org/10.1016/0042-6989(86)90078-7

Kreidler, S. M., Muller, K. E., Grunwald, G. K., Ringham, B. M., Coker-Dukowitz, Z. T., Sakhadeo, U. R., Barón, A. E., & Glueck, D. H. (2013). GLIMMPSE: Online power computation for linear models with and without a baseline covariate. *Journal of Statistical Software*, *54*(10), i10.

Krüger, H. M., & Hunt, A. R. (2013). Inhibition of return across eye and object movements: The role of prediction. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(3), 735–744. https://doi.org/10.1037/a0030092

Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, *49*(4), 764–766. https://doi.org/10.1016/j.jesp.2013.03.013

Maylor, E. A., & Hockey, R. (1985). Inhibitory component of externally controlled covert orienting in visual space. *Journal of Experimental Psychology: Human Perception and Performance*, *11*(6), 777–787. https://doi.org/10.1037/0096-1523.11.6.777

Miura, T., Shinohara, K., & Kanda, K. (2002). Shift of attention in depth in a semi-realistic setting. *Japanese Psychological Research*, *44*(3), 124–133. https://doi.org/10.1111/1468-5884.00015

Pertzov, Y., Zohary, E., & Avidan, G. (2010). Rapid formation of spatiotopic representations as revealed by inhibition of return. *Journal of Neuroscience*, *30*(26), 8882–8887. https://doi.org/10.1523/JNEUROSCI.3986-09.2010

Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. *Attention and Performance X*, *32*, 531–556.

Posner, M. I., Rafal, R. D., Choate, L. S., & Vaughan, J. (1985). Inhibition of return: Neural basis and function. *Cognitive Neuropsychology*, *2*(3), 211–228. https://doi.org/10.1080/02643298508252866

Reppa, I., Fougnie, D., & Schmidt, W. C. (2010). How does attention spread across objects oriented in depth? *Attention, Perception, & Psychophysics*, *72*(4), 912–925. https://doi.org/10.3758/APP.72.4.912

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

Song, J., Bennett, P. J., Sekuler, A. B., & Sun, H.-J. (2021). The effect of apparent distance on peripheral target detection. *Journal of Vision*, *21*(10), 1–8. https://doi.org/10.1167/jov.21.10.8

Taylor, J. E. T., Chan, D., Bennett, P. J., & Pratt, J. (2015). Attentional cartography: Mapping the distribution of attention across time and space. *Attention, Perception, & Psychophysics*, *77*(7), 2240–2246. https://doi.org/10.3758/s13414-015-0943-0

Wang, A., Liu, X., Chen, Q., & Zhang, M. (2016). Effect of different directions of attentional shift on inhibition of return in three-dimensional space. *Attention, Perception, & Psychophysics*, *78*(3), 838–847. https://doi.org/10.3758/s13414-015-1055-6

Chapter 5: General Discussion

What does it mean to visually attend to a specific location or object? James (1890) claimed that spatial attention was the explicit act of selecting certain areas of space on which to focus, concentrate, and place into consciousness while dropping others. He insisted that inhibition could be produced by "fixing the eyes on vacancy." Gibson (1966) would have said that it involves looking at something with one's eyes and simply extracting information from a set of stimuli. Marr (1982) would have gone beyond this bottom-up approach and insist that the internal representation of information—how the onset of light or change in shape of a structure—affects incoming visual information by processing it in a top-down manner. Hommel et al. (2019) suggested on dropping the term "attention" altogether and instead urged researchers to refer to the exact behavioural selection process or effect being studied. Considering this philosophical approach to studying attention, this dissertation only makes new claims about one specific behavioural effect—inhibition of return (IOR)—within a paradigm of exogenously cued covert detection in a virtual three-dimensional (3D) scene. This dissertation proposes that the IOR effect partly arises because an observer tags a previously cued area of space—an area of space they are told to ignore or disregard during a trial— as irrelevant or insignificant. The author of this dissertation is aware that the IOR effect likely occurs due to an output-based motor component acting in conjunction with an input-based cognitive component of attentional inhibition. Nevertheless, this dissertation does not make any claims about any output-based IOR effects, which are commonly investigated with saccades or by requiring arbitrary (as opposed to spatially congruent) response-mappings for spatial locations (Howard et al., 1999; Ivanoff & Klein, 2001; Lyons et al., 2006; T. L. Taylor & Klein, 2000; Tipper et al., 1998). Because covert and exogenously cued IOR, as implemented in our study, may not actually require the planning of saccades and may be entirely functionally independent of them (Smith et al., 2012), this dissertation makes the claim that any IOR generated in the experiments occurred mostly (or at least partly) within the input-based stream of visuospatial attention. In other words, the IOR observed within our experiments is considered to have been generated because of observers inhibiting perceived depth planes and objects at the level of the brain where the internal representation of visual space is formed and processed.

In 2D space, IOR has shown to monotonically decrease as a function of cue-to-target distance (Bennett & Pratt, 2001; Maylor & Hockey, 1985; J. E. T. Taylor et al., 2015), with peak IOR occurring at the original cue or at the population vector's centre of gravity of multiple cues (Christie et al., 2013; R. M. Klein et al., 2005; Tipper et al., 1998). In 2D space, IOR has also shown to peak at around 600 ms CTOA, eventually starting to dissipate at 1200 ms until entirely disappearing around 3000 ms. IOR is also affected by whether cues and targets belong to a single object or are structurally associated with two different objects. When a cue and target share two different spatial locations, IOR decreases the farther the cue and target are apart from one another. However, if the cue and target's distance remains the same but they are now placed within a single object, the IOR tends to increase, becoming more pronounced with the increased salience and continuity of an object's boundaries and area (Jordan & Tipper, 1999; Leek et al., 2003). This occurs because of a lasting inhibitory tag placed on the entire object that encapsulates the cue and target, observed even when the object moves in space (Tipper et al., 1991, 1994, 1997; Weaver et al., 1998).

Most of the research looking at IOR in 3D space has only employed stereopsis depth (Bourke et al., 2006; Casagrande et al., 2012; Theeuwes & Pratt, 2003; A. Wang et al., 2016). The one main difference with IOR in 3D space is that IOR decreases further when cues and targets are located at different depths but share the same x,y retinal locations. When cues and targets are not structurally associated with any objects but occupy feature-less spatial locations along the depth axis, a difference in IOR along the depth axis is seen compared to the horizontal axis (Casagrande et al., 2012). This is referred to as the depth-specificity of IOR and occurs because observers consider the distance between cue and target along the depth axis as instrumental in separating the two stimuli. Observers cognitively represent the world coordinates of the cue and target in addition to their retinal coordinate differences (M. D. Hilchey et al., 2012; Krüger & Hunt, 2013; Satel & Wang, 2012). A depth asymmetry in IOR has also been exhibited, where an depth switch from far-to-near space creates an IOR of a magnitude different from that observed for an depth switch made from near-to-far space (A. Wang et al., 2016). IOR has also shown to exhibit object-based properties, with IOR increasing with cues and targets belonging to a single object than to two different objects when separated by the same distance along the depth axis (Bourke et al., 2006). This dissertation extends this research by showing that IOR can indeed be depth-specific and depth-asymmetric in a 3D space composed only of pictorial depth cues, a finding that was previously shown to be unlikely (A. Wang et al., 2016). Furthermore, whether the IOR generated in 3D space is viewer-centered (depends on where the observer is in relation to the cue and target) or world-centered (depends on where the cue is in relation to the target) is investigated, a notion unexplored by previous IOR research.

In Chapter 2, two peripheral cued target detection experiments were conducted, which used a modified IOR paradigm in a 3D setting. Peripheral cues and targets appeared upon placeholder boxes placed on the surface of a textured ground plane with linear perspective and relative size as the main pictorial depth cues. Both experiments found that in an environment without stereopsis but with rich pictorial depth cues, IOR showed a depth-specific effect. The depth-specific effect found in our 3D condition (factoring both 3D and 2D results) was not entirely explained by 2D parameters. The IOR difference between same versus different depths (for the far cue and near target condition) was much greater than any other conditions. A large source of this difference was because far-to-near depth switch trajectories in the 3D setting resulted in a reduced IOR magnitude than near-to-far depth switches. The near depth plane of space likely has a higher baseline of attentional focus, favouring facilitation rather than inhibition.

Chapter 3 clarified whether the depth-asymmetric IOR effect observed in Chapter 2 was mostly due to location- or object-based components. A cued target detection experiment was conducted, which again employed a modified IOR paradigm in a 3D scene. Cues and targets appeared within placeholder boxes situated on the surface of a textured ground plane presenting linear perspective, texture gradient, and relative size as the main pictorial depth cues. When the cue appeared in a farther location than that of the subsequent target, the IOR effect was depth-specific. Thus, IOR was smaller when the cue and target appeared in different depth planes than when cued and target appeared in the same depth plane (when compared with the 2D control). This depth effect was only present when the cue and target appeared in two placeholder boxes separated in depth, but not when cue and target appeared in one placeholder box while all other spatial properties remained the same. Our results suggest that IOR can be depth-specific, but the depth effect is limited to the specific direction of the depth switch across depth and object-based inhibition can override this depth-specific effect. The main issue with the experiments conducted in Chapters 2 and 3 was that when the target was farther than the cue, it was also farther from the observer (and vice versa when the target was nearer). Therefore, world-centered IOR effects were never isolated from viewer-centered IOR effects. This contrast was investigated in Chapter 4.

A viewer-centered frames of reference typically goes hand in hand with variance in the structure of a scene—as an observer experiences a change in their environment, either through simulated locomotion or a spatial change in luminance, the input of a scene changes. However, it is when the viewer experiences change and the structure of a scene remains the same, that invariant features of a scene may also arise. Thus, with Chapter 4, the variant viewer-centered frame of reference was manipulated to make it possible to study the invariant world-centered frame of reference. Chapter 4 isolated the relative contributions of viewer-centered IOR and world-centered IOR by controlling for the viewer-to-cue and viewer-to-target distance versus the cue-to-target distance along the depth axis. Previous studies in Chapters 2 and 3 found evidence for a depth-specific IOR—IOR was reduced most drastically for far-to-near depth switches. However, it is not clear what frame of reference is involved in this depth-specific IOR since viewer-centered and world-centered frame of references commonly specify the same depth relation. In Chapter 4, the spatial frame of reference for the depth-specific effect was investigated by creating a discrepancy between depth information specified by viewer-centered and world-centered coordinate systems. Target detection was tested in a spatial cueing paradigm in a simulated 3D space illustrated through pictorial depth cues (such as linear perspective and optic flow). In the control condition, the cue and target were always presented in the same depth plane relative to a stationary viewpoint. In the experimental conditions, the cue and target appeared at different viewer-centered distances (Experiment 1 and 2) or different world-centered depth planes (Experiment 2), with the competing frame of reference controlled between cue and target onset. Results showed that in Experiment 1, the IOR only changed when targets located very near to the viewer's apparent peripersonal space were cued from farther away relative to the viewer. The near depth plane of space likely has a higher baseline of attentional focus, promoting a sense of behavioural urgency when the viewer approaches the task-relevant placeholder wall. In Experiment 2, IOR was drastically reduced only when cues and targets were located on placeholder walls separated in depth, indicating that the world-centered reference frame strongly contributes to how IOR is generated.

This dissertation aimed to fill a gap in IOR research involving exogenous covert spatial cueing in 3D scenes. Previous studies have suggested that the depth-specificity of IOR might only occur in 3D scenes composed of stereopsis depth cues, with some recommending to only study IOR in stereo space to produce meaningful results (Casagrande et al., 2012; A. Wang et al., 2016). However, by incorporating 3D scenes composed only of pictorial depth cues, the experiments from this dissertation suggest that observers can infer a depth axis rather than just use the x,y retinal coordinates in generating a depth-specific and depth-asymmetric IOR. Specifically, when cues and targets were located at different depths in a 3D scene, the IOR was shortened further than what was seen in the comparable 2D scene. Additionally, this reduction in IOR was only prominent for depth switches occurring from far-to-near space and only when the cue and target appeared between different objects. In effect, the IOR measured in our 3D scenes exhibited a depth-asymmetry that depended on the cues and targets belonging to differentiated boundaries along the depth axis. The average observer in our experiments likely experienced a speed advantage for targets occurring in near space, with inhibition being unable to spread from far to near space. This near advantage was also observed when an observer experienced simulated forward motion in 3D space with the cue and target appearing at the same world depth plane—the closer the observer was to the placeholder, the faster the reaction time. However, when the viewer-to-cue and viewer-to-target depth distance was equal, but the world depth planes different, the IOR dropped drastically. When the viewer-to-cue and viewer-to-target depth difference is the same, the viewer more so relies on the difference between cue and target in world coordinates when generating IOR. Therefore, both frames of reference are likely used but at different times. The viewer-centered frame of reference is probably used when the cue and target are on the same world depth plane, with reaction times being shortened the closer a viewer is to a task-relevant placeholder. The world-centered frame of reference likely dominates when the cue and target are on two different depth planes, leading to the massive reduction in IOR since the observer may find it more useful to react quickly to stimuli appearing on new objects or placeholders.

All in all, this dissertation provides the building blocks towards investigating scenarios that impact pedestrian and driver safety. Vehicles can be re-designed to promote driver-to-vehicle, pedestrian-to-vehicle, and vehicle-to-vehicle communication. There very well may be an area of space in each unique environmental situation where inhibition is the highest or where location- versus object-based components each have their own role to play in visuospatial attention. Inhibitory tags placed on objects may be useful when a driver inhibits the trajectories of cars going in the opposite direction on a highway, but this inhibition may become life-threatening when one of those cars ends up veering into the driver's lane. Furthermore, with the onset of autonomous vehicles, how do we prevent people from inhibiting their attention to everything outside of their car—spatial locations and objects alike? What sorts of interfaces can we build to motivate drivers to attend to highly relevant locations and objects of interest, particularly at high-risk intersections or neighbourhoods with high-crash rates? This question carries more importance as we enter an age where vehicles of varying levels of autonomy will be driving on the same roads. It will be necessary to change pieces of our road and city infrastructure to facilitate this transition. The results from this dissertation could better inform professionals of the most appropriate interfaces to design and develop. Interfaces that work alongside computer vision algorithms to promote human focus to specific regions of a scene outside and inside a car cabin (and the transition between the two), can aid drivers in responding to relevant items of interest (Breitschaft et al., 2019; Schroeter & Steinberger, 2016).

References

Bennett, P. J., & Pratt, J. (2001). The spatial distribution of inhibition of return. *Psychological Science*, *12*(1), 76–80. https://doi.org/10.1111/1467-9280.00313

Bourke, P. A., Partridge, H., & Pollux, P. M. J. (2006). Additive effects of inhibiting attention to objects and locations in three-dimensional displays. *Visual Cognition*, *13*(5), 643–654. https://doi.org/10.1080/13506280544000309

Breitschaft, S. J., Clarke, S., & Carbon, C.-C. (2019). A theoretical framework of haptic processing in automotive user interfaces and its implications on design and engineering. *Frontiers in Psychology*, *10*, 1–18. https://doi.org/10.3389/fpsyg.2019.01470

Casagrande, M., Barbato, M., Mereu, S., Martella, D., Marotta, A., Theeuwes, J., & Collinson, S. (2012). Inhibition of return: A "depth-blind" mechanism? *Acta Psychologica*, *140*, 75–80. https://doi.org/10.1016/j.actpsy.2012.02.011

Christie, J., Hilchey, M. D., & Klein, R. M. (2013). Inhibition of return is at the midpoint of simultaneous cues. *Attention, Perception, & Psychophysics*, *75*(8), 1610–1618. https://doi.org/10.3758/s13414-013-0510-5

Gibson, J. J. (1966). *The senses considered as perceptual systems*. Houghton Mifflin.

Hilchey, M. D., Klein, R. M., Satel, J., & Wang, Z. (2012). Oculomotor inhibition of return: How soon is it "recoded" into spatiotopic coordinates? *Attention,* *Perception, & Psychophysics*, *74*(6), 1145–1153. https://doi.org/10.3758/s13414-012-0312-1

Hommel, B., Chapman, C. S., Cisek, P., Neyedli, H. F., Song, J.-H., & Welsh, T. N. (2019). No one knows what attention is. *Attention, Perception, & Psychophysics*, *81*(7), 2288–2303. https://doi.org/10.3758/s13414-019-01846-w

Howard, L. A., Lupiáñez, J., & Tipper, S. P. (1999). Inhibition of return in a selective reaching task: An investigation of reference frames. *The Journal of General Psychology*, *126*(4), 421–442. https://doi.org/10.1080/00221309909595374

Ivanoff, J., & Klein, R. M. (2001). The presence of a nonresponding effector increases inhibition of return. *Psychonomic Bulletin & Review*, *8*(2), 307–314. https://doi.org/10.3758/bf03196166

James, W. (1890). *The Principles of Psychology*. London, England: Dover Publications.

Jordan, H., & Tipper, S. P. (1999). Spread of inhibition across an object's surface. *British Journal of Psychology*, *90*(4), 495–507. https://doi.org/10.1348/000712699161576

Klein, R. M., Christie, J., & Morris, E. P. (2005). Vector averaging of inhibition of return. *Psychonomic Bulletin & Review*, *12*(2), 295–300. https://doi.org/10.3758/BF03196375

Krüger, H. M., & Hunt, A. R. (2013). Inhibition of return across eye and object movements: The role of prediction. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(3), 735–744. https://doi.org/10.1037/a0030092

Leek, E. C., Reppa, L., & Tipper, S. P. (2003). Inhibition of return for objects and locations in static displays. *Perception & Psychophysics*, *65*(3), 388–395. https://doi.org/10.3758/bf03194570

Lyons, J., Glazebrook, C. M., Keetch, K. M., Dhillon, V. P., & Elliott, D. (2006). Influence of endogenous and exogenous orientations of attention on inhibition of return in a cross-modal target-target aiming task. *Journal of Motor Behavior*, *38*(3), 219–228. https://doi.org/10.3200/JMBR.38.3.219-228

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. W.H. Freeman.

Maylor, E. A., & Hockey, R. (1985). Inhibitory component of externally controlled covert orienting in visual space. *Journal of Experimental Psychology: Human Perception and Performance*, *11*(6), 777–787. https://doi.org/10.1037/0096-1523.11.6.777

Satel, J., & Wang, Z. (2012). Investigating a two causes theory of inhibition of return. *Experimental Brain Research*, *223*(4), 469–478. https://doi.org/10.1007/s00221-012-3274-6

Schroeter, R., & Steinberger, F. (2016). Pokémon DRIVE: Towards increased situational awareness in semi-automated driving. *Proceedings of the 28th Australian Conference on Computer-Human Interaction - OzCHI '16*, 25–29. https://doi.org/10.1145/3010915.3010973

Smith, D. T., Schenk, T., & Rorden, C. (2012). Saccade preparation is required for exogenous attention but not endogenous attention or IOR. *Journal of*

Experimental Psychology. Human Perception and Performance, *38*(6), 1438–1447. https://doi.org/10.1037/a0027794

Taylor, J. E. T., Chan, D., Bennett, P. J., & Pratt, J. (2015). Attentional cartography: Mapping the distribution of attention across time and space. *Attention, Perception, & Psychophysics*, *77*(7), 2240–2246. https://doi.org/10.3758/s13414-015-0943-0

Taylor, T. L., & Klein, R. M. (2000). Visual and motor effects in inhibition of return. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(5), 1639–1656. https://doi.org/10.1037/0096-1523.26.5.1639

Theeuwes, J., & Pratt, J. (2003). Inhibition of return spreads across 3-D space. *Psychonomic Bulletin & Review*, *10*(3), 616–620. https://doi.org/10.3758/BF03196523

Tipper, S. P., Driver, J., & Weaver, B. (1991). Object-centred inhibition of return of visual attention. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, *43*(2), 289–298. https://doi.org/10.1080/14640749108400971

Tipper, S. P., Howard, L. A., & Houghton, G. (1998). Action-based mechanisms of attention. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *353*(1373), 1385–1393.

Tipper, S. P., Rafal, R., Reuter-Lorenz, P. A., Starrveldt, Y., Ro, T., Egly, R., Danzinger, S., & Weaver, B. (1997). Object-based facilitation and inhibition from visual orienting in the human split-brain. *Journal of Experimental Psychology. Human*

Perception and Performance, *23*(5), 1522–1532. https://doi.org//10.1037//0096-1523.23.5.1522

Tipper, S. P., Weaver, B., Jerreat, L. M., & Burak, A. L. (1994). Object-based and environment-based inhibition of return of visual attention. *Journal of Experimental Psychology. Human Perception and Performance*, *20*(3), 478–499. https://doi.org/10.1037/0096-1523.20.3.478

Wang, A., Liu, X., Chen, Q., & Zhang, M. (2016). Effect of different directions of attentional shift on inhibition of return in three-dimensional space. *Attention, Perception, & Psychophysics*, *78*(3), 838–847. https://doi.org/10.3758/s13414-015-1055-6

Weaver, B., Lupiáñez, J., & Watson, F. L. (1998). The effects of practice on object-based, location-based, and static-display inhibition of return. *Perception & Psychophysics*, *60*(6), 993–1003. https://doi.org/10.3758/BF03211934

Appendix

Table A1. Error rates for Chapter 2, Experiment 1

Exp 1

Independent Variables			Error Rate (%)
Target Depth	Perspective	Depth or Elevation	Side
NEARER (OR LOWER) THAN CUE	3D	SAME	SAME (0.88)
			DIFF (1.02)
		DIFF	SAME (0.88)
			DIFF (1.09)
	2D	SAME	SAME (1.99)
			DIFF (1.86)
		DIFF	SAME (1.80)
			DIFF (1.70)
FARTHER (OR HIGHER) THAN CUE	3D	SAME	SAME (0.88)
			DIFF (1.02)
		DIFF	SAME (1.04)
			DIFF (0.90)
	2D	SAME	SAME (1.99)
			DIFF (1.86)
		DIFF	SAME (2.05)
			DIFF (1.64)

Table A2. Error rates for Chapter 2, Experiment 2

Exp 2

Independent Variables			Error Rate (%)
Target Depth	Perspective	Depth or Elevation	Side
NEAR	3D	SAME	SAME (3.59)
			DIFF (4.0)
		DIFF	SAME (1.95)
			DIFF (2.56)
	2D	SAME	SAME (2.42)
			DIFF (2.92)
		DIFF	SAME (2.72)
			DIFF (2.12)
FAR	3D	SAME	SAME (2.97)
			DIFF (3.08)
		DIFF	SAME (2.87)
			DIFF (4.10)
	2D	SAME	SAME (2.52)
			DIFF (2.32)
		DIFF	SAME (2.62)
			DIFF (2.72)

Table A3 Error rates for Chapter 3, Experiment 1

One garden bed on either side of fixation

Target Depth	Perspective	Depth or Elevation	Side (Error Rate (%))
NEAR	3D	SAME	SAME (0.87)
			DIFF (0.65)
		DIFF	SAME (1.19)
			DIFF (1.03)
	2D	SAME	SAME (1.41)
			DIFF (2.62)
		DIFF	SAME (1.41)
			DIFF (1.09)
FAR	3D	SAME	SAME (1.25)
			DIFF (0.76)
		DIFF	SAME (0.86)
			DIFF (1.57)
	2D	SAME	SAME (1.62)
			DIFF (1.62)
		DIFF	SAME (1.35)
			DIFF (1.14)

Two garden beds on either side of fixation

Target Depth	Perspective	Depth or Elevation	Side (Error Rate (%))
NEAR	3D	SAME	SAME (0.97)
			DIFF (0.75)
		DIFF	SAME (1.29)
			DIFF (0.97)
	2D	SAME	SAME (0.75)
			DIFF (0.75)
		DIFF	SAME (0.54)
			DIFF (0.70)
FAR	3D	SAME	SAME (0.91)
			DIFF (0.97)
		DIFF	SAME (0.38)
			DIFF (1.19)
	2D	SAME	SAME (0.75)
			DIFF (0.86)
		DIFF	SAME (0.91)
			DIFF (0.81)

Table A4. Error rates for Chapter 4, Experiment 1

Exp 1

Viewer-to-Target Distance	Viewer-Centered Distance Validity	Side (Error Rate (%))
MIDDLE	SAME	SAME (4.67)
		DIFF (4.78)
	DIFF	SAME (2.52)
		DIFF (3.86)
NEAR	SAME	SAME (4.61)
		DIFF (6.11)
	DIFF	SAME (6.37)
		DIFF (8.04)

Table A5. Error rates for Chapter 4, Experiment 2

Exp 2

Condition	Error Rate (%)
VsWs	2.21
VdWs	1.98
VsWd	1.94