
THE UNIVERSITY
*of* EDINBURGH

# *The role of foveal vision in static and dynamic environments*


THE UNIVERSITY
*of* EDINBURGH

**Adam Christopher Clayden**

**Doctor of Philosophy**

**School of Philosophy, Psychology and Language Sciences**

**College of Arts, Humanities and Social Sciences**

**University of Edinburgh**

**May 2019**

*Stay a while and listen*

- Deckard Cain

# Abstract

The visual field has multiple regions, with visual acuity being highest in the centre before declining rapidly outward toward the periphery. This central region, otherwise known as the fovea, is typically defined as the central 2.0° of vision. Although comparatively small with respect to other visual field regions, being able to discern objects of interest in fine detail is only possible in this region. Due to this, people make ballistic eye movements (saccades) towards the fine details and depending on the task, may stabilise their gaze in the form of a fixation to discriminate parts of this newly attended area. A typical everyday task that can exhibit such behaviours is visual search (scanning a visual environment for objects or features among distractors), and much of this thesis is concerned with the importance of foveal vision with respect to visual search tasks. Seven experiments are presented in this thesis, with chapters 1 – 4 containing a review of the literature, methodologies and a glossary. In Chapter 5, search experiments with targets of varying sizes were conducted to assess the role of foveal vision on search performance. This chapter builds on a debate on whether foveal vision was necessary for the successful completion of a visual search task. In addition, a novel algorithm was developed to embed targets at a specified change in local contrast to automate target placement. The presented results show that the time taken to find targets with and without foveal vision is nearly identical even when target size is small. Chapter 6 modulates search difficulty by manipulating salience to investigate the effect of target size and salience on search performance. Coined the Compensation Effect, the results of Chapter 6 show that the above-mentioned variables were able to compensate for one another, resulting in an improved search performance. This effect occurred regardless of visual field degradation. In the same chapter, another experiment revealed the stage in the search process where performance costs originated from, which was the final stage of search concerned

with target verification. Finally, Chapter 7 transitions from static images to dynamic scenes which simulate self-motion. Additional algorithms were developed, including an extension to the existing gaze-based decomposition of search time (Malcolm & Henderson 2009). Chapter 7 investigated the role of foveal vision in visual search whilst optical flow was present. Unlike previous results, the final experiments of this thesis revealed the necessity of foveal vision for the attainment of a normal search performance.

The results of this thesis demonstrate that the importance of foveal vision with respect to visual search is modulated by the stimulus environment, with it being relatively unimportant for static scenes, but important for dynamic scenes.

# Lay Summary

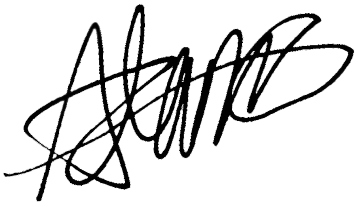
Most of what you can see is split up into visual field regions, with the world around you being at its highest quality in the centre, before gradually blurring the further out your visual field gets. The dead-centre of vision is home to the fovea, or foveal vision, which is responsible for discerning objects of interest in fine detail. If something catches your eye, you may make an eye movement toward it to see what it was. Without the high resolution of the fovea, the world around you will never be seen at a high quality; a trait that is essential for everyday tasks such as visual search. Imagine needing to look for your car keys without foveal vision - a task that is already a nuisance has just been made more difficult. Now imagine this loss of foveal vision being part of your everyday life, as is the truth for those with foveal impairment. This thesis is concerned with the importance of foveal vision with respect to visual search tasks. Seven experiments are presented in this thesis, with chapters 1 – 4 containing a review of the literature, methodologies and a glossary. In Chapter 5, search experiments with targets of varying sizes were conducted to investigate the role of foveal vision on search performance. Foveal vision loss was simulated and then compared with normal scene viewing. The process of looking for targets was found to be nearly identical between the two conditions, meaning that even if the target was very small, foveal vision did not matter. In Chapter 6, we manipulated search difficulty by placing either small or large targets in regions that are either less likely or more likely to be looked at. This was achieved via an algorithmic process. By having two search regions, and two target sizes, we were able to compare both groups to see whether their interactivity affected search performance. Again, foveal vision was degraded and compared with normal scene viewing. When targets were in locations that were less likely to be looked at, larger targets in said regions helped to improve search time. Similarly, targets in regions that were more likely to be looked at, but small

in size were found at a similar time. This interactivity occurred regardless of foveal vision being intact. Chapter 6 also contains results on the stage of search where performance costs originate. For example, does the target and where it is located affect your ability to start searching, the actual process of looking for it, or the process of verifying that it is the correct one? The presented results show the latter stage as being where performance costs originate; however, this was only found when a region just beyond (and encompassing) foveal vision was degraded. Foveal vision loss did not produce such costs for finding and verifying targets. Chapter 7 investigated foveal vision when searching for targets on the move. When moving, the world around you also looks like it's moving which is known as optical flow. The results showed that targets were harder to find when moving. Taking away foveal vision had an even worse effect on search performance, demonstrating that the role of foveal vision differs in accordance to an individuals' motion state (still vs moving), even when the task is the same.

# Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or processional qualification except as specified.

Chapter 5 is based on experiments that I designed and data that I collected and analysed (all with the advice and feedback from my supervisors). The original draft of the chapter was written by myself, but it was then substantially rewritten by A. Nuthmann to submit for publication.



# Acknowledgements

There are a number of people that I would like to thank, so sit tight and prepare yourself. Starting with my fiancée, Monica. You have always been there for emotional support and encouraged me to always strive for improvement. You will always be my inspiration. To my family, who have been the bedrock of everything. Without your support throughout my education, I would not have gotten this far. My friends also deserve my thanks, perhaps for being the constant reminder to finish this PhD.

I give special thanks to my undergraduate supervisor, Dr. Ian van der Linde, for believing in me, and giving me the opportunity to flourish within the academic environment. Also, to one of my mentors Steven Harris, an academic whose wisdom I admire, and who I look up to.

Finally, I would like to give thanks to my supervisors. Professor Robert B. Fisher, who always took a professional stance toward me and made my life at Edinburgh pleasant and enjoyable. And, Dr. Antje Nuthmann, who has been considerate, patient and thought-provoking.

# Contents

# List of Figures

# List of Abbreviations and Acronyms

AA – Anti Aliasing

AMD – Age- related macular degeneration

CE – Compensation Effect

FXAA – Fast Approximate Anti Aliasing

INI / IT – Initiation time

SCAN / ST – Scanning time

T.E.A. – Target Embedding Algorithm

T.E.A.M.S – Target Embedding Algorithm for Moving Scenes

VERI / VT – Verification time

# List of Appendices

# 1 Introduction

This thesis is concerned with the importance of different visual field regions with respect to visual search, salience, and target size; and how these factors interact with one another. These factors are important to study, as knowledge about them and their relation to visual field regions can be utilised to positively affect visual search performance. Additionally, it is important to study these factors in relation to eye movement, as many people lose part of their vision because of ophthalmic diseases. Understanding more of these factors in relation to eye movements can bolster the development of visual aid applications designed specifically to improve a person's quality of life.

Visual search is the process of finding items, which may or may not involve the use of eye movements. A typical search task with eye movements can be broken down into three steps: the initiation phase in which a person begins to move their eyes; the scanning phase, which relates to the process of finding the target; and the verification phase, where a person validates or clarifies that the item they are looking at is the right one. A modulating factor of search is salience, which is

categorised as either top-down or bottom-up. Top-down factors relate to variables such as scene context and objects, which can affect search by guiding a person's eyes towards areas of the scene that seem the most logical (i.e., looking for an aeroplane in the sky rather than on the floor). Bottom-up salience on the other hand, relates to image features and properties that make up objects, such as their colour, orientation, contrast, edges and so on. If an item has a high contrast with respect to the rest of the scene, it is more likely to pop out, which makes it more likely for a person to notice it in their periphery and to then attend to its location. The targets used in all experiments of this thesis are context free. This is particularly important for eliminating most top-down factors such as scene guidance. However, it is important to consider that even context free targets still rely on top-down influence through target template guidance (Malcolm & Henderson, 2009).

In this thesis we claim that foveal vision is not necessary to achieve normal search performance in static scenes when the item to find is context free. We also claim search time is influenced through two interacting factors: target size and target salience. We hypothesised that search time should increase as target size decreases, but if the target is highly salient, this should act as a countermeasure and therefore improve search performance for small targets. Furthermore, we claim that when transitioning to dynamic scenes, the presence of foveal vision becomes a necessity for achieving normal search performance.

The investigations reported in this thesis begin with analysing the effect of target size on visual search. At first, we were interested in the interaction between target size and visual salience. At what object size does visual salience begin to influence search towards a target? To investigate this, a study on the effect of target size

alone was needed to firstly obtain a suitable range of sizes to incorporate. Target size itself was also interesting with respect to foveal vision. At what size does a target become less detectable outside of the fovea, to then require the use of extra-foveal processing? Similarly, are small detected targets identifiable without foveal vision? Chapter 5 presents the experiments done to answer these questions and analyses.

The motivation behind investigating the importance of foveal vision was conceptualised by an ongoing debate in the literature. On the one hand, foveal vision was seen to be important for visual search in alphanumeric displays (Bertera & Rayner, 2000), and not as important for contextually relevant, medium sized objects (Nuthmann, 2014). As such, a paradigm was developed for Chapter 5 that incorporated factors from both studies, to then be used as a foundation for all experiments to follow. The paradigm involved the use of letter targets embedded onto naturalistic backgrounds, causing the letter targets to be naturally context free.

The total search time measurement in all chapters of this thesis is divided into smaller, behaviourally defined epochs known as Initiation Time (IT), Scanning Time (ST), and Verification Time (VT) to more easily identify the origin of any costs or facilitations to search performance. If a reduction in target size affects search performance, it is useful to know where in the search process this originates from. Is it during the very first eye movement (IT), the actual search process for finding the target (ST), the final stage of verifying whether that target is the correct one (VT), or all three? This gaze-based decomposition of search time was first developed by Malcolm and Henderson (2009) and although it is a very useful tool for analysing gaze data, the algorithm does not apply so well when observer

motion is concerned. As a result, a proposed extension to this algorithm has been implemented to accommodate for such scenes (see Chapter 7). Furthermore, Chapter 5 also introduces a novel method for automatically embedding targets into 2D images, the algorithm of which is then used continually throughout this thesis for target insertion.

Given the results from Chapter 5, we hypothesised that there was an interaction between target size and visual salience during visual search. By altering the salience to be high or low, it might be possible to influence search for better or worse. However, upon crossing this with target size a range of outcomes could be possible. If a target is small but highly salient, is it found in the same amount of time as a target that is very large, but lowly salient? It appears that search is very much a top-down task, and thus reliant on factors such as scene context and congruency for an efficient search performance. Therefore, it is reasonable to assume that salience may not have any effect at all. It is also possible that a large target would simply pop out and skew the results. We claim salience affects search more so for smaller targets than large ones, and Chapter 6 investigates this issue.

Target salience was not a manipulated factor in Chapter 5, but we include this factor in Chapter 6. This was not only to assess the effects it had on visual search for context free targets, but also to see whether it interacted with target size in some way. The interaction could shed light on what leads to an efficient search even with degraded vision. However, to draw meaningful conclusions on this part of the investigations, Chapter 6 also includes analyses on the effects of target salience within other regions of the visual field. This was accomplished by obscuring the observer's peripheral vision. The absence of peripheral vision leaves foveal vision intact, which can then be compared against a central scotoma. Due to Chapter 5

already incorporating a foveal scotoma, Chapter 6 extended this by increasing the scotoma's size to encompass a larger radius. The use of a larger central scotoma over a foveal one is motivated by the results obtained in Chapter 5. The effect on search performance without foveal vision in 2D naturalistic scenes is explained in section 5.4.

Chapters 5 and 6 detail the investigation on the importance of foveal vision, as well as the importance of target factors affecting the visual search task within standard naturalistic scene images. Due to these images being 2D, search exists solely on an x,y plane, which means that factors such as image depth, object motion, and self-motion are occluded. Omitting such factors is useful for reducing noise by isolating specific variables for hypothesis testing. However, by omitting them, an experimenter is also neglecting the effects that they may have on the task itself. Moving on from 2D naturalistic scenes, Chapter 7 investigates the importance of foveal vision to a visual search task in 3D dynamic scenes. We claim that search should be harder in such scenes, as interference from observer induced optical flow could obfuscate the target. Searching without foveal vision should also be detrimental, as such obfuscation could cause an increased reliance on high acuity vision.

These dynamic scenes contain no object motion for observer detection, but rather incorporate a simulated self-motion along a predefined path for a fixed length of time. As observers move along the path, they were also instructed to actively search for a specific target, with and without their foveal vision. The experiments in Chapter 7 not only assess search performance along a straight line, but also along a curved one. The purpose of moving on from 2D scenes was partly to increase the

ecological validity of the experiments, and to see whether the results of previous chapters in 2D scenes translated to 3D.

As the scenes operate in 3D space, the original method for target insertion defined in Chapter 5 needed to be extended. This extension controlled for several variables such as the size of the world, target size in screen and world space, positions of other 3D objects and so on. See section 6.1 for a full account on this extension.

Curved scenes were introduced to replicate any effects found along a straight line, and to also test the path's influence on an observer's search performance. Manipulating the path of self-motion was the natural next step in immersive scene study and is the final experimental section of this thesis. Finally, this thesis ends with a concluding chapter (8) summarising the achievements, success and limitations, and future work.

## 1.1 Original Contributions

*Five discoveries* were made over the course of this research. The first was demonstrating the fact that foveal vision is relatively unimportant for localising targets in 2D space, however, search accuracy was impaired when searching for context free targets (Chapter 5) as opposed to contextually relevant objects (Nuthmann, 2014). The second was the apparent increase of saccade amplitude when searching for smaller targets (Chapter 5). Thirdly, an interaction effect between target size and target salience was found which demonstrated that salience can compensate for the difficulty of finding targets as they become small. Similarly, targets that are lowly salient can benefit from being very large (Chapter 6). A major contribution of this thesis is the discovery of where, amongst all three epochs

of search time, search costs originate from, which is the stage of search associated with verifying the identity of targets (Chapter 6). Subsequently, this contribution provides knowledge on the region of the visual field that is necessary to be intact for a normal search performance.

The fifth contribution (see Section 7.6) is evidence for an increased reliance on foveal vision due to a reduced visual acuity of moving targets in 3D immersive scenes. As the scenes used in Chapter 7 include observer induced optical flow, the current hypothesis is that this interference on the visual field caused optic flow conforming targets to blend in with the scene. As a result, the target's visual salience is reduced. This reduction, in combination with the naturally degrading acuity of the visual field, emphasises the importance of the high acuity vision provided by foveal vision.

The results of Chapter 7 are replicated in a final study incorporating immersive scenes along a curved path. The curved path provided a way to optimise target location by having the target move towards, or away from the observer, depending on the side it is on. Although the optimised location improved search performance, it did not improve it enough to eliminate the reliance on high acuity vision. At the time of this writing, immersive scenes constructed in this way have not been used to study regions of the visual field with respect to visual search. Therefore, the results presented here, and the scenes generated can be used as a foundation for subsequent experiments in the area.

*Two algorithms* and *one algorithmic extension* were developed over the course of this research. A novel method for target insertion coined the Target Embedding Algorithm (T.E.A.) was first developed to automatically embed targets into scenes

based on contrast difference values (Chapter 5). An extension to the original gaze-based decomposition of search time (Malcolm & Henderson, 2009) was developed to accommodate for immersive scenes. This extension involved a more sophisticated approach to decomposition via a dead-zone radius around the target. This checked for factors such as fixation locations and saccade direction (see Chapter 7 for a full overview). The T.E.A. algorithm was extended to encompass 3D immersive scenes. The extension provided a method for placing targets spatiotemporally, with an equal target size across all scenes at the halfway point. Furthermore, the algorithm was optimised for target placement along curved lines by calculating equal and opposite signs of optical flow. This adjustment was made to reduce potential noise that could have been generated from unbalanced target placement.

Finally, the 2D static images and 3D immersive stimuli that were developed using the T.E.A. are freely available for subsequent research (see Appendix C).

# 2 Glossary

*Collider* – An invisible object in virtual space moulded to the shape of other objects for collision detection.

*Context free target* – A target based on criteria independent from the semantic and syntactic information of a given scene.

*Foveal Vision* – The central $2^{o}$ of vision responsible for high acuity processing.

*Optical Flow* – The pattern of the apparent motion of objects as a result of relative motion between an obsever and the environment.

*Parafoveal* – The region of the visual field subtending out from the fovea to about $5^{o}$.

*Raycast* – A beam that intersects with virtual mshes to return information about them.

*Rigidbody* – A component of a virtual object that controls its physics.

*Saccade* – A rapid movement of the eye between fixation points.

*Sandbox* – A large open virtual space for content creation.

*Scene guidance* – A combination of a scene's visual features and contextual information in addition with semantic and syntactic information to guide eye movements during a visuo-cognitive task.

*Scotoma* – A partial loss of vision or blind spot in an otherwise normal visual field.

*Spatiotemporal* – Belonging to both space and time.

*Target template* – An observer's mental representation of a target to guide their eye movements during search tasks.

*Visual Salience* – the perceptual quality of an item that makes it stand out relative to its neighbours.

*Z-Dimension* – a 3-Dimensional axis that describes scene depth.

# 3 Methodology and Infrastructure

This chapter describes the common methods that are present throughout each experiment in this thesis.

## 3.1 Participants

The participants used in all the experiments that follow did not differ in any systematic way. All participants were drawn from the student population of the University of Edinburgh, who were aged between 18 and 26 and studied a variety of academic disciplines. The mean age varied slightly from study to study, however, gender remained relatively equal. Furthermore, this thesis will not consider any effects of gender.

All participants reported to having normal or corrected-to-normal vision. No eye makeup nor mascara was worn during experimental procedures and those that yielded very high error rates and/or missing data were replaced. In such events, the participant was paid an inconvenience allowance for taking part, which was equivalent to £7 per hour (with the exception of the shorter experiments 5, 6, and 7 that yielded £5 per 30 minutes).

Participants signed a written consent form prior to the experiment. In this form, they were made aware that they were permitted to leave at any time without

question. All studies were approved by the Psychology Research Ethics Committee under reference number 138-1415/1 and conformed to the tenets of the Declaration of Helsinki.

## 3.2 Stimuli

Experiments 1 – 4 used the greyscale equivalent of naturalistic photographs of scenes in order to study the role of various regions of the visual field in relation to a real-world environment. The target itself was embedded via an algorithmic process that is discussed in Chapter 5.

Experiments 5 – 7 used 3D, dynamic virtual world environments created using the Unity3D game engine. The targets were embedded spatiotemporally using an extension of the algorithm created for the first four experiments (see Chapter 7). Chapter 7 of this thesis elaborates on the development of the stimuli and the manipulations used.

## 3.3 Eye tracking apparatus and methodology

### 3.3.1 Apparatus

The eye tracker used was the SR Research EyeLink 1000/2K Desktop mount. This particular piece of hardware allowed for binocular tracking. This was important because all of the experiments reported here simulated central (and foveal) visual field loss; therefore, it was crucial that the simulation moved concomitantly with the participant's gaze. Taking the average coordinates of both eyes as was adopted in previous research (Nuthmann 2013, 2014, 2015) was the opted for approach in

all experiments presented in this thesis, as opposed to recording one eye. By recording only one eye the simulation may appear slightly offset to the participant, which can lead to inaccurate and unnatural behaviour. This is particularly important if an observer's calibration reduces in accuracy for a given trial. Taking the average of both eyes will keep the gaze contingent scotoma relatively in the centre, as opposed to being locked to one particular eye. A 21-inch CRT monitor was used for stimulus presentation, which had a refresh rate of 140Hz.

### 3.3.2 Calibration Procedure

Participants had their eyes calibrated using the standard 9 dot calibration routine with the EyeLink 1000. The procedure consists of a series of 9 dots that are displayed at each corner, side and centre of the screen. The participants needed to fixate each dot in turn once for calibration, and then a second time to validate the calibration's accuracy. If the validation procedure produced a mean error that was greater than 0.5°, then the whole calibration procedure was repeated. Corneal and pupil reflection thresholds and biases could be manipulated in order to optimise the camera for tracking.

During the experiment, if the pupils showed signs of drift due to head movements or other means, the participant was brought back to the calibration screen to adjust for this.

### 3.3.3 Data Analysis

The company supporting the EyeLink 1000 also manufacture the Data Viewer software; which can convert the raw data obtained by the eye tracker into a fixation sequence matrix. Other reports were also provided, but the fixation report was

primarily used. For all analyses, a custom interest period was defined within Data Viewer in correspondence with trial start and end points defined within the Matlab experimental code. The reason being to accommodate for factors such as fixations prior to scene onset (central fixation cross check), recalibrations throughout a given experiment, and any fixations that may get recorded after scene termination. Finally, adding the newly defined 'interest period start time' into the fixation report, the current start and end fixation times can be calculated for accurate analysis.

Data was analysed primarily by two-way and one-way (see Chapter 7) ANOVAs (analysis of variance) with repeated measures implemented in the R statistical computing software (version 3.2.3; R Development Core Team, 2015) using the ez package (Lawrence, 2015). Partial Eta-Squared effect sizes are reported throughout, with Partial Omega-Squared values available in Appendix G. Single and multiple regression analyses were also performed in combination with post hoc pairwise comparisons (see Chapter 6).

When at times Mauchly's test indicated that the assumption of sphericity had been violated, a Greenhouse-Geisser correction was then implemented to adjust the degrees of freedom. Error bars presented for all graphs were constructed using the method proposed by Cousineau (2005) to correct for within-subject designs.

# 4 Literature Review

The review of the literature that follows demonstrates that no one has investigated in a systematic way how the importance of foveal vision during a search task varies according to the stimulus environment; and how foveal vision is influenced by visual salience. Although there have been numerous studies involving foveal vision and visual search, conclusions remain contradictory. In this thesis, we claim that foveal vision, as studied in the traditional sense with 2D static scenes, is relatively unimportant when it comes to finding targets. Furthermore, we also claim that visual salience can influence search performance when scene context is unavailable. Finally, the role of foveal vision is argued to become more apparent when changing the stimuli from 2D naturalistic scenes, to 3D immersive scenes, despite having no change in the task itself. Therefore, in this chapter we will review the literature on attention and the visual field, visual search, dynamic scene studies, and visual degradation.

Historically, visual search has been viewed as a top-down task, i.e., an individual largely relies on context and semantics to find what they are looking for. For example, if you are looking for a frying pan in your house, you wouldn't start by searching in the bathroom. Aside from context and semantics, people engaging in visual search may also rely on built up representations of what the target may look like in their mind. Such target templates (Malcolm & Henderson, 2009) can guide search toward areas that conform with what the target is (e.g. a mug on a desk) and help the observer to filter out distractor targets. This filtering is relevant as search studies may not necessarily involve contextually relevant targets (e.g. letters

in random displays) but are still considered top-down due to factors such as target template guidance. Using context free targets is advantageous, as they minimise the amount of top-down information available to the observer, but they do not eliminate them. Furthermore, if someone were to undertake search tasks under visual impairment, one may find the task more difficult as one can no longer rely on the different regions of the visual field to the same calibre as if one was normally sighted.

Central vision, for example, is useful for verifying objects by bringing them from the periphery into central view. Everyday tasks such as buttering bread requires such verification, as one must first locate the items using peripheral vision before performing the task of spreading the butter. Naturally, once one has found the butter, the use of central vision is needed in order to spread it over the bread effectively. It seems intuitive that the impairment of this region will lead to elongated time spent on the task, and perhaps a reduced quality in its output. To aid this assumption, the change in eye movement and fixation location behaviour has been investigated previously in natural tasks (Hayhoe et al., 2012; Hayhoe & Ballard, 2005). Locating the items would involve context about the environment, and some logical sense. From this, search looks to be very much driven by top-down procedures.

What role would bottom-up features have in this case? If we associate certain items with certain environments, then we only need to rely on information such as scene knowledge and context to find them. Research in favour of top-down control over bottom-up salience such as Malcolm and Henderson (2009) claim that visual salience is overwritten due to cognitive relevance. However, the results of this thesis

provide evidence for bottom-up factors facilitating search performance in naturalistic scenes (see Chapter 6).

## 4.1 Attention and the Visual Field

The eye constantly samples information from the environment to project onto the retina, giving us a visual field in which we see the world. We can divide this visual field into different regions, starting with the point of highest visual acuity, foveal vision. This centrally-located region encompasses about 2° of your vision and is the part of the retina with the highest cone density (Palmer, 1999). From this, subtending out to about 5° is the parafoveal region and everything beyond this is referred to as the periphery. Several experiments in this thesis are concerned with the role of foveal vision in visual search (see chapters 5, 6, and 7). For experiments related to other visual field regions, see Chapter 6.

The fovea is comparatively small with respect to the rest of the visual field, and so much of what we want to see is within our peripheral vision. To view these areas of interest in fine detail, we direct our eyes towards them via ballistic eye-movements known as saccades. Depending on one's current task, after a saccade, the newly attended area may contain some aspect that requires the eyes to remain relatively stable in order to discriminate it; this state is defined as a fixation. The longevity of fixation durations is reflective of our current cognitive processing, and it is apparent that complex or difficult to see objects require longer fixations (Henderson, Weeks, & Hollingworth, 1999; Snowden, Thompson, & Troscianko, 2006).

Saccades happen often, with a frequency of 3-4 per second (Rayner, 1998) and are extremely fast, with peak velocities of around $500^{o}s^{-1}$ (Becker, 1991). Furthermore,

their properties vary depending on the task (M. S. Castelhano, Mack, & Henderson, 2009) and are constrained by various oculomotor regularities (Tatler, 2007; Unema, Pannasch, Joos, & Velichkovsky, 2005). The ability to decide where to saccade to next and when are two very important aspects of eye-movement research, and how the visual system guides these movements is a fundamental goal of real-world search. The 'when' and 'where' of saccades are well investigated topics (see Rayner 1998 for a review), with many different views arising as to what drives these spatio-temporal elements; Although, it has been shown that these decisions within the oculomotor system are largely independent (Findlay & Walker, 1999). The experiments in this thesis are focused on where we look within each scene, and the areas of the visual field that contribute to this process.

Although eye-movements happen frequently, they are not directed randomly. In fact, research shows that they are guided in accordance with the current task and environment, and several theories of attention have been developed over the years in attempts to demonstrate this (Julesz, 1984; Treisman & Gelade, 1980; Wolfe, 1994).

Scanning for an item in a scene requires attentional mechanisms to shift from one region to the next. The extraction of item information has been debated to be either serial, where items are selected sequentially and processed one at a time; or parallel, where items are processed by the visual field globally. The idea that information is processed serially has been investigated through methods such as reading (Reichle, Vanyukov, Laurent, & Warren, 2008), and the processing of targets amongst distractors of varying set sizes (Treisman & Gelade, 1980). The hypothesis for a serial deployment of attention is that the more demanding a scene is to process, the longer it will take to detect the current item of interest upon

inspection of items in an accept/reject manner. A classic lab controlled visual search task to test serial shifts of attention typically utilises object arrays with the target object and a set of distractor items. Observers are normally required to state whether the target is present or absent within each array. Supporting evidence for a serial search is the finding that reaction time increases linearly with set size and that there is a 2:1 ratio of target-absent slopes to target-present slopes (Treisman & Gelade, 1980). This makes sense since absence requires that an observer looked at all targets, whereas presence means that an observer will look at half before finding the target. This is not a cause for concern with respect to the experiments in this thesis however, as there are no target-absent trials. Target absent trials are usually included to prevent instantaneous detection of the target, but the experiments in this thesis control for such cases through algorithmic insertion of the targets (see Section 5.3.5).

Typical serial search displays, however, are not truly representative of the real world. Not every object we see is homogeneous with the target in question; if we are looking for a person in a park setting we would not search all the trees one at a time throughout the process. Models of parallel processing state that basic features about the environment such as colour, orientation, and motion can be extracted in parallel across the visual field (see Figure 1).

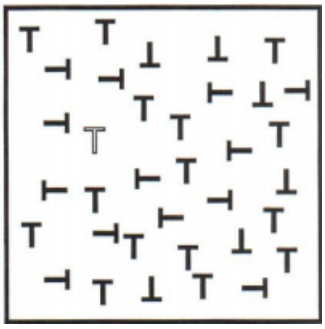


***Figure 1.*** **Parallel feature search array (Wolfe, 1992)**

We can see from Figure 1 that searching for a target that is defined by a single feature is simple enough to find regardless of the distractor count. All of the items can be processed simultaneously. Other arrays are more complex and require observers to search for a target among similar distractors, such as a 'T' among 'L's. Even though at times the 'T' is easily distinguishable from the 'L', the search process is random as the observer moves from item to item, dismissing incorrect targets in a serial, self-terminating manner. However, it has been shown that in most cases, the use of parallel processing guides the search process (Wolfe, 1992). As shown in Figure 2, we see that the search for a white letter 'T' is augmented by a parallel process for colour. Due to this, we simply ignore all the black items, effectively guiding search to the white items. Knowledge of search guidance in contrast to a target's features is of general importance for this thesis, as all the experiments conducted use context free targets embedded into static and dynamic environments.

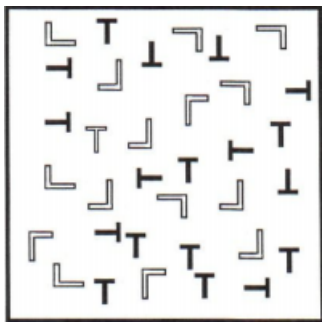


***Figure 2.*** **Guided search for a white letter 'T' by parallel processing colour (Wolfe, 1992).**

It is relevant to note that linear search functions with a variety of slope ratios for serial search is also found in parallel search models (Townsend, 1990). The 2:1 ratio pattern also hasn't always been found (Wolfe, 1998). It has been shown, however, that by varying the orientation of targets and distractors and by changing their locations periodically, parallel processing is disrupted thereby forcing observers

to engage in serial search. There have been studies for (Horowitz & Wolfe, 2003) and against this (Kristjánsson, 2000); however, it is likely the case that there is indeed direct evidence for serial search as there was a successful replication under an improved randomised search paradigm that accommodated for the experiments by Kristjánsson (2000) (See Horowitz & Wolfe, 2003, for discussion).

This debate between serial and parallel processing has been greatly influenced by the *Feature Integration Theory* (FIT) (Treisman & Gelade, 1980). The theory relies on a two-stage process, the first of which states that there is an automated registration of features that happens early and in parallel. Stage two occurs later, whereby objects are identified separately. Due to this, a debate has risen between being in favour of either serial or parallel search. This was an unintended consequence of the model; as there are not two clear classes of search. As explained by Wolfe (2000), there is no value that can be derived from the continuum of slopes from efficient to inefficient in order to place them into two groups.

A later model known as the Guided Search Model (Wolfe, 1994) extended the two-stage paradigm of the FIT by including the appearance of a target as a factor of search guidance during the serial stage. Wolfe's search model gave rise to a significant interest in target template guidance (Malcolm & Henderson, 2009) with the suggestion that a target's visual features gathered through parallel processing would then be matched with the target template's visual features during the serial stage. This is the basis for search guidance and implies that top-down processes facilitate search.

With regards to the division of attention in the literature, experimenters have made claims to the evidence of either category in the form of covert vs. overt attention.

Covert attention is the processing of a location or object that occurs whilst the eyes remain fixated elsewhere. The ability to covertly attend to peripheral locations of interest has been investigated from as early as 1890 (James, 1890). On the contrary, overt attention is the ability to process a location or object through eye movements. These lead to the distinction between exogenous attention, which is bottom-up or stimulus-driven, and endogenous which is more goal-driven or top-down (MacLean et al., 2009).

Carrasco & McElree (2001) found that the use of covert attention enables an efficient extraction of information from a dynamic environment and that peripheral cueing improves discriminability. They state that their findings and pattern of results provide evidence for parallel processing and that the speed of which observers shifted their attention from one location to another could not be attributed to serial processing. This as they put it, is due to peripheral cueing causing attention to be deployed at the relevant location on display onset rather than shifting attention and then processing the target.

Other than in lab controlled experiments, it has been acknowledged that covert shifts of attention could be rare in the real world, and even when they occur they would precede an eye movement; except in scenarios such as complex object search as argued by Findlay, Brown, and Gilchrist (2001) who suggested that likely target locations are obtained during the first fixation using parafoveal preview. Furthermore, attending to a spatial region covertly and then planning a subsequent eye-movement has been suggested to be one and the same by theories such as the premotor theory of attention (Rizzolatti, Riggio, Dascola, & Umiltá, 1987). This is a controversial theory, however, and has been questioned as a result (Smith & Schenk, 2012). It is important to acknowledge the covert attention literature

described here; however, the experiments in this thesis all utilise eye movements, and therefore pay more tribute to the overt attention literature.

Turning now to the types of visual-cognitive experiments in the lab, a common task to analyse eye-movements is the free viewing task. This requires the subject to naturally gaze about the scene for a finite set of time. These tasks are effective in analysing gaze patterns and their relation to the environment. For example, there have been a number of studies that have looked at the role of low-level visual features during a free viewing task, and have demonstrated that a higher magnitude of such features are present at fixated, as opposed to non-fixated locations (Baddeley & Tatler, 2006; Foulsham & Underwood, 2008; Mannan, Ruddock, & Wooding, 1995; Tatler, Baddeley, & Gilchrist, 2005).

Smith and Mital (2013) performed two experiments whereby the participant was instructed to either view a scene freely, or to identify the location depicted in each stimuli. Their results provided evidence for gaze clustering across multiple viewers using a term they dubbed as attentional synchrony.

Viewing a scene freely has disadvantages. Without a specific goal in mind, it is often difficult to temporally measure oculomotor behaviour. Other scenarios such as visual search is useful for this, as the goal of finding the target must be completed within a certain time; people engaging in the task will also seemingly adapt their gaze behaviour to aid their search.

With visual search, we can also measure the role and importance of various visual field regions by masking them. I now turn towards visual field degradation studies,

starting with the foundational concepts which gave rise to gaze-contingency; and, which in turn, is used in all experiments presented in this thesis.

### 4.1.1 Artificial degradation of the visual field

Individual regions of the visual field make different contributions to a person's everyday tasks. By studying these regions in isolation, we can identify these contributions, and their relative importance to tasks, such as reading and visual search. This was the foundation for the development of the gaze-contingent moving window (McConkie & Rayner, 1975) and moving mask paradigms (Rayner & Bertera, 1979). The paradigms can be defined as a method of degrading a defined region of vision to the extent that the desired information is no longer available for the task (Loschky & McConkie, 2002; Larson & Loschky, 2009).

Caldara, Zhou & Miellet (2010) applied the moving window paradigm to face recognition, demonstrating that an observer's fixation strategy changes depending on the visibility of facial features. They described it as viewing the scene through a 'spotlight', whereby a region centred on the fovea moves concomitantly with an observer's gaze, with outside content visually degraded.

The moving mask paradigm has also been applied to tasks such as reading (Fine & Rubin, 1999), and scene and object perception (Henderson, Mcclure, Pierce, & Schrock, 1997; Larson & Loschky, 2009; van Diepen, Ruelens, & d'Ydewalle, 1999), with its name commonly referred to as a 'blindspot' (Miellet, Zhou, He, Rodger & Caldara, 2010) or 'scotoma' (Cornelissen, Bruin, & Kooijman, 2005). The terms 'blindspot' and 'scotoma' will be synonymous throughout this thesis, with gaze-contingency being a core experimental technique applied to all experiments in subsequent chapters.

Moving mask paradigms were not applied much over recent decades due to technical limitations. Developing a blindspot that could move with gaze without any noticeable lag was difficult, and not to mention retinal instability. Collectively, poorly implemented moving masks would result in a distracting, jittering scotoma. Methodological papers have addressed these problems (Aguilar & Castet, 2011; Geringswald, Baumgartner, & Pollmann, 2013), with eye-tracking technology advancing over the years to accommodate for these pitfalls. As explained by Nuthmann (2014), the SR Research EyeLink 1000 used in this thesis is suitable for gaze contingency due to its binocular recording, and high sampling rate of 1,000 Hz per eye (see Chapter 3). A more detailed review of foveal analysis can be viewed at section 4.2.

One of the first research fields to incorporate the above paradigms was reading (see Rayner, 1998, for a review), but it was not until 1994 when van Diepen, De Graef & Van Rensbergen pushed for the application of such paradigms to real-world scene perception. One purpose of this thesis is to identify the importance of visual field regions to the sub-processes of search. The next section summarises areas of both basic, and real-world visual search to provide a rationale for the thesis.

## 4.2 Foveal Vision and Age-related Macular Degeneration (AMD)

AMD is a medical condition whereby the central part of a person's retina (macula) has been damaged. This leads to a severe degradation or loss of central vision (scotoma). Because of this, the accomplishment of everyday tasks becomes difficult which can reduce the quality of life of those affected. In relation to this thesis, by

simulating the loss of foveal (see Chapters 5, 6, 7) or central (see Chapter 6) vision [The simulated loss of foveal vision can be achieved through the moving-mask and moving-window methodologies introduced in section 4.1.1], we can further understand the importance of visual salience to search.

A similar hereditary condition called macular dystrophy, or juvenile macular degeneration (JMD) also exists, but typically affects children and young adults. Due to having the condition from a young age, those affected typically develop gaze strategies that accommodate for their vision loss. Patients who utilise their peripheral vision through what is known as Eccentric Viewing are able to shift their vision slightly as a means of 'looking around' the degraded central scotoma region. Adopting this new Preferred Retinal Location (PRL) or pseudofovea, of which has been investigated extensively by Barraza-Bernal, Rifai and Wahl (2018), gives observers an entirely new oculomotor reference frame that provides a sharper, better vision than the diseased fovea.

In 2013, Van der Stigchel and colleagues investigated the eye movement behaviours of observers with JMD during a visual search task. Throughout the task, participants were required to locate either a circle or a dot amongst a series of c-shaped distractor items. The search task was split into serial and parallel conditions (see Section 4.1) to analyse the search behaviour of patients with a PRL. Their results revealed that searching with a PRL-based reference frame lead to an increased number of eye movements before target location, and decreased saccade amplitudes toward the scotoma. The authors reasoned that this decrease in search performance could be related to the time taken to peripherally process the target (as opposed to foveal processing). Performance loss with foveal and central scotomas has been demonstrated elsewhere (Bertera & Rayner, 2000; Cornelissen,

Bruin, & Kooijman, 2005; Lingnau, Schwarzbach, & Vorberg, 2010; Lingnau et al., 2014), in addition to comparisons with peripheral loss (Le Callet et al., 2018), however, as mentioned by Foulsham et al., (2009), patients with visual agnosia based their gaze behaviour more upon raw saliency than target knowledge. Could the influence of visual salience affect search performance for the visually impaired? This question is further investigated in Chapter 6 by simulating the degradation of visual field regions.

## 4.3 Visual Search

When eye-tracking technology was made available to researchers, it was almost immediately applied to pictures of scenes (Buswell, 1935; Yarbus, 1967). Since then, over a million trials have been analysed on the topic (see Wolfe, 1998) making it one of the most investigated tasks in cognitive science.

A typical search task would normally present a series of items with a target to find. A participant is then required to report on whether the target was present or absent. To distinguish the target from its distractors, participants are required to match their knowledge about the target against the features within the search display. Thanks to the wealth of knowledge discovered over the years by the visual search community, out of all the behaviours of visual cognition, search has become one of the best understood.

Visual search is a very useful tool in understanding the underlying mechanisms behind behavioural tasks such as selection and recognition. A considerable amount of literature has been published on visual search theories, the most relevant to this thesis will be discussed here.

When looking for an item, we may glance over objects in the process, however, previous studies have reported that the way in which we interact with each item is contingent on the task at hand. For example, if I were to assemble a bed I would recognise many of the tools required around me, but I would select the right one for my current situation. This is a key aspect to the development of theories on selection for action (Allport, 1987; Neumann, 1987; Norman & Shallice, 1986). Breaking this down further, since selection of items changes over time, the task associated with it should be a sub-task of the current assignment. Separating the assignment out in this manner gives each sub-task its own selection requirements (Hayhoe & Ballard, 2005; Land & Hayhoe, 2001).

Theories on recognition, however, tend to neglect the above theory by instead focusing on covert processes of attention for the development of high-speed serial search models (see Ward, Duncan, & Shapiro, 1996, for discussion). Selection in these models serves more as an internal director that points to locations in space, rather than an oculomotor one. This is supported by research that has demonstrated the fact that attention can shift covertly (e.g., Klein, 1980; Klein & Farrell, 1989; Murthy, Thompson, & Schall, 2001; Posner, 1980).

Despite the above view, people almost always move their eyes as they search whilst also directing their gaze to a target to verify their decision (Geisler & Chou, 1995; Zelinsky, 2008). There is an unambiguous relationship between the search process and motor behaviour (Henderson, 2007; Findlay & Gilchrist, 2003), with saccadic eye movements acting in support of visual search (see Findlay & Gilchrist, 2003; Rayner, 1978, 1998; and Viviani, 1990, for reviews).

The act of searching covertly has been challenged in recent years by highlighting the importance of retinal acuity limitations in the shaping of visual search behaviour (Zelinsky 2008). The natural decline in peripheral acuity can be compensated for by the inclusion of eye movements (Geisler & Chou, 1995). It has been suggested that future models of covert search should adopt a simulated retina as a standard component.

Turning now to overt attention, researchers have been able to study its effect on search performance by designing object arrays that isolate items from a semantically coherent background to a simple display with no background. The next section discusses the benefits of overt search tasks within such stimuli.

### 4.3.1 Basic stimuli and Overt Search

Simple displays with basic stimuli have paved the way for understanding the underlying mechanisms that govern attention and search performance. Many of these search tasks have been concerned with factors that affect target acquisition and detection (Treisman, 1988; Wolfe, 1998). Overt search tasks allow participants to bring areas of interest into the high-resolution view of the fovea. This not only allows for fine detail processing, but it can also facilitate the decision-making process by the participant (Malcom & Henderson, 2009). This is partly because detecting targets that are further away from the starting fixation becomes increasingly difficult (Carrasco & Frieder, 1997; Engel, 1977; Scialfa & Joffe, 1998; Yeshurun & Carrasco, 1998) This is commonly referred to as the eccentricity effect. This thesis concentrates solely on overt search tasks, whilst acknowledging the covert attention literature.

Since searching overtly includes a motor component, such tasks allow us to better understand the mechanisms behind eye movement control. Characteristics such as where we look (fixation location) and when we look (fixation durations) at any given time are some factors that can be measured through overt search tasks.

Different search tasks can relate to the eccentricity effect in numerous ways (Carrasco et al., 1995; Scialfa & Kline, 1988; Scialfa et al., 1987). Scialfa and Joffe (1998) showed through a series of studies that conjunction search required a narrower attentional focus than feature search, and that eye movements were not needed in a feature search task but where beneficial in conjunction search. Their results were consistent with the hypothesis that the relationship between conjunction search and the eccentricity effect is linked either to the lower spatial resolution or increased lateral inhibition of non-central vision.

The above highlights the importance of target features in visual search and the same seems intuitive for real world objects. For example, identifying the serrated edge of a bread knife amongst other cutlery before slicing the bread. But even this can prove difficult if the vicinity is very messy, our attention could be shifted elsewhere, causing our eyes to go somewhere we did not intend them to go (Bridgeman & Stark, 1991; Theeuwes, Kramer, Hahn, & Irwin, 1998).

Basic stimuli have some advantages over complex real-world scenes for analysing target acquisition due to the targets typically being harder to discern from one another. For example, a simple display may have a series of O's among Q's (Treisman & Souther, 1985) that are only distinguished by the tail of the Q, whereas real-world objects are rather different with respect to other items in the scene. Highly similar simple targets force the participant to make fine visual

discriminations in order to locate the correct target. Due to this, as the number of distractors in the array (set size) increases, the longer it takes to find the target. This is commonly referred to as the set size effect (Palmer, 1995). Furthermore, these targets are context free, meaning that if they were to be embedded in a real-world scene, a participant cannot rely on the surrounding environment to locate them. The targets used throughout all experiments in this thesis are context free.

To accompany the above, I now turn your attention back to Wolfe's Guided Search Model. There have been numerous revisions to the model, with the latest being its fifth iteration (Wolfe et al., 2015). However, for the time being I will focus on its third (GS3) (Wolfe & Gancarz, 1997) and fourth (GS4) (Wolfe, 2010) iterations. GS3 integrated eye movements and eccentricity effects to make the model more comparable to human search behaviour. The model was set against a series of search tasks such as feature search, using basic stimuli, and was found to mimic human data. GS4 expands on this with the hope that the results of the model scale up to natural and artificial search tasks performed by people in the real world. According to Wolfe (2007), a good search model should account for the following phenomena:

- Set size
- Presence/absence
- Features and target-distractor similarity
- Distractor heterogeneity
- Flanking/linear 'separability'
- Search asymmetry
- Categorical processing
- Guidance

The important point here is that each of the above require some form of object recognition, whether to accept the current target as the desired one, reject it, or to confirm the absence of a target. This recognition does not happen instantaneously when the target is selected, as previous versions of the GS model suggest. Instead, parallel processes in early vision transfer information to object recognition processes. Verifying an object against a set of stored representations takes time and it is important to consider search time overall as a set of sub-processes (see section 4.5). Experiments 2, 4, and 6 in this thesis investigate the verification process in more detail.

In the above paragraphs, I have explained how basic stimuli have been used to measure the behaviours of visual search and how target features, as well as the type of search task can make for an efficient or inefficient search. However, we do not search through basic stimuli in the real world. Often, the item to find is contingent with the task at hand, such as wondering when your plane is going to arrive. The next section will discuss visual search in relation to real-world scenes.

### 4.3.2 Real-world Scenes

Real-world scenes comprise of background elements in conjunction with discrete objects that are spatially organised in a semantically coherent manner (Henderson & Ferreira, 2004; Henderson & Hollingworth, 1999). These scenes might include the 'whole picture' (e.g. a back garden), or a subset of that picture (e.g. the shed in the garden). Because of this hierarchical structure, one can zoom in and out to generate a set of scenes from just one environment. Previous literature has expressed a degree of control in this regard however, by using the notion of human scale to exclude unnatural vantage points that would not typically be encountered in day to day interactions (Henderson & Ferreira, 2004; Henderson & Hollingworth,

1999). For ease of reading, the terms 'real-world' and 'naturalistic' shall be used interchangeably throughout this thesis.

Naturalistic scenes are usually captured in photograph form to produce static 2D representations of the real-world. Other depictions such as drawings and computer-generated graphics are also considered to be reasonable substitutes. All such scenes should conform to a set of general rules (see Biederman et al., 1982):

- Objects must abide by the laws of physics, for example a person cannot float unsupported in the air.
- Objects should be constrained by semantics, for example a cooker is not generally found in the shower, nor would you find a chair on top of a wardrobe.

Static representations capture important properties of scene perception such as scene semantics, and visual complexity, whilst omitting factors that are harder to control for (e.g. scene depth and motion). Such omissions could be disadvantageous, but it very much depends on the question being asked (see Chapter 6). A wealth of information has been gained from using such stimuli but previous research has investigated eye movements in relation to the image, rather than the immersive structure of the real world. Experiments 1 – 4 in this thesis use naturalistic photographs of scenes whilst Experiments 5 – 7 use computer generated 3D environments.

The use of real-world scenes has shown that where we place our eyes in a given scene is heavily dependent on the task. Castelhano and Heaven (2010) examined the contributions of target features and scene context to search performance. Their results show that the locating of targets or *attentional guidance* is affected by both

variables. In their study, scene context was given either by word or picture cue prior to scene onset. They predicted that search performance should not differ across conditions if sufficient information was acquired through the name-cue. What they found was an additive effect on attentional guidance through scene context and target features. Additional knowledge about the target features also showed to improve search performance. They concluded that both scene information and target information must be considered when assessing components that affect overall search performance.

Prior knowledge about the target that contribute to improved search performances is the result of target template guidance. It has been shown that searching repeatedly within an object array leads to shorter search times and a more efficient search (Chun & Jiang, 1998; Jiang & Wagner, 2004). However, upon searching within real-world scenes, guidance is in fact derived from global features (Brockmole et al., 2006; Brockmole & Henderson, 2006). It has been demonstrated that search performance can change simply by making the background of a scene available to an observer (Wolfe, 1994; Wolfe et al., 2002). It has also been demonstrated that combining a target template with the availability of scene context can greatly increase search performance (Malcolm & Henderson, 2009).

Previous literature has shown that visual search is dominated by top-down processes (Ehinger et al., 2009; Einhäuser et al., 2008; Foulsham & Underwood, 2007; Henderson et al., 2009; Turano et al., 2003; Zelinsky, 2008; Zelinsky et al., 2006). Although it feels intuitive to believe that where we look is guided by objects themselves, it was later suggested that they are not better predictors of fixations over early saliency (Borji, Sihite, & Itti, 2013), however, objects were still reported to play a role in guiding attention. The debate on whether attention is object-

driven (top-down) or stimulus-driven (bottom-up) has been ongoing for some time, and is hard to resolve due to the many confounding effects that arise from various forms of scene guidance (Biederman, Mezzanotte, & Rabinowitz, 1982). Although there have been some recent attempts at reconciling the two views (Stoll et al., 2015).

If visual salience is a good predictor of search, then what are the supporting mechanisms for this claim? The next section will detail visual search studies in support of the visual salience hypothesis.

## 4.4 Visual Salience

Could low-level features guide eye movements? Measurements of image statistics have shown that local measurements such as contrast are predictive of fixation location (Mannan, Ruddock, & Wooding, 1996; Zetzsche, 2005) due to those local measurements being higher at fixated regions as opposed to randomly selected regions. Other investigations of fixations and image features have also shown correlates (Reinagel & Zador, 1999; Parkhurst & Neibur, 2003; Tatler, Baddeley, & Gilchrist, 2005; Frey, König, & Einhauser, 2007). Of course, correlations do not equal causation. Searching for a target whose location contains a high number of visual features, in an environment that is heavily influenced by top-down control, will almost certainly produce a correlation. The alternative to analysing image statistics is to compare model-predicted fixations against human fixated locations.

The most widely used computational model for such an approach was developed by Itti & Koch (2000). This model of visual salience was first introduced as a model of orienting in search. The performance of the model when searching for targets among distractors was similar, and sometimes better than human searchers. For

this reason, the act of defining a task for participants to complete over simply free viewing the scene is a good method for analysing the interaction between bottom-up saliency and top-down control.

How visually salient something is can be defined as how well it stands out relative to its neighbouring regions. According to the visual salience hypothesis, properties of the stimulus such as intensity, colour, and edge orientation aid determining where we look in a given scene. The concept, which was first introduced theoretically, by Koch and Ullman (1985), has given rise to many computational models of visual attention that specialise in bottom-up salience (see Borji & Itti, 2013 for a review).

Evidence for the role of visual salience in visual search tasks using naturalistic-scenes has shown that highly salient objects are fixated early and often [in a memory-encoding task] (Underwood et al., 2006). Underwood et al., (2006) also showed that search for specific non-salient targets resulted in fewer fixations to the above regions. Underwood et al., (2006) suggested that top-down control supersedes visual salience in search by bypassing the computation in a term they coined, 'cognitive override'. Zelinsky et al, (2005) also suggested the same by zero-weighting regions in each scene.

A comparison between medium and low salience targets has also shown that low salience targets are fixated less often in an encoding task (Foulsham & Underwood, 2007). Such studies usually manipulate the saliency of targets which has the potential to introduce confounds. Image artefacts introduced because of post-hoc editing as is the case in studies such as Henderson et al., 2007 could be identifiable by an observer, and interfere with the overall naturalness of the scene. Other studies

that do not embed targets into scenes but instead calculate salient positions often do not account for the change in salience that occurs once the target has been inserted. Foulsham & Underwood (2007) calculate medium and low regions of salience based on salience map peaks and use this information to place targets (a piece of fruit) in the scene themselves (without image editing). This point is discussed in more detail in Chapter 5, along with a proposed solution to this problem.

There have been studies that analyse the role of target salience without such manipulations (Foulsham & Underwood, 2011), however, areas of the scenes were manipulated in their second and third experiments. Foulsham and Underwood (2011) concluded that visual salience is a good predictor for knowing which targets are found the quickest during visual search. They reasoned that this result could be caused by other properties of the target, and that scene knowledge about what and where targets are located could be responsible for the advantages salient regions give in search.

The role of visual salience could be better measured by eliminating factors such as scene knowledge and contextual guidance. One such method would be to use context free targets through post-hoc editing of the scene. A disadvantage to this approach is that it makes visual salience of the target hard to control for, as the saliency of a region would change upon insertion of the target. See Chapter 5 for a solution to this problem, and Chapter 6 for experiments involving visual salience manipulation.

There have been several authors who have argued that top-down control dominates visual search, and that visually salient, but irrelevant, items rarely distract observers

(Chen & Zelinsky, 2006). Similarly, with everyday tasks such as walking (Jovancevic, Sullivan, & Hayhoe, 2006; Turano, Geruschat, & Baker, 2003), visual salience has been shown to explain little of the variance of fixation locations (See Chapter 6 for arguments against this).

The above discussion relates to normally sighted observers, with no interfering degradation of their visual field. However, studies have shown that eye movement strategies change when parts of the visual field are obstructed. Foulsham et al., (2009) investigated task performance with a patient afflicted with visual agnosia. Their results revealed that, due to visual impairment, the patient inspected areas of the scene that were based more upon raw saliency, as opposed to prior knowledge about the target. Their study demonstrated that a change in eye movement behaviour also changes the importance of visual salience.

The above finding is particularly relevant for clinical studies not just with visual agnosia, but all manner of visual impairment. The next section will introduce Age-related Macular Degeneration (AMD), and how it affects the everyday tasks of those with it.

## 4.5 Dynamic Scenes and Virtual Environments

Everything mentioned so far about visual attention has been attributed in some way to static environments. Although static environments are useful for many experiments, they are limited with respect to depth, motion, and immersion. Viewing a picture of a kitchen is different from being in the kitchen as you are limited to one vantage point. Similarly, the loss of motion causes a change in eye movement behaviour as motion drives attention, however, this can be advantageous depending on the hypothesis being tested. Since most cases can be translated just

as well to static scenes, the issues with dynamic scenes are the questions themselves. Why use dynamic scenes at all? Some may argue that they are more ecologically viable, which is true from a certain point of view. However, static naturalistic scenes provide just as much information unless you are specifically analysing the effects of motion. This thesis provides a series of experiments using dynamic environments to explore what effects they may have on visual search performance.

The first objective with dynamic scenes is to figure out what question to ask. Smith and Mital (2013) analysed the effect of viewing task during dynamic scene viewing. In a term coined, 'attentional synchrony' (Smith & Henderson, 2008), they claimed that areas of high motion influenced gaze allocation during a free viewing task. Upon changing the task to an identification one, they found that gaze was biased away from flicker and people. They concluded that the viewing task can significantly influence gaze allocation, and that exogenous (stimulus- driven) factors such as motion naturally draw saccades towards them, which in turn prolongs the initiation of endogenous (cognitively driven) control. The interesting point about their experiment was that they compared gaze behaviour between static and dynamic scenes.

The general premise is that dynamic scenes should contain motion of some form. One simple way of capturing this would be to create a finite recording of the natural world. Trees billowing, cars moving, and people walking are all forms of motion that occur daily. These scenes are very useful since they are fixed at a certain angle with no transitions throughout. For visual search purposes, this allows the experimenter to keep the motion without having to worry about multiple video cuts that could affect the results.

Although useful, these scenes do not involve self-motion, which in terms of this thesis is extremely important. This means studies like the one by Smith and Mital (2013) omit observer induced optical flow and all depth information about the scene, which reduces their ecological validity. But, can optical flow affect search performance? To justify the inclusion of self-motion then optical flow must first be isolated and examined.

It was first shown in 1950 that movement through a stationary environment by an observer produces a radial pattern of optical flow (Gibson, 1950). As an observer moves, flow funnels out from a singularity known as the focus of expansion (FoE). As an observer rotates about a scene, a rotational component of 'solenoidal flow' arises (having no FoE or contraction), decomposing the optical flow pattern into rotational and translational components. The FoE specifies the direction of self-motion, otherwise known as 'heading'. Flow patterns also contain a rich array of information such as the time-to-contact with objects (Pepping & Grealy, 2007), the motion of other objects and the 3D shape and layout of environmental surfaces (Domini & Caudek, 2003; Todd, 1995). Computer Vision research have also used locally consistent optical flow to segment image regions which usually relate to either surfaces at different depths or independently moving objects. Flow patterns can also be used to control locomotion (Warren & Fajen, 2004). The concept of an optical flow played an important role in the development of an ecologically valid approach to perception (see, Gibson, 1979).

In 1988, Warren and Hannon investigated observers' abilities to distinguish their direction of moving from where they are looking. They wanted to find out whether optical flow is sufficient enough for perceiving the direction of self-motion. In their

first experiment, they presented the participants with either a flow field that only contained a translational component (stationary fixation point) or one contained both translational and rotational via movement of the fixation point. Their results showed that observers were very accurate with making heading judgements, even during pursuit eye movements. One hypothesis to explain an accurate judgement of heading is that oculomotor signals such as an efference copy, could compensate for eye movements. To test this, the authors conducted a second experiment of two parts whereby participants tracked a moving fixation point in part one and a stationary fixation point in part two. The notable difference here though is that part two's optic flow field simulated translational and rotational components on the screen whilst requiring the participant to fixate at a stationary point.

The results of experiment two also showed accurate heading judgements, with their mean percentage of correct judgements being significantly above chance level in both conditions. If the hypothesis proved correct, then performance level would simply be by chance.

They concluded that optical flow alone is sufficient for observers to distinguish where they are heading from where they are looking. Factors such as oculomotor signals, multiple fixations, and edge parallax were seen to be not needed for the successful completion of the task.

The above demonstrates that optical flow is important. Bringing this into the context of this thesis, how does optical flow relate to visual search and object detection? In terms of processing optical flow itself by the brain, several hypotheses have been put forward. Control of posture, gaze stabilisation during ego movement (motion relative to a rigid scene), and the recovery of independent surfaces and

surface shapes are a few. However, the above study as well as Gibson's original work in 1950 present the more dominant theory that the optic flow is used for the guidance of locomotion toward a specific target.

All the above involve computer generated displays of optical flow. These simulations allow for a pure isolation of the flow field for analysis. However, another study by Rushton et al (1998) challenged the guidance of locomotion by analysing self-motion on foot. They argued that moving on foot allows for one's own knowledge of their body orientation with respect to surrounding objects, which is more economical and should affect a person's trajectory to a target.

In their experiment, participants were required to wear prism glasses to displace their view and then walk toward where they thought the target was. They argued that a more 'perceived- direction' strategy whereby a participant walks forward, rotates their gaze and body to the target's direction, and then moves toward that target should be more effective than using a flow-based strategy (relying on the relative position of the FoE to control locomotion). Their results show that locomotion was controlled in a continuous manner 'on-line', which generated trajectories similar to the ones predicted using their perceived-direction model.

Despite this contradiction, recent evidence has been in favour of a hypothesis known as 'flow- parsing'. Coined by Rushton and Warren (2005), the 'flow- parsing' hypothesis states that since the brain is sensitive to patterns of optic flow, it uses them to identify components of retinal motion that occur from self- movement. From this, the brain can then separate this motion from the actual movement of objects within the scene. A later study by Warren and Rushton (2008) further demonstrated evidence for flow-parsing by simulating self-movement on a screen.

These patterns were generated using moving dots upon a black background (although a red filter was placed in front of the screen to increase contrast) and in one condition, the pattern represented a radial flow with the opposing condition representing a shuffled radial flow. Participants were required to track a moving dot before reporting on their perception of its trajectory.

Their results showed that displaying a probe to track in conjunction with an asymmetric radial flow field, illusory motion (perceived movement of the probe was not real) was not perceived. However, the systematic degradation of the flow field increased the perception of illusory motion. The same pattern of results was also present for the shuffled condition. Collectively, these results were taken as evidence for the reliance of different optic flow components for the successful perception of object movement during self-movement.

Warren and Rushton (2008) also compared their results against current theories and models, with the theories of induced motion (see Reinhardt-Rutland, 1988 for a review) and the motion contrast detector model (Murakami & Shimojo, 1995) in particular. They stated that the above theory and models do not predict or account for their findings. A suggestion by Reinhardt-Rutland (1988), however, provides some compatibility with their findings; he suggested that further research in this area could in fact derivate from Gibson's earlier work (1950, 1979, 1983) on the analysis of ecological validity and visual sensation. His suggestion states that a stationary environment is often filled with objects that have the potential to move. For a forward-moving observer, stationary objects produce visual movements that conform to the presented optical flow, which can be determined by the eccentricity of the objects and how close they are to the observer. See Chapter 6 for research into the detection and recognition of such objects.

As demonstrated above, both object movement and self-movement are necessary for determining whether or how objects have moved within scenes. Due to self-movement being present, Warren and Rushton (2009) wanted to investigate the necessity of stereo disparity and other sources of depth information for flow parsing. They asked whether the presence of disparity was necessary for the occurrence of flow parsing, but if static depth information is sufficient then is there a way to grade performance with the amount of available depth information?

In their first experiment, participants were required to watch as a probe moved amongst binocularly visible cubes on the screen and then report on its perceived lateral direction. The objects in the display remained static for two seconds before moving for a further two seconds. In experiment two the cubes were only presented to one eye (monocularly) but the probe was still visible across both. This experiment tested combinations of different monocular depth cues (see Figure 3), including:

- Motion parallax (Figure 3a)
- Relative size (Figure 3b)
- Linear perspective (Figure 3c)
- Occlusion (Figure 3d)

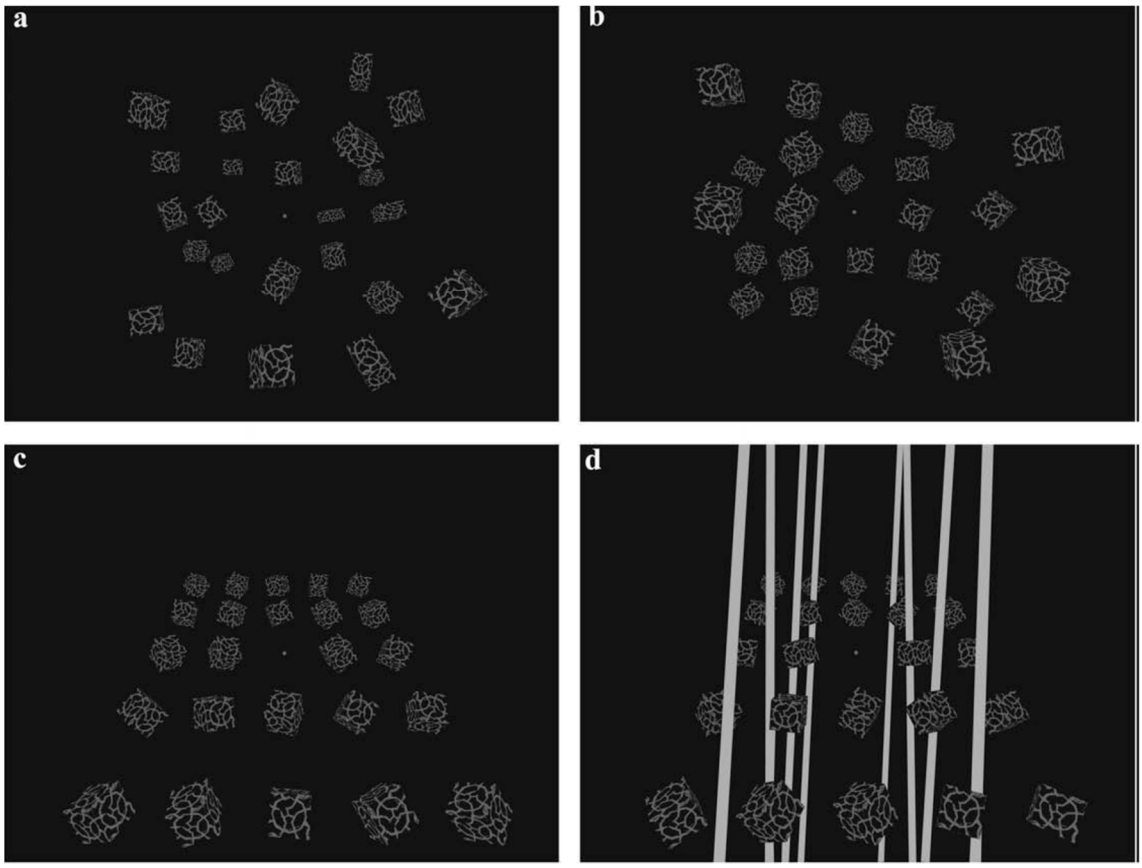



***Figure 3.*** **Combinations of different monocular depth cues**

Their results demonstrate that stereoscopic viewing of a scene is not necessary for flow parsing, and that other available depth information can be used to compensate. The quantity of depth information available also influenced flow parsing performance, and that perceived movement of an object depends upon that object's depth in the scene.

Previous experiments on the detection of object movement (Rushton & Warren, 2005), perception of object trajectory (Warren & Rushton, 2007), and object 'pop-out' (Rushton et al., 2007), together with the above, provide a body of evidence in favour of the flow-parsing hypothesis.

The above relates primarily to object motion within flow fields, however, it was important to understand Gibson's original work on ecological validity, and the flow-parsing hypothesis before continuing to visual search in such displays. In 2001, Royden, Wolfe, and Klempen performed a series of experiments to demonstrate the efficiency of search for stationary and moving targets within structured and random flow fields. Notably, their conditions contained distractors moving in uniform, random and Brownian motions. Similar to the previous studies, the stimuli were not within natural environments.

Their results showed that visual search for object motion is similar across flow fields, and that radial flow fields have no particularly special status. Search for stationary targets among moving distractors was found to be difficult, and that search within a structured flow field was more efficient than search within randomly moving distractors. The authors made three conclusions:

- Search asymmetry makes search for motion more efficient over search for its absence.
- Local differences in feature search is more important than global properties of the display for feature search.
- Within the confines of visual search, the visual system does not treat radial flow fields differently from other more structured motion fields.

Despite their conclusions, they also mentioned that their results could also be due to the stimuli used, as their stimuli was not particularly large or rich in environmental information. As such, they proposed that different results may be obtained if the stimuli were made using virtual reality. See Chapter 7 for a range of experiments involving radial flow fields within a set of virtual reality stimuli.

Speaking directly to virtual reality, there have been seldom few studies that involve such scenes for testing visual search hypotheses, mainly due to the computational resources required and the time taken to put together enough stimuli to create an impactful experiment.

However, several studies have been conducted to investigate navigation. Fortenbaugh et al (2007a), took inspiration from Geisler and Perry's (2002) gaze-contingent multiresolution display concept, and created a new technique for masking regions of the visual field within virtual environments. Using this technique, they could study navigation within virtual reality under different visual field loss constraints.

In another study of the same year, Fortenbaugh et al (2007b), examined the peripheral visual field during a navigational task whereby participants were required to move to a series of statues within virtual space. The statues were then removed from the environment after which participants were required to move to where they thought each statue used to be. Field of view was manipulated in restrictions of $40^{o}$, $20^{o}$, and $10^{o}$ (diameter) using the above technique.

Their results showed that the greater the loss of the peripheral visual field, the less accurate participants were at remembering target locations. It appeared from their results that the availability of global spatial information from the periphery was important for accurate spatial representations. This premise is backed by the visual search literature, as parafoveal/peripheral preview is needed to attend to areas of interest (Henderson, Pollatsek, & Rayner, 1989; Kowler & McKee, 1987; Peterson, Kramer, & Irwin, 2004). However, Fortenbaugh et al's (2007b) results

demonstrated that field of view sizes of 10° and 40° produced comparable results in terms of performance.

From the above literature survey, we believe that the following issues have not been investigated:

- The importance of different visual field regions for the localisation of targets of varying sizes
- The effect of target size and salience in search with relation to visual field regions
- The importance of foveal vision within 3D dynamic scenes with simulated self-motion

The next chapter begins with investigations into the first issue.

# 5 The relative (un)importance of foveal vision

## 5.1 Introduction

How important is the availability of high-acuity foveal vision to visual search? This question has been investigated with different search tasks, ranging from letter search in alphanumeric displays (Bertera & Rayner, 2000) to object-in-scene search (Nuthmann, 2014), either highlighting the relative importance (letter search) or unimportance (scene search) of foveal vision. The aim of the present work was to combine design features from both search paradigms to reconcile the seemingly conflicting findings reported in the literature. In Experiment 1, observers searched for the letter "T" embedded in greyscale pictures of real-world scenes, with or without foveal vision. In Experiment 2, we added a letter recognition component to the search task ("Is it a T or an L?"). In both experiments, we also varied the size of the letter target while controlling its visual salience.

Visual acuity is highest at the fovea before declining rapidly as it approaches the periphery (Strasburger, Rentschler, & Juttner, 2011). Whereas the foveal region is typically defined as the central 2.0º of vision, the parafoveal region extends from the foveal region out to about 5.0º from fixation; the fovea and parafovea together

are commonly referred to as central vision. The peripheral region is everything beyond the parafoveal region.

Research into the (un)importance of foveal and/or central vision has clinical relevance due to the rising number of older adults who develop age-related macular degeneration (AMD) (Congdon et al., 2004; Friedman et al., 2004). AMD is a medical condition in which central vision is impaired or lost, leading to a scotoma. Because visual perception and action involves active information seeking via eye movements, it is important to study the eye-movement behaviour of AMD patients in everyday visual-cognitive tasks. This, however, is challenging, such that very little research has been published on the topic (Boucart et al., 2015; Van der Stigchel et al., 2013). An alternative approach is to *simulate* the loss of foveal or central vision in age-matched *normally sighted* subjects by using an advanced eye-tracking methodology dubbed the gaze-contingent *Moving Mask* technique (Bertera, 1988; Rayner & Bertera, 1979).

The importance of foveal vision was first studied in sentence reading, which is a well investigated task (see Rayner, 1998, for a review). In a now classic study, Rayner and Bertera (Rayner & Bertera, 1979) aligned a visual mask with the reader's gaze to wipe out the text in view. The size of the mask ranged between 1 and 17 characters, where three characters equalled 1º of visual angle. Simulating reading without a fovea that way reduced the reader's reading speed by increasing the number of fixations, fixation duration, and reducing saccade length. Moreover, reading comprehension suffered.

In a subsequent study, the same authors investigated the importance of foveal vision in visual search (Bertera & Rayner, 2000). Owing to their background in

reading research, they had participants search for the target letter "y" within a randomly arranged array of alphanumeric characters (i.e., letters and digits), with or without a simulated scotoma. Five different scotoma sizes were compared to a no-mask control condition: 0.3º, 1º, 1.7º, 2.3º, and 3º. In the scotoma conditions, the fixated character and all characters falling within the mask region were replaced by a +. With the smallest mask size of 0.3º, only the character fixated was replaced. The results showed that as the mask size increased, the lower the search accuracy, the longer the search time, and the more fixations were made. In summary, foveal vision was necessary for making the types of discriminations needed to successfully complete the task.

Interestingly, visual search studies involving naturalistic scenes have found rather different results (McIlreavy, Fiser, & Bex, 2012; Nuthmann, 2014). In the study by Nuthmann (Nuthmann, 2014), participants were engaged in an ecologically valid search task in which they looked for a specific object in a coloured image of a real-world scene (e.g., a blender in a kitchen scene). Search was cued with a word label and search objects had an average size of 2.5º × 2.5º (medium size). When searching the scene with artificially impaired foveal or central vision, search performance was surprisingly unimpaired. Importantly, foveal vision was not necessary to attain normal search performance (i.e., search accuracy, search time). When searching without central vision, participants' gaze data revealed that they were not impaired in locating the search object in the scene, but in verifying the identity of the target.

In the study by Nuthmann (Nuthmann, 2014), the scene image contained contextually relevant search targets (cf. Torralba, Oliva, Castelhano, & Henderson, 2006). McIlreavy et al. (McIlreavy et al., 2012) excluded such contextual guidance

towards the target by asking observers to look for spatial distortions, which were embedded at random places in greyscale images of natural scenes. The Gaussian distortions in the images are described by their magnitude (0.5º) and spatial period (2º); the magnitude of distortion beyond a 2º radius approached 0. McIlreavy et al.'s results for search times were similar to the ones by Nuthmann (Nuthmann, 2014). When searching with a foveal scotoma ($\sigma_{x,y} = 1º$), participants had no problem finding the distortion in the scene, with performance being identical to that of the no-scotoma control condition. Only the largest central scotoma condition ($\sigma_{x,y} = 4º$) led to a significant increase in mean search time.

Both McIlreavy et al. (McIlreavy et al., 2012) and Nuthmann (Nuthmann, 2014) discuss that target size could be an important mediating factor for their findings on the (un)importance of foveal vision. Before elaborating on this argument, we briefly review research on size and eccentricity effects in (normal) visual search. A common paradigm is to use fairly small simple displays which observers search covertly in the absence of eye movements. Using this approach, Duncan and Humphreys (1989) investigated the effect of stimulus size and showed that search is more difficult for small letters than for large letters. A related finding is the eccentricity effect: search performance deteriorates as the target is presented at farther peripheral locations (Carrasco et al., 1995; Geisler & Chou, 1995). This reduction in search efficiency may be due to the poorer spatial resolution in the periphery. Consistent with this view, enlarging the stimuli according to the cortical magnification factor (Rovamo & Virsu, 1979) eliminated the eccentricity effect (Carrasco & Frieder, 1997; Carrasco, Mclean, Katz, & Frieder, 1998; Wolfe, O'Neill, & Bennett, 1998). The eccentricity effect is also observed in the presence of eye movements (Scialfa & Joffe, 1998; Zelinsky, 2008).

In the context of visual search in real-world scenes, the effect of target size has received little systematic investigation. Wolfe, Alvarez, Rosenholtz, Kuzmova, and Sherman (2011, Experiment 1) had observers repeatedly search for objects in photographs of real-world scenes. Their approach was to exhaustively label a set of images. Reflecting the design of the world, the annotated objects showed a natural variability in size and eccentricity. In the experiment, eye movements were permitted but not recorded; the initial fixation position was assumed to be central but this was not explicitly controlled for. The data showed a strong effect of target size such that search times increased steadily as targets became smaller. In addition, there was a rather small eccentricity effect.

Miellet, Zhou, He, Rodger, and Caldara (Miellet, Zhou, He, Rodger, & Caldara, 2010) asked observers to search for animals in photographs of zoo enclosures. The size of the search target was either 2º, 5º, or 8º. There was a main effect of target size with better search performance for larger targets. The main objective of the study by Miellet et al. was to investigate whether culture had an impact on extrafoveal information use in visual search. This is why the authors parametrically manipulated both target size and the size of a gaze-contingent moving mask (diameter: 2º, 5º, or 8º) and tested both Eastern and Western observers. No effects of cultural diversity were found. However, there was a main effect of mask size such that performance increasingly suffered as the simulated scotoma got larger (McIlreavy et al., 2012; A Nuthmann, 2014). Importantly, there was an interaction between mask size and target size such that the deleterious effect of mask size was more pronounced for smaller targets. The 2º-*Blindspot* condition in the study by Miellet et al. approximates the unavailability of foveal vision. The averages presented in their Table 1 suggest that such a foveal scotoma led to a reduction in search performance for 2º targets but not for 8º targets, for both observer

groups. Although suggestive, any findings involving target size in this study need to be treated cautiously because target salience (Itti & Koch, 2000) was not controlled for. Other potential confounds are target eccentricity (i.e., distance from scene center) and contextual guidance.

The goal of the present research was to further investigate the importance of foveal vision to visual search. Stimuli were greyscale pictures of real-world scenes which included a target letter (Experiment 1: T, Experiment 2: T or L). Four letter sizes, ranging from 0.25º to 1.5º in width, were crossed with the presence vs. absence of foveal vision. To control for visual salience, the letter was algorithmically placed for each scene in a location for which there was a medium change in local contrast when inserting the letter. Letter targets were used for a number of reasons. The small to large animal targets in Miellet et al. (Miellet et al., 2010) were all part of different scenes. Our approach allowed us to place letter targets of variable size at the same location within a given scene. In addition, using context-free letter targets rather than contextually relevant search targets prevents observers from using their knowledge about the likely positions of targets to guide their eye movements (McIlreavy et al., 2012). Moreover, we chose to embed letters into images of real-world scenes to combine design features from letter search and scene search tasks that have yielded seemingly conflicting results (Bertera & Rayner, 2000; Nuthmann, 2014). In Experiment 1, on each trial participants were asked to look for the letter "T". In Experiment 2, we added a recognition component to the task. The target was either a "T" or an "L", and—once they found the letter—participants had to indicate which one it was. We chose these two letters because they share exactly the same features (strokes) and differ only in their spatial arrangement (Duncan & Humphreys, 1989). Because we used participants' eye-

movement data to verify that targets had indeed been found, there were no target-absent trials (Nuthmann, 2013, 2014; Nuthmann & Malcolm, 2016).

If foveal vision is necessary to achieve normal search performance during letter-in-scene search, then we should observe a reduction in performance—lower search accuracy and longer search time—when searching the scene with a simulated foveal scotoma, compared with a normal vision control condition. Moreover, we expected to find effects of target size, with better performance for larger targets. Critically, the experimental design allowed us to investigate whether the importance of foveal vision depended on the size of the search target (Miellet et al., 2010). Why would size matter? Here, our hypotheses concern two separate sub-processes of search: scanning for the target and verifying that the fixated object is the target. The scanning process involves the localisation of the target in space, the duration of which (scanning time) is indexed as the time between the first saccade and the first fixation on the target (Malcolm & Henderson, 2009). Similarly, verification time is the elapsed time between the beginning of the first fixation on the target and search termination.

The question arises whether the actual search process, indexed by the scanning time, is slowed down without foveal vision. This what not the case when observers searched for medium-sized real-world objects in scenes (Nuthmann, 2014). However, observers may not be able to resolve the smaller letter targets in parafoveal vision, and they cannot rely on contextual guidance. Therefore, a search strategy involving sequential foveation of potential target sites may be required, prolonging scanning times. Alternatively, blocking out foveal vision may only affect the verification process, as explained next.

Upon fixation with a foveal scotoma, all of the target—or some part of it—will be covered by the scotoma. The extent of this masking depends on both the size of the target and the initial fixation position on the search target (Nuthmann, 2014). The smaller the target, the more likely it is that the target will be completely masked during on-target fixations. This may lead to an increased tendency to move the eyes off the target to unmask the letter once again and to process it further in parafoveal or peripheral vision (Nuthmann, 2014). Such behaviour would increase verification times. Taken together, regarding the verification process we hypothesised that any detrimental effect of the foveal scotoma may only occur for smaller targets, or may be more pronounced for those. Moreover, in Experiment 2 we changed the task to involve not only target detection but also target identification. At least for small letters, letter identification may require the extraction of fine detail via foveal analysis. Therefore, we reasoned that any adverse effect of the foveal scotoma, and its interaction with target size, may be stronger in Experiment 2 than in Experiment 1.

## 5.2 Experiments One and Two

## 5.3 Methods

### 5.3.1 Participants

Thirty-two participants (12 males) between the ages of 18 and 27 (mean age 20 years) participated in Experiment 1. Thirty-two different participants (8 males) between the ages of 18 and 27 (mean age 22 years) participated in Experiment 2. All participants had normal or corrected-to-normal vision by self-report. They gave their written consent prior to the experiment and either received study credit or were paid at a rate of £7 per hour for their participation.

### 5.3.2 Apparatus

Working with gaze-contingent displays requires minimising the latency of the system. This was achieved by using (a) an eye tracker with high temporal resolution, (b) modern graphics hardware, and (c) a monitor with a high refresh rate. Stimuli were presented on a 21-inch CRT monitor with a refresh rate of 140 Hz at a viewing distance of 90 cm, taking up a 24.8º × 18.6º (width × height) field of view. A chin rest was used to keep the participants' head position stable. During stimulus presentation, the eye movements of the participants were recorded binocularly with an SR Research EyeLink 1000 Desktop mount system. It was equipped with the 2000 Hz camera upgrade, allowing for binocular recordings at a sampling rate of 1000 Hz per eye. The experiment itself was programmed in MATLAB 2013a (The MathWorks, Natick, MA) using the OpenGL-based Psychophysics Toolbox 3 (Brainard, 1997; Kleiner, Brainard, & Pelli, 2007) which incorporates the EyeLink Toolbox extensions (Cornelissen, Peters, & Palmer, 2002). A controller was used to record participants' behavioural responses.

### 5.3.3 Stimuli

In Experiment 1, stimuli consisted of 120 (+ 8 practice) greyscale images of naturalistic scenes (800 × 600 pixels), which came from a variety of categories (example scenes are shown in Figures 4, 6 and 10). We used image processing techniques to insert the letter T in four sizes at the same location within a given scene, such that the chosen location was of median salience, as explained below. Note that in the experiment, each participant viewed a given scene only once, in one of the four target size conditions (and either with or without foveal vision).

In Experiment 2, 128 (+ 8 practice) greyscale images of real-world scenes were used, 120 of which were from experiment 1 with 8 new images. The new images were chosen because the experimental design required an equal number of T- and L-scenes in each target-size condition. The search target was either a letter T or L that was algorithmically placed into the scene at a median salience location.


***Figure 4.*** **Example scenes depicting locations of median change in local contrast for T and L targets. Target size represented in all four is Intermediate (0.41º).**

### 5.3.4 Design

Both experiments used a 2 × 4 within-subjects design with 2-level factor scotoma (on vs. off) and 4-level factor target size. The factor scotoma refers to the availability of foveal vision. In the scotoma condition, foveal vision was blocked by a gaze-contingent moving mask (scotoma on, or foveal vision absent). This was

contrasted with a normal-vision control condition (scotoma off, or foveal vision present).

In both experiments, the presence or absence of foveal vision was crossed with four target sizes. In Experiment 1, they were equally spaced as follows: S - Small (letter width 0.25º), M - Medium (0.66º), L - Large (1.08º), and XL - Extra Large (1.5º) (see Figure 5). The XL target size was chosen such that the letter was still completely obscured when fixated with a foveal scotoma, which had a radius of 1º. In Experiment 2, we removed the XL targets; instead, we added targets of intermediate size (0.41º) halfway between the small and medium targets (see Figure 4). These adjustments were informed by the results obtained in Experiment 1: search efficiency was much worse for small targets compared with medium-sized targets, while performance differences between large and extra-large targets were much less pronounced.


***Figure 5.*** **Target size depictions for Experiment 1 and 2. The XL target size was only used in Experiment 1, with the I size in Experiment 2**

In Experiment 1, the 120 T-scenes were assigned to eight lists of 15 scenes each. The scene lists were rotated over participants, such that a given participant was exposed to a list for only one of the eight experimental conditions created by the 2 × 4 design. There were eight groups of four participants, and each group of participants

was exposed to unique combinations of list and experimental condition. To summarise, participants viewed each of the 120 scene items once, with 15 scenes in each of the eight experimental conditions. Across the 32 participants, each scene item appeared in each condition four times.

For Experiment 2, each of the 128 original scene images was submitted to the Target Embedding Algorithm (hereafter T.E.A.) to produce four T-scenes and four L-scenes, one for each target size. In the experiment, half of the original scenes were used as T-scenes, the other half as L-scenes. The decision about which scenes to use in either category was guided by visual inspection. We then created eight scene lists, each comprising eight T-scenes and eight L-scenes. Apart from that, the same counterbalancing procedure as in Experiment 1 was used to control for item effects.

The foveal vision manipulation was blocked so that participants completed two blocks of trials in the experiment: in one block observers' foveal vision was available, in the other block it was obstructed by a gaze-contingent scotoma. Each block started with four practice trials, one for each target size condition. The order of blocks was counterbalanced across subjects. Within a block, scenes with targets of different sizes and types (Experiment 2 only) were presented randomly.

### 5.3.5 The Target Embedding Algorithm – T.E.A.

One variable that we manipulated in the present experiments was the size of the letter target. To this end, we developed an algorithm to place letter targets of variable size at the same location within a given scene while controlling for visual salience. Salience describes how much an element of the scene stands out from its neighbouring regions. While there are many methods of constructing salience maps

for images of real-world scenes (Borji, Sihite, & Itti, 2013), it is widely held that simple stimulus features such as colour, orientation and intensity (luminance contrast) contribute to the computation of visual salience (Laurent Itti & Koch, 2000). Using the output of a computational model of visual salience as input for our algorithm would be prohibitively computationally expensive. As a viable alternative, we used a version of root-mean-square (RMS) contrast: when stepping through the scene, the standard deviation of luminance values of all pixels in the evaluated region was divided by the mean luminance of the image. Calculating luminance contrast this way is consistent with measures of detectability in natural scenes (Bex & Makous, 2002), and with filter properties of early vision (Moulden, Kingdom, & Gatley, 1990). Moreover, it has been used in experimental studies on fixation selection in scenes (e.g., Nuthmann & Einhäuser, 2015; Reinagel & Zador, 1999).

The target was placed at an image position that caused a median RMS contrast change. To compute this, a rectangular region that was slightly larger than the target moved pixel-by-pixel through the image. The RMS contrast $M_o$ was calculated at each position. Afterwards, the target was inserted and the RMS contrast $M_w$ was computed at each position. By computing $\Delta C = M_w - M_0$ at each pixel, we obtained an image map comprising of the contrast difference values within the image. After calculating the contrast difference map for each target size, the four resultant maps were summed together to obtain a final summed difference map. This summing acted as a way for the algorithm to compute a single location for all target sizes, as the values of each individual difference map varied slightly.

This final map was then probed by our algorithm to locate all pixel (i.e., potential target) positions with the median change in contrast before comparing the location

against the following two criteria. First, locations within 5° of visual angle from the center were excluded from evaluation due to the central region being the initial location of both the participant's gaze and the gaze-contingent scotoma. Second, locations at the boundaries of the image were also excluded to avoid truncation of the letter. From all remaining possible target positions, one was selected at random as the location of the target for that stimulus.

For Experiment 2, the algorithm was extended to handle multiple target letters. In this case, a new 'TL' contrast difference map was generated by computing:

$$\Delta C(r,c) = \sum_{\mathcal{L}_s} \left| \Delta C_{\mathcal{L}}^{[s]}(r,c) - t_{\mathcal{L}_s} \right|$$

***Equation 1.*** **T.E.A. formula for generating contrast difference maps comprising of multiple letters.**

where $\Delta C_{\mathcal{L}}^{[s]}$ is the difference map for a given font size [s] and letter $\mathcal{L} \in \{T,L\}$, with [r,c] denoting the map's rows and columns. Each of its values were then subtracted by the median contrast of a given map, denoted by $t_{\mathcal{L}_s}$. This process was repeated for both letters and all four scales before adding the resultant image maps together. Due to subtracting $t_{\mathcal{L}_s}$, the lowest value in this new map (with a minimum of zero) is the pixel closest to the target value $t_{\mathcal{L}_s}$, and the coordinates of this pixel defined the target position for that image. Figure 6 provides an illustration by depicting the contrast difference map and the algorithmic probing.

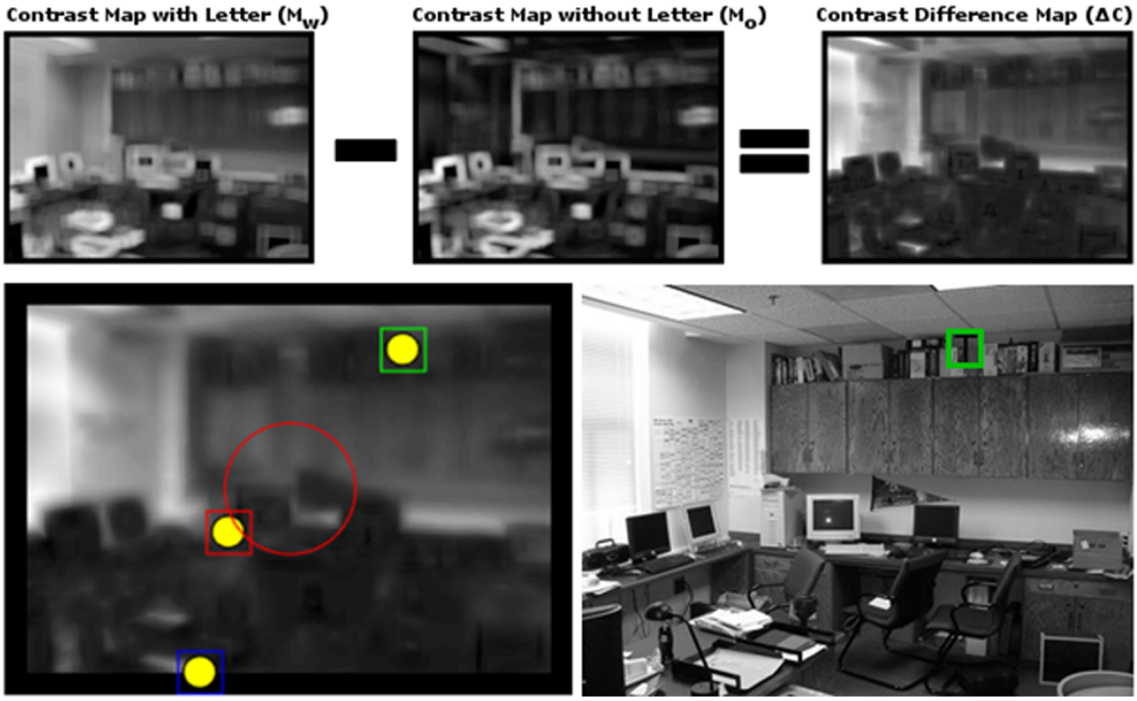



***Figure 6.*** **Illustration of the T.E.A depicting the initial creation of the contrast difference map. The T.E.A creates 3 contrast maps (from left to right): contrast with the letter, contrast without the letter, and the difference between them. Using the contrast difference map, the algorithm then probes the scene for locations by checking against the boundary of the screen (blue boxed dot), the boundary of the central circle (red boxed dot), and the value of the location. If multiple positions are found (yellow dots), one is chosen at random (green boxed dot) before target insertion.**

#### 5.3.5.1 Salience model comparison

Evaluating the T.E.A. against the Adaptive Whitening Saliency model (Garcia-Diaz et al., 2012), we can see the difference between the two. In Figure 7 (row 2), we can see that on first pass, both the T.E.A. (left) and the A.W.S. (right) chose locations that, due to differing target sizes, our exclusion criteria would ignore. This is because the largest size at this location would either overlap with the central circular zone or truncate off the edge of the image. Row 3 of Figure 7 depict the values between the two maps in viable locations chosen by the T.E.A. and A.W.S. Calculating salience on an image with the target inserted means that the target's location has its salience already manipulated. This suggests that a comparison between locations chosen by the T.E.A. and a standard salience model is unfair. In the context of this chapter's experiments, this means that the pixels in that region (now completely black) no longer represent median values with respect to the whole image. A salience map applied to an image without any target inserted also means that it does not consider the change in salience upon target insertion. The T.E.A on the other hand, considers targets at each pixel region, and then calculates the contrast difference values to produce a final map for location selection. In effect, comparing against a standard salience model means that we're taking a global measurement of salience, choosing the median value of that (at which point the salience changes), and then comparing it to the T.E.A which considers that change in salience across all regions. In addition, if we observe the original image, we can see several objects such as the lamps, the window and the painting, all of which are regions where one may not want to place targets as they are likely to already draw attention. For these reasons, we consider the T.E.A to be a practical alternative to salience models, but to enhance both its credibility and future performance gains, additional model comparisons would be needed.

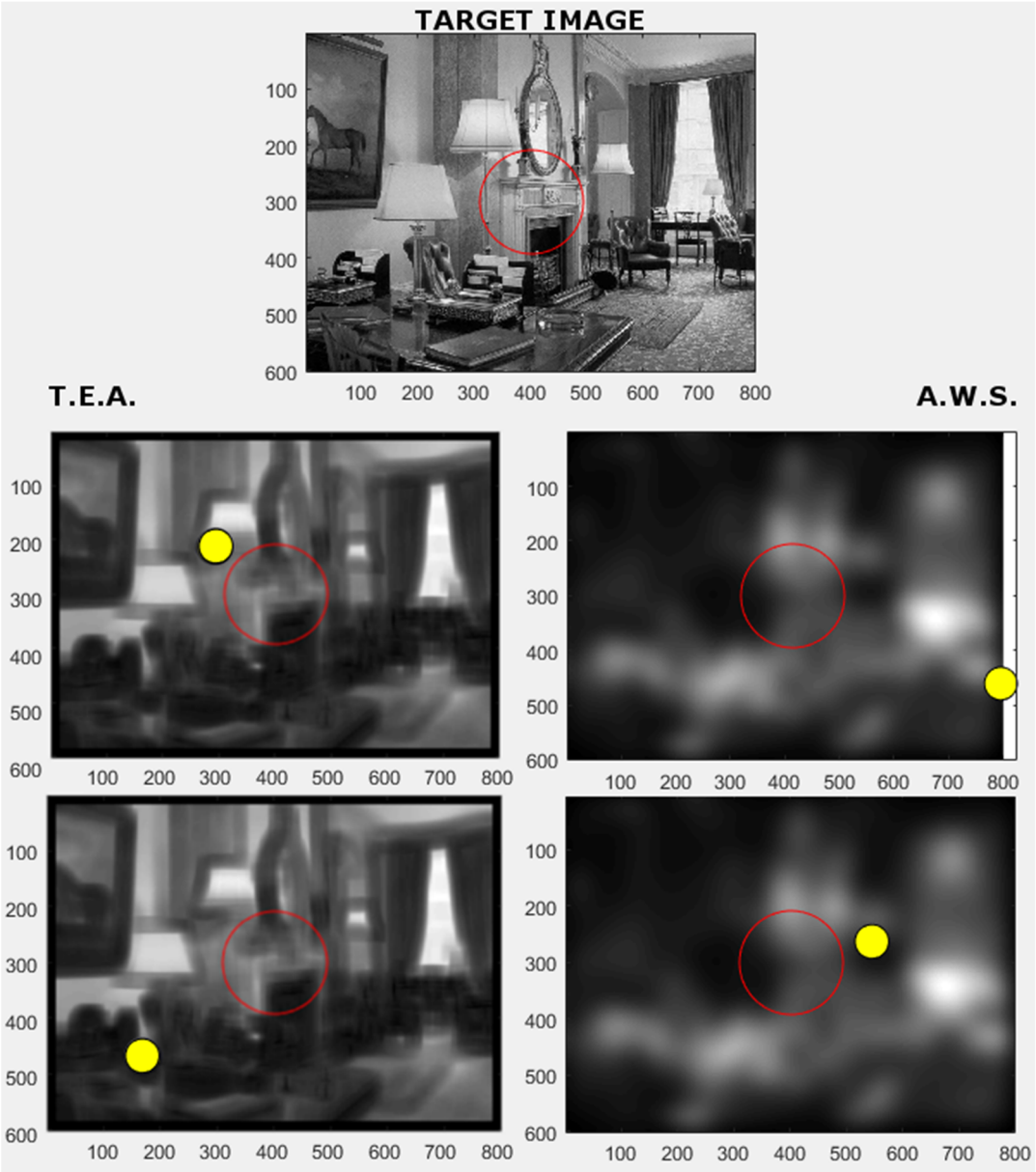



***Figure 7*** **T.E.A. (left) vs A.W.S. (right). The top image depicts the original target image. The middle rows depict initial target insertion and the bottom row shows viable location value comparisons.**

### 5.3.6 Creation of gaze-contingent scotoma

The foveal scotoma was created using texture-mapping and OpenGL (Open Graphics Library). This technique provides various blending operations that enable simple image combinations to take place via an image's alpha channel (see

Duchowski & Çöltekin, 2007, for details on the general technique). The scotoma was a symmetric circular mask with a radius of 1°, that is the scotoma size was chosen to completely obscure foveal vision. The foveal mask moved concomitantly with the participant's gaze. To this end, the average horizontal and vertical position of the two eyes (Nuthmann, 2013) was continuously evaluated online. Updating the display contingent on the viewer's gaze required 1 ms to receive a sample from the eye tracker, less than 1 ms to draw the image textures and up to 7 ms to refresh the screen. Thus, the display was updated depending on observers' gaze position in close to real time. A detailed account of the gaze-contingent implementation is provided in Nuthmann (2013, 2014).

There are some subtle differences between the implementation of the foveal scotoma in a previous study from our lab (Nuthmann, 2014) and here. Nuthmann (2014) used full-colour images, and foveal vision was degraded by applying a very strong low-pass filter to the currently fixated scene region. Moreover, a Gaussian mask was used, and the size of the scotoma was defined as the standard deviation of the two-dimensional Gaussian distribution (1.6° for the foveal scotoma, or small *Blindspot*). In the present experiments, using greyscale images, we used a circular mask drawn in grey. To avoid a sharp-boundary mask, the perimeter of the circular mask was slightly faded through low-pass filtering, while the interior remained untouched. When investigating the importance of foveal vision (i.e., a relatively small region of the visual field), it seems more appropriate to define the size of the moving mask as the radius of a circle rather than the standard deviation of a Gaussian.

### 5.3.7 Procedure

At the beginning of the experiment, a 9-point calibration procedure was performed, followed by a 9-point calibration accuracy test (validation). At the beginning of each trial a fixation cross was presented at the center of the screen for 600 ms, and acted as a fixation check. The fixation check was deemed successful if gaze position, averaged across both eyes, continuously stayed within an area of 40 × 40 pixels (1.24º × 1.24º) for 200 ms. If this condition was not met, the fixation check timed out after 500 ms. In this case, the fixation check procedure was either repeated or replaced by another calibration procedure. If the fixation check was successful, the scene image appeared on the screen. Once subjects had found the target letter, they should fixate their gaze on it and press a button on the controller to end the trial (Glaholt, Rayner, & Reingold, 2012; Nuthmann, 2014). In experiment 1, participants could press any button to indicate that they had found the T. Upon identifying the target in Experiment 2, observers pressed one of two triggers on the controller corresponding to either "T" or "L". Trials timed-out 15 s after stimulus presentation if no response was made. There was an inter-trial interval of 1 s before the next fixation cross was presented.

### 5.3.8 Data Analysis

Saccades were defined with a 50º/s velocity threshold using a nine-sample saccade detection model. The SR Research Data Viewer software was used to convert the raw data obtained by the eye tracker into a fixation sequence matrix. Analyses of fixation durations and saccade lengths excluded fixations that were interrupted with blinks. Analysis of fixation durations disregarded fixations that were the first or last fixation in a trial. Fixations that had durations of less than 50 ms or greater

than 750 ms were also excluded, based on the assumption that they are not determined by on-line cognitive processes.

The behavioural and eye-movement data were further processed and analysed using the R system for statistical computing (version 3.2.3; R Development Core Team, 2015) under the GNU General Public License (Version 2, June 1991). Data were analysed with two-way repeated measures analyses of variance (ANOVAs), unless otherwise stated. ANOVAs were conducted using the *ez* package (Lawrence, 2015) supplied in R. For target size (within-subjects factor with more than two levels), a Greenhouse-Geisser correction was used to adjust the degrees of freedom if Mauchly's test indicated that the assumption of sphericity had been violated.

Figures were created using MATLAB (Figures 6, 7 and 10) or the *ggplot2* package (Wickham, 2009) supplied in R (remaining figures). The T.E.A was programmed in MATLAB. When using the T.E.A to prepare the stimulus material for Experiment 1, due to an input error the target was not inserted into an adequate scene location for eight of the scenes; these scenes were subsequently removed from statistical analysis.

## 5.4 Results and Discussion

The results of the two letter-in-scene search experiments are presented in three main sections. First, different measures of search accuracy were analysed as indicators of search efficiency. Second, the time to find the target was analysed. Behavioural search times were then decomposed based on participants' gaze data to illuminate disruptions in specific sub-processes of search (Malcolm & Henderson, 2009; Nuthmann, 2014). Third, we examined saccade amplitude and fixation duration across the viewing period as general eye-movement measures.

### 5.4.1 Search Accuracy

The first set of analyses examined the likelihood of finding the target letter in the scene. Performance for each experimental condition was divided into probabilities of "hit," "miss," and "timeout" cases (Nuthmann, 2014). If the participant had not responded within 15 s, the trial was coded as a "timeout." A response was scored as a "hit" if the participant indicated to have located the target by button press and his or her gaze was within the target interest area. The interest area was rectangular and made to encompass the XL target letter. For gaze scoring, we implemented an additional 0.5º of padding to either side to accommodate for (a) the potential spatial inaccuracy of the eye tracker, and (b) the inaccuracy of the visuo-oculomotor system when targeting relatively small objects (Pajak & Nuthmann, 2013). A response was scored as a "miss" if viewers had not fixated their gaze on the target when they pressed the button. Those cases include incidents where the target was not located, but also trials where observers' eyes did not fixate on a correctly located target when the overt response was made. All trials that were algorithmically classified as miss trials were therefore manually inspected by me to distinguish incidents where the target was mislocated in space (true misses) from off-target fixations; those latter cases were re-coded as hit trials. It is important to note that the qualitative pattern of results for hit trials was no different with or without the recoding (See Figure 8). Prior to this, other methods were implemented and tested as ways of distinguishing "true misses" from off-target fixations in an automatic manner; none of the solutions were satisfactory as they all included an arbitrary criterion of some kind. The advantage of recoding trials manually is that I could evaluate the entire gaze path for each trial to make an informed judgement. The search accuracy results for both experiments are depicted in Figure 9.

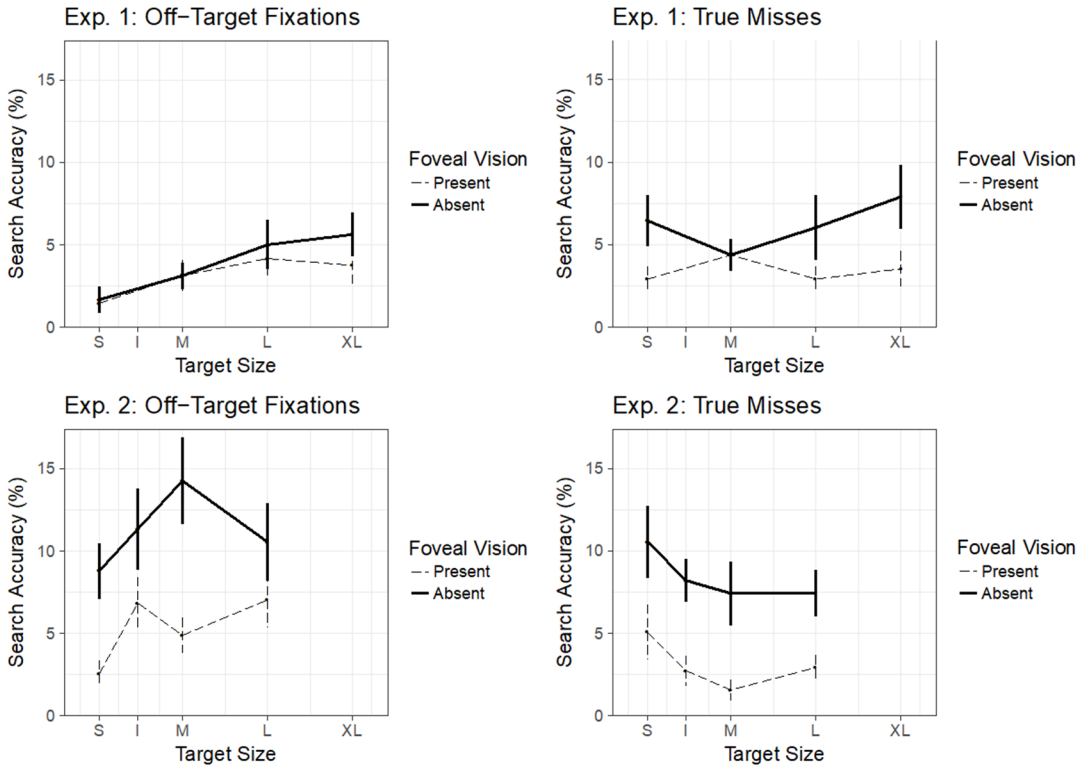



***Figure 8*** **Figure depicting off target fixations from true misses across experiments 1 and 2. The x-axis describes target size with all measures included across both experiments for ease of comparison. The y-axis refers to search accuracy which reflects the percentage of either off target fixations or true misses. Dotted lines represent the presence of foveal vision.**

## Experiment 1

There was a significant main effect of scotoma on the probability of "hitting" the target such that participants were less likely to locate and correctly identify the target when foveal vision was not available, $F(1, 31) = 38.15$, $p < .001$. There was also a significant main effect of target size with lower search accuracy for smaller targets, $F(2.21, 68.6) = 87.29$, $p < .001$; clearly, this effect was driven by the small targets (Figure 9a). There was also a significant interaction between scotoma and target size, $F(1.77, 54.93) = 21.99$, $p < .001$, indicating that the

detrimental effect of the foveal scotoma was most pronounced for small targets. The drop in performance for small targets was due to an increase in timed out trials (Figure 9b). Timeout probability was low for all other target sizes, with or without a foveal scotoma. It was substantially elevated for small targets in the normal-vision control condition, and even more so in the foveal scotoma condition. The probability of missing the target was low, with and without a scotoma (Figure 9c).

### Experiment 2

Experiment 2 included an additional letter recognition component (is the target a "T" or an "L"?). The probability of making an incorrect response (i.e., pressing the wrong button) was very low in all experimental conditions (Figure 9g). For hit trials (Figure 9d), only trials with correct recognition responses were analysed further. There was a significant main effect of scotoma on the probability of "hitting" the target such that participants were less likely to locate and correctly identify the target without foveal vision, $F(1, 31) = 65.09$, $p < .001$. There was also a significant main effect of target size with lower search accuracy for smaller targets, $F(1.87, 57.99) = 104.26$, $p < .001$ (Figure 9d). The interaction between scotoma and target size was also significant, $F(2.03, 62.84) = 7.04$, $p = .002$. The drop in performance for search without foveal vision also shows in increased probabilities of missing the target (Figure 9f) and not responding within 15s (Figure 9e).

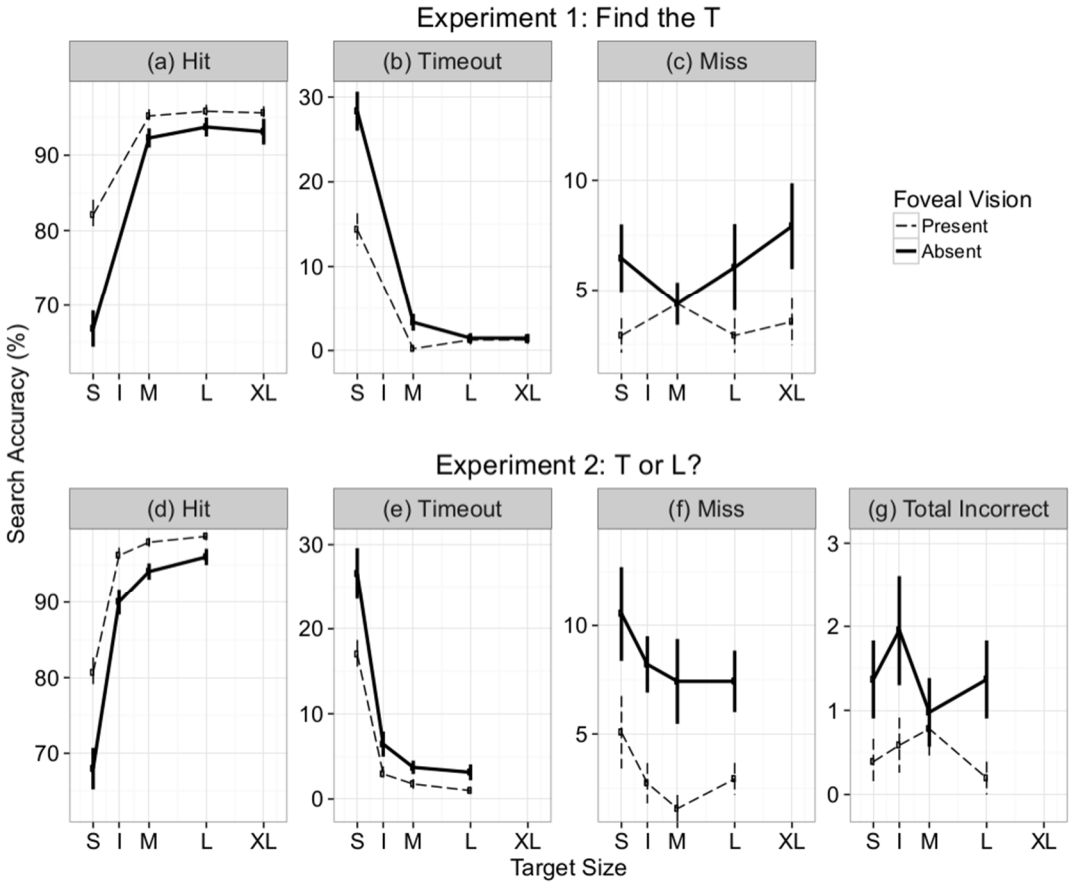



***Figure 9.*** **Measures of search accuracy for Experiment 1 (top row) and Experiment 2 (bottom row). Each column presents means obtained for a designated dependent variable, which is specified in the panel title (see text for definitions). For a given dependent variable, the *y*-axis has been normalised across plots for ease of comparison between the two experiments; but note the different *y*-axis scales for the different measures. Target sizes on the *x*-axis are described by letters (S: Small, I: Intermediate - Experiment 2 only, M: Medium, L: Large, XL: Extra Large - Experiment 1 only; see text for actual sizes in degrees of visual angle). The *x*-axis is scaled to show all target sizes across both experiments; the spacing on the *x*-axis preserves the relative distances between target sizes. Error bars are within-subjects standard errors (Cousineau, 2005).**

### 5.4.2 Search Time and its Subcomponents

Search guidance was analysed further for correct trials ("hits") only. Search time is the overall time taken from scene onset to a user response terminating the search. We then used participants' gaze data to divide search time into three behaviourally defined epochs: search initiation time, scanning time, and verification time (Malcolm & Henderson, 2009; Nuthmann, 2014; Nuthmann & Malcolm, 2016; Spotorno, Malcolm, & Tatler, 2015). This was done to test how the availability of foveal vision as well as the size of the target would affect different sub-processes of search. Search initiation time is the interval between scene onset and the initiation of the first saccade (i.e., the initial saccade latency). This epoch measures the time needed to begin search. Scanning time is the time from the first eye movement until the participant's gaze enters the target's area of interest (minus the first saccade). The scanning time measure reflects the process of localising the target in space (Malcolm & Henderson, 2009). Finally, the verification process is indexed by the time taken from first entering the target interest area until the participant confirms their decision via button press. The segmentation of search time by oculomotor behaviour is visualised in Figure 10.

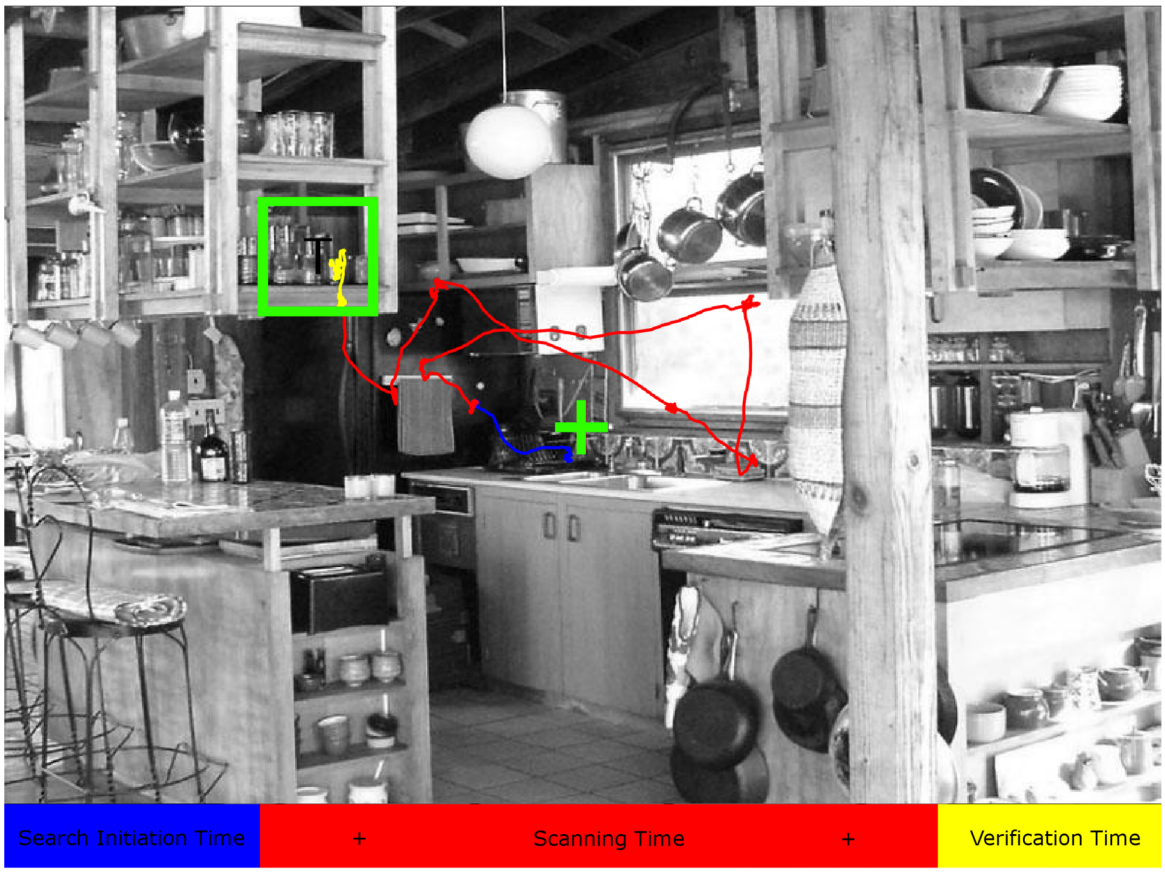



***Figure 10.*** **Gaze-based decomposition of search time. For an example search trial, the scene image is presented together with the raw gaze data from one observer (curvy lines are saccades, clustered data points are fixations). Visualising the division of search time, blue represents search initiation (i.e., initial saccade latency); red, scanning time; and yellow, verification time. When summed, they yield the total search time. The green box is the interest area around the target letter "T".**

Separating oculomotor behaviour in this way provides clean measures to analyse in relation to our hypotheses. Although we had no hypotheses relating to initiation time, it is important to separate it from scanning time to measure an observer's ability to find targets. This approach to dividing search time could be considered simple, as it may neglect other behaviours. Verification time for instance occurs when gaze first lands within the interest area, which ignores any verification that

could be occurring during the scanning phase via parafoveal processing. Although an extension to this algorithm has been offered in Chapter 7 (see also Figure 7), additional factors such as refixation counts, and off target fixation behaviours could be analysed to aid in creating a more systematic measure of verification time.

### Experiment 1

Critically, the foveal scotoma had no significant effect on search time, $F(1, 31) = 1.69$, $p = .203$ (Figure 11a). However, there was a significant effect of target size with faster search times for larger targets, $F(1.57, 48.73) = 190.86$, $p < .001$. The interaction between scotoma and target size was not significant, $F(2.21, 68.61) = 0.25$, $p = .798$.

For search initiation time, the effect of scotoma was not significant, $F(1, 31) = 2.95$, $p = .096$. Importantly, scanning time was not prolonged when searching with a foveal scotoma, $F(1, 31) = .19$, $p = .664$, but verification time was, $F(1, 31) = 7.15$, $p = .012$. The effect of target size on initiation time did not reach significance, $F(3, 93) = 2.48$, $p = .066$. However, target size did affect both scanning time, $F(1.5, 46.49) = 130.31$, $p < .001$, and verification time, $F(1.39, 42.98) = 41.4$, $p < .001$. Observers went through both sub-processes of search, the larger the target the faster the search. For verification time, the theoretically relevant interaction between scotoma and target size was not significant, $F(1.34, 41.61) = 2.05$, $p = .155$. Interactions for initiation and scanning times were not significant.

### Experiment 2

There was again no significant main effect of scotoma on total search time, $F(1, 31) = 2.44$, $p = .128$, but there was a main effect of target size, $F(2.22, 68.87) = 88.3$, $p < .001$ (Figure 11e). The interaction between foveal scotoma and target size was not significant, $F(2.2, 68.28) = 0.66$, $p = .533$.

Different to Experiment 1, the effect of scotoma on search initiation time was now significant, $F(1, 31) = 4.24$, $p = .048$, and so was the effect of target size, $F(3, 93) = 2.99$, $p = .035$. The interaction between scotoma and target size was not significant, $F(3, 93) = 1.15$, $p = .332$. For scanning time, there was no significant effect of scotoma, $F(1, 31) = 3.42$, $p = .074$. There was, however, a significant effect of target size, $F(3, 93) = 98.75$, $p < .001$. The interaction between scotoma and target size was not significant, $F(1.95, 60.31) = 1.74$, $p = .185$. For verification time, the main effect of scotoma was significant, $F(1, 31) = 27.11$, $p < .001$. There was also a significant effect of target size, $F(1.79, 55.41) = 18.16$, $p < .001$, but the interaction between scotoma and target size did not reach significance, $F(1.92, 59.66) = 3.03$, $p = .058$.

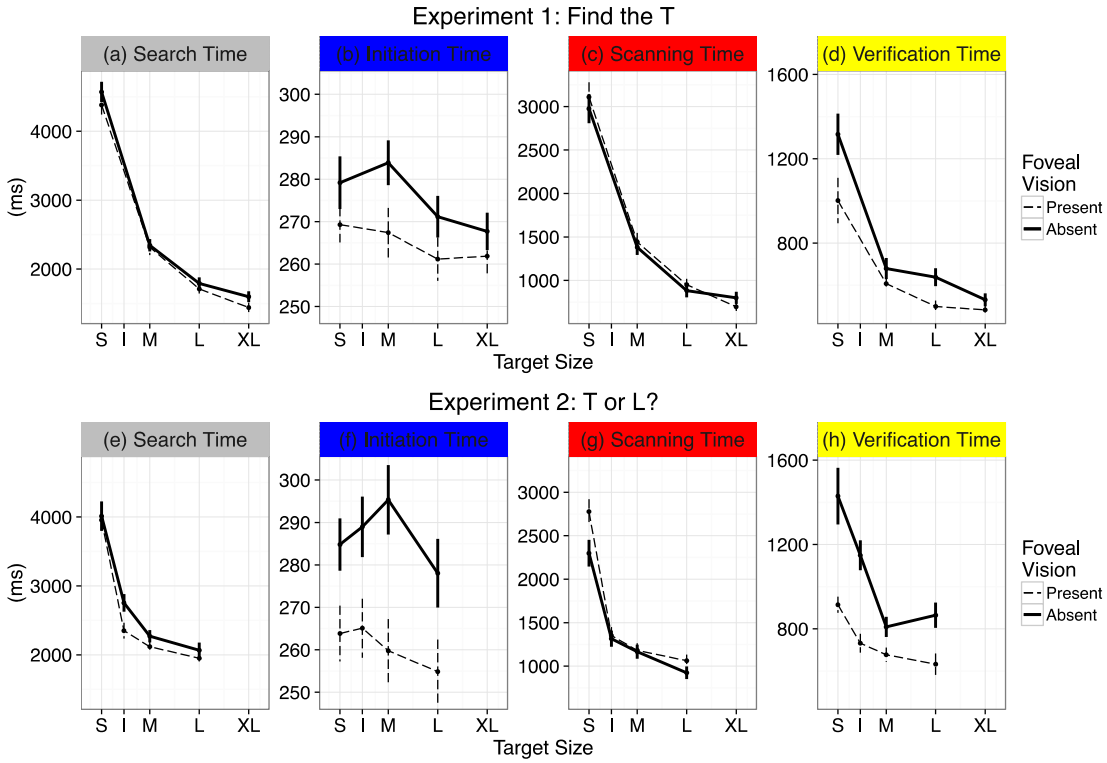



***Figure 11.*** **Search time and its three epochs for Experiment 1 (top row) and Experiment 2 (bottom row). Each column presents means obtained for a designated dependent variable (see panel title). For a given dependent variable, the *y*-axis has been normalized across plots for ease of comparison between experiments. For the three sub-processes of search (initiation, scanning, verification), the background color of the facet strips is using the color scheme from Figure 10. Solid bold lines represent the foveal scotoma condition, dashed lines the normal-vision control condition. Target sizes on the *x*-axis are described by letters (S: Small, I: Intermediate - Experiment 2 only, M: Medium, L: Large, XL: Extra Large - Experiment 1 only). The *x*-axis is scaled to show all target sizes across both experiments; the spacing on the *x*-axis preserves the relative distances between target sizes. Error bars are within-subjects standard errors (Cousineau, 2005).**

### Comparison of verification times across experiments

Whereas Experiment 1 required observers to search for the letter T, Experiment 2 included an explicit letter recognition component such that observers additionally indicated whether the letter was a T or an L. We expected that any adverse effect of searching with a foveal scotoma as well as the interaction between foveal scotoma and target size should be stronger in Experiment 2 than in Experiment 1. This prediction was tested by means of a three-way mixed-design ANOVA with one between-subjects factor (experiment) and two within-subjects factors (scotoma and target size). Due to the slight differences in target sizes across experiments, only the three common sizes were included (S, M, L).

There were significant main effects of scotoma, $F(1, 62) = 22.93$, $p < .001$, and target size, $F(1.34, 83.2) = 60.78$, $p < .001$. Interestingly, there was neither a significant main effect of experiment, $F(1, 62) = 1.52$, $p = .223$, nor a significant interaction between experiment and scotoma, $F(1, 62) = 1.47$, $p = .231$. Thus, contrary to our expectations the detrimental effect of a foveal scotoma proved not to be stronger in Experiment 2 than in Experiment 1.

However, in the mixed ANOVA the interaction between scotoma and target size was significant, $F(1.37, 85.12) = 5.33$, $p = .014$. This result was different from the separate analyses in which the interaction was either not significant (Experiment 1) or only marginally significant (Experiment 2). Obtaining a significant result in the mixed ANOVA could be due to increased statistical power by pooling data from the two experiments. Another possibility is that this analysis only considered the three target sizes that were common to both experiments. For each experiment separately, we therefore ran an additional 2 × 3 repeated-measures

ANOVA, now including the three common target sizes only. There was still no significant interaction between scotoma and target size when XL-targets were excluded from Experiment 1, $F(1.24, 38.47) = 1.7$, $p = .201$. When I-targets were excluded from Experiment 2, the previously marginally significant interaction was now significant, $F(1.45, 44.88) = 3.76$, $p = .044$.

Returning to the mixed ANOVA results, the interaction between experiment and target size was not significant, $F(1.34, 83.2) = 1.32$, $p = .265$; the three-way interaction between experiment, scotoma, and target size was also not significant, $F(1.37, 85.12) = 0.28$, $p = .675$.

### 5.4.3 Saccade Amplitudes and Fixation Durations

Saccade amplitudes and fixation durations were analysed to characterise eye-movement behaviour during visual search (Figure 12). For both experiments, results for mean saccade amplitudes showed a significant main effect of scotoma, with longer saccades when searching with a foveal scotoma than without (Experiment 1: $F(1, 31) = 58.07$, $p < .001$, Experiment 2: $F(1, 31) = 52.95$, $p < .001$). There was also a significant main effect of target size, with saccade amplitudes somewhat increasing with a reduction in target size (Experiment 1: $F(3, 93) = 12.74$, $p < .001$, Experiment 2: $F(3, 93) = 10.87$, $p < .001$). The interaction between scotoma and target size was significant in Experiment 1 ($F(3, 93) = 3.67$, $p = .015$) but not in Experiment 2 ($F(3, 93) = 0.92$, $p = .433$).

Fixation durations also showed a significant main effect of scotoma, with longer fixation durations when searching with a foveal scotoma than without (Experiment 1: $F(1, 31) = 14.79$, $p = .001$, Experiment 2: $F(1, 31) = 12.59$, $p = .001$). There was also a significant main effect of target size, such that a reduction in target

size was associated with longer fixation durations (Experiment 1: $F(3, 93) = 21.66$, $p < .001$, Experiment 2: $F(3, 93) = 6.81$, $p < .001$). The interaction between scotoma and target size was not significant (Experiment 1: $F(3, 93) = 1.1$, $p = .352$, Experiment 2: $F(3, 93) = 0.62$, $p = .605$).

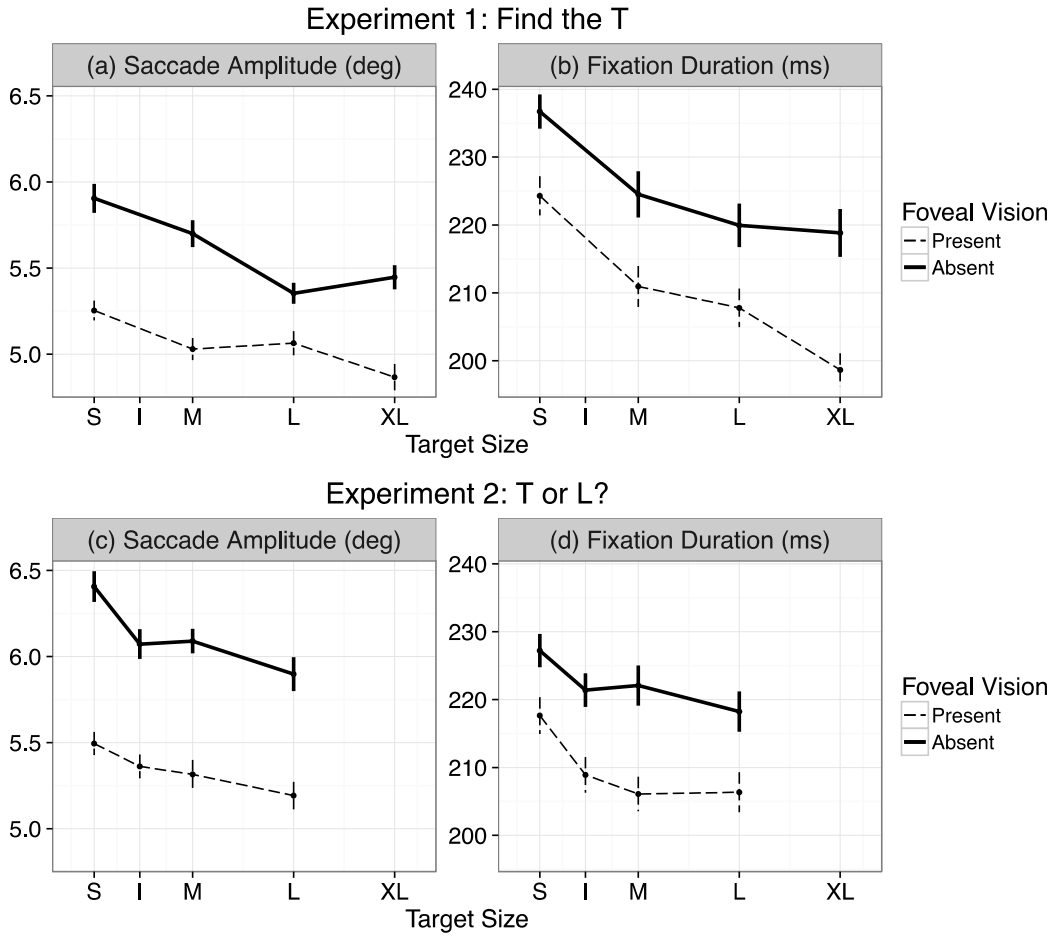



***Figure 12.*** **Mean saccade amplitudes and fixation durations for both experiments. Solid bold lines represent the foveal scotoma condition, dashed lines the normal-vision control condition. Target sizes on the *x*-axis are described by letters (S: Small, I: Intermediate - Experiment 2 only, M: Medium, L: Large, XL: Extra Large - Experiment 1 only). The *x*-axis is scaled to show all target sizes across both experiments; the spacing on the *x*-axis preserves the relative distances between target sizes. Error bars are within-subjects standard errors (Cousineau, 2005).**

## 5.5 General Discussion

Two experiments were conducted to test the degree to which foveal vision was necessary for the localisation, identification, and verification of context-free target letters in naturalistic scenes. A gaze-contingent moving mask (Rayner & Bertera, 1979) was used to simulate the absence of foveal vision. In Experiment 1, observers searched for the letter "T" which could occur at four different sizes. In Experiment 2, the target was either a "T" or an "L", and participants had to indicate which letter it was. If foveal vision was necessary to achieve normal search performance, the time taken to find the target should be significantly longer without foveal vision than with. Moreover, we reasoned that the importance of foveal vision may depend on the size of the search target such that foveal vision loss may be more detrimental for smaller targets. The addition of a letter recognition component to the task in Experiment 2 was expected to highlight the importance of foveal vision to the final process of target verification in particular.

While searching for the target without foveal vision, observers were significantly less likely to find the target than with normal vision, and this was most pronounced for the smallest targets. Our main analyses considered all correct trials ("hits"), for which we analysed search times along with three sub-processes of search (cf. Nuthmann, 2014). In both experiments, without foveal vision participants were not impaired in locating the search target in the scene (indexed by scanning time), but the process of positively identifying the target and responding was delayed (indexed by verification time). Moreover, search initiation times were significantly prolonged with a foveal scotoma in Experiment 2, but not in Experiment 1. During scene search, average verification times are typically shorter than scanning times, and initiation times are shorter still. This is why the button-press search times—the

sum of these three epochs—did not significantly differ between the foveal-scotoma and the normal-vision control conditions, despite a significant effect of foveal scotoma on verification time. We suggest that target acquisition with a foveal scotoma occurred in an all-or-nothing fashion. In trials where the target was found, search time was similar to controls; otherwise the target was not found at all.

How does this pattern of results compare to existing research on scene search? Our key finding was that a simulated foveal scotoma did not significantly prolong search times, which is compatible with results from two studies in which observers were asked to look for either spatial distortions (McIlreavy et al., 2012) or medium-sized real objects (Nuthmann, 2014). Beyond that, foveal vision appeared to be more important in the present letter-in-scene search tasks than during object-in-scene search (Nuthmann, 2014). Neither search accuracy, nor search time or any of its components were affected by a simulated foveal scotoma in Nuthmann (Nuthmann, 2014). In contrast, search accuracy was significantly lower, and target verification time significantly prolonged in the present experiments, in which the target was a context-free letter rather than a contextually relevant object.

Recent results regarding the unimportance of foveal vision for visual search in naturalistic scenes (McIlreavy et al., 2012; Nuthmann, 2014) were surprising, given the importance of foveal vision in both reading (Rayner & Bertera, 1979) and visual search within alphanumeric displays (Bertera & Rayner, 2000). For the present experiments, we therefore combined design features from letter search (Bertera & Rayner, 2000) and scene search tasks (McIlreavy et al., 2012; Nuthmann, 2014) by embedding letters into images of real-world scenes. Our findings for search time are different than those by Bertera and Rayner (Bertera & Rayner, 2000) for a task in which observers searched for the target letter "y" in displays consisting of

letters and numbers. This type of search task required the extraction of fine detail via foveal analysis. Observers had to analyse each element of the display to make an accept or reject decision. In contrast, our naturalistic scenes contained exactly one target letter to be analysed against the scene background. In Experiment 1, observers searched for the same letter on each trial (a "T"). In Experiment 2 we added a letter recognition component to the task (is the target a "T" or an "L"?). Qualifying the search-time findings, we found that the verification process took longer to complete with a foveal scotoma than without. We hypothesised that this effect should be more pronounced in Experiment 2 than in Experiment 1, but this was not the case. In sum, the present results point towards the relative unimportance of foveal vision to visual search.

Moreover, the experiments were designed to test whether target size was an important mediating factor for previous findings on the (un)importance of foveal vision during scene search (McIlreavy et al., 2012; Nuthmann, 2014). Not surprisingly, the data from both experiments showed a main effect of target size with better search performance for larger targets, in keeping with previous research (Miellet et al., 2010). The critical question was whether the size of the search target would affect the importance of foveal vision to the task (Miellet et al., 2010). Specifically, we hypothesised that any detrimental effect of the foveal scotoma on the target verification process may only occur for smaller targets, or may be more pronounced for smaller than for larger targets. Moreover, we reasoned that any adverse effect of the foveal scotoma and the interaction with target size should be stronger in Experiment 2 than in Experiment 1. We found very limited, if any, support for these predictions.

The present results replicate the finding that fixation durations and saccade amplitudes are both elevated in the presence of an artificial scotoma (Bertera & Rayner, 2000; Cornelissen, Bruin, & Kooijman, 2005; Miellet et al., 2010; Nuthmann, 2014). The saccade amplitude adjustment reflects a tendency to fixate more locations in the non-degraded scene area than the degraded area (Nuthmann, 2014). Both global eye-movement parameters were also affected by target size; a reduction in target size was associated with both larger saccade amplitudes (Miellet et al., 2010) as well as longer fixation durations. These findings were unexpected, because participants had no way of knowing which target size would be displayed next, due to the randomised presentation of scenes. Over the course of scene viewing, there is a tendency for fixation durations to increase and saccade amplitudes to decrease (Pannasch, Helmert, Roth, Herbold, & Walter, 2008; Unema, Pannasch, Joos, & Velichkovsky, 2005). In our experiments, search time equates to viewing time, such that the longer search times for small targets could potentially explain the longer fixation durations (but not the larger saccade amplitudes). However, time course analyses (not reported here) provided no evidence for this. Previous research has demonstrated that the "story," or gist of a scene can be gleaned from it within around 100 ms of the onset of a scene (Oliva, 2005; Potter, 1975). Scene gist is typically perceived without recognising any individual object. Moreover, research using simple displays suggests that object size is a property which can be processed preattentively and in parallel across the visual field (Stuart, Bossomaier, & Johnson, 1993). Therefore, the possibility exists that observers, during the first glance of the scene, form a hypothesis about the scene's search difficulty in terms of target size, and globally adjust their fixation durations and saccade amplitudes accordingly. Of course, fixation durations in particular are known to be influenced by more complex online visual and cognitive processing

(Nuthmann, 2016) such that the present speculations require additional systematic research.

Theories of visual search have largely been built on search for targets in arbitrary 2D arrays of items which observers searched without moving their eyes (Wolfe & Horowitz, 2017, for review). However, most real-world search takes place in structured scenes which observers explore through eye movements. The adoption of more ecologically valid stimuli has led to a new brand of image-based search theory (Eckstein, 2011, for review). Most of these models ignore that visual acuity declines systematically from the central fovea into the periphery (Nuthmann, 2014). The present results, together with our previous findings, inform future model building by specifying how (un)important the different regions of the visual field are for different sub-processes of search.

One contribution of the present work was to systematically investigate the effects of target size on sub-processes of search. A related variable is target salience (Laurent Itti & Koch, 2000). The current approach was to manipulate target size whilst controlling for target salience by probing the scene for locations of median salience. Future studies may assess the independent effects of size and salience (see Chapter 6). Our Target Embedding Algorithm provides researchers with the necessary tools to manipulate such variables.

As was pointed out in the introduction of this article, the simulated loss of foveal vision provides clinical relevance as it is closely related to the vision loss experienced by patients with AMD. In the present experiments, foveal vision was relatively unimportant for visual search and only beneficial for target verification. In future

investigations, it might be possible to contribute towards visual aid applications for the visually impaired.

## 5.6 Conclusions

The current study demonstrates that by combining aspects of the two search paradigms of artificial scene viewing (Bertera & Rayner, 2000) with naturalistic scene viewing (Nuthmann, 2014), localisation of a target was unaffected by the absence of foveal vision. The pattern of results observed by the two experiments provides support to the notion that foveal vision is not necessary in order to attain normal search performance. However, it is beneficial for target verification and identification.

# 6 The effect of target size, salience, and degraded vision

## 6.1 Introduction

The aim of this chapter is to assess the role of visual salience on visual search; as there is currently a debate on whether attention is stimulus driven, or object driven. Although this chapter does not directly speak to this debate, understanding the role of visual salience is crucial for identifying regions of the visual field where visual salience is most important in relation to different sub-processes of search (see Chapter 5).

In the last chapter, I investigated the importance of foveal vision to the achievement of normal search performance. If foveal vision was necessary, then searching without it would lead to lower search accuracy and longer search time. One variable that was thought to influence foveal necessity was target size. We hypothesised that small targets would be unresolvable in parafoveal vision, which could lead to sequential foveation of potential target sites. The results showed that the role of foveal vision on the finding of these targets was relatively unimportant. However, the visual salience of the target was not manipulated but instead controlled for by algorithmically locating target areas of median salience. The

experiments presented in this chapter manipulate visual salience to demonstrate its importance on visual search.

## 6.2 Stimulus-driven vs object driven attention

There is a current debate between whether overt attention in natural scenes is guided more by objects (Einhäuser, Spain, & Perona, 2008; Nuthmann & Henderson, 2010) or low level image features (Stoll, Thrun, Nuthmann, & Einhäuser, 2015). The aim of the present work was to investigate the effect of visual salience and target size on visual search performance in natural scenes. In Experiment 3, *we predicted that performance without foveal vision should be similar to that of the control as was the case with Chapter 5, but should be worse when verifying targets in lowly salient regions.* Although it has been shown that foveal vision loss does not affect the localisation process, extending beyond this region may introduce localisation costs. Therefore, in Experiment 4, *we predict that visual salience would have a more prominent effect on search with a central scotoma, whilst having little to no effect with a peripheral scotoma.*

Visual search on contextually relevant objects has been investigated with respect to foveal vision (Nuthmann, 2014), showing that foveal vision loss does not reduce search accuracy or localisation. In contrast, it has been demonstrated that search for medium salience context free targets under foveal vision loss reduced search accuracy, but not localisation (see Chapter 5). Peripheral vision on the other hand has been shown to be essential for obtaining an efficient search performance (Loschky & McConkie, 2002).

Visual search is a goal-oriented task, which means that top-down factors are useful for facilitating the search process. Top-down models such as the one proposed by

Rao et al., (2002) utilise task-dependant templates alongside an area-based search. Other models such as Ehinger et al., (2009) compare top-down aspects such as scene context against salience to conclude that a scene context model is better for search guidance. Both, among others, treat the entire visual field as having equally high resolution, and therefore are not representative of the naturally degrading acuity that occurs in a real visual field (Nuthmann, 2014, for discussion). Nevertheless, a top-down approach to modelling eye movements is effective for when top-down task requirements are involved (Einhäuser, Rutishauser, & Koch, 2008; Hayhoe & Ballard, 2005; Buswell, 1935; Yarbus, 1967). It seems intuitive that searching for your coffee cup would involve looking for something cylindrical in shape, relatively small, and on a surface such as a desk rather than spending time searching a scene's intrinsically interesting features (Kanan et al., 2009).

Stimulus-driven, or bottom-up factors on the other hand have been studied to investigate eye guidance to regions that are independent of the observer's knowledge or task. Koch and Ullman (1985) introduced the concept of a "saliency map" for the accomplishment of attentional selection. Their representation provided an explanation of how attention moves over a scene, however, combining information from different features into a single map proved to be difficult. Since then, the bottom-up model of visual attention by Itti and Koch (2000) has been implemented, which in turn has spawned several bottom-up salience models (see Borji, 2013b for a review).

Despite the above, integrating multiple processes, rather than studying them in isolation has been suggested to be a more effective method for guiding visual attention; especially in the real-world, where the environmentally rich information available caters to both top-down and bottom-up processes. Integration of different

processes has been demonstrated in the Contextual Guidance Model (Torralba et al., 2006), which constrains local saliency by global scene context. In counting search tasks, the model was accurate at predicting the first few eye movements made by the participant. This model suggests the overall benefit of integrating multiple sources of information as opposed to focusing on one aspect.

Empirical evidence against the above model has since been revealed, with results that demonstrate the visual system relying less on low-level saliency (Einhauser, Spain, & Perona, 2008; Foulsham & Underwood, 2007; Henderson, Brockmole, Castelhano, & Mack, 2007; Henderson, Malcolm, & Schandl, 2009; Zelinsky et al., 2006). Einhäuser et al., (2008) for example, explored the role of objects in scenes as a measure of fixation prediction over visual salience. Their study required participants to inspect a series of photographs before rating them on a scale from 1 to 5. In another condition, participants were also asked to search for a specific target and report on whether it was present in the scene. It was hypothesised that visual salience mainly acts through objects, and that it is the objects themselves that are important units of attentional selection in scenes. Their results showed that objects are a better predictor of fixations over early saliency. An object map for each image was computed, which provided a number of results supporting their hypothesis. They concluded that once objects were known; early salience did not contribute as much to fixation prediction. In attention guidance, early salience was shown to only have a small impact whereas "interesting" objects to the viewer provided more of an impact.

A counterargument was later published by Borji, Sihite, and Itti (2013c) in which the authors tested the alternate hypothesis against 11 saliency models including the ones used by Einhäuser, Spain, and Perona (2008). A shuffled area under curve

score was used in place of a regular AUC to discount central bias within the fixation data (Tatler, Baddeley, & Gilchrist, 2005). Three analyses were performed which investigated how evaluation scores may be affected by central bias, model selected, and the smoothness of object and saliency maps. The results of all three showed that objects do not predict fixations better than early saliency, however, objects were still reported to play a role in guiding attention.

Other top down factors such as task goal and scene semantics also play a role in guiding attention, with existing studies utilising these (among others) to do away with image-based salience maps in representing how attention is directed (Henderson et al., 2009; Underwood & Foulsham, 2006; Underwood, Foulsham, van Loon, Humphreys, & Bloyce, 2006).

In 2009, Henderson, Malcolm and Schandl developed their cognitive relevance hypothesis, which relies on the interaction of semantic knowledge of the scene with the goal of the current task. This hypothesis attempts to utilise a more visuospatial representation of how attention is directed, whilst simultaneously dismissing image-based salience maps. To test their framework, they conducted an experiment which analysed the number of fixations made on a lowly salient target object, against a highly salient region. Their results showed that participants were unlikely to saccade to the most salient scene region but were more likely to saccade towards the lowly salient target.

One criticism to this approach is that the task required participants to search for contextually relevant objects that were known to them beforehand. Because of this, participants would have built up a target template in their mind to guide their search process. It has been shown that having a target template generated via

picture cue can highly facilitate search performance (Malcolm & Henderson, 2009), which would suggest why participants didn't fixate the highly salient region as frequently as they may have otherwise.

These studies also contain target objects that have had their salience manipulated with an artistic software. This form of post-hoc editing of the stimuli can introduce artefacts in the image which could be identifiable by an observer. It is also unclear where the effects originate from, as confounding effects can arise from various forms of scene guidance (Biederman, Mezzanotte, & Rabinowitz, 1982).

The above summarises the importance of both image-based and object-based selection. It also highlights the fact that previous models favouring object-based selection neglect early salience by assuming a uniform distribution of fixations on objects, and that state-of-the-art salience models can outperform them. These two conflicting views referred to as the "salience-view" and the "object-view" have become a matter of heavy debate within the literature. Rather than favouring a particular side, the experiments of this chapter will instead acknowledge that objects differ in their visual saliency, and propose that saliency may be used to select between objects for fixations.

To limit the impact of these factors, this chapter proposes the following approach:

- To use context-free letter targets instead of contextually relevant objects to reduce confounding effects of scene guidance.
- To algorithmically identify locations for target placement to generate within-scene manipulations of target salience over post-hoc editing via an artistic software.

### 6.2.1 Scene Guidance

The target to find can affect your ability to acquire it due to semantic and syntactic guidance (Biederman, Mezzanotte, & Rabinowitz, 1982; Wolfe & Horowitz, 2017). If searching for a person, the likelihood of looking in the sky or on top of a wall as opposed to the ground would be reduced. Due to the observer's understanding on the world around them and their knowledge of the target, they are simply unlikely to look in such places. On the contrary, if the target were a bird, it is quite possible, and also natural, for it to be in the sky or on a wall. It has been shown that targets that violate one's understanding of semantics (e.g. a bar of soap next to a desktop computer) can produce different neural signatures to syntactic violations (e.g. a computer mouse on top of a laptop screen) (Võ & Wolfe, 2013). Similarly, when searching based on scene gist information, guidance is influenced through a combination of semantic and syntactic information (Oliva, 2005; Rensink, 2000). In addition, knowing information about the target prior can aid in one's knowledge of where in the scene that target could be located (Castelhano & Heaven, 2011; Malcolm & Henderson, 2010). In real world scenes, the use of real-world objects as search targets would provide additional search guidance that would act as a confounding factor in the experiments presented in this chapter.

Since real world objects are confounded with semantic knowledge, the experiments in this chapter use letter targets. Using context free targets, we eliminate scene guidance, which gives us better control over visual salience across a set of stimuli. The results from chapter 5 revealed that search times for such targets were longer the smaller the target became, but a comparison of search time between scotoma conditions did not differ. These targets were placed in locations of median salience using the Target Embedding Algorithm (T.E.A.). Using this algorithm, we can now

assess the effects of size and salience by placing small or large targets in locations of low or high saliency.

If we as observers utilise visual salience during a visual search task, then the performance of finding targets in lowly salient regions should reveal to be worse than finding targets in highly salient regions. Similarly, performance should be at its worst when finding small targets in lowly salient regions, and at its best when finding large targets in highly salient regions. What is not yet clear is how target size and salience interact with one another. We predict that by increasing target size, visual salience should have less of an effect on search performance, as the target is more likely to stand out regardless of salience region; whereas a reduction in target size should lead to search performance costs in general, with highly salient regions having less of a contributing factor.

Furthermore, this chapter assesses the effects of size and salience against foveal vision (Experiment 3) and central vs peripheral vision (Experiment 4). As established in Chapter 4, the loss of foveal vision did not change the time taken to find context free targets at median salience locations, which suggests that peripheral vision is more important for scene search than previously thought. This chapter presents two experiments. In Experiment 3, observers searched for the letter 'T' embedded at high or low salient regions of greyscale scenes of the real-world. This task was performed with or without foveal vision. In Experiment 4, we increased the scotoma size to beyond foveal vision, and introduced an inverse peripheral scotoma. The size of the letter varied in Experiment 3 but remained constant in Experiment 4 (see Section 6.4.4).

## 6.3 Experiments Three and Four

## 6.4 Methods

### 6.4.1 Participants

Thirty-two participants (15 males) between the ages of 18 and 27 (mean age 20) participated in the experiment, all of which had normal or corrected-to-normal vision. They gave their written consent prior to the experiment and were paid at a rate of £7 per hour for their participation.

### 6.4.2 Apparatus

Stimuli were presented on a 21-inch CRT monitor with a refresh rate of 140 Hz at a viewing distance of 90 cm, taking up a 24.8º × 18.6º (width × height) field of view. Throughout the experiment, the eye movements of the participants were recorded binocularly with an SR Research EyeLink 1000 Desktop mount system. More information on the apparatus is detailed in Chapter 3.

### 6.4.3 Stimuli

In Experiment 3, a total of 120 (+ 8 practice) greyscale images of real-world scenes were used. The search target was a letter 'T' at two different target sizes. Experiment 4 used the same 120 (+ 8 practice) stimuli, however, the size of the target remained constant throughout the experiment (see 5.3.4).

Target locations in both experiments were chosen as being of low or high salience. The T.E.A. was used (see Chapter 5) to accomplish this, the locations of which

relate to lower and upper quartile changes of local contrast in each image (see Figure 13 b and c). Target locations also differed between experiments.

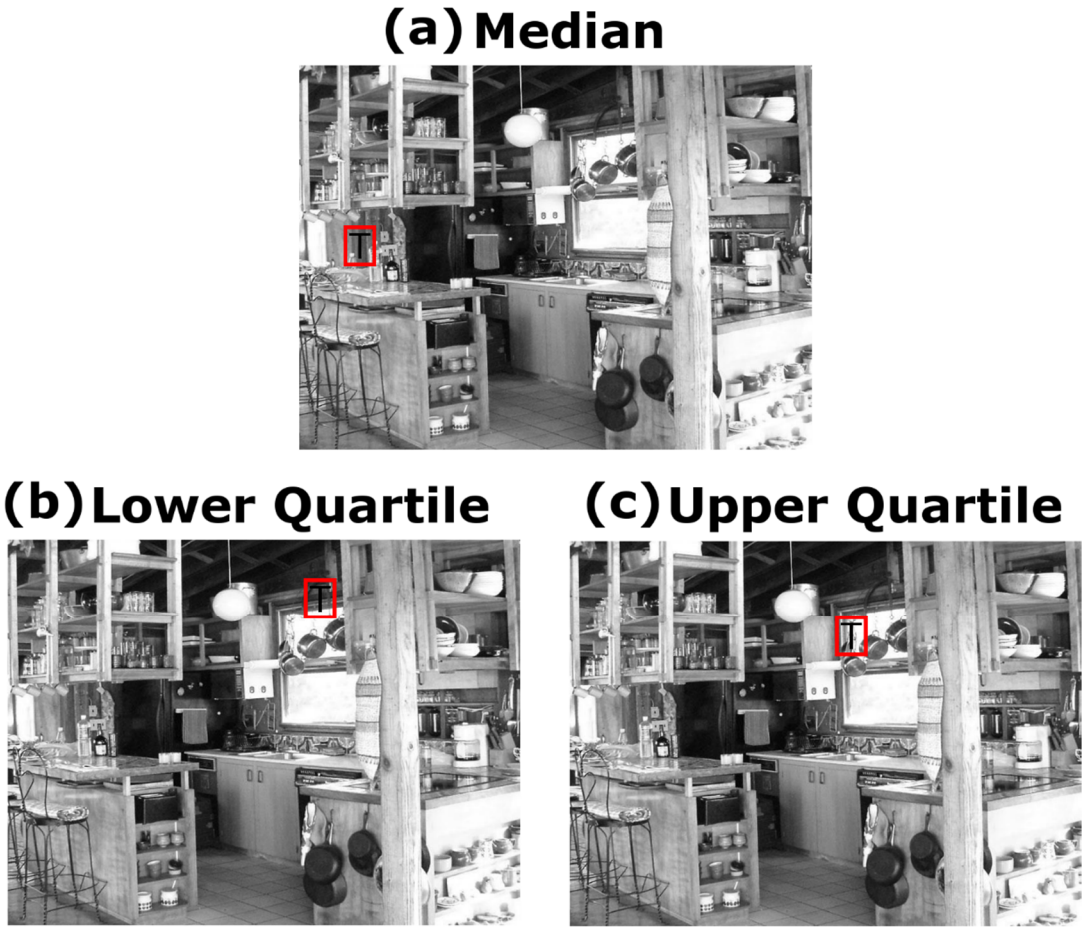



***Figure 13.*** **(a) Median, (b) Lower Quartile and (c) Upper Quartile salient regions produced with the T.E.A.**

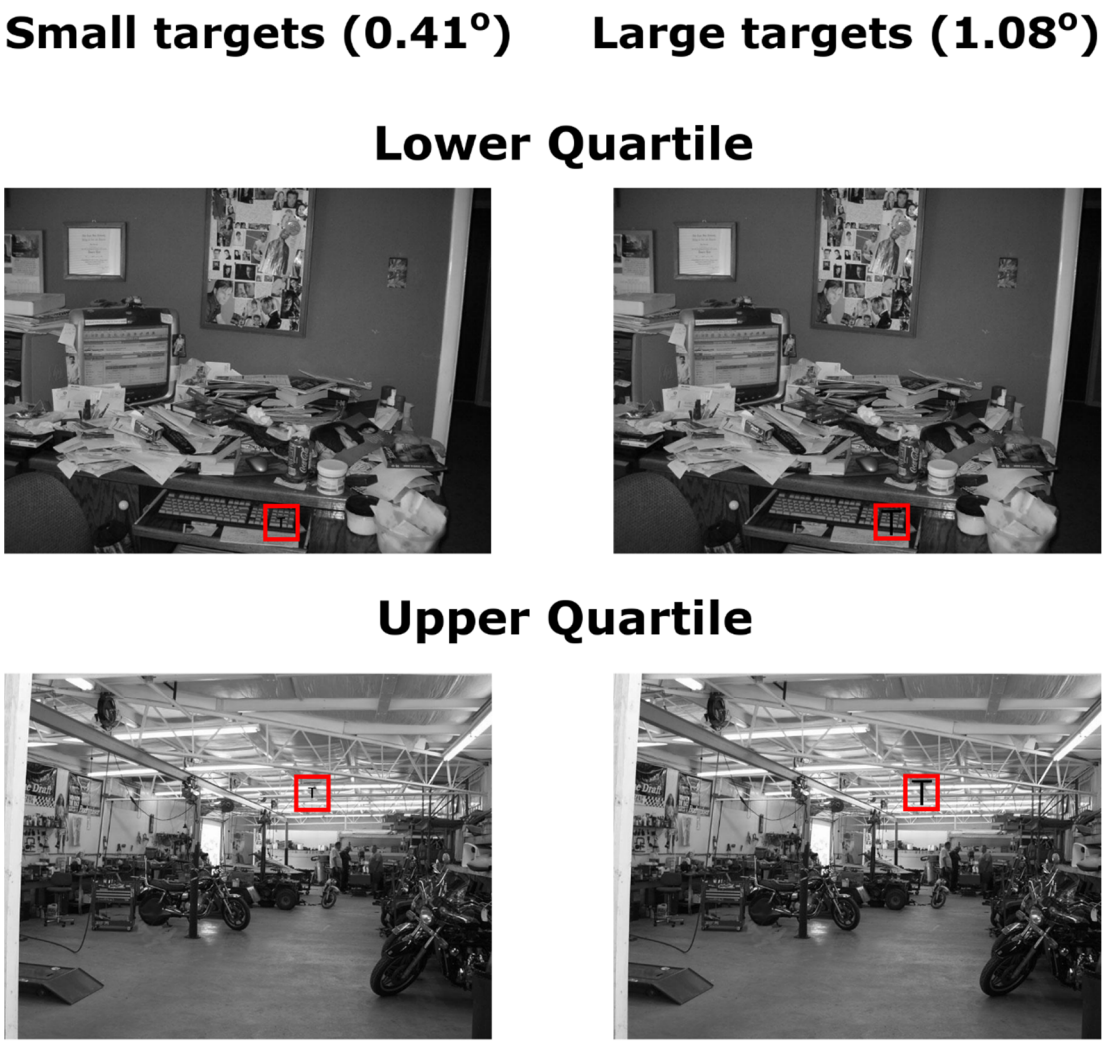



***Figure 14.*** **Example stimuli depicting small and large targets at their respective upper and lower quartile locations.**

### 6.4.4 Design

Experiment 3 used a 2×2×2 within-subjects design with 2-levels factor target size, scotoma, and salience. The factor scotoma refers to the availability of foveal vision. In the scotoma condition, foveal vision was blocked by a gaze-contingent moving mask (scotoma on, or foveal vision absent). This was contrasted with a normal-vision control condition (scotoma off, or foveal vision present). In Experiment 3,

the presence or absence of foveal vision was crossed with two target sizes valued at 0.41° and 1.08° in width.

Experiment 4 used a 3×2 within-subjects design with 3-levels factor scotoma, and 2-levels factor salience. The scotoma factor refers to the availability of central or peripheral vision. In one scotoma condition, a larger scotoma compared to Experiment 3 was used with a radius of 2.5°. This scotoma was then inversed to obscure peripheral vision, which resulted in a 2.5° radius viewing hole. This experiment also only used the smaller of the two sizes present in Experiment 3. The target sizes used throughout both experiments were chosen based on the data obtained from Chapter 5.

### 6.4.5 Procedure

At the start of the experiment, the participant's eyes were aligned with the eye tracker via the standard 9 dot calibration routine used with the Eyelink 1000. For a more detailed explanation of the calibration procedure, see Chapter 3 section 3.3.2.

### 6.4.6 Data analysis

Data were analysed by a two-way ANOVA (analysis of variance) with repeated measures implemented in the R statistical computing software (R Core Team, 2015). When at times Mauchly's test indicated that the assumption of sphericity had been violated, a Greenhouse-Geisser correction was then implemented to adjust the degrees of freedom.

Effect sizes are reported using the partial eta squared statistic with sample sizes calculated through the experimental design and further justified using G*Power (Faul, Erdfelder, Lang, & Buchner, 2007). For Experiment 3, in order to predict a medium effect size (f = 0.252), a minimum of 28 participants was needed. This number was obtained using an alpha of 0.05, a power of 0.95 and a $\eta^2p$ value of 0.06. For Experiment 4, In order to predict a medium effect size (f = 0.252), a minimum of 31 participants was needed. This number was obtained using an alpha of 0.05, a power of 0.95 and a $\eta_p^2$ value of 0.06. Partial omega-squared values are also given (Albers & Lakens, 2018; Yiğit & Mendes, 2018) and can be viewed in Appendix G.

Experiment 3 includes further analyses in the form of single and multiple regressions, and post-hoc comparisons via the Tukey method. Experiment 4 was firstly analysed by a two-way ANOVA but was then subject to post-hoc comparisons via the Tukey method.

## 6.5 Results and Discussion

The results presented here are firstly on search accuracy and search time, followed by results on saccade amplitudes and fixation durations (see Chapter 5 for general formatting).

### 6.5.1 Search Accuracy, Search Time and its Subcomponents

As in Chapter 5, performance for each experimental condition was divided into probabilities of 'hit', 'miss', and 'timeout' cases. The participant had 15 seconds to respond, otherwise the trial was coded as a 'timeout'. If the participant made a

response by button press within the allotted time, and had their gaze within the target interest area, then the response was scored as a 'hit'. Responses made outside of the interest area were scored as 'misses'. The size of the interest area in this chapter derives from the experiments in Chapter 5.

### Experiment 3

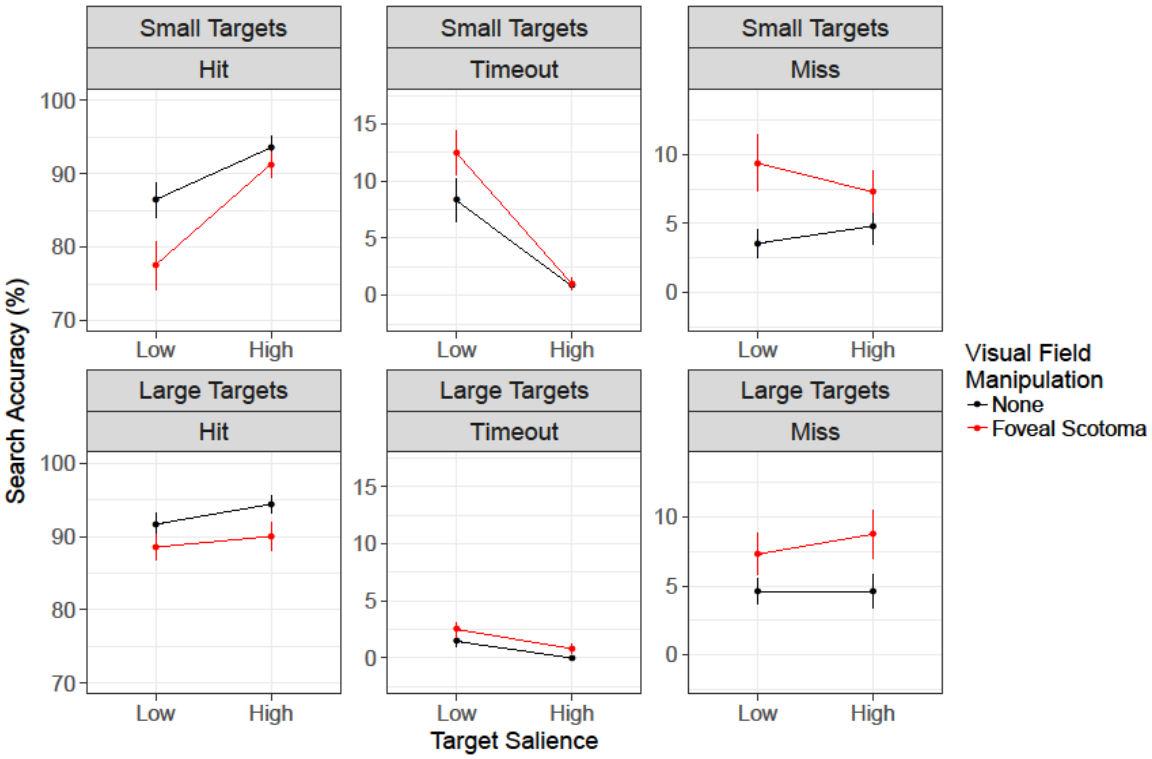



***Figure 15.*** **Results for Experiment 3. Left: hit probabilities. Middle: timeout cases. Right: target missed cases. Top: small targets. Bottom: large targets. Error bars are within-subjects standard errors (Cousineau, 2005).**

There were main effects of scotoma $F(1, 31) = 8.99$, $p = .005$, $\eta_p^2 = .22$, target salience $F(1, 31) = 29.31$, $p < .001$, $\eta_p^2 = .49$ and target size $F(1, 31) = 10.26$, $p = .003$, $\eta_p^2 = .25$ on hit percentage. The only interaction to reach significance was target size and target salience $F(1, 31) = 13.30$, $p < .001$, $\eta_p^2 = .30$.

For timeout cases, there were main effects of scotoma $F(1, 31) = 5.52$, $p = .03$, $\eta_p^2 = .15$, salience $F(1, 31) = 47.98$, $p < .001$, $\eta_p^2 = .61$, and target size $F(1, 31) = 36.03$, $p < .001$, $\eta_p^2 = .54$, with an interaction between target size and salience $F(1, 31) = 40.69$, $p < .001$, $\eta_p^2 = .57$.

For misses, there was a main effect of scotoma only $F(1, 31) = 8.68$, $p = .006$, $\eta_p^2 = .22$ with all other measures failing to reach significance.

There was no main effect of having a foveal scotoma during search time overall $F(1, 31) = 1.38$, $p = 0.249$, $\eta_p^2 = .04$, however, there were main effects on target size $F(1, 31) = 104.5$, $p < .001$, $\eta_p^2 = .77$ and salience $F(1, 31) = 235.51$, $p < .001$, $\eta_p^2 = .88$. The interactions between scotoma and salience $F(1, 31) = 0.58$, $p = 0.452$, $\eta_p^2 = .02$, as well as scotoma and target size were non-significant $F(1, 31) = 0$, $p = 0.968$, $\eta_p^2 < .001$. The interaction between salience and target size was significant $F(1, 31) = 31.37$, $p < .001$, $\eta_p^2 = .50$. The three way interaction was non-significant $F(1, 31) = 2.6$, $p = 0.117$, $\eta_p^2 = .08$.

For Initiation Time there were main effects scotoma $F(1, 31) = 9.83$, $p = 0.004$, $\eta_p^2 = .24$, salience $F(1, 31) = 24.7$, $p < .001$, $\eta_p^2 = .44$, and target size $F(1, 31) = 5.09$, $p = 0.031$, $\eta_p^2 = .14$. All interactions apart from salience and target size $F(1, 31) = 5.74$, $p = 0.023$, $\eta_p^2 = .16$ were non-significant.

There was no main effect of scotoma for Scanning Time $F(1, 31) = 0.65$, $p = 0.427$, $\eta_p^2 = .02$ but there were main effects of salience $F(1, 31) = 217.13$, $p < .001$, $\eta_p^2 = .88$ and target size $F(1, 31) = 88.13$, $p < .001$, $\eta_p^2 = .74$. The only significant interaction was between salience and target size $F(1, 31) = 24.38$, $p < .001$, $\eta_p^2 = .44$.

With Verification Time there were significant main effects of scotoma $F(1, 31) = 4.88$, $p = 0.035$, $\eta_p^2 = .14$, salience $F(1, 31) = 60.05$, $p < .001$, $\eta_p^2 = .66$, and target size $F(1, 31) = 28.86$, $p < .001$, $\eta_p^2 = .48$, with the only significant interaction being between salience and target size $F(1, 31) = 14.3$, $p = 0.001$, $\eta_p^2 = .32$.

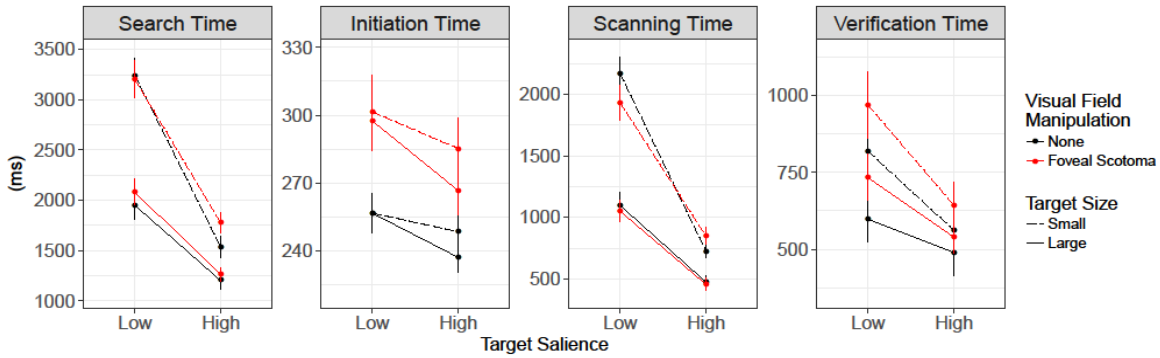



***Figure 16*** **Search time and its three epochs for Experiment 3. Dashed lines represent small targets whereas solid lines represent large targets. The red lines indicate the presence of a foveal scotoma, whereas the black lines represent normal scene viewing. Target salience is described along the x-axis as being low (lower quartile) and high (upper quartile) with the y-axis described in milliseconds. Error bars are within-subjects standard errors (Cousineau, 2005).**

## Experiment 4

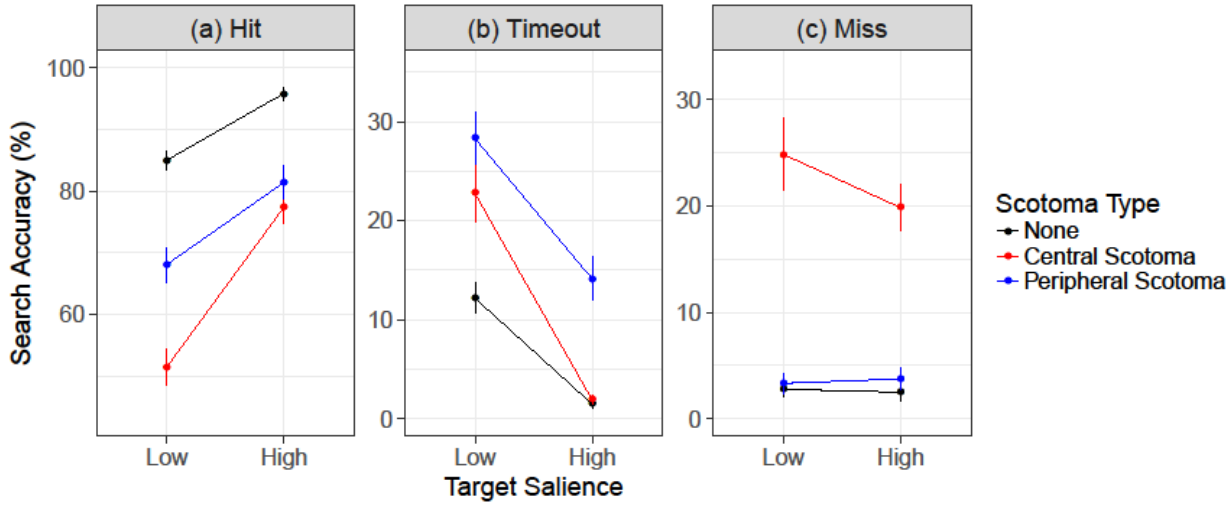



***Figure 17*** **Results for Experiment 4. Left: hit probabilities. Middle: timeout cases. Right: target missed cases. Note here that misses with a central scotoma were far more common as opposed to misses with a foveal scotoma. Error bars are within-subjects standard errors (Cousineau, 2005).**

There were main effects of scotoma for hit percentage $F(1.93, 67.46) = 50.37$, $p < .001$, $\eta_p^2 = .59$, timeouts $F(1.81, 63.43) = 22.85$, $p < .001$, $\eta_p^2 = .40$, but not for misses. There were main effects of target salience for hit percentage $F(1, 35) = 180.75$, $p < .001$, $\eta_p^2 = .84$, and timeouts $F(1, 35) = 141.01$, $p < .001$, $\eta_p^2 = .80$, but not for misses. There were interactions between scotoma type and target salience for hits $F(1.60, 56.08) = 14.46$, $p < .001$, $\eta_p^2 = .29$ and timeouts $F(1.93, 67.38) = 6.44$, $p = .003$, $\eta_p^2$ .16, but not for misses.

A Post-hoc Tukey test for hit percentage showed that within the low salience group, the control and peripheral scotoma conditions differed significantly at $p < .001$ as did the control and central scotoma conditions at $p < .001$ and the peripheral and central scotoma conditions at $p < .001$. In the high salience group,

the control and central scotoma conditions differed at $p < .001$, as did the control and peripheral scotoma conditions at $p < .001$. The central and peripheral scotoma condition did not differ significantly at $p = .764$. A Post-hoc Tukey test for timeouts showed that within the low salience group, the control and peripheral scotoma conditions differed significantly at $p < .001$, as did the control and central scotoma conditions at $p = .001$ but not the central and peripheral conditions at $p = .268$. In the high salience group, the control and central scotoma conditions did not differ at $p = 1.0$ but both the control and peripheral and the central and peripheral conditions did at $p < .001$. For misses, Post-hoc tests show that in the low salience group, the control and central scotoma conditions differed significantly at $p < .001$, but the control and peripheral condition did not at $p = .999$. The central and peripheral conditions differed at $p < .001$. In the high salience group, both the control and central scotoma condition and the central and peripheral condition differed at $p < .001$. The control and peripheral condition did not differ at $p = .994$.

### *Search Time and its Subcomponents*

For Search Time (Figure 18, panel 1) there were main effects of scotoma type $F(1.77, 61.83) = 361.49$, $p < .001$, $\eta_p^2 = .91$, salience type $F(1, 35) = 164.77$, $p < .001$, $\eta_p^2 = .82$, with an interaction between the two $F(1.9, 66.42) = 13.26$, $p < .001$, $\eta_p^2 = .27$. A Post-hoc Tukey test showed that within the low salience group, the control and peripheral scotoma conditions differed significantly at $p < .001$ as did the peripheral and central scotoma conditions at $p < .001$. The control and central scotoma conditions did not differ significantly at $p = 0.093$. Within the high salience group, the control and peripheral scotoma conditions differed significantly at $p < .001$ as did the central and peripheral conditions at $p < .001$. The control and central scotoma conditions did not differ significantly at $p =$

0.512. Other post hoc comparisons are not relevant for this chapter's hypotheses, but the full tables can be viewed in Appendix B.

For Initiation Time (Figure 18, panel 2) there were main effects of scotoma type $F(1.8, 63.02) = 43.86$, $p < .001$, $\eta_p^2 = .56$ and salience type $F(1, 35) = 8.12$, $p = .007$, $\eta_p^2 = .19$ with no significant interaction $F(1.73, 60.43) = 0.27$, $p = .73$, $\eta_p^2 = .008$. As there was no interaction, and that initiation time is irrelevant to this chapter's hypotheses, post hoc comparisons are not reported here (but can be seen in Appendix B).

For Scanning Time (Figure 18, panel 3) there were main effects for scotoma type $F(1.46, 51.27) = 536.57$, $p < .001$, $\eta_p^2 = .94$ and salience type $F(1, 35) = 116.12$, $p < .001$, $\eta_p^2 = .77$, with an interaction between the two $F(1.78, 62.3) = 10.22$, $p < .001$, $\eta_p^2 = .23$. A post hoc Tukey test showed that for the low salience group, the control and peripheral conditions differed significantly at $p < .001$, as did the central and peripheral conditions at $p < .001$. The control and central conditions did not differ at $p = .70$.

For Verification Time (Figure 18, panel 4) there were main effects for both scotoma type $F(1.81, 63.3) = 23.4$, $p < .001$, $\eta_p^2 = .40$, and salience type $F(1, 35) = 44.58$, $p < .001$, $\eta_p^2 = .56$, with an interaction $F(1.23, 42.88) = 10.84$, $p < .001$, $\eta_p^2 = .24$. A post hoc Tukey test showed that for the low salience group, the control and central scotoma conditions differed significantly at $p < .001$ as did the control and peripheral at $p < .001$ and the central and peripheral at $p < .001$. All relevant comparisons within the high salience group did not differ significantly: control and central at $p = .132$, control and peripheral at $p = .999$, central and peripheral at $p = .241$.

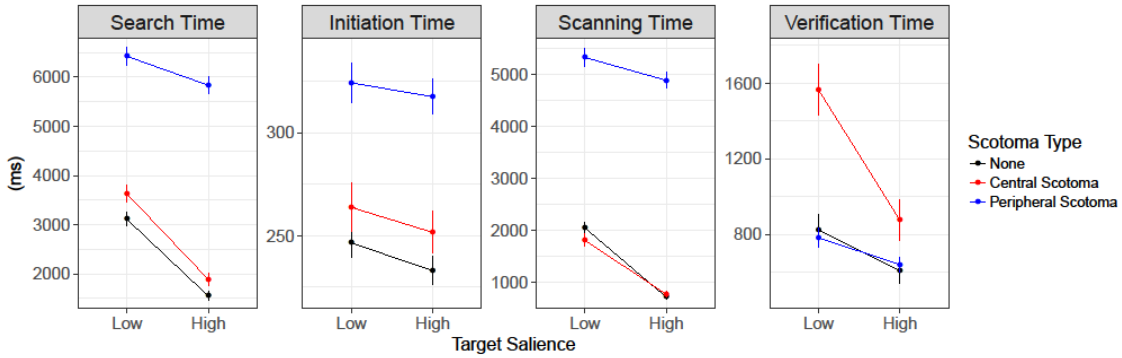



***Figure 18*** **Search time and its three epochs for Experiment 4. Target salience is represented along the x-axis with milliseconds along the y-axis. The black line represents normal scene viewing. The red line represents the central 2.5º radius scotoma. The blue line represents the peripheral scotoma, which provides a 2.5º radius spotlight. Error bars are within-subjects standard errors (Cousineau, 2005).**

### 6.5.2 Saccade Amplitudes and Fixation Durations

#### Experiment 3

There were main effects of having a foveal scotoma $F(1, 31) = 25.77$, $p < .001$, $\eta_p^2 = .45$, salience $F(1, 31) = 70.38$, $p < .001$, $\eta_p^2 = .69$, and target size $F(1, 31) = 27.47$, $p < .001$, $\eta_p^2 = .47$ for saccade amplitudes. There were also interactions between scotoma and target size $F(1, 31) = 4.55$, $p = 0.041$, $\eta_p^2 = .13$, as well as salience and target size $F(1, 31) = 4.3$, $p = 0.046$, $\eta_p^2 = .12$.

For fixation durations, there were main effects of scotoma $F(1, 31) = 12.77$, $p < .001$, $\eta_p^2 = .29$, salience $F(1, 31) = 59.94$, $p < .001$, $\eta_p^2 = .66$, and target size $F(1, 31) = 14.97$, $p < .001$, $\eta_p^2 = .33$. There was an interaction between scotoma and target size $F(1, 31) = 7.62$, $p = 0.01$, $\eta_p^2 = .20$ but no others.

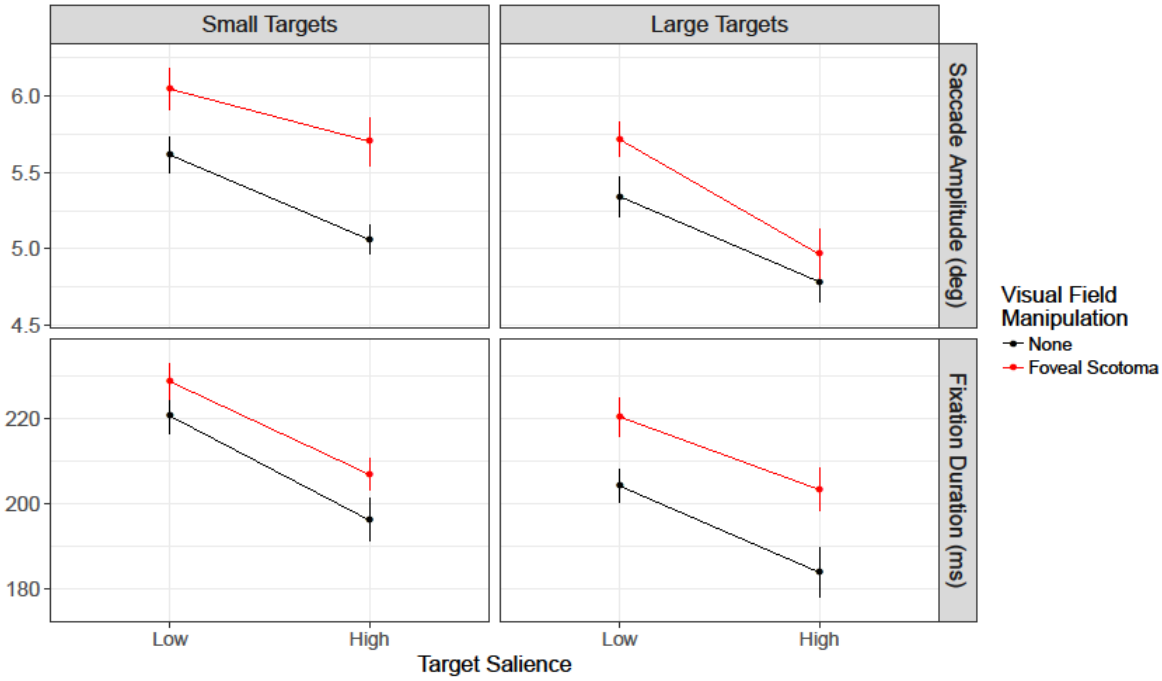



***Figure 19*** **Mean saccade amplitudes and fixation durations for Experiment 3. The columns represent target size with target salience desribed along the x-axis. The y-axis is shown in degrees for saccade amplitude, and milliseconds for fixation duration. Error bars are within-subjects standard errors (Cousineau, 2005).**

## Experiment 4

For Saccade Amplitudes (Figure 20, panel 1) there were main effects of scotoma type $F(1.49, 52.22) = 410.07$, $p < .001$, $\eta_p^2 = .92$ and salience type $F(1, 35) = 37.7$, $p < .001$, $\eta_p^2 = .52$ with an interaction between the two $F(1.33, 46.42) = 7.58$, $p = .005$, $\eta_p^2 = .18$. A post hoc Tukey test showed that for the low salience group the control and central scotoma conditions differed significantly at $p < .001$, as did the control and peripheral conditions at $p < .001$ and finally, the central and peripheral conditions at $p < .001$. As for the high salience group, the control and central conditions differed significantly at $p < .001$ as did the control and

peripheral at $p < .001$ and finally the central and peripheral conditions at $p < .001$.

For Fixation Durations (Figure 20, panel 2) there were main effects of scotoma type $F(1.68, 58.66) = 24.74$, $p < .001$, $\eta_p^2 = .41$ and salience type $F(1, 35) = 46.38$, $p < .001$, $\eta_p^2 = .57$ with an interaction between the two $F(1.84, 64.45) = 25.54$, $p < .001$, $\eta_p^2 = .42$. A post hoc Tukey test showed that for the low salience group, the control and peripheral conditions differed significantly at p = .003 as did the central and peripheral at p = .037 but the control and central did not at p = .963. All three relevant comparisons within the high salience group did differ significantly, with the control and central at p < .001, control and peripheral at p < .001 and finally the central and peripheral at p < .001.

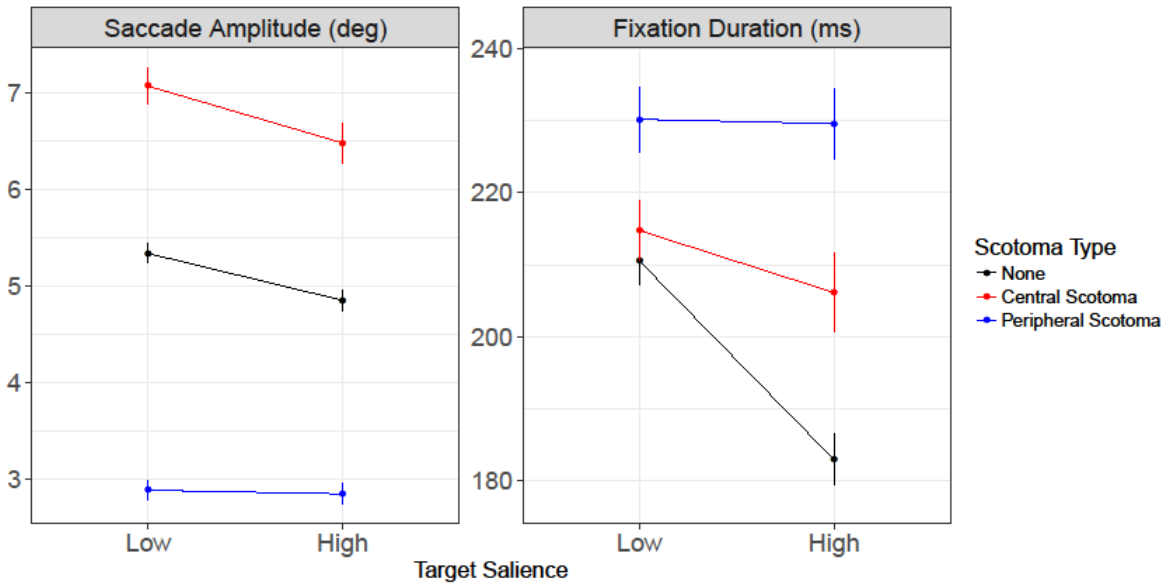



***Figure 20*** **Mean saccade amplitudes and fixation durations for Experiment 4. The x-axis describes target salience with the y-axis represented in degrees for saccade amplitude, and milliseconds for fixation duration. The surprising result here is that fixations are longest with a peripheral scotoma. Error bars are within-subjects standard errors (Cousineau, 2005).**

## 6.6 The Compensation Effect

The results from Experiment 3 showed that salience can compensate for small targets to aid in search performance. Similarly, an increased size can compensate for when there is a lack of salience. This phenomenon, for the purpose of this thesis, shall be coined the Compensation Effect (CE). Why does this occur? Let us first examine the target in question, a capital letter 'T'. What information can we gather from this? We know that it contains high spatial frequency information (Chung, Legge, & Tjan, 2002; Majaj, Pelli, Kurshan, & Palomares, 2002) and resolving the identity of such information requires high acuity processing (typically from the fovea) (Rayner 1998). We also know from the search literature that targets can easily pop out for numerous reasons (Bauer, Jolicoeur, & Cowan, 1998, 1999; Bauer, Jolicoeur, & Cowan, 1996; D'Zmura, 1991; Wyszecki & Stiles, 2000) and that gist processing can allow an observer to gather relevant information to better understand the scene (Oliva, 2005; Potter, 1975).

Without taking salience into account, it is possible that the combination of the target's features and its size will contribute to pop out effects, which is evident when we examine the data points of Figure 16 corresponding to high salience and large target size. With the addition of salience, particularly locations of low salience, pop out effects are less likely. Gist processing is also unlikely to be of much use, as the targets in lowly salient regions would be less likely to be processed compared with targets in highly salient regions. However, salience is not the issue when it comes to larger targets, which suggests that possible top down influences are contributing to the CE. The most notable of which is the Target Template Paradigm (Malcolm & Henderson, 2009) where observers are shown what to find prior to search.

In combination with the target's features, it does not matter if the target is placed in a lowly salient location; as it would be large enough to locate in the periphery at a reasonable performance. If we observe the graphs on saccade amplitudes in Figure 19, this hypothesis is further supported from the interaction effect seen between salience and target size. Upon inspection of small targets at highly salient regions, the saccade amplitudes were on average similar to observers searching for lowly salient but larger targets. I hypothesise that the Compensation Effect occurs when bottom-up salience overrides the cognitive processing induced through top-down factors, such as target templates, but only when targets are small. Similarly, the opposite occurs when targets are large.

### 6.6.1 Post-Hoc Comparisons

As reported above in Experiment 3, there were interactions between target size and salience for reaction time overall, initiation, scanning and verification times, and saccade amplitude. A series of single and multiple linear regressions were calculated to predict search performance based on target size and salience. Target size was coded as small = 1, and large = 2, and salience was coded as low = 1, and high = 2. In addition, post-hoc pairwise comparisons were also investigated. It should be noted that the most important pair for the validity of the Compensation Effect is the cross comparison of *large* but *lowly salient* targets vs *small* but *highly salient* targets (abbreviated to L.L.S.H. targets). The hypothesis for the Compensation Effect is that bottom-up salience will override top-down factors if the target to find is small. If the target is large, the target's size overrides bottom-up salience. For this hypothesis to appear valid, there should be a significant difference within the above group for scanning time. In addition, as

saccade amplitudes were similar on average in the cross comparison, the difference between saccade amplitudes should not differ greatly.

### Total reaction time

A significant regression equation was found ($F(2, 253) = 11.3$, $p < .001$), with an adjusted $R^2$ of 0.463. Participants' predicted total reaction time was equal to 5016.23 – 813.19 (target size) - 1175.36 (salience). Participant's search performance decreased by 813.19 ms as target size increased, and by a further 1175.36 ms for each increase in salience. Both target size and salience were significant predictors of search performance. Single linear regressions showed significant regression equations for both target size ($F(1, 254) = 45.35$, $p < .001$) with an adjusted $R^2$ of 0.148; and salience ($F(1, 254) = 117.6$, $p < .001$) with an adjusted $R^2$ of 0.316. A Post-hoc Tukey test showed that the cross comparison of Large – Low vs Small – High differed significantly at $p = 0.009$.

Although the above analyses provide significant differences between L.L.S.H targets, it is important to note that this is for total search time. Without decomposing into smaller epochs, it is hard to tell whether these results have any significance towards the Compensation Effect. They do, however, show that both target size and salience are able to influence search performance as a whole. The sections below will investigate this result further with respect to the three subprocesses of search outlined by Malcolm and Henderson (2009).

### Initiation Time

It was hypothesised that gist processing and target pop out were unlikely to be contributors towards the Compensation Effect. If in fact they were contributing,

then we should see a significant difference in the cross comparison of L.L.S.H. targets.

A significant regression was found $F(2, 253) = 3.154$, $p = 0.04$), with an adjusted $R^2$ of 0.016. Participant's predicted initiation time was equal to 309.427 – 8.393 (target size) - 18.712 (salience). Participant's initiation time decreased by 8.393 ms as target size increased, and again by 18.712 ms for each increase in salience. Salience was a significant predictor of faster initiation times. Single linear regressions showed a significant regression equation for salience ($F(1, 254) = 5.25$, $p = 0.02$), with an adjusted $R^2$ of .016, but not for target size. A Post-hoc Tukey test showed the cross comparison of Large – Low vs Small – High *did not differ significantly at p = 0.21*.

The above pairwise comparison demonstrates that initiation time has little to no impact on the Compensation Effect. Salience contributed more to faster initiation times than target size, but collectively they were not enough to produce initiation times that were fast enough initiation times to warrant evidence towards gist processing or pop out effect contribution (Bacon-Macé, Macé, Fabre-Thorpe, & Thorpe, 2005; Larson & Loschky, 2009; Loschky et al., 2007).

### Scanning Time

For the Compensation Effect to be valid, the cross comparison of L.L.S.H. targets should differ significantly. This would mean that the millisecond difference that target size or salience has on the other during scanning time is large enough to produce compensatory effects.

A significant regression equation was found ($F(2, 253) = 134.9$, $p < .001$), with an adjusted $R^2$ of 0.512. Participants' predicted scanning time was equal to 3468.84 – 646.85 (target size) – 935.78 (salience). Participant's search performance decreased by 646.85 ms as target size increased, and by a further 935.78 ms for each increase in salience. Both target size and salience were significant predictors of faster scanning times. Single linear regressions showed significant regression equations for both target size ($F(1,254) = 50.87$, $p < .001$) with an adjusted $R^2$ of 0.163; and salience ($F(1,254) = 136.3$, $p < .001$) with an adjusted $R^2$ of 0.346. A Post-hoc Tukey test showed the cross comparison of Large – Low vs Small – High *differed significantly at $p = 0.015$*.

As the above pairwise comparison has shown, the L.L.S.H. targets differ significantly during scanning time. Individually, target size and salience contributed to faster scanning times, with their cross comparison being large enough to produce compensatory effects. This is evidence in support of not only the Compensation Effect, but also that bottom-up factors have an influence on search performance.

### Verification Time

Although there were no prior hypotheses with regards to verification time on the Compensation Effect, it is still useful to know whether the effect could be contributory during this phase of search. If there is an effect, then it would demonstrate that compensatory effects aided in a participant's ability to verify targets.

A significant regression equation was found ($F(2, 253) = 12.04$, $p < .001$), with an adjusted $R^2$ of 0.079. Participants' predicted verification time was equal to

1237.96 – 157.94 (target size) – 220.87 (salience). Participant's search performance decreased by 157.94 ms as target size increased, and by a further 220.87 ms for each increase in salience. Both target size and salience were significant predictors of faster verification times. Single linear regressions showed significant regression equations for both target size ($F(1, 254) = 7.69$, $p = .005$) with an adjusted $R^2$ of 0.025; and salience ($F(1, 254) = 15.49$, $p < .001$) with an adjusted $R^2$ of 0.053. A Post-hoc Tukey test showed the cross comparison of Large – Low vs Small – High *did not differ significantly at p = 0.42*.

Although target size and salience were significant predictors of verification time (as is also evident in Section 6.5.1.1) the cross comparison of L.L.S.H targets demonstrates that the millisecond difference was not enough to produce compensatory effects. This non-significant finding makes sense, given that the hypotheses encompassing the Compensation Effect derive from target localisation.

### Saccade Amplitude

To further aid in the validity of the Compensation Effect, a pairwise comparison between L.L.S.H. targets was conducted for saccade amplitudes to investigate whether the effect was not simply due to participants making larger saccades toward the target. Saccade amplitudes between L.L.S.H. targets should not differ significantly for compensatory effects to have a significant contribution to search performance.

A significant regression equation was found ($F(2, 253) = 25.04$, $p < .001$), with an adjusted $R^2$ of 0.158. Participants' predicted saccade amplitude was equal to 6.83 – 0.40 (target size) – 0.55 (salience). Participant's saccade amplitudes decreased by 0.4° as target size increased, and by a further 0.55° for each increase

in salience. Both target size and salience were significant predictors of saccade amplitude. Single linear regressions showed significant regression equations for both target size ($F(1, 254) = 15.55$, $p < .001$) with an adjusted $R^2$ of 0.053; and salience ($F(1,254) = 30.62$, $p < .001$) with an adjusted $R^2$ of 0.104. A Post-hoc Tukey test showed the cross comparison of Large – Low vs Small – High *did not differ significantly at p = 0.46*.

Saccade amplitudes did not differ significantly in the cross comparison of L.L.S.H. targets. This means that saccade amplitudes remained similar in the cross comparison, suggesting that the influence of target size or salience for L.L.S.H. targets did not contribute to saccade amplitude adjustment throughout search time.

## 6.7 General Discussion

Two experiments were conducted to investigate the effect of target size and salience on visual search with respect to various regions of the visual field. In Experiment 3, participants were instructed to search for a capital letter T within 2D static naturalistic scenes. The T varied between two sizes (0.41°, 1.08°) and embedded at either a low or high change of local contrast. Foveal vision was obscured in one condition, with vision intact in the other. It was hypothesised that if we utilise salience in visual search, then performance should be at its worst with lowly salient small targets and at its best with highly salient large targets. How these two factors interacted with one another was unclear prior to this investigation. In Experiment 4, we increased the scotoma size to beyond foveal vision, and introduced an inverse peripheral scotoma. The size of the letter varied in Experiment 3 but remained constant in Experiment 4 (see Section 6.4.4). It was

also hypothesised that localisation costs could be produced by increasing the scotoma size, as Chapter 5 showed that foveal vision loss did not affect the localisation process. Furthermore, by extending the scotoma beyond the fovea, we should see salience having a more prominent effect on search performance, whilst having little to no effect with the peripheral scotoma.

As established in Chapter 5, the loss of foveal vision did not change the time taken to find context free targets at median salience locations. Due to the manipulation of salience in the current chapter, the results indicate that visual salience affected search performance by interacting with target size. For small but highly salient targets, search time was comparable to large but lowly salient targets. Also, both times were considerably shorter than the harder conditions, e.g., small and low salience targets.

In figure 16, the solid and dotted lines represent target size, with salience on the x-axis. To observe the interaction effect between target size and salience, the same coloured lines but opposite points should be compared. That is, the leftmost point of the solid against the rightmost point of the dotted. Despite the size difference between targets, it appears that an increase in salience compensates for the reduction in size. This effect is mimicked in reverse, with increased size seemingly compensating for the lack of high salience (see Section 6.6). Due to the targets being context free, the only factor that observers could utilise was their own knowledge of what the target looked like. Other than this, there were no additional top-down influences that could affect their search, although it is important to consider the processing of scene semantics as the scenes used were naturalistic (Cornelissen & Võ, 2017). While Cornelissen and Võ (2017) reported only target absent trials, this was only to avoid instantaneous detection of the target, which

would yield too little eye movement data. Although their assumption was reasonable, instantaneous detection could be mitigated in the future via use of the T.E.A. to control the perceived salience of the target. The authors have since conducted additional studies that include target present trials (Cornelissen, Holmqvist, & Võ, 2017) the results of which replicate their previous findings.

Aside from the interaction between target size and salience, currently coined the Compensation Effect, the results have also replicated the non-significant effect on search performance with a foveal scotoma. Despite the influence of salience, participants were able to complete the task in a similar time to the control. The factors of target size and salience can easily make a task easier or harder, but with regards to localising targets, foveal degradation seems to have little detrimental effect.

The results of Experiment 3 are in support of both bottom-up and top-down views of visual search and has shown that target size is a factor that should be controlled for in such studies. As the Compensation Effect has demonstrated, target size becomes increasingly more relevant as salience decreases, and vice-versa. Post-hoc comparisons are in support of this hypothesis, showing that search performance between lowly salient large targets, and highly salient small targets differed significantly. The Compensation Effect should be reproduced in future studies that include target size and salience to further aid in its validity.

Experiment 4 was more concerned with identifying the region of the visual field where search costs originated, and the subprocess of search associated with that cost. The results of all experiments so far have demonstrated that performance with and without a foveal scotoma does not differ much with the exception of

search accuracy. Collectively, these results acted as a motivation to increase the scotoma size to a radius of 2.5°. The results on search time (see Figure 18) show very clearly that search without peripheral vision leads to prolonged search times but quick verification times. This is consistent with previous findings (Loschky & McConkie, 2002) and is unsurprising. What is surprising is the effect salience has in this condition as one would think that the two blue data points (Figure 18, panel 3) would be connected via a horizontal line, as the targets to find are completely obscured until they are brought into the 2.5° viewing region.

It is possible that participants took advantage of what little salient information was available in their immediate field of view, and that once highly salient targets came into view, they were easier to spot. As evident in Figure 20, participants adjusted their gaze pattern when their periphery was obscured, resulting in lower saccade amplitudes and longer fixation durations. Through experimenter observation, I could see that participants were slowly scanning the scene; and it is possible that this prolonged exposure to smaller regions allowed highly salient targets to be found faster. In the peripheral scotoma condition, the few millisecond difference salience produced on scanning time and also initiation time was a surprising result, but should not be overinterpreted as the difference was small.

The most prominent finding from Experiment 4 is the difference in verifying lowly salient targets between the control and central scotoma conditions. This result is entirely different to verification time results of previous experiments as the central scotoma condition results in a very long verification time on average compared with the control. Like before, the saccade amplitude and fixation duration results are consistent with previous findings (Bertera & Rayner, 2000; Cornelissen et al., 2005; Miellet et al., 2010; Nuthmann, 2014). As it stands, the detrimental effect

on verification time solely via the central scotoma is evidence of search cost origination. It could be reasoned that targets took longer to verify as the scotoma's size was very large in comparison. If that were the only case, then we should see similar effects across both low and highly salient targets. Instead, high verification times seem to be driven only for lowly salient targets.

Experiment 4 used relatively small targets (0.41°) which, accompanied with low salience makes them harder to detect overall. The processing of fine details is all but eliminated when fixating the target given the scotoma's size, although as scanning time was similar to the control, participants had little trouble locating the target regardless of salience. It is evident through the elevated saccade amplitudes and time spent verifying the target that participants are spending time fixating back and forth in order to verify their decision within their periphery. Participants were not necessarily fixating over the target any longer than the control (Figure 20, panel 2); which suggests that participants were actively trying to identify the target as opposed to passively thinking about it. In previous experiments, participants had parafoveal information to aid them during verification; whereas now, not only is this region obscured but the targets were also lowly salient. Other researchers claim that top-down factors dominate the search process (Chen & Zelinsky, 2006; Jovancevic, Sullivan, & Hayhoe, 2006; Turano et al., 2003), but the results of this Chapter demonstrate that salience does not only influence search, in particular with verification time, it can also aid in facilitating search for small targets.

## 6.8 Conclusion

The results of this chapter argue in favour of bottom-up features playing a role in visual search by demonstrating that salience can improve search performance when

targets are small and vice-versa through the Compensation Effect. Salience also seems to be more important outside the foveal region as Experiment 4 suggests; with highly salient targets having shorter verification times than their low counterparts. Finally, we now have evidence to suggest that significant search costs do not arise from foveal degradation, but also not as wide as central vision. Instead, a point extending just beyond foveal vision was enough to reduce search performance, with costs originating within the verification period.

# 7 The relative importance of foveal vision

## 7.1 Introduction

This chapter explores visual search performance within 3D immersive scenes. Three experiments are presented here which involve visual search with simulated self-motion. Due to the inclusion of optical-flow, the following questions were asked:

- Does the inclusion of optical-flow produce costs to search performance with foveal vision loss?
- Is target recognition also affected for objects that conform with the optical flow?
- If there is a cost, can we improve search performance by manipulating optic-flow trajectory and target placement?

In this chapter, we show that target motion that moves in the same direction and same speed with the flow field (optic flow conformity), requires observers to utilise their high acuity vision by fixating near the target for successful detection. The results of this chapter suggest that the optical flow interfered with the target localisation process, which results in an increased reliance on high acuity vision.

How could optical-flow produce search costs? It has been shown in studies of optic flow that judgements about object motion may require more focused attention (Royden & Hildreth, 1999). The targets used throughout this chapter are all non-moving relative to the scene, they will still conform with the optic flow induced by the observer motion, which may influence search performance. Secondly, targets that conform with optical flow may blend with the environment, making search without foveal vision more difficult.

If the above proves likely, then searching without the high visual acuity of foveal vision should also be more difficult than in previous chapters. It is known that recognition is a high acuity task (Strasburger et al., 2011), and targets that have their visual appearance degraded by motion will need this acuity for an effective search. If optical flow can influence search performance, then we hypothesise that manipulating the path of self-motion can positively or negatively affect search. Why do we need to do this? Previous chapters have shown that foveal vision is not needed in order to achieve normal search performance in static scenes. If experiments with moving scenes reveals the opposite, thereby forcing an increased reliance on foveal vision, then there must be a way to reduce this reliance such that the detrimental effect disappears. Experiment 7 investigates this by changing the path of motion from linear to curved, and observing search performance on optimised vs unoptimised target locations (see section 7.6).

Unlike previous chapters, the scenes presented in this chapter were constructed using the Unity game engine, to produce stimuli comprising of virtual 3D space with a z depth plane. The scenes were created in this way to allow for simulated movement through the scene and the presence of depth information. Furthermore, the scenes were constructed in 3D space as opposed to recordings of the natural

world as this allowed for the careful manipulation the search object's properties, most notably its visual salience and target size. Aside from the target, the environment can also be manipulated to the experimenter's needs; whereas the real world would be unpredictable, as there could be motion from objects (e.g. cars), as well as forces such as the wind affecting the environment (e.g. trees), all of which could draw unnecessary attention.

Visual search using immersive scenes that induce optic flow, in relation to the role of foveal vision, has not been investigated (see Section 4.5 for an overview of the virtual environment literature). Over the course of this thesis, and in a lot of the visual search literature, search has been studied in relatively artificial situations. This ranges from simplistic stimuli that are static frames of either the real world or an artificial display comprising of several artificial components (see Sections 4.2.1 and 4.2.2). 3D stimuli have also been used in eye-tracking studies (Drew et al., 2013; Kit et al., 2014; Pomplun, Garaas, & Carrasco, 2013), with such scenes being recently integrated with Virtual Reality (VR) headsets (Meißner, Pfeiffer, Pfeiffer, & Oppewal, 2017). Additionally, mobile eye tracking has been used (Foulsham et al., 2014; Kugler et al., 2014; Mack & Eckstein, 2011) but the study of optic flow fields themselves has usually been investigated by isolating the optic flow using basic stimuli (see figure 3, section 4.5). Such research refers to detection of object motion, distinguishing observer from object motion, and judgements of heading. A study by Royden, Wolfe, and Klempen (2001) did investigate visual search using flow fields, but they used basic stimuli, and they did not investigate search with degraded vision.

Visual search within dynamic scenes is normally conducted by simulating an environment that contains object motion with an onlooking observer. This usually

results in a forward-facing camera to act as the observer's field of view, which captures object motion (e.g. a person walking or a car moving). Although the scenes are dynamic, the camera is static and so lack the necessary depth information that would otherwise be available to the observer if they were walking through the environment. As a result, these scenes can easily be translated into their 2D counterparts for creating static controls. Furthermore, a study by Mital et al., (2011) has shown that such dynamic scenes can affect the results due to attention being drawn to the motion itself.

The scenes generated for experiments 5 and 6 use a forward-facing camera that travels along a straight path, with no object motion present within the scenes. To create the stimuli, the Unity Game Engine was chosen as it allowed for the creation of the environments as well as camera manipulation for scene traversal. For target placement, the original T.E.A. algorithm was used and extended to incorporate 3D space. This extension scanned each scene to separate out solid objects from the background. It then recorded all suitable locations that would remain present in the scene from start to finish. These locations were then processed in MATLAB to return locations a median contrast change.

## 7.2 T.E.A.M.S. Algorithm

The algorithm described here has several steps that use different software packages, and so is divided and explained accordingly.

The algorithm also uses a selection of terminology unique to the Unity Game Engine and are defined below:

- Mesh Collider
    - An invisible object moulded to the shape of virtual objects (meshes). These colliders allow collision detection between meshes and primitives.
- Raycast
    - A 'ray' or 'beam' that intersects with in-scene meshes. The ray is cast in one direction at a set distance from the camera.
- Rigidbody
    - A component that controls the attached object's position through physics simulation.
- Continuous Dynamic Collision
    - A form of collision detection available in Unity. It provides a more sophisticated form of collision detection for dynamic objects colliding with static objects.
- Sandbox
    - A large open space to place content for a scene.
- Position handles
    - Position handles are user interface (UI) elements that allow a user to move, rotate, and resize objects in a scene.

The length of each scene was determined by the camera linearly interpolating along a vector towards an invisible object in the scene, over 20 seconds. As both the start and end points were known, the equation of slopes was used:

$$\frac{y - y_0}{x - x_0} = \frac{y_1 - y_0}{x_1 - x_0}$$

This was because the size of each scene's sandbox differed (see Figure 21). The size of the sandbox does not matter, as only a small section of each scene was used for the experiments. These small sections were generated in two parts. First, the camera was manually placed in an area large enough for it to move in a straight line without colliding with other objects (e.g. houses). Secondly, the end point for the camera to reach was calculated by:

- Placing an invisible object some distance away from the camera;
- Linearly interpolating the camera's position between its initial start position and that invisible object's position whilst recording the traversal time taken in seconds;
- If the camera reached the invisible object's location before traversing for 20 seconds, the invisible object automatically repositioned itself a set distance away so that the camera could continue moving. In Experiments 5 and 6, this path generated was completely straight. In Experiment 7, the path was curved (see Section 7.2.4).
- Step 3 was repeated until the straight path was long enough to last 20 seconds of traversal.
- Once completed, the initial position of the camera and the newly obtained end position of the camera's path were stored.

Adjusting the length of the path in this way allowed for the same traversal time across scenes.

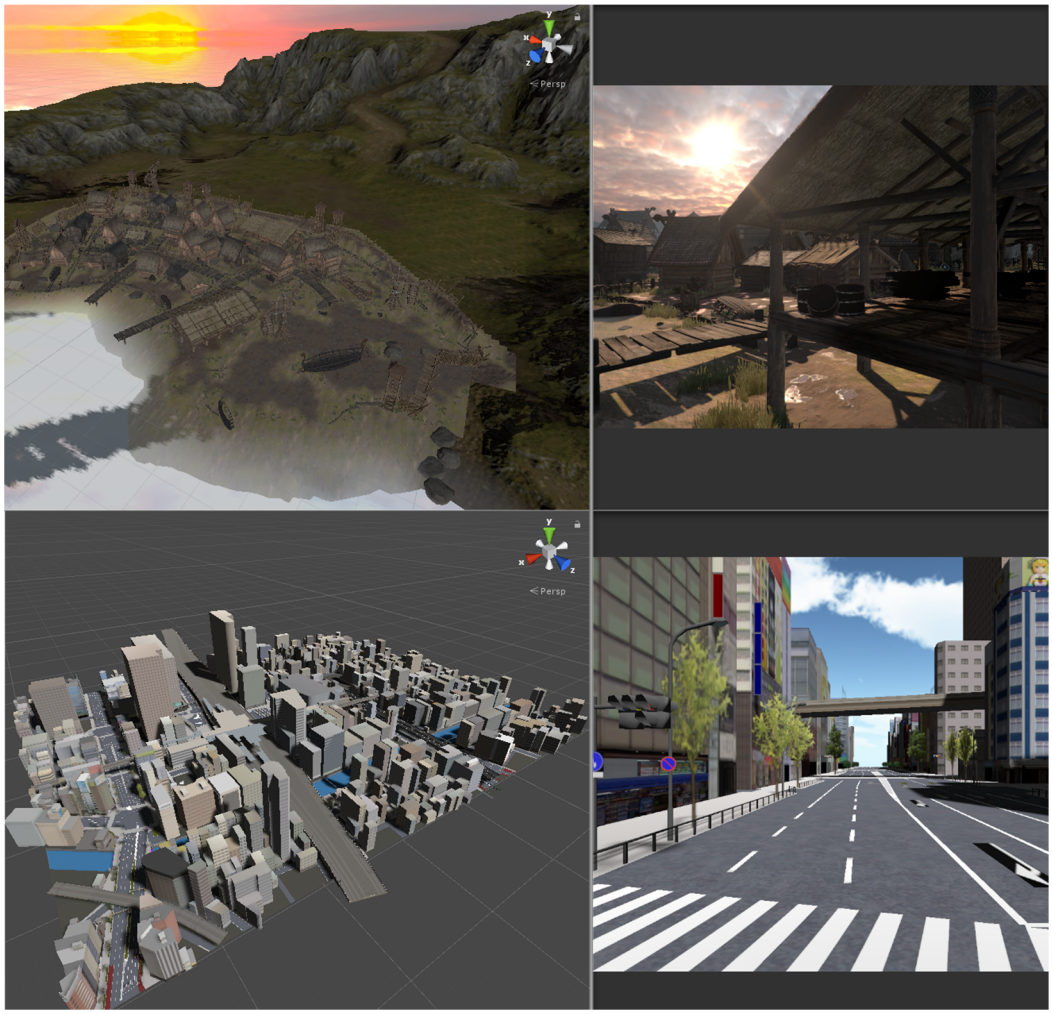


***Figure 21.*** **Sandbox comparison between two Unity Scenes. The top scene shows a zoomed out view of a village whereas the bottom scene shows a zoomed out view of an entire city. The scenes on the righthand side show the camera's perspective in each scene, which demonstrates that although the sandboxes are large, only a portion of them were used.**

Due to the size of each scene's sandbox being naturally different from one another, the perceived size of the target object in each scene differed. To accommodate for

this, an additional calculation was made to resize the target to be of an equal size across all scenes halfway through stimulus presentation:

$$diameter = target.bounds.extents.magnitude$$
$$pixelLimit = 24.8\, pixels\, per\, .8^o$$
$$target = (x_1, y_1) \rightarrow camera = (x_2, y_2)$$
$$distance(target, camera) = \sqrt{(\Delta x)^2 + (\Delta y)^2} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$
$$angularSize = 2 * \text{atan}\left(\frac{\frac{diameter}{2}}{distance}\right) * \left(\frac{180}{\pi}\right)$$
$$pixelSize = \left(\frac{angularSize}{camera.FOV}\right) * screenHeight$$
$$newDiameter = (2 * distance) * \tan(\left(\frac{1}{2}\right) * (\frac{pixelLimit}{screenHeight})$$
$$scale = \frac{newDiameter}{diameter}$$
$$targetNormalized = normalize(target.transform.localScale)$$
$$target.transform.localScale = (targetNormalized *= scale)$$

These steps altered the size of the target such that it was the same perceived size across all scenes at the midpoint, although it could be larger or smaller at the start and end points depending on the distance travelled in each experiment. The size to which the target grows however is minimal, with it only being relatively large once it reaches the observer, at which point the scene would have timed out. The target never moves past the observer either, as locations chosen (see 7.2.1) underwent a filtering process such that unsuitable locations were discarded. Furthermore, the size of the target always began at the same size, with its size always being the same at the midpoint across all scenes. As all scenes were 20 seconds long, and as the observer had a fixed speed, the natural growth of the target was also the same across all scenes. The target itself within the sandbox's

world space greatly differed, however, the perceived size in screen space was always the same.

### 7.2.1 Target location determination

2D images are (m x n) matrices, therefore moving a target pixel by pixel across the entire image is one way of generating appropriate locations for a visual search study. However, 3D scenes have depth and geometry (see Figure 23), therefore moving pixel by pixel across 2 dimensions would not make sense. This is because some scene regions do not have geometry such as the sky, and more importantly adding a target in this way gives no spatial information about the target's z coordinate. This means that the target could potentially float in 3D space and therefore generate its own salience due to non-optic flow conformity.

To prevent this, the following sequence was pre-processed:

- Scan entire scene for all objects located within
    - If any object does not contain a mesh renderer or mesh collider, add both to the object to make it a solid object that can be recognised by Unity's Raycast system.

After this, a raycast began scanning from the bottom left corner of the camera to the top right, covering the whole scene. If the ray intersected with a piece of geometry, it would calculate the distance and direction to the object and record the coordinates at that position. This process was done to algorithmically obtain valid target locations. As the scenes are in 3D space, it is important to place the targets against geometry, so that they move with the scene rather than floating in mid-air. Floating in mid-air, even if its position relative to the camera looked like it was against geometry, would cause the target to pop out. This process was

repeated until all azimuth and elevation positions had been scanned. The azimuth and elevation criteria moved at 1 degree in both directions.

During target placement there was a chance that portions of the target would be obscured by the scene geometry (see Figure 22), which could make visual search more difficult. To prevent this, a collision detection system was implemented. Using the most intelligent form of collision detection available by Unity known as Continuous Dynamic Collision, whenever the target collided with geometry, it was forced outside that geometry as expected. The target needs to be forced outside so that the entire target is visible to the observer. However, the algorithm for the velocity of the target and the distance the target moved from one frame to the next created a scenario whereby the target passed through a piece of geometry without ever overlapping with the geometry's collider. In the experiments, the targets have no velocity, but the algorithm uses velocity to push the target onto geometry in the scene. To fix the collision issue, both the timestep of each collision, and a custom collision detection system was implemented to generate more intermediate steps.


***Figure 22.*** **Target being partly obscured by scene geometry before applying collision detection**

This new system considered the point of contact between the target and the polygon it collided with, so that we can determine the direction in which the target was moving originally. After this, the direction was normalised and reversed such that the traversal vector for the target moved in the opposite direction, relative to the world space of the two objects. Finally, force was applied to the target's rigidbody component in the form of an impulse (instant force that is reliant on the object's mass) to force it to float beside the collided object rather than inside it. The target needs to float beside collided objects so that none of the object gets obscured, and so that when the camera moves, the target's optical flow conforms with the rest of the scene.

A list structure then stored all the positions of potential 3D target positions that the raycast discovered. Targets that were not acceptable (i.e. targets that did not persist throughout the camera's traversal) were filtered out by placing the camera at the end point of the scene, recording the positions of all remaining targets, then discarding the rest.

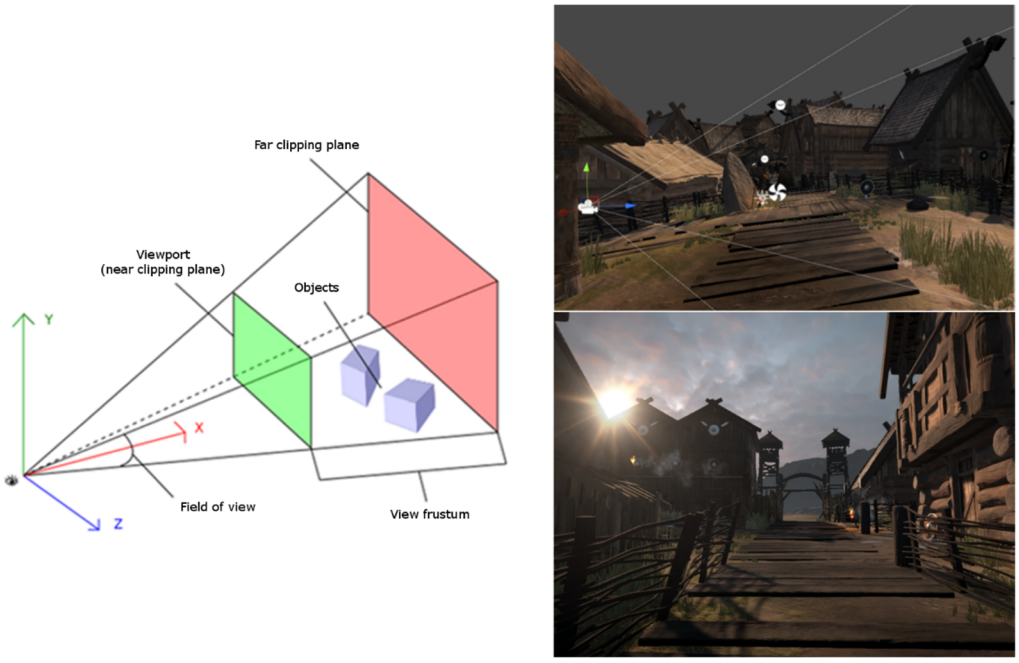



***Figure 23.*** **Example of a frustum plane rendered as part of an observer's field of view and the camera's field of view in a 3D generated scene**

To record all possible target positions from the above batch, the camera was then repositioned to the midpoint of the scene. The target then also positioned itself at each coordinate within the filtered list before allowing Unity to create an image of each position. Both the image and its respective target coordinates (x,y,z) were exported as *.png* and *.txt* files. This entire process was automated across all 40 scenes that were created prior to obtain coordinates for all positions across all stimuli.

### 7.2.2 Matlab Difference Map Processing

Using Matlab and the T.E.A. (see Chapter 5), contrast difference maps were then generated for each potential target position calculated in the previous step. Each map was then placed into a distribution and the median contrast difference location was calculated. The map with this value was then analysed to extract the indexed target coordinate of the current scene that best matches the value obtained. This process was repeated for all scenes to obtain the locations for the targets that approximately satisfied the contrast criterion.

### 7.2.3 Unity Target Placement and Film Creation

Each target location was loaded from a text file into Unity scene by scene, with its perceived size readjusted as before. The internal render path of Unity was manipulated to give priority to the target object. This meant that out of all possible objects in the scene, Unity will draw the target at all times. Not giving the target priority could cause it to not get drawn each frame, which would interfere with the experiment. The occlusion culling feature was also disabled, because occlusion culling stops objects from being drawn if they are not directly within the observer's field of view. The camera then linearly translated through the scene from start to finish whilst capturing the screen at each frame. A total of 600 frames were captured per scene to allow for a 20 second clip at 30 frames per second.

Collectively, these frames were then loaded into Matlab for each given scene. To maintain the correct order that the frames should be viewed in, they were then sorted based on their number that was given when initially saved in Unity. The frames were then converted to greyscale and compiled into the *.avi* format. The

next subsection details an algorithmic extension needed for experiment seven. See section 7.3 for the continuation of experiments five and six.

### 7.2.4 Quadratic Beziér Curve System

The final Unity experiment presented in this thesis incorporated a curved trajectory. To implement this, a Beziér curve was used as it allowed for the creation of curved splines for the camera to traverse. This is because they are parametric, meaning that they store a value as an input between zero and one to represent a point along the curve. As this value increases towards one, the camera moves from the first point of the curve to the last point.

First, a line was drawn between two points in the scene. These points were then transformed from local space of the line, to world space. This means that the axis of each point follows the same direction of the world, and not localised directions. Unity's pivot rotation was then set to the drawn line (see Figure 24) for the correct alignment of the position handles. The handles were then converted to the local space of the line to allow for manipulation of the points. The points of the curve were then stored within an array.

Linear interpolation was then used to create a middle point. The process was performed between the first point and the middle of the line to create the middle point, and lastly between the middle and last point.

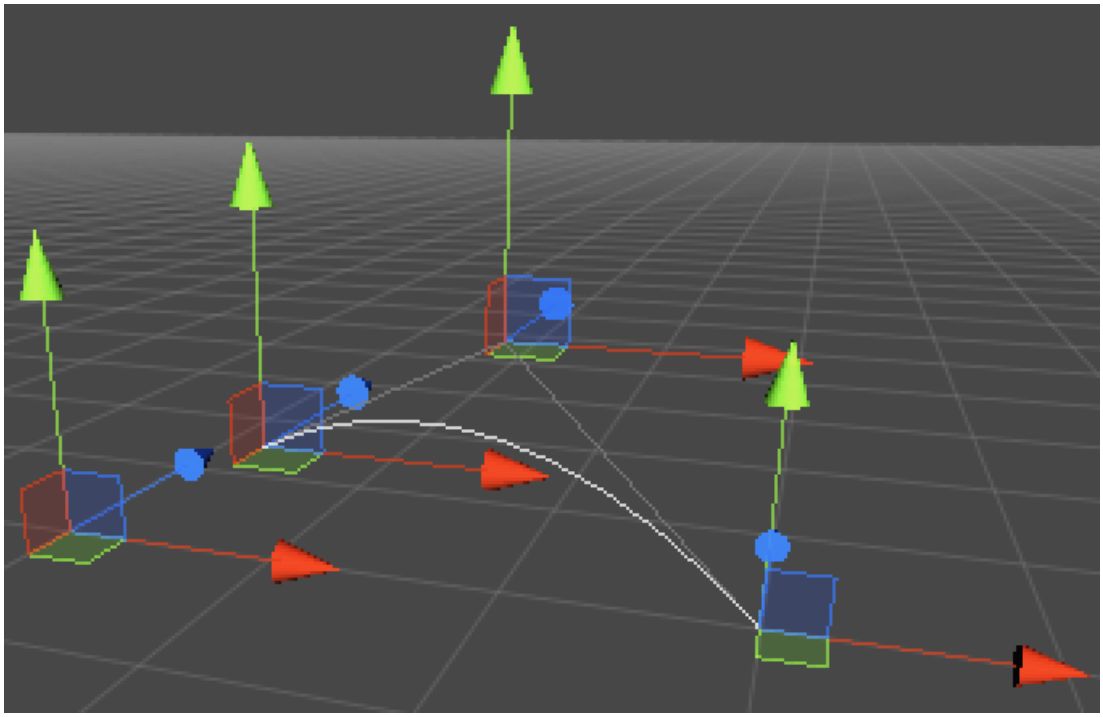


***Figure 24.*** **Quadratic Beziér Curve represented within the Unity game engine**

This linear curve is written as:

$$B(t) = (1 - t)P_0 + tP_1$$

Due to the linear curve being replaced by two new linear curves (first and middle / middle and last), we obtain:

$$B(t) = (1 - t)^2 P_0 + 2(1 - t)tP_1 + t^2 P_2$$

The first derivative of the polynomial function (first derivative of the quadratic Beziér curve) is:

$$B'(t) = 2(1 - t)\,(P_1 - P_0) + 2\,t(P_2 - P_1)$$

This produces lines that are tangent to the curve. This is to control the velocity vector for when moving along the curve (the speed in which the camera moves).

The second derivative is defined as:

$$B''(t) = 2(P_2 - 2\,P_1 + P_0)$$

which is responsible for the acceleration along the curve. However, acceleration always remained constant throughout the experiment.

### 7.2.5 Temporal Reprojection – Anti-aliasing and super-sampling

3D scenes, and scenes that contain movement especially, produce aliasing. This is where the scene generates a series of artefacts on the screen such as jittering objects, half-rendered items, and jaggy lines. Aliasing has been a problem ever since the dawn of 3D modelling and there are several ways to reduce the problem. Normal Anti-Aliasing techniques such as NVIDIA's FXAA (Fast Approximate Anti-Aliasing) (see Figure 26) are useful for smoothing edges and improving the overall quality of the scene. However, deploying this by itself, no matter how much multi-sampling is used it will not affect the temporal aliasing effect that arises through motion. It only improves the appearance of polygon edges so that they are not jagged. However, jagged lines (stair-stepping) (see Figure 25) is not the issue when it comes to temporal stability, but rather the sub-pixel aliasing that normal spatial filters can't reconstruct.

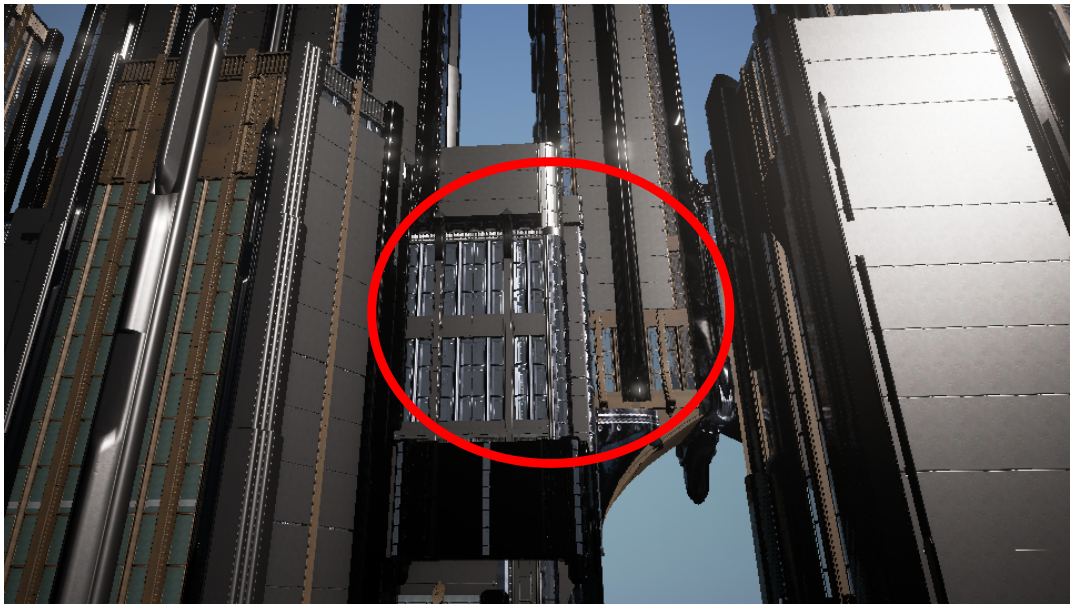


***Figure 25.*** **Example of an image without AA. Notice the broken lines and fragmented shapes**

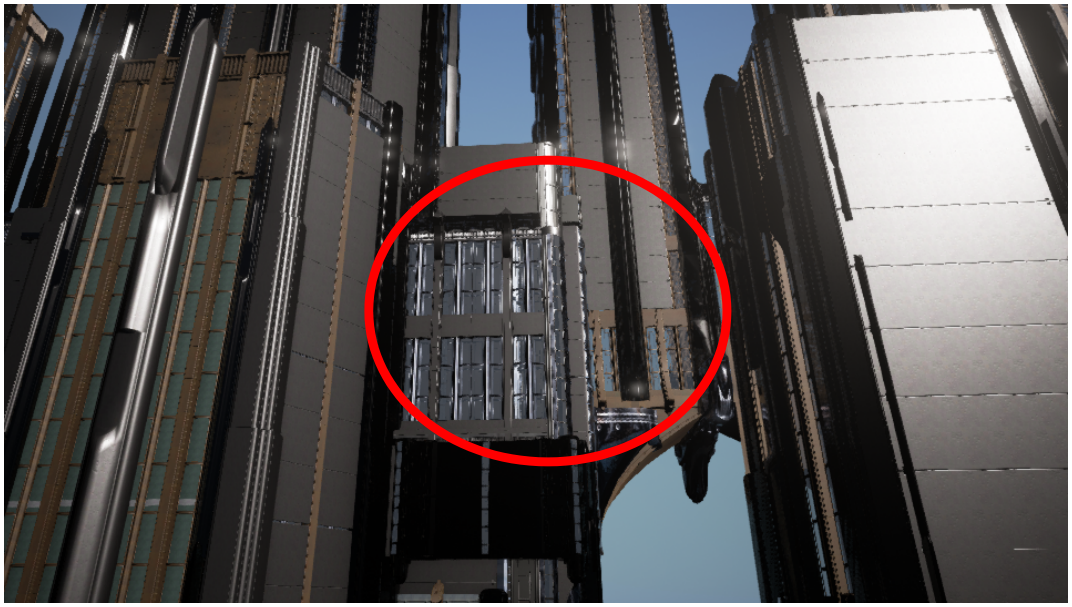


***Figure 26.*** **Example of an image with the standard FXAA. Notice some improvement**

***Figure 27.*** **Example of an image with temporal super-sampling. Notice much reduced artefacts**

The current process for reducing temporal aliasing is with what is known as Temporal Reprojection (see Figure 27). This method utilises data from the current and previous frame in a filter known as a temporal filter. A representation of the pixel's history is then stored within an internal history buffer. Invalid samples are then removed through velocity weighting to eliminate object ghosting (see Figure 28).

***Figure 28.*** **Example of object ghosting. Notice several duplications of the moving character.**

Velocity weighting is a process in which portions of a pixel's history are rejected when the velocity differs. Due to the history of a pixel being very large, these calculations can be expensive, and so the history's range is restricted to the range of the current frame's local neighbourhood. There are different models of temporal reprojection that allow for lossy or lossless pixel history. For our experiments, we used a YCoCg model due to its lossless nature.

## 7.3 Experiments Five and Six

The experiments presented in this section are highly similar in design, and so they are presented together and address the following questions:

- Does the inclusion of optical-flow produce costs to search performance with foveal vision loss?
- Is target recognition also affected for objects that conform with the optical flow?

Experiment 5 investigates the first bullet point through a visual search task that required participants to search for the letter 'T' within 3D, simulated self-motion scenes, made with the Unity Game Engine. As mentioned previously, the detection of objects during self-motion relies on optical flow (Warren & Rushton, 2008). If this result translates to visual search, then simulated self-motion scenes should produce search performance costs.

### 7.3.1 Participants

Thirty-two participants (6 males) between the ages of 18 and 27 (mean age 20 years) participated in Experiment 5. Thirty-two different participants (12 males) between the ages of 18 and 27 (mean age 22 years) participated in Experiment 6. All participants had normal or corrected-to-normal vision by self-report. They gave their written consent prior to the experiment and were paid at a rate of £7 per hour for their participation.

### 7.3.2 Stimuli

Stimuli consisted of 40 (+ 4 practice) greyscale virtual world film clips (800 × 600 pixels) that simulated self-motion in a straight line. The algorithm above was used to insert the letter T within a given scene. The target did not differ in size. Experiment six followed the same procedure, but only 50% of the stimuli contained the letter T, with the other 50% containing the letter L. The target size was the same as the larger letter size used in Chapter 6 (1.08° in width). Below are example frames from several scenes, with the target circled in red:

***Figure 29.*** **Example frames of the Unity stimuli at the 10 second interval. Target is highlighted in red.**

### 7.3.3 Design

A 1 × 2 within-subjects design with 2-levels factor scotoma (present vs. absent) and a 1-level factor target size. The scotoma's size encompassed the observer's foveal vision using the same gaze-contingent moving mask method described in chapter 5.

## 7.4 Dead zone extension to search time epochs

Previously, search time was broken down into three subcomponents (initiation, scanning, and verification time) as described by Malcom and Henderson (2009). Search initiation time is the time taken from stimulus onset, to the first saccade;

scanning time is the duration until the eyes first land within the interest area; verification time is then the gaze duration on the target. This approach is disadvantageous, because determining verification time is more complex than simply recording from when the eyes first gaze into the interest area. There are times where the eyes could simply glance over the interest area without finding the target. But since this will count as verification, we are now losing a lot of what should be scanning time. I address this problem by making some extensions to the break-down process:

Assume that the eyes have gazed into the interest area for the first time before moving out again. What factors can we use to determine whether the observer is still scanning, or deploying some strategy for verification? Since observers may saccade back into the interest area, we should only analyse the current state of this fixation, and not the initial one that first moves out. Below are a series of steps that represent the algorithm in a typical scenario with the order reflecting the numbers in Figure 30:

1. Gaze lands within interest area
2. Gaze moves out of interest area
3. Gaze moves to a new fixation location
4. Check if the fixation is not in the direction of the interest area
5. Check if the fixation coordinates are outside a circular zone around the interest area
6. If both are true, then the participant is still scanning
7. If false, then they are verifying.

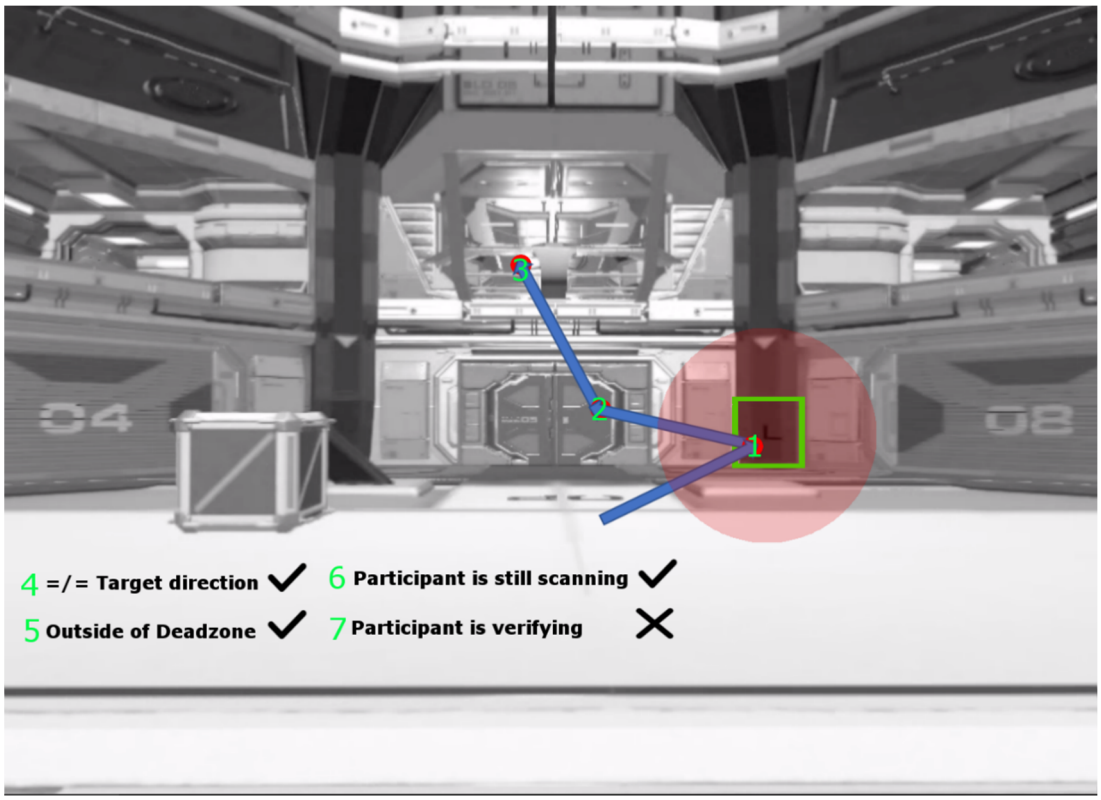



***Figure 30.*** **Depiction of the Dead-Zone algorithm. Numbers are in reference to the numbered sequence outlined in the text prior.**

The key points to consider are the circular zone around the interest area, and the second saccade after initial gaze on the target. The circular area acts as a dead zone that filters out scanning fixations. All fixations made within this zone during algorithmic processing can be assumed as contributions to verification time. The size of the dead zone was defined as being three times the size of the interest area to cover enough space for fixation capture, but not too much as to be counter-intuitive.

Further justification to creating this algorithm is due to observer motion. As the observer moves through a given scene, the target to find will also shift position because the target is fixed to the scene and the scene expands as the observer

approaches it. It may begin in a region that the observer glances over early on but does not recognise the target. In this particular case, it may be some time before the target is found. A period that should be scanning time but is instead counted as verification time. This effect is not present in 2D static images, as the target is superimposed on top of the image, in addition to the image having no depth information.

Finally, the Dead Zone algorithm was applied to Experiments 1 – 4 to see if it would influence the results, but it made no difference.

## 7.5 Results

The results are presented in three main sections:

- Search accuracy as an indication of search efficiency
- Reaction time in terms of its three epochs as explained in previous chapters
- Saccade amplitude and fixation duration across the viewing period as general eye movement measures

In order to predict a medium effect size (f = 0.252), a minimum of 27 participants was needed. This number was obtained using an alpha of 0.05, a power of 0.80 and a $\eta_p^2$ value of 0.06. Partial omega-squared values are also given (Albers & Lakens, 2018; Yiğit & Mendes, 2018) and can be viewed in Appendix G.

### 7.5.1 Search Accuracy

Search accuracy for each experimental condition was divided into probabilities of 'hit', 'miss', and 'timeout' cases (Nuthmann, 2014). If the participant had not responded within 20 seconds, the trial was coded as a 'timeout'. 'Hits' were

recorded if the participant's gaze was within the target interest area followed by a button press. The interest area was rectangular and was the same size as in the previous chapters. For gaze scoring, we implemented an additional 0.5° of padding either side to accommodate for (a) the potential spatial inaccuracy of the eye tracker, and (b) the inaccuracy of the visuo-oculomotor system when targeting relatively small objects (Pajak & Nuthmann, 2013). A response was scored as a "miss" if viewers had not fixated their gaze on the target when they pressed the button. Those cases include incidents where the target was not located, but also trials where observers' eyes did not fixate on a correctly located target when the overt response was made. The total number of incorrect hits are referred to as when the observer's gaze is within the interest area and they have confirmed their selection, but incorrectly.

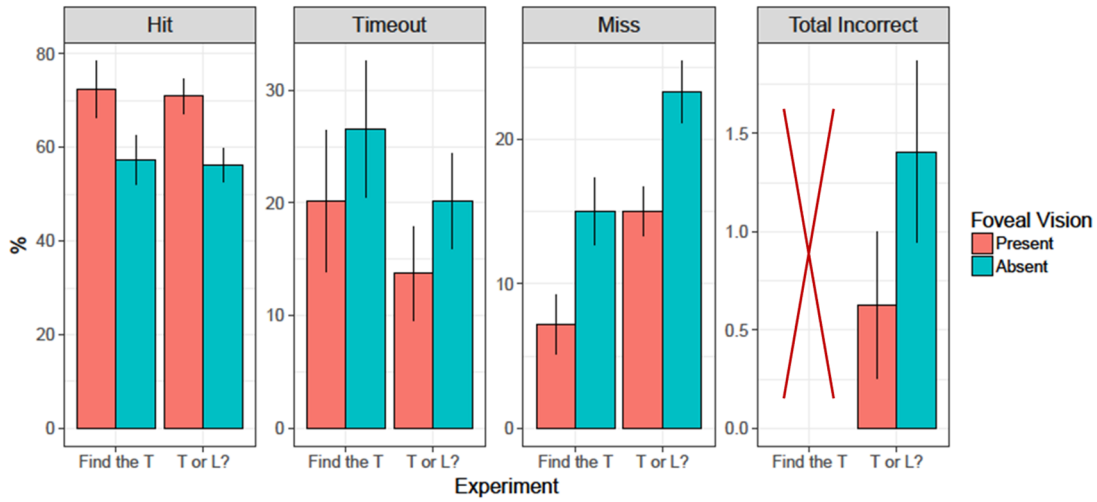



***Figure 31.*** **Search accuracy for Experiments 5 (Find the T) and 6 (T or L?). The red cross within the final panel indicates that there is no recognition component in Experiment 5. Error bars are within-subjects standard errors (Cousineau, 2005).**

For Experiment 5 (see Figure 31, 'Find the T'), there were main effects for hit percentage $F(1, 31) = 26.97$, $p < .001$, $\eta_p^2 = .47$, timeout cases $F(1, 31) = 14.23$,

$p < .001$, $\eta_p^2 = .31$, and misses $F(1, 31) = 13.65$, $p < .001$, $\eta_p^2 = .31$. Experiment 6 (see Figure 31, 'T or L?'), there were main effects for hit percentage $F(1, 31) = 26.74$, $p < .001$, $\eta_p^2 = .46$, timeout cases $F(1, 31) = 7.25$, $p = 0.01$, $\eta_p^2 = .19$, and misses $F(1, 31) = 12.41$, $p < .001$, $\eta_p^2 = .29$. The foveal scotoma had no effect on the total number of incorrect hits (where the observer validated the target to be a T but was in fact an L, and vice versa).

### 7.5.2 Search time and its subcomponents

For Experiment 5 (see Figure 32), there were main effects of scotoma on reaction time $F(1, 31) = 18.58$, $p < .001$, scanning time $F(1, 31) = 14.76$, $p < .001$, and verification time $F(1, 31) = 10.64$, $p = 0.003$. However, there was no main effect of scotoma on initiation time $F(1, 31) = 3.81$, $p = 0.06$.

For Experiment 6 (see Figure 32), main effects of scotoma were again present on reaction time $F(1, 31) = 18$, $p <.001$ and scanning time $F(1, 31) = 9.65$, $p = 0.004$. Initiation time now had a main effect of scotoma $F(1, 31) = 24.09$, $p = < .001$ but verification time did not $F(1, 31) = 3.29$, $p = 0.079$.

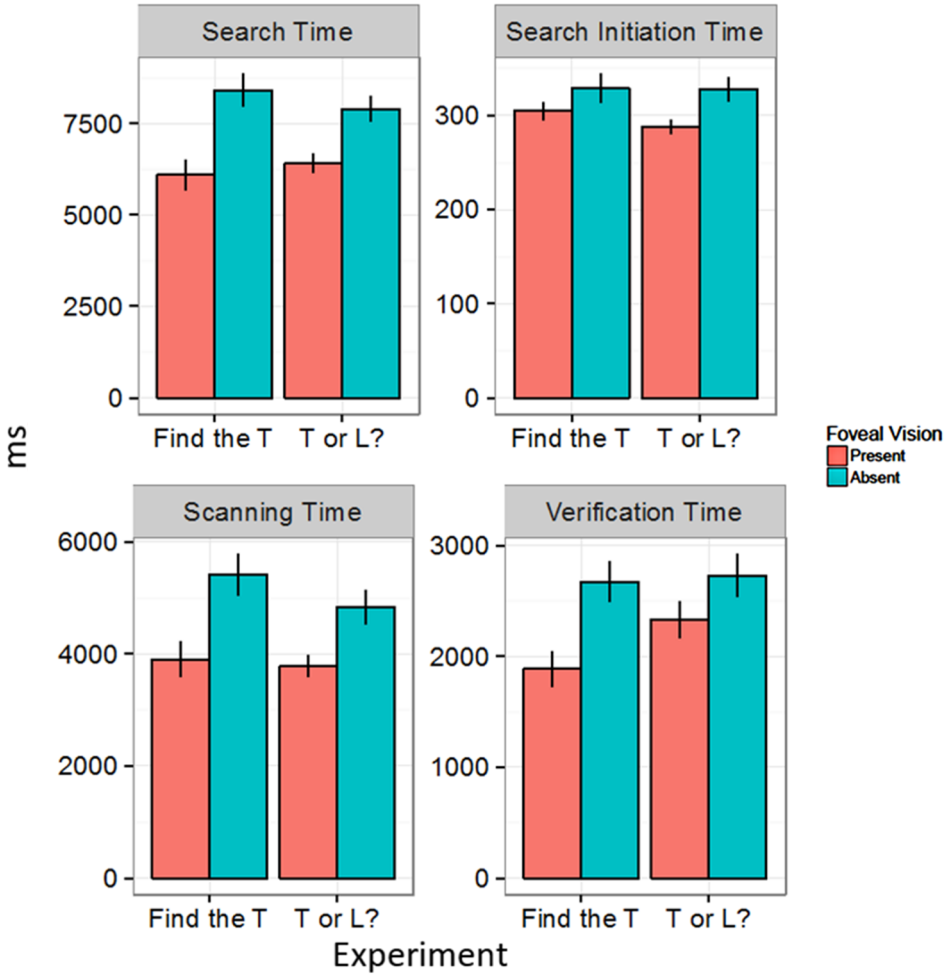



***Figure 32.*** **Search time and its three epochs for Experiments 5 (Find the T) and 6 (T or L). Each column presents means obtained for a designated dependent variable (see panel title). Error bars are within-subjects standard errors (Cousineau, 2005).**

### 7.5.3 Eye movement behaviours

For Experiment 5, there was a main effect of scotoma on saccade amplitude $F(1, 31) = 69.91$, $p < .001$, but not for fixation duration $F(1, 31) = 0.02$, $p = 0.896$. These results were replicated in Experiment 6, with a main effect of scotoma on saccade amplitude $F(1, 31) = 87.74$, $p < .001$, but no effect on fixation duration $F(1, 31) = 1.87$, $p = 0.181$.

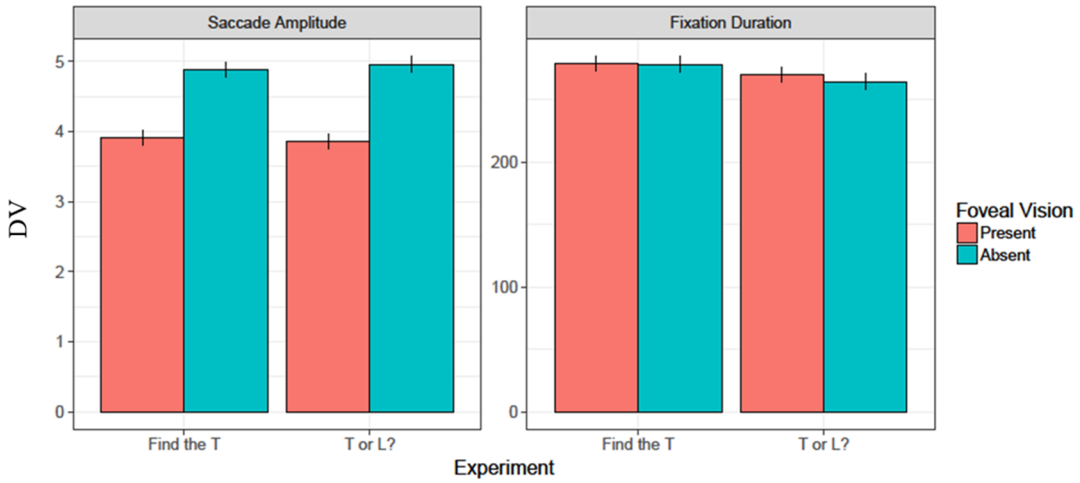



***Figure 33.*** **Mean saccade amplitudes and fixation durations for experiments 5 and 6. Saccade amplitude is measured in degrees, with Fixation duration measured in milliseconds. Error bars are within-subjects standard errors (Cousineau, 2005).**

## 7.6 Experiment Seven – Find the T on a curved path

The experiments presented in this chapter so far have investigated the role of foveal vision to a visual search task within simulated self-motion scenes. In conflict with the rest of this thesis, the results of this chapter have revealed that foveal impairment leads to a significant reduction in search performance. Why is this suddenly the case? The newest factor introduced with the stimuli of this chapter has been the simulated self-motion. It seems that the act of moving through a scene, as opposed to observing a snapshot of the real world, has somehow hindered our search capabilities. To investigate this further, we conducted a third experiment that manipulated the observer motion. Rather than traversal along a straight path, observers would now move along a curved path, with the target being present on either side of the curve (but not simultaneously).

Varying the target location in this way creates two types of location: optimal and suboptimal. Imagine you are walking along a curved path in a park. On one side, the trees are seemingly moving away from your field of view, whereas the trees on the other side are moving directly into your vision. We hypothesise that if the target's location is optimised by being on the latter side, search performance should be greater than if the target was positioned sub-optimally. If so, then we should see an interaction between searching with a foveal scotoma and target position, which may eliminate the foveal necessity result found from experiments five and six.

### 7.6.1 Participants

Thirty-two participants (10 males) between the ages of 18 and 27 (mean age 20 years) participated in Experiment 7. All participants had normal or corrected-to-normal vision by self-report. They gave their written consent prior to the experiment and were paid at a rate of £7 per hour for their participation

### 7.6.2 Stimuli

Stimuli consisted of 40 (+ 4 practice) greyscale virtual world film clips (800 × 600 pixels) that simulated self-motion along a curved path. The algorithm in section 7.2 was used to insert the letter T within a given scene, with the letter being the same target size as in Experiments 5 and 6. Quadratic Beziér curves were drawn within the Unity Game Engine to represent the curved path (see section 7.2.4 of this chapter). The target did not differ in size.

### 7.6.3 Design

A 2 × 2 within-subjects design with 2-levels factor scotoma (present vs. absent) and 2-levels factor target location. The target was situated on either side of the curved path with the location [with respect to the camera's movement] defined as optimal or suboptimal. The scotoma's size encompassed the observer's foveal vision using the same gaze-contingent moving mask method described in chapter 5. The list of 40 stimuli were assigned to two lists of 20 scenes each. The scene lists were rotated over participants, such that a given participant was exposed to a list for only one of the eight experimental conditions created by the 2 × 2 design. There were eight groups of four participants, and each group of participants was exposed to unique combinations of list and experimental condition.

The foveal vision manipulation was blocked so that participants completed two blocks of trials in the experiment: in one block observers' foveal vision was available, in the other block it was obstructed by a gaze-contingent scotoma. The order of blocks was counterbalanced across subjects. Within a block, there were an equal number of scenes with target location (optimal vs suboptimal), as well as curve side (curving to the left/right) presented randomly.

### 7.6.4 Target Location on a Curved Path

Aside from using the extended T.E.A. algorithm for target insertion, the optical flow of the target when on a curved path was also controlled for. As the target's location could be on either side of the curve for a given scene, the optical flow at that location may make the target easier to locate, and therefore introduce unnecessary noise.

The equations and methods below were designed by Professor Robert B. Fisher, and then implemented by me to identify target positions having optical flow with an equal and opposite sign on either side of a given curve.

### Optical flow on a straight path

The steps presented here gives the equivalent optical flow of a target when motion is on a straight trajectory, with a fixed camera speed $w$, and if the radius $\rho \rightarrow +\infty$.

$$optical\ flow(u,v) = \frac{fuw}{(wt-v)^2}$$

Below are the steps required to reach this formula (refer to Figure 34 and 35 for the motion and camera model respectfully):

- A camera moves on a circular path (see Figure 34), such that it is at position $\vec{b}(t) = \rho(1-\cos(\sigma t), \sin(\sigma t))$ at time t, starting from (0,0) and moving with angular velocity $\sigma$ on a circle of radius $\rho$ with centre at $(\rho,0)^{\mathsf{T}}$.
- The camera faces forward along its +Z axis, so its facing direction is $\vec{d}(t) = R(t)\ (0,1)^T$, where

$$\begin{bmatrix} \cos(\sigma t) & sin(\sigma t) \\ -sin(\sigma t) & cos(\sigma t) \end{bmatrix}$$

And positive $\sigma t$ is in a clockwise direction.

- There is a target observed at point $\vec{x} = (u,v)$ .
- A pinhole camera moves so that it is at point $\vec{b}(t)$ relative to the global (initial) origin at $(0,0)^T$ and its optical axis points in the direction $\vec{d}(t)$.
- In the camera's local coordinates, the target $\vec{x}$ is seen at $(g,h)^T$, where

$$R(t)^{-1}\left(\vec{x}-\vec{b}(t)\right) = \binom{g}{h} = \begin{bmatrix} \cos(\sigma t) & -sin(\sigma t) \\ sin(\sigma t) & cos(\sigma t) \end{bmatrix}\left(\binom{u}{v} - \begin{pmatrix} \rho(1-\cos(\sigma t)) \\ psin(\sigma t) \end{pmatrix}\right)$$

- $(g,h)^T$ projects to $(a,f)^T = (\frac{fg}{h},f)^T$ on the observer's image plane (see Figure 35), where $f$ is the camera focal length.
- The observed optical flow is $\frac{da}{dt} = \frac{f}{h^2}\left(h\frac{dg}{dt} - g\frac{dh}{dt}\right)$.
- Letting $c = \cos(\sigma t)$ and $s = sin(\sigma t)$, we simplify the calculation of $(g,h)^T$ to

$$\begin{pmatrix} g \\ h \end{pmatrix} = \begin{pmatrix} uc - vs - \rho c + \rho \\ us + vc - \rho s \end{pmatrix}$$

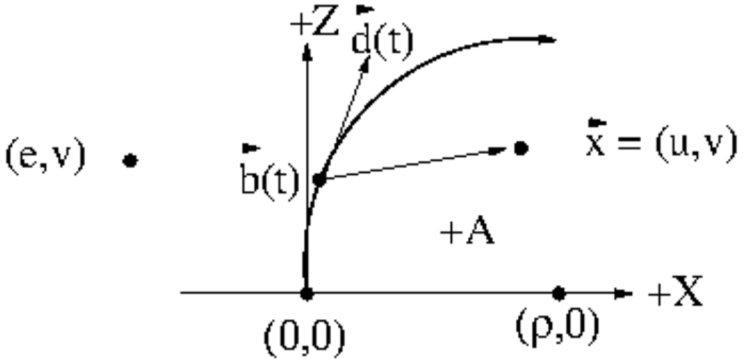



**Figure 34. Camera motion geometry**

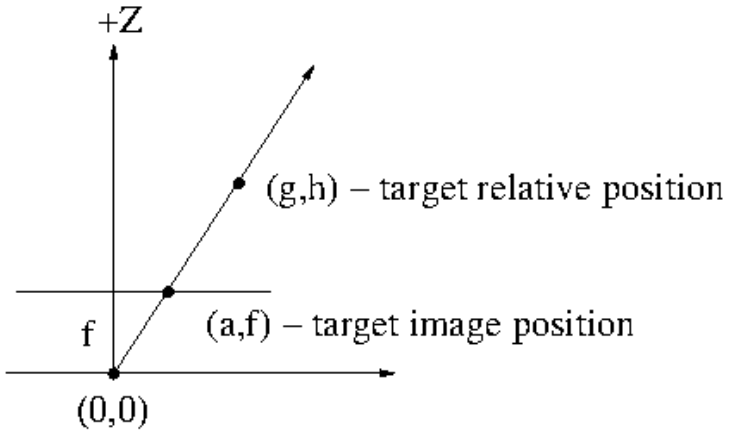



**Figure 35. Camera projection geometry**

By computing the derivatives for the optical flow:

$$\frac{dg}{dt} = -\sigma(us + vc - \rho s)$$
$$\frac{dh}{dt} = \sigma(uc + vs - \rho c)$$

And substituting this into the formula for optical flow, we get the following equation:

$$optical\ flow(u,v) = -\sigma f\frac{(u-\rho)^2 + \rho c(u-\rho) + v^2 - \rho vs}{(s(u-\rho)+vc)^2}$$

Since, for $\rho \to +\infty$, then $\sigma \to 0$. As $\rho t$ is small, we can approximate the above $s$ to be $s = \sin(\rho t) \approx \sigma t$ and approximate c to be $c = 1$. We also know $\sigma\rho = w$. This simplifies the equation to give:

$$optical\ flow(u,v) = \frac{fuw}{(wt-v)^2}$$

Which is equal to the optical flow of a trajectory that begins with a straight path.

### Obtaining the initial optical flow

By assuming that $c = 1$ and $s = 0$, we can obtain the initial optical flow at $t = 0$. Consider the following equation:

$$optical\ flow(u,v) = -\frac{\sigma f}{v^2}((u-\rho)^2 + \rho(u-\rho) + v^2) = \frac{\sigma f}{v^2}(u^2 + v^2 - \rho u)$$

The above implies that as a trajectory transitions from straight to curved, there is a critical radius $\hat{p}$ on that path where the optical flow goes from negative to positive. Using this information, we can now determine the geometric interpretation of the critical radius $\hat{p}$ where $optical\ flow(u,v) = 0$ at $t = 0$. The above equation implies this to be $u^2 + v^2 = \hat{\rho}u$.

Let us assume that our critical $\hat{\rho}$ gives a path to the right of a target $\vec{x}$. Conceptually this would mean that an observer would see the target on their left side as they pass it. This implies $\hat{\rho} < \| (u,v) - (\hat{\rho},0) \|$. Expanding this to $\hat{\rho}^2 < \| (u,v) - (\hat{\rho},0) \|^2 = u^2 - 2u\hat{p} + \hat{\rho}^2 + v^2$ , then substituting our geometric interpretation $u^2 + v^2 = \hat{p}u$ before simplifying gives us $\hat{p}u < 0$. Since $u$ and $p$

are both positive, the expression is contradictory, which implies that the path goes to the left of $\vec{x}$.

From this we can now gather that if $\rho > \hat{\rho}$ (the observer passes to the left of the target), then the initial optical flow at $(u, v)$ is greater than 0. Similarly, if $\rho < \hat{\rho}$ (the observer passes to the right of the target), then the initial optical flow at $(u, v)$ is less than 0.

### Calculating the second target position

Determining whether the path is curving towards the left or to the right is useful for calculating the second target position $(e, v)$ which is at the same distance $v$ as the first target, only here has the opposite optical flow. If we look again at the above equation we can see that the critical radius also interacts with the camera's speed (see $\rho$). Therefore, we also need to consider the second target position as a function of the speed.

Before we find the position $(e, v)$ with $-OF$, we must first look at some constraints:

1. We need $s(u - \rho) + vc \neq 0$, otherwise the projected target position moves to $\pm\infty$. In other words, the first target must not be in a position that allows it to cross over the curve at any point in time. At $t = 0$ this quantity is initially positive, so the constraint becomes $s(u - \rho) + vc > 0$. Solving for the time $t$ we get:

$$t < \frac{t}{\sigma} tan^{-1}(\frac{v}{\rho - u})$$

2. A similar constraint arises from the left target $(e, v)$:

$$t < \frac{t}{\sigma} tan^{-1}(\frac{v}{\rho - e})$$

3. The observer should also not travel past the target, so

$$t < \frac{1}{\sigma} sin^{-1}\left(\frac{v}{\rho}\right)$$

We now bound $t$ by the smaller of these two constraints to prevent $t$ from getting close to the limit.

4. As the trajectory needs to travel to the left of $\vec{x}$, there is a constraint on $\rho$. We need $\rho > \| (u,v) - (\rho,0) \|$ (ie. The path radius has to be longer than the distance from the path centre to $\vec{x}$. Solving we get:

$$\rho > \frac{u^2 + v^2}{2u}$$

Considering the constraint on $p$ from the above equation, this now gives:

$$\rho > \frac{u^2 + v^2}{u}$$

5. We now want to solve for $e$ as a function of $u, v, t, \sigma, \rho$. Assume we pass to the left of the target $\vec{x}$. We want the optical flow to be equal in magnitude but with the opposite sign, i.e:

$$optical\ flow(e,v) = -optical\ flow(u,v)$$

***Equation 2.*** **For the determining of magnitude equivalence of optical flows that are of opposite signs**

We can now apply equation (2) to Matlab's symbolic equation solver, to solve for $e$ in the equation: $optical\ flow(e,v) + optical\ flow(u,v) = 0$:

```
syms u v e r c s
% s = sin(\sigma*t)
% c = cos(\sigma*t)
numu = (u-r)^2 + r*c*(u-r) + v^2 - r*v*s;
denu = (s*(u-r)+v*c)^2;
nume = (e-r)^2 + r*c*(e-r) + v^2 - r*v*s;
dene = (s*(e-r)+v*c)^2;
solutions = solve(numu / denu + nume / dene == 0, e);
```

[r, common] = subexpr(solutions);

The result is a quadratic equation with two solutions, one for the left-hand side of the curve and the other for the right. For a given $u, v, p, \sigma, t$, the solution $(e, v)$ is given by:

$$c = \cos(\sigma t)$$
$$s = \sin(\sigma)$$
$$\begin{aligned} X = (&c^4\rho^2v^2 - 6c^3\rho^3sv + 6c^3\rho^2suv + 4c^3\rho^2v^2 - 4c^3\rho uv^2 + c^2\rho^4s^2 \\ &- 2c^2\rho^3s^2u + 4c^2\rho^3sv + c^2\rho^2s^2u^2 - 4c^2\rho^2s^2v^2 - 8c^2\rho^2suv \\ &- 4c^2\rho^2v^2 + 4c^2\rho su^2v + 12c^2\rho sv^3 + 8c^2\rho uv^2 - 4c^2u^2v^2 \\ &- 8c^2v^4 - 4c^{\rho^3}s^3v + 4c\rho^2S^3uv - 4c\rho^2s^2v^2 + 4c\rho s^2uv^2 \\ &+ 8c\rho sv^3 - 8csuv^3 + 8\rho^3s^3v - 4\rho^2s^4v^2 - 16\rho^2s^3uv - 8\rho^2s^2v^2 \\ &+ 8\rho s^3u^2v + 8\rho s^3v^3 + 16\rho s^2uv^2 - 8s^2u^2v^2 - 4s^2v^4)^{\frac{1}{2}} \end{aligned}$$
$$\begin{aligned} A = \ &c^3\rho v^2 - 4c^2\rho^2sv + 4c^2\rho suv + c\rho^3s^2 - 2c\rho^2s^2u - 2c\rho^2sv + c\rho s^2u^2 \\ &- 2c\rho s^2v^2 - 4c\rho suv + 2csu^2v + 2csv^3 \end{aligned}$$
$$\begin{aligned} D = 2(&c^2v^2 - c\rho^2s^2 + c\rho s^2u - 2c\rho sv + 2csu + 2\rho^2s^2 - \rho s^3v - 4\rho s^2u \\ &+ 2s^2u^2 + s^2v^2) \end{aligned}$$
$$B = X(cv - \rho s + su)$$
$$e = (A - B)/D$$

6. Using some example values: $u = 5, v = 5, \rho = 20, \sigma = 1$, we obtain the following plot:

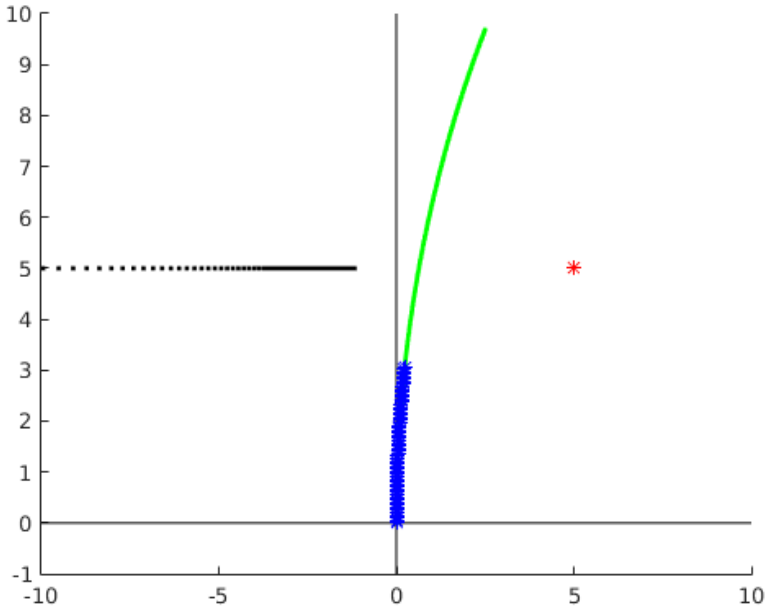



***Figure 36.*** **Calculated target positions along a trajectory**

With the target $(u, v)$ in red, the camera's trajectory in green, the observation points in blue, and the mirror target positions $(e, v)$ in black. Note that there is no one point that gives the opposite optical flow for all positions of the camera. Because of this, the locations chosen were points that worked reasonably well for many positions along the trajectory. Please see Appendix D for a full list of figures relating to each stimuli.

### 7.6.5 Results

The results presented here follow the same format as all previous experiments.

In order to predict a medium effect size (f = 0.252), a minimum of 23 participants was needed. This number was obtained using an alpha of 0.05, a power of 0.80 and a $\eta_p^2$ value of 0.06. Partial omega-squared values are also given (Albers & Lakens, 2018; Yiğit & Mendes, 2018) and can be viewed in Appendix G. More information on data analysis procedures can be found in Chapter 3 section 3.3.3.

### Search Accuracy

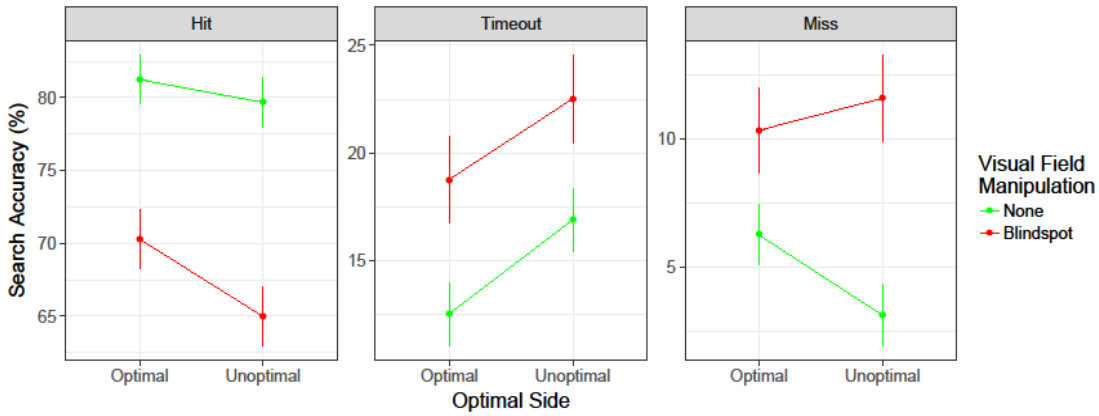



***Figure 37.*** **Search accuracy measures for experiment 7. The x-axis indicates the side of the curve the targets were on. Error bars are within-subjects standard errors (Cousineau, 2005).**

There was a main effect of scotoma for hit percentage $F(1, 31) = 23.78$, $p < .001$, $\eta_p^2 = .43$, timeouts $F(1, 31) = 10.37$, $p = .003$, $\eta_p^2 = .25$ and miss trials $F(1, 31) = 12.92$, $p = .001$, $\eta_p^2 = .29$.

The side of the curve for target location, with the optimal side having targets move towards the observer's field of view, there was a main effect for timeouts $F(1, 31) = 6.62$, $p = .02$, $\eta_p^2 = .18$. There were no main effects of target side for hit percentage or percentage of miss hits. There were no interactions between scotoma and target side for all three measures of search accuracy.

### Search Time and its Subcomponents

There were main effects of scotoma for reaction time $F(1, 31) = 18.08$, $p < 0.001$, scanning time $F(1, 31) = 12.17$, $p = 0.001$, and verification time $F(1, 31) = 5.52$, $p = 0.025$. Initiation time had no main effect of scotoma $F(1, 31) = 0.49$, $p = 0.49$.

In terms of the side of the curve for target location, with the optimal side having targets move towards the observer's field of view, there were main effects for reaction time $F(1, 31) = 12.08$, $p = 0.002$ and scanning time $F(1, 31) = 12.85$, $p = 0.001$. There were no main effects of target side for initiation time and verification time. There were no interactions between scotoma and target side.

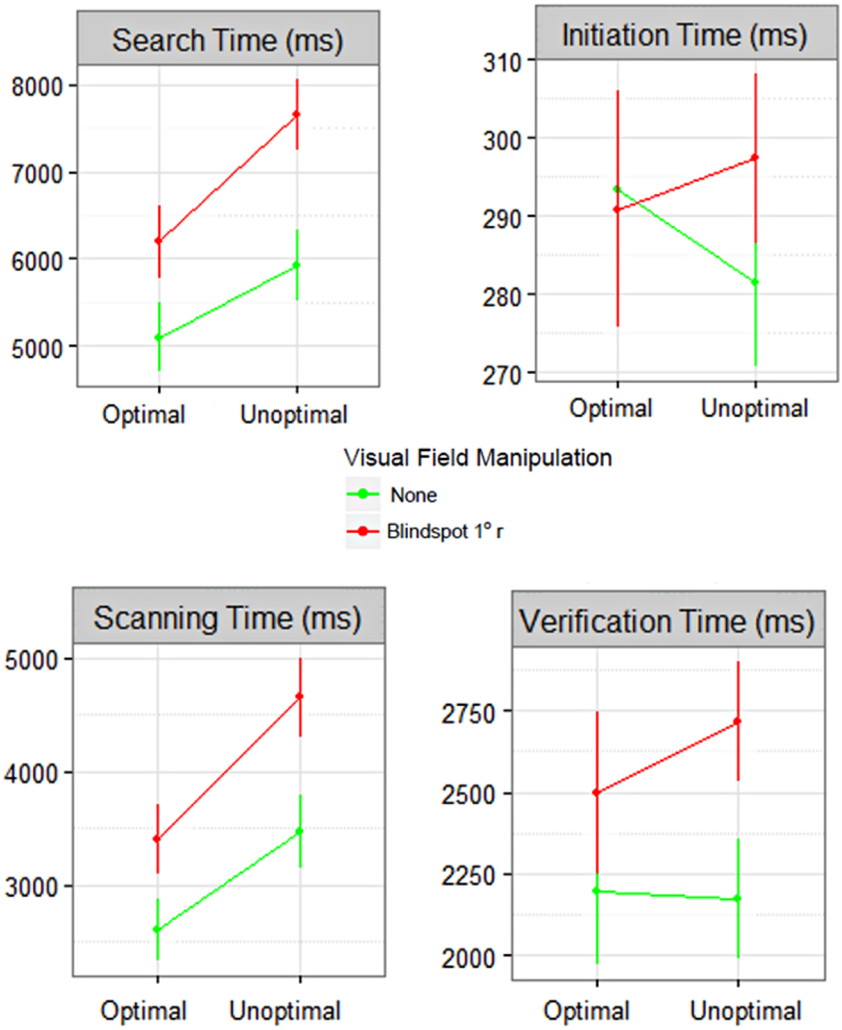



***Figure 38.*** **Search time and its three epochs for Experiment 7. Each column presents means obtained for a designated dependent variable (see panel title). Error bars are within-subjects standard errors (Cousineau, 2005).**

### Eye movement behaviours

There was a main effect of scotoma on saccade amplitude $F(1, 31) = 150.18$, $p < 0.001$. There were no main effects of target side nor any interactions between scotoma and target side for saccade amplitude and fixation duration. There was also no main effect of scotoma on fixation duration.

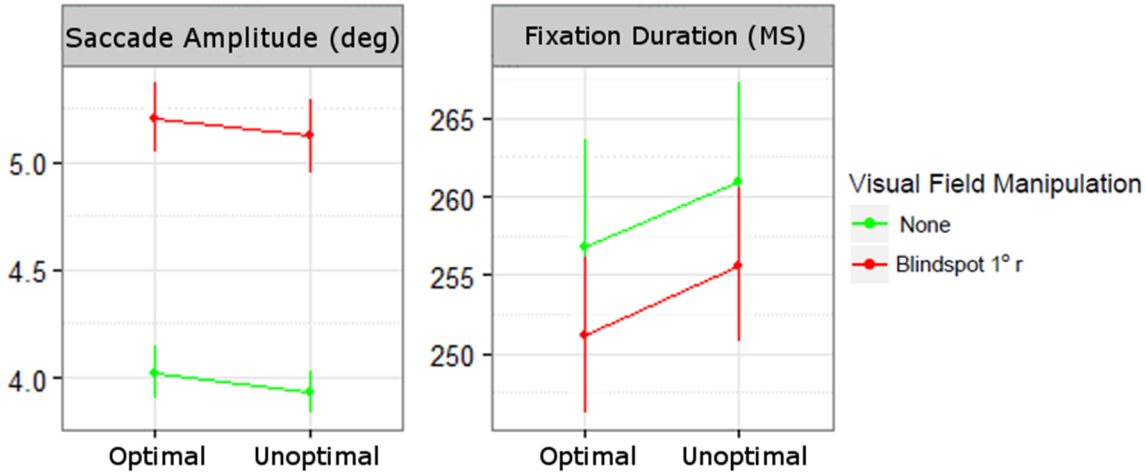



***Figure 39.*** **Mean saccade amplitudes and fixation durations for experiment 7. Saccade amplitude is measured in degrees, with Fixation duration measured in milliseconds. Error bars are within-subjects standard errors (Cousineau, 2005).**

## 7.7 General Discussion

Three experiments were conducted to assess search performance in 3D immersive scenes, with and without foveal vision. In Experiment 5 observers were tasked with searching for a letter 'T', whereas in Experiment 6, observers searched for the letter 'T' or 'L'. Due to similarities in search task between these experiments and those in Chapters 5 and 6, it seemed logical at first to not expect any new effect from the change in stimulus. It was reasoned, however, that the inclusion of an optical flow (motion induced through self-movement) could somehow influence search performance. This is backed by previous research which demonstrates that search for targets that conform with an optical flow requires more focused attention (Royden & Hildreth, 1999). Studies involving optical flow primarily focus on judgements of heading and object detection. However, very few studies have

investigated search, and those that have do not use complex 3D stimuli (Royden, Wolfe, & Klempen, 2001).

Due to discerning an object's motion from optic flow requiring a more focused attention within the object's vicinity, it was reasoned that perhaps the same logic could translate to visual search, regardless of the object moving. Thus, the first experiment of this chapter was designed to investigate an observer's abilities in finding context free targets as a forward movement was simulated. If there was no interference from optic flow, then the effect on search performance should be similar if not the same as visual search in 2D static scenes. The results of this experiment reveal that participants do indeed take longer in searching for targets during simulated self-movement. This is consistent with previous investigations in the optic flow literature. It appears that as the participant seemingly walks through an environment, the optic flow generated somehow interfered with the task.

From the results obtained in previous chapters, we know that search is primarily a peripheral task, utilising peripheral and parafoveal processing to find targets without needing to foveate directly in its vicinity. In 2D scenes it is reasonable to suggest that this result is found because the scene itself is a snapshot of the real world, omitting factors that could otherwise influence search. With 3D dynamic scenes, the observer must also deal with the everchanging peripheral environment. Environmental objects such as trees and city buildings are constantly moving out of view in a radial fashion, giving the observer new areas to attend to as they proceed further into the stimulus. Furthermore, the target itself does not pop out easily due to its optic flow conformity. As a result, the target moves in unison with the entire scene, and so blending with the generated optic flow. If targets did not conform with the optic flow, they would either be present in the same pixel

locations across all frames, or they would be embedded in the scene but in its own local space. The former approach sacrifices depth information for clarity. The target will always be present in a single location but completely isolated from the scene itself. Think of a fly that has gotten stuck on a car's window. The outside world continues as normal, but the fly will always be there. The latter approach allows the target to be present within the scene, benefiting from the scene's depth plane. Unless the targets are placed on a physical part of the scene (e.g. a building), sections of the environment will move behind it, as it would be floating in mid-air. This constant change of background against the target would cause it to pop out more frequently.

We can infer from Figure 32 (c) that the effect on scanning time without foveal vision is far more detrimental than previous experiments. This suggests that without foveal vision, the ability to find targets in a 3D immersive scene is harder because of an observer's inability to discriminate targets in their parafovea. In static scenes it was easier, as the environment in their periphery was not being modified, with only the naturally declining visual acuity of their vision modulating difficulty. Now, we have a natural decline in combination with a visual disturbance through optic flow, which further degrades acuity in the periphery. Therefore, I suggest that the reason why objects require more focused attention for detection along an optic flow, is because participants must rely on the high acuity vision available in the fovea. Without it, search becomes increasingly harder as the optic flow proceeds to blur the environment.

If the optic flow is indeed a manipulating factor, then it should also be possible to take advantage of its properties to facilitate search. Experiment 7 extended the paradigms of Experiments 5 and 6 by manipulating the path of self-motion from a

straight line to a curve. The change to a curved motion creates a unique situation for the observer. First, the optical flow on one side of the curve is drastically different from the side opposite, meaning that targets on one side are seemingly moving away from their field of view, whereas objects on the opposite side are moving towards their field of view. Experiment 7 investigated the effect of target location along a simulated self-motion curved path for the facilitation of search performance. If targets are harder to find as participants move in a straight line, then making targets move towards a participant's centre field of view should facilitate search performance, reducing the foveal load required for an efficient search. Similarly, targets moving away from their field of view would be unoptimal, and therefore be harder to find.

The results of this chapter demonstrate that search can indeed be optimised by placing targets on the optimal side of the curve; and search can also be made more difficult by placing them in unoptimal locations. However, the notion that foveal load could be reduced does not seem reasonable, as the search performance along this curve without foveal vision still proves to be far too difficult in comparison to the control (see Figure 38). After a successful replication of the effect of having a foveal scotoma, it seems that this effect is not an artefact of the scenes used. In fact, the results collectively indicate that optic flow does interfere with visual search performance through implicit inhibition of a participant's visual acuity. The whole notion that foveal vision is important for visual search seems to be most relevant for 3D immersive stimuli, which for future work could be further investigated through a virtual reality toolkit that allows real movement in a virtual environment (e.g. HTC Vive).

Why would the stimulus matter? There have been discussions about this throughout the history of eye tracking research. Basic stimuli were argued to be less ecologically valid than naturalistic backgrounds. It was also argued that such stimuli were too artificial, as observers do not usually search for targets such as letters or dots. Chapter 6 has shown that search for such targets is useful for the visual salience literature, due to the elimination of top-down factors. The difference, say, between natural scenes and scrambled scenes is the control of visual information. Scrambled scenes are highly cluttered and so contain a large amount of high spatial frequency information. Searching for a target that is already full of high spatial frequencies will lead to interference, thus making the search task unnecessarily hard. Natural scenes not only attribute themselves to environments that we're used to, but they also have the capacity to contain a mixture of frequency information.

The transition from static to dynamic is different. Either the observer moves, objects in the environment move, or both. Motion of objects is a well-known causing factor of pop-out effects (Smith & Mital, 2013); whereas self-motion and direction of heading are useful for navigational tasks. Most optic flow studies research flow fields in isolation, and so use blank backgrounds. The experiments presented here combine scene search with simulated self-motion as a means of pushing the field forward. The most striking result from Experiment 5 is the fact that search time increased significantly without foveal vision. An effect that was not present from the 2D studies from Chapter 5.

Decomposing the search time into its 3 epochs by applying our epoch extension algorithm, we can see clearly that the costs originate from scanning time. That is, the localisation process of the target. By adding a recognition component, it was

reasoned that verification time could also see an increase, which was not the case as indicative on figure 32. It seems that searching for targets that have optic-flow conformity are somehow harder than targets on static backgrounds. Why? Thinking back to visual acuity (Section 4.1.1), we know that vision is highest at the fovea before dropping rapidly toward the periphery. Consequently, we make saccades to interesting areas bringing them into high resolution. In static scenes, the target's local acuity remains constant. However, its global acuity is affected by how close an observer's gaze is to it; self-motion drastically alters this, as the local acuity of the target is being affected by the generated optical-flow. Targets moving along an optical flow are harder to detect in general. How does this relate to foveal vision loss? High acuity tasks such as reading require foveal vision to be completed effectively. In search, verification is also seen as a high acuity task. As the target now has more blur than normal, detection may only be possible once gaze is within a certain detection range from the target. Removing foveal vision makes detection harder as the target is more likely to be covered once in range, making that detection range obsolete.

At first glance, it might appear that search within such scenes is simply more difficult but in reality, there are other factors to consider. As stated above, when moving through a scene the optic flow generated interferes with the scene itself. 2D images do not have any real depth information or movement, which makes search within such stimuli easier as the structural integrity of the image is maintained. In the real world, people naturally move around and more importantly, move through environments which will consequently blur the target to find with the environment. Being able to detect such targets will require the use of high acuity vision, which is not necessarily the case with 2D images. This may be one reason why the results change so drastically between static and dynamic scenes.

Another potential cause for the difference is that optical flow salience may dominate contrast salience.

Targets that have their own motion, and consequently their own optic-flow would be easier to spot as recent evidence suggests (Rushton, Niehorster, Warren, & Li, 2018). Not only due to potential pop-out effects but also due to flow parsing. Collectively, they provide support to the argument that targets with optic-flow conformity are harder to find and may require the observer's gaze to first land within a certain range of the target. Determining this range, however, is for future work, as perhaps it is not as binary as an in/out range around the target. Perhaps efficient detection is only possible through the high acuity vision provided by the fovea. If that is the case, the range could be considered the extent of foveal vision, but this is pure speculation and therefore more research is needed.

# 8 Conclusion

## 8.1 Achievements

This thesis presented five original contributions to knowledge, as well as two algorithms and one algorithmic extension. The research analysed foveal vision and visual search performance in 2D naturalistic scenes. Firstly, experiments showed that the time taken to localise context free targets was nearly identical with and without foveal vision. As stated in Chapter 5, there was a debate in the literature (Bertera & Rayner, 2000; Nuthmann, 2014) as to whether foveal vision was important for attaining normal search performance. Secondly, saccade amplitudes were also found to increase when searching for smaller targets (see Chapter 5). Thirdly, target size and target salience was found to interact with one another such that salience can compensate for the difficulty of finding targets as they become small (and vice-versa). Fourthly, and one of the more significant contributions of this thesis is the discovery of how search costs are split amongst the three epochs of search time outlined by Malcolm and Henderson (2009) (See Chapter 6). Finally, Chapter 7 includes three experiments that use 3D immersive scenes to assess the importance of foveal vision under simulated self-motion. There was evidence for an increased reliance on foveal vision due to a reduced visual salience of the target and although search performance was improved (see Section 7.6) it was not enough to eliminate the reliance on foveal vision within such scenes.

The results of Chapter 5 show that although foveal vision loss impacted search accuracy, the time taken to scan the scene and find targets did not differ. One

could argue that naturalistic scenes and context free targets are separate, and therefore do not explain anything related to search in the real world. However, it was shown recently by Cornelissen and Võ (2017) that people do not suppress semantic processing when the target is semantically inconsistent. This means that the presence of a naturalistic background influences oculomotor behaviour due to semantic processing, regardless of the target being semantically inconsistent (a letter T or L). Importantly, when searching for a context-free letter target the scene is more than just a patterned background. Processing of scene and object relationships appears to be obligatory, in a sense that it is hard to suppress (Cornelissen & Võ, 2017). To extend the present findings, it would be useful to systematically explore the role played by various forms of scene guidance, using manipulations like scene inversion (Foulsham & Underwood, 2011), scene scrambling (Foulsham, Alan, & Kingstone, 2011), or pseudo-scene viewing (Luke & Henderson, 2016). Although the point in time where search performance costs originated was not found in Chapter 5 (see Chapter 6), the debate on whether foveal vision was important for attaining normal search performance was largely resolved.

Chapter 5 also includes work on a novel Target Embedding Algorithm (T.E.A.). This algorithm was designed to embed context free targets into naturalistic backgrounds based on contrast difference values (see Section 5.3.5). Using this algorithm allowed careful embedding of targets based on varying levels of visual salience. Due to an algorithm automating this process, it had the added benefit of artefact reduction, which would not have been the case had the stimuli been created through manual methods such as photoshopping. Not photoshopping the targets also meant that the target locations were objectively based on a set of criteria, rather than subjective.

Finally, Chapter 5 also contains specific criteria on how the scotoma was formed to accomplish full degradation of foveal vision. These scotoma criteria were then used throughout the entire research to degrade various regions of the visual field. A number of visual search studies that use scotomas tend to use ones that are suboptimal (Nuthmann, 2014). Therefore, the criteria for generating scotomas from this research could be used as the preferred method for future studies that involve visual field loss.

Chapter 6 investigated possible interaction effects between target size and salience. As Chapter 5 indicated, the decrease in target size increases search times, and reduces search accuracy regardless of foveal vision presence. Simply put, making targets smaller makes the task harder in general. Chapter 6 contains results for two questions: at what search stage does search performance costs originate? Do target size and salience interact with one another and if so, how does this impact search performance?

There was an interaction between target size and salience to the point where salience starts to compensate when target size is small. The opposite is also true, size starts to compensate when salience is very low. This effect has been coined the Compensation Effect. Since the targets are always semantically inconsistent with respect to the background, participants could not rely on top-down knowledge of the scene to guide their search. See Section 6.6 for a detailed discussion on this, but Chapter 6 should be treated as evidence in favour of search not solely being a top-down task.

The main search performance costs were found to be within the verification time period but were not affected by foveal vision loss. It appears that the loss of foveal vision does little to impact search performance. Accuracy decreases but finding and verifying targets is almost identical to having normal or corrected-to-normal vision. Increasing this visual field loss to just beyond foveal vision (larger radius but not encompassing all central vision) then we found a significant difference. Scanning time was the same for both conditions, which emphasises the hypothesis that search is primarily a peripheral task, but verification time dramatically increases when searching for low salience targets (see Section 6.5.1, Figure 18). Due to these findings, the previous assertion that foveal vision was necessary for the attainment of normal search performance is less credible, as it is not within the fovea where these search costs are originating.

Chapter 7 investigated the questions in 3D immersive scenes over its 2D static counterparts. The experiments of Chapter 7 were also concerned with search performance, only this time the scenes used simulated self-motion, which produced an optical flow field. The chapter investigated search performance, target recognition, and methods to improve search performance under foveal vision loss. Unlike the results of previous chapters, Chapter 7 demonstrated that the inclusion of optical flow contributes to a significant drop in search performance when searching without foveal vision. The original motivation behind Chapter 7's experiments was from a study by Royden and Hildreth (1999) who explained that a more focused attention was needed when judging an object's motion from an optic flow field. If optic flow can invoke such a response from the observer, then searching with an optic flow present should simply be harder overall. Experiment 5 certainly suggested this but it was not sufficient to explain why searching without foveal vision was so detrimental. If it were only the case that search was harder,

then we should see a similar pattern for both conditions rather than a stark difference. The results of experiment 5 became the motivation for the other two experiments in Chapter 7. The ability to recognise targets was not affected much from optic flow inclusion (see the general discussion of Chapter 7). The final experiment of this thesis demonstrates that search performance can be improved upon manipulation of motion from the observer's point of view, but not enough to eliminate the detrimental effect of not having foveal vision.

The results of Chapter 7 are important, as they show that results can change when moving from 2D to immersive 3D scenes. It might well be the case that we are neglecting important factors by using 2D static images, which in turn stops observers from utilising capabilities (like optic flow) that would otherwise have been available to them. 3D immersive scenes also grant experimenters full advantage of the scene distance, which can then be used to give a sense of depth that would be unavailable otherwise. At the time of this writing, immersive scenes constructed via a game engine and the methodologies outlined in Chapter 7 have not been previously used to study regions of the visual field with respect to visual search. Therefore, the scenes generated could be used as a foundation for subsequent experiments in the area.

Furthermore, Chapter 7 outlines both the extension to the Target Embedding Algorithm for 3D stimuli (dubbed T.E.A.M.S. to include moving scenes. See Section 7.2) and the proposed Dead-Zone modification to the Malcolm and Henderson (2009) algorithm on search time decomposition (see Section 7.4). The experiments from Chapter 7 act as an introduction to a new way of investigating visual search, with results that, I hope, are interesting enough for others in the field to pursue.

Finally, the entire range of stimuli both 2D and 3D that were developed using the T.E.A/M.S. are freely available for subsequent research (see Appendix C).

## 8.2 Successes and Limitations

The Target Embedding Algorithm, although novel and useful for target insertion, is limited in terms of computational performance. The algorithm processed images pixel by pixel to use contrast difference values in order to obtain a map of viable target positions. This was already computationally expensive and although adding more features to encompass all factors of salience (such as colour and orientation) would make the T.E.A. more representative, the computation time needed would increase. Methods were devised to improve the T.E.A.'s performance, but they were dismissed as they would produce contrast difference maps that were less accurate. Now that we have sets of stimuli made from the T.E.A., future iterations could be investigated to make computation time more efficient.

Aside from this limitation, the T.E.A. was very effective at providing a broad range of target positions that collectively across all stimuli were not biased towards any particular quadrant. The algorithm was also flexible in its ability to adjust contrast values to obtain stimuli of varying levels of perceived difficulty. Its transferability from 2D to 3D scenes was also met with success. In addition, the scotoma used throughout this research has been developed and outlined in sufficient detail for others to use in subsequent research. It could be very useful in future research as it fulfils the purpose of blocking out the required region entirely and blends with the scene as to not cause distraction.

Chapter 7 suffered from the lack of photorealism with the 3D immersive scenes. The Unity game engine is very powerful, and capable of displaying high quality 3D environments, but the lack of available artistic skills meant that all scenes had to be made from existing assets from Unity's asset store and pieced together by hand. This did not seem to cause too much of an issue as all scenes were enhanced using temporal anti-aliasing, but it is certainly something to improve upon in the future.

The preliminary results of Chapter 7 called for an algorithmic extension to the gaze-based decomposition of search time by Malcolm and Henderson (2009). The proposed approach (see Section 7.4) was successful in balancing the verification and scanning time epochs and is open to discussion for further improvement. The new decomposition of search time now considers additional fixations, saccades and saccade directions to distinguish false from true verification.

## 8.3 Future Work

The experiments contained in this thesis lay the groundwork for a lot of future research. To investigate 3D immersive scenes in more detail, a future experiment could involve a peripheral scotoma. As speculated in Chapter 7, the optic flow makes targets blend with the background, increasing an observer's reliance on high acuity processing. If the same experiment was conducted with a peripheral scotoma, we may be able to measure the vicinity in which the target becomes visible. As mentioned in Section 8.1, 3D immersive scenes provide depth via the Z dimension, but the observer is still analysing stimuli on a screen whilst sitting down. The natural progression of this work would be to involve virtual reality headsets such as the HTC Vive, or Oculus Rift. We have already seen that search performance can improve by manipulating motion on the screen, but can this be taken a step further by incorporating physical movement of the observer? A paper

by Rushton et al (1998) suggests this when investigating navigation, which leads to the hypothesis that the results will change again for visual search due to factors such as head movements and immersion. Ecological validity has been encouraged in recent years (Brennan et al., 2017; Foulsham & Kingstone, 2017; Risko et al., 2012; Risko, Richardson, & Kingstone, 2016) however, the use of virtual reality headsets (in particular the HTC Vive) to aid in ecological validity are unsuitable for experiments that involve simulated self motion (Niehorster, Li, & Lappe, 2017). One way or another, by moving to more ecologically valid scenarios, this line of research makes visual aid applications all the more possible and will hopefully contribute to the quality of life of ophthalmic disease sufferers.

# 9 References

Aguilar, C., & Castet, E. (2011). Gaze-contingent simulation of retinopathy: some potential pitfalls and remedies. *Vision Research*, *51*(9), 997–1012. https://doi.org/10.1016/j.visres.2011.02.010

Albers, C., & Lakens, D. (2018). When power analyses based on pilot data are biased: Inaccurate effect size estimators and follow-up bias. *Journal of Experimental Social Psychology*, *74*, 187–195. https://doi.org/10.1016/j.jesp.2017.09.004

Allport, D. A. (1987). Selection for Action: Some Behavioral and Neurophysiological Considerations of Attention and Action. In H. Heuer & H. F. Sanders (Eds.), *Perspectives on Perception and Action* (pp. 395–419). Lawerence Erlbaum.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Bacon-Macé, N., Macé, M. J.-M., Fabre-Thorpe, M., & Thorpe, S. J. (2005). The time course of visual processing: Backward masking and natural scene categorisation. *Vision Research*, *45*(11), 1459–1469. https://doi.org/10.1016/j.visres.2005.01.004

Baddeley, R. J., & Tatler, B. W. (2006). High frequency edges (but not contrast) predict where we fixate: A Bayesian system identification analysis. *Vision Research*, *46*(18), 2824–2833. https://doi.org/10.1016/J.VISRES.2006.02.024

Barraza-Bernal, M. J., Rifai, K., & Wahl, S. (2018). The retinal locus of fixation in simulations of progressing central scotomas. *Journal of Vision*, *18*(1), 7. https://doi.org/10.1167/18.1.7

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using {lme4}. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bauer, B., Jolicoeur, P., & Cowan, W. B. (1996). Visual search for colour targets that are or are not linearly separable from distractors. *Vision Research*, *36*(10), 1439–1466. https://doi.org/10.1016/0042-6989(95)00207-3

Bauer, B., Jolicoeur, P., & Cowan, W. B. (1998). The linear separability effect in color visual search: ruling out the additive color hypothesis. *Perception & Psychophysics*, *60*(6), 1083–1093. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9718965

Bauer, B., Jolicoeur, P., & Cowan, W. B. (1999). Convex hull test of the linear separability hypothesis in visual search. *Vision Research*, *39*(16), 2681–2695. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10492830

Becker, W. (1991). Saccades. *R.H.S. Carpenter (Ed.), Eye Movements*, 95–137.

Bertera, J. H. (1988). The effect of simulated scotomas on visual search in normal subjects. *Investigative Ophthalmology & Visual Science*, *29*(3), 470–475.

Bertera, J. H., & Rayner, K. (2000). Eye movements and the span of the effective stimulus in visual search. *Perception & Psychophysics*, *62*(3). Retrieved from http://doi.org/10.3758/BF03212109

Bex, P. J., & Makous, W. (2002). Spatial frequency, phase, and the contrast of natural images. *Journal of the Optical Society of America A-Optics Image Science and Vision*, *19*(6), 1096–1106. https://doi.org/10.1364/JOSAA.19.001096

Biederman, I., Mezzanotte, R. J., & Rabinowitz, J. C. (1982). Scene perception: Detecting and judging objects undergoing relational violations. *Cognitive Psychology*, *14*(2), 143–177. https://doi.org/10.1016/0010-0285(82)90007-X

Borji, A., & Itti, L. (2013). State-of-the-art in visual attention modeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *35*(1), 185–207. https://doi.org/10.1109/TPAMI.2012.89

Borji, A., Sihite, D. N., & Itti, L. (2013). Objects do not predict fixations better than early saliency: a re-analysis of Einhauser et al.'s data. *Journal of Vision*, *13*(10), 18. https://doi.org/10.1167/13.10.18

Borji, A., Sihite, D. N., & Itti, L. (2013). Quantitative analysis of human-model agreement in visual saliency modeling: a comparative study. *IEEE Transactions on Image Processing*, *22*(1), 55–69. https://doi.org/10.1109/TIP.2012.2210727

Boucart, M., Delerue, C., Thibaut, M., Szaffarczyk, S., Hayhoe, M., & Tran, T. H. C. (2015). Impact of wet macular degeneration on the execution of natural actions. *Investigative Ophthalmology & Visual Science*, *56*(11), 6832–6838. https://doi.org/10.1167/iovs.15-16758

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

Brennan, A. A., Bruderer, A. J., Liu-Ambrose, T., Handy, T. C., & Enns, J. T. (2017). Lifespan changes in attention revisited: Everyday visual search. *Canadian Journal of Experimental Psychology*, *71*(2), 160–171. https://doi.org/10.1037/cep0000130

Bridgeman, B., & Stark, L. (1991). Ocular proprioception and efference copy in registering visual direction. *Vision Research*, *31*(11), 1903–1913. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1771774

Brockmole, J. R., Castelhano, M. S., & Henderson, J. M. (2006). Contextual cueing in naturalistic scenes: Global and local contexts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(4), 699–706. https://doi.org/10.1037/0278-7393.32.4.699

Brockmole, J. R., & Henderson, J. M. (2006). Short Article: Recognition and Attention Guidance during Contextual Cueing in Real-World Scenes: Evidence from Eye Movements. *Quarterly Journal of Experimental Psychology*, *59*(7), 1177–1187. https://doi.org/10.1080/17470210600665996

Caldara, R., Zhou, X., & Miellet, S. (2010). Putting Culture Under the 'Spotlight' Reveals Universal Information Use for Face Recognition. *PLoS ONE*, *5*(3), e9708. https://doi.org/10.1371/journal.pone.0009708

Carrasco, M., Evert, D. L., Chang, I., & Katz, S. M. (1995). The eccentricity effect: Target eccentricity affects performance on conjunction searches. *Perception & Psychophysics*, *57*(8), 1241–1261. https://doi.org/10.3758/BF03208380

Carrasco, M., & Frieder, K. S. (1997). Cortical magnification neutralizes the eccentricity effect in visual search. *Vision Research*, *37*(1). Retrieved from http://doi.org/10.1016/S0042-6989(96)

Carrasco, M., & McElree, B. (2001). Covert attention accelerates the rate of visual information processing. *Proceedings of the National Academy of Sciences*, *98*(9), 5363–5367. https://doi.org/10.1073/pnas.081074098

Carrasco, M., Mclean, T. L., Katz, S. M., & Frieder, K. S. (1998). Feature asymmetries in visual search: {Effects} of display duration, target eccentricity, orientation and spatial frequency. *Vision Research*, *38*(3). Retrieved from http://doi.org/10.1016/S0042-6989(97)

Castelhano, M. S., & Heaven, C. (2010). The relative contribution of scene context and target features to visual search in scenes. *Attention, Perception & Psychophysics*, *72*(5), 1283–1297. https://doi.org/10.3758/APP.72.5.1283

Castelhano, M. S., & Heaven, C. (2011). Scene context influences without scene gist: Eye movements guided by spatial associations in visual search. *Psychonomic Bulletin and Review*, *18*(5), 890–896. https://doi.org/10.3758/s13423-011-0107-8

Castelhano, M. S., Mack, M. L., & Henderson, J. M. (2009). Viewing task influences eye movement control during active scene perception. *Journal of Vision*, *9*(3), 6–6. https://doi.org/10.1167/9.3.6

Chen, X., & Zelinsky, G. J. (2006). Real-world visual search is dominated by top-down guidance. *Vision Research*, *46*(24), 4118–4133. https://doi.org/10.1016/j.visres.2006.08.008

Chun, M. M., & Jiang, Y. (1998). Contextual Cueing: Implicit Learning and Memory of Visual Context Guides Spatial Attention. *Cognitive Psychology*, *36*(1), 28–71. https://doi.org/10.1006/cogp.1998.0681

Chung, S. T. L., Legge, G. E., & Tjan, B. S. (2002). Spatial-frequency characteristics of letter identification in central and peripheral vision. *Vision Research*, *42*(18), 2137–2152. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12207975

Congdon, N., O'Colmain, B., Klaver, C. C. W., Klein, R., Munoz, B., Friedman, D. S., ... Grp, E. D. P. R. (2004). Causes and prevalence of visual impairment among adults in the United States. *Archives of Ophthalmology*, *122*(4), 477–485. https://doi.org/10.1001/archopht.122.4.477

Cornelissen, F. W., Bruin, K. J., & Kooijman, A. C. (2005). The influence of artificial scotomas on eye movements during visual search. *Optometry and Vision Science*, *82*(1), 27–35.

Cornelissen, F. W., Peters, E. M., & Palmer, J. (2002). The Eyelink Toolbox: Eye tracking with MATLAB and the Psychophysics Toolbox. *Behavior Research Methods, Instruments, & Computers*, *34*(4), 613–617. https://doi.org/10.3758/BF03195489

Cornelissen, T. H. W., & Võ, M. L.-H. (2017). Stuck on semantics: Processing of irrelevant object-scene inconsistencies modulates ongoing gaze behavior. *Attention Perception & Psychophysics*, *79*(1), 154–168. https://doi.org/10.3758/s13414-016-1203-7

Cornelissen, T., Holmqvist, K., & Vo, M. (2017). I couldn't help but notice: Irrelevant object-scene inconsistencies influence search for highly visible gabor patches. *Journal of Vision*, *17*(10), 565. https://doi.org/10.1167/17.10.565

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology*, *1*(1), 42–45. https://doi.org/10.20982/tqmp.01.1.p042

D'Zmura, M. (1991). Color in visual search. *Vision Research*, *31*(6), 951–966. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1858326

David, Callet P., L., Silva M., P. Da, & Lebranchu P. (2018). How are ocular behaviours affected by central and peripheral vision loss? A study based on artificial scotomas and gaze-contingent paradigm. *Electronic Imaging*, *2018*(14), 1–6. https://doi.org/10.2352/issn.2470-1173.2018.14.hvei-504

Dixon, P. (2008). Models of accuracy in repeated-measures designs. *Journal of Memory and Language*, *59*(4), 447–456. https://doi.org/10.1016/j.jml.2007.11.004

Domini, F., & Caudek, C. (2003). 3-D structure perceived from dynamic information: a new theory. *Trends in Cognitive Sciences*, *7*(10), 444–449. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14550491

Drew, T., Vo, M. L.-H., Olwal, A., Jacobson, F., Seltzer, S. E., & Wolfe, J. M. (2013). Scanners and drillers: Characterizing expert visual search through volumetric images. *Journal of Vision*, *13*(10), 3–3. https://doi.org/10.1167/13.10.3

Duchowski, A. T., & Çöltekin, A. (2007). Foveated gaze-contingent displays for peripheral LOD management, 3D visualization, and stereo Imaging. *ACM Transactions on Multimedia Computing Communications and Applications*, *3(4):24*, 1–18. https://doi.org/10.1145/1314303.1314309

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, *96*(3), 433–458. https://doi.org/10.1037//0033-295X.96.3.433

Eckstein, M. P. (2011). Visual search: A retrospective. *Journal of Vision*, *11(5):14*, 1–36. https://doi.org/10.1167/11.5.14

Ehinger, K. a., Hidalgo-Sotelo, B., Torralba, A., & Oliva, A. (2009). Modelling search for people in 900 scenes: A combined source model of eye guidance. *Visual Cognition*, *17*(6–7), 945–978. https://doi.org/10.1080/13506280902834720

Einhäuser, W., Rutishauser, U., & Koch, C. (2008). Task-demands can immediately reverse the effects of sensory-driven saliency in complex visual stimuli. *Journal of Vision*, *8*(2), 2.1-19. https://doi.org/10.1167/8.2.2

Einhäuser, W., Spain, M., & Perona, P. (2008). Objects predict fixations better than early saliency. *Journal of Vision*, *8*(14), 18.1-26. https://doi.org/10.1167/8.14.18

Engel, F. L. (1977). Visual conspicuity, visual search and fixation tendencies of the eye. *Vision Research*, *17*(1), 95–108. https://doi.org/10.1016/0042-6989(77)90207-3

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/17695343

Findlay, J. M., Brown, V., & Gilchrist, I. D. (2001). Saccade target selection in visual search: the effect of information from the previous fixation. *Vision Research*, *41*(1), 87–95. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11163618

Findlay, J. M., & Walker, R. (1999). A model of saccade generation based on parallel processing and competitive inhibition. *The Behavioral and Brain Sciences*, *22*(4), 661-74; discussion 674-721. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11301526

Fine, E. M., & Rubin, G. S. (1999). Reading with central field loss: number of letters masked is more important than the size of the mask in degrees. *Vision Research*, *39*(4), 747–756. https://doi.org/10.1016/S0042-6989(98)00142-4

Fortenbaugh, F. C., Hicks, J. C., Hao, L., & Turano, K. A. (2007a). A technique for simulating visual field losses in virtual environments to study human navigation.

*Behavior Research Methods*, *39*(3), 552–560. https://doi.org/10.3758/BF03193025

Fortenbaugh, F. C., Hicks, J. C., Hao, L., & Turano, K. A. (2007b). Losing sight of the bigger picture: peripheral field loss compresses representations of space. *Vision Research*, *47*(19), 2506–2520. https://doi.org/10.1016/j.visres.2007.06.012

Foulsham, T., Alan, R., & Kingstone, A. (2011). Scrambled eyes? Disrupting scene structure impedes focal processing and increases bottom-up guidance. *Attention Perception & Psychophysics*, *73*(7), 2008–2025. https://doi.org/10.3758/s13414-011-0158-y

Foulsham, T., Barton, J. J. S., Kingstone, A., Dewhurst, R., & Underwood, G. (2009). Fixation and saliency during search of natural scenes: The case of visual agnosia. *Neuropsychologia*, *47*(8–9), 1994–2003. https://doi.org/10.1016/j.neuropsychologia.2009.03.013

Foulsham, T., Chapman, C., Nasiopoulos, E., & Kingstone, A. (2014). Top-down and bottom-up aspects of active search in a real-world environment. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *68*(1), 8–19. https://doi.org/10.1037/cep0000004

Foulsham, T., & Kingstone, A. (2017). Are fixations in static natural scenes a useful predictor of attention in the real world? *Canadian Journal of Experimental Psychology*, *71*(2), 172–181. https://doi.org/10.1037/cep0000125

Foulsham, T., & Underwood, G. (2007). How Does the Purpose of Inspection Influence the Potency of Visual Salience in Scene Perception? *Perception*, *36*(8), 1123–1138. https://doi.org/10.1068/p5659

Foulsham, T., & Underwood, G. (2008). What can saliency models predict about eye movements? Spatial and sequential aspects of fixations during encoding and recognition. *Journal of Vision*, *8*(2), 6. https://doi.org/10.1167/8.2.6

Foulsham, T., & Underwood, G. (2011). If visual saliency predicts search, then why? Evidence from normal and gaze-contingent search tasks in natural scenes. *Cognitive Computation*, *3*(1), 48–63. https://doi.org/10.1007/s12559-010-9069-9

Friedman, D. S., O'Colmain, B., Tomany, S. C., McCarty, C., de Jong, P., Nemesure, B., ... Grp, E. D. P. R. (2004). Prevalence of age-related macular degeneration in the United States. *Archives of Ophthalmology*, *122*(4), 564–572. https://doi.org/10.1001/archopht.122.4.564

Garcia-Diaz, A. (2011). Modeling Early Visual Coding and Saliency Through Adaptive Whitening: Plausibility, Assessment and Applications. Doctoral Thesis, University of Santiago de Compostela.

Garcia-Diaz, A., Fdez-Vidal, X. R., Pardo, X. M., and Dosil, R. (2012a). Saliency from hierarchical adaptation through decorrelation and variance normalization. Image Vis. Comput. 30, 51–64. doi: 10.1016/j.imavis.2011.11.007

Geisler, W. S., & Chou, K. L. (1995). Separation of low-level and high-level factors in complex tasks: Visual search. *Psychological Review*, *102*(2), 356–378. https://doi.org/10.1037/0033-295X.102.2.356

Geringswald, F., Baumgartner, F. J., & Pollmann, S. (2013). A behavioral task for the validation of a gaze-contingent simulated scotoma. *Behavior Research Methods*, *45*(4), 1313–1321. https://doi.org/10.3758/s13428-013-0321-6

Gibson, J. J. (1950). The perception of the visual world. The perception of the visual world. Oxford, England: Houghton Mifflin.

Gibson, J. J. (James J. (1979). *The ecological approach to visual perception*. Houghton Mifflin.

Gibson, J. J. (James J. (1983). *The senses considered as perceptual systems*. Greenwood Press.

Glaholt, M. G., Rayner, K., & Reingold, E. M. (2012). The mask-onset delay paradigm and the availability of central and peripheral visual information during scene viewing. *Journal of Vision*, *12(1):9*, 1–19. https://doi.org/10.1167/12.1.9

Hayhoe, M. M., McKinney, T., Chajka, K., & Pelz, J. B. (2012). Predictive eye movements in natural vision. *Experimental Brain Research*, *217*(1), 125–136. https://doi.org/10.1007/s00221-011-2979-2

Hayhoe, M., & Ballard, D. (2005). Eye movements in natural behavior. *Trends in Cognitive Sciences*, *9*(4), 188–194. https://doi.org/10.1016/j.tics.2005.02.009

Henderson, J. M., Brockmole, J. R., Castelhano, M. S., & Mack, M. L. (2007). Visual saliency does not account for eye movements during visual search in real world scenes. In *Eye Movements: A Window on Mind and Brain* (pp. 537–562). https://doi.org/10.1167/9.3.6.

Henderson, J. M., & Ferreira, F. (2004). Scene Perception for Psycholinguists. In *The interface of language, vision, and action: Eye movements and the visual world.* (pp. 1–58). Henderson, John M.: Michigan State University, MI, US: Psychology Press.

Henderson, J. M., & Hollingworth, A. (1999). High-level scene perception. *Annual Review of Psychology*. US: Annual Reviews. https://doi.org/10.1146/annurev.psych.50.1.243

Henderson, J. M., Malcolm, G. L., & Schandl, C. (2009). Searching in the dark: cognitive relevance drives attention in real-world scenes. *Psychonomic Bulletin & Review*, *16*(5), 850–856. https://doi.org/10.3758/PBR.16.5.850

Henderson, J. M., Mcclure, K. K., Pierce, S., & Schrock, G. (1997). Object identification without foveal vision: Evidence from an artificial scotoma paradigm. *Perception & Psychophysics*, *59*(3), 323–346. https://doi.org/10.3758/BF03211901

Henderson, J. M., Pollatsek, A., & Rayner, K. (1989). Covert visual attention and extrafoveal information use during object identification. *Perception & Psychophysics*, *45*(3), 196–208. https://doi.org/10.3758/BF03210697

Henderson, J. M., Weeks, P. A., & Hollingworth, A. (1999). The effects of semantic consistency on eye movements during complex scene viewing. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(1), 210–228. https://doi.org/10.1037/0096-1523.25.1.210

Horowitz, T., & Wolfe, J. (2003). Memory for rejected distractors in visual search? *Visual Cognition*, *10*(3), 257–298. https://doi.org/10.1080/13506280143000005

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, *40*(10–12), 1489–1506. https://doi.org/10.1016/S0042-6989(99)00163-7

J. Zelinsky, G., Zhang, W., Yu, B., Chen, X., & Samaras, D. (2005). The Role of Top-down and Bottom-up Processes in Guiding Eye Movements during Visual Search. (Vol. 18).

James, W. (1890). *The principles of psychology*. New York : Henry Holt and Co., 1890.

Jiang, Y., & Wagner, L. C. (2004). What is learned in spatial contextual cuing—configuration or individual locations? *Perception & Psychophysics*, *66*(3), 454–463. https://doi.org/10.3758/BF03194893

Jovancevic, J., Sullivan, B., & Hayhoe, M. (2006). Control of attention and gaze in complex environments. *Journal of Vision*, *6*(12), 9. https://doi.org/10.1167/6.12.9

Julesz, B. (1984). A brief outline of the texton theory of human vision. *Trends in Neurosciences*, *7*(2), 41–45. https://doi.org/10.1016/S0166-2236(84)80275-1

Kanan, C., Tong, M. H., Zhang, L., & Cottrell, G. W. (2009). SUN: Top-down saliency using natural statistics. *Visual Cognition*, *17*(6–7), 979–1003. https://doi.org/10.1080/13506280902771138

Kenward, M. G., & Roger, J. H. (1997). Small Sample Inference for Fixed Effects from Restricted Maximum Likelihood. *Biometrics*, *53*(3), 983–997. Retrieved from http://www.jstor.org/stable/2533558

Kit, D., Katz, L., Sullivan, B., Snyder, K., Ballard, D., & Hayhoe, M. (2014). Eye movements, visual search and scene memory, in an immersive virtual environment. *PLoS ONE*, *9*(4), 94362. https://doi.org/10.1371/journal.pone.0094362

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

Klein, R., & Farrell, M. (1989). Search performance without eye movements. *Perception & Psychophysics*, *46*(5), 476–482. https://doi.org/10.3758/BF03210863

Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in psychtoolbox-3. *Perception*, *36*(14), 1–16.

Kowler, E., & McKee, S. P. (1987). Sensitivity of smooth eye movement to small differences in target velocity. *Vision Research*, *27*(6), 993–1015. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3660658

Kristjánsson, Á. (2000). In Search of Remembrance: Evidence for Memory in Visual Search. *Psychological Science*, *11*(4), 328–332. https://doi.org/10.1111/1467-9280.00265

Kugler, G., Huppert, D., Schneider, E., & Brandt, T. (2014). Fear of heights freezes gaze to the horizon. *Journal of Vestibular Research : Equilibrium & Orientation*, *24*(5–6), 433–441. https://doi.org/10.3233/VES-140529

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Land, M. F., & Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities? *Vision Research*, *41*(25–26), 3559–3565. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11718795

Larson, A. M., & Loschky, L. C. (2009). The contributions of central versus peripheral vision to scene gist recognition. *Journal of Vision*, *9*(10), 6–6. https://doi.org/10.1167/9.10.6

Lawrence, M. A. (2015). ez: Easy analysis and visualization of factorial experiments (R package version 4.3).

Loschky, L. C., & McConkie, G. W. (2002). Investigating spatial vision and dynamic attentional selection using a gaze-contingent multiresolutional display. *Journal of Experimental Psychology. Applied*, *8*(2), 99–117. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12075694

Loschky, L. C., Sethi, A., Simons, D. J., Pydimarri, T. N., Ochs, D., & Corbeille, J. L. (2007). The Importance of Information Localization in Scene Gist Recognition. https://doi.org/10.1037/0096-1523.33.6.1431

Luke, S. G., & Henderson, J. M. (2016). The influence of content meaningfulness on eye movements across tasks: Evidence from scene viewing and reading. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2015.00257

Mack, S. C., & Eckstein, M. P. (2011). Object co-occurrence serves as a contextual cue to guide and facilitate visual search in a natural viewing environment. *Journal of Vision*, *11*(9), 9–9. https://doi.org/10.1167/11.9.9

Majaj, N. J., Pelli, D. G., Kurshan, P., & Palomares, M. (n.d.). *The role of spatial frequency channels in letter identification*. Retrieved from www.elsevier.com/locate/visres

Malcolm, G. L., & Henderson, J. M. (2009). The effects of target template specificity on visual search in real-world scenes: Evidence from eye movements. *Journal of Vision*, *9*(11), 1–13. https://doi.org/10.1167/9.11.8

Malcolm, G. L., & Henderson, J. M. (2010). Combining top-down processes to guide eye movements during real-world scene search. *Journal of Vision*, *10*(2), 1–11. https://doi.org/10.1167/10.2.4

Mannan, S. K., Ruddock, K. H., & Wooding, D. S. (1996). The relationship between the locations of spatial features and those of fixations made during visual examination of briefly presented images. *Spatial Vision*, *10*(3), 165–188. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9061830

Mannan, S., Ruddock, K. H., & Wooding, D. S. (1995). Automatic control of saccadic eye movements made in visual inspection of briefly presented 2-D images. *Spatial Vision*, *9*(3), 363–386. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8962841

McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception & Psychophysics*, *17*(6), 578–586. https://doi.org/10.3758/BF03203972

McIlreavy, L., Fiser, J., & Bex, P. J. (2012). Impact of simulated central scotomas on visual search in natural scenes. *Optometry and Vision Science*, *89*(9), 1385–1394. https://doi.org/10.1097/OPX.0b013e318267a914

Meißner, M., Pfeiffer, J., Pfeiffer, T., & Oppewal, H. (2017). Combining virtual reality and mobile eye tracking to provide a naturalistic experimental environment for shopper research. *Journal of Business Research*. https://doi.org/10.1016/J.JBUSRES.2017.09.028

Miellet, S., Zhou, X., He, L., Rodger, H., & Caldara, R. (2010). Investigating cultural diversity for extrafoveal information use in visual scenes. *Journal of Vision*, *10*(6), 1–18. https://doi.org/10.1167/10.6.21

Mital, P. K., Smith, T. J., Hill, R. L., & Henderson, J. M. (2011). Clustering of Gaze During Dynamic Scene Viewing is Predicted by Motion. *Cognitive Computation*, *3*(1), 5–24. https://doi.org/10.1007/s12559-010-9074-z

Moulden, B., Kingdom, F., & Gatley, L. F. (1990). The standard deviation of luminance as a metric for contrast in random-dot images. *Perception*, *19*(1), 79–101. https://doi.org/10.1068/p190079

Murakami, I., & Shimojo, S. (1995). Modulation of motion aftereffect by surround motion and its dependence on stimulus size and eccentricity. *Vision Research*, *35*(13), 1835–1844. https://doi.org/10.1016/0042-6989(94)00269-R

Murthy, A., Thompson, K. G., & Schall, J. D. (2001). Dynamic Dissociation of Visual Selection From Saccade Programming in Frontal Eye Field. *Journal of Neurophysiology*, *86*(5), 2634–2637. https://doi.org/10.1152/jn.2001.86.5.2634

Neumann, O. (1987). Beyond capacity: A functional view of attention. In *Perspectives on perception and action.* (pp. 361–394). Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.

Niehorster, D. C., Li, L., & Lappe, M. (2017). The Accuracy and Precision of Position and Orientation Tracking in the HTC Vive Virtual Reality System for Scientific Research. *I-Perception*, *8*(3), 204166951770820. https://doi.org/10.1177/2041669517708205

Norman, D., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. Cognitive Neuroscience: A reader.

Nuthmann, A. (2013). On the visual span during object search in real-world scenes. *Visual Cognition*, *21*(7), 803–837. https://doi.org/10.1080/13506285.2013.832449

Nuthmann, A. (2014). How do the regions of the visual field contribute to object search in real-world scenes? Evidence from eye movements. *Journal of Experimental Psychology. Human Perception and Performance*, *40*(1), 342–360. https://doi.org/10.1037/a0033854

Nuthmann, A. (2016). Fixation durations in scene viewing: Modeling the effects of local image features, oculomotor parameters, and task. *Psychonomic Bulletin & Review*, 1–23. https://doi.org/10.3758/s13423-016-1124-4

Nuthmann, A., & Einhäuser, W. (2015). A new approach to modeling the influence of image features on fixation selection in scenes. *Annals of the New York Academy of Sciences*, *1339*(1), 82–96. https://doi.org/10.1111/nyas.12705

Nuthmann, A., & Malcolm, G. L. (2016). Eye guidance during real-world scene search: The role color plays in central and peripheral vision. *Journal of Vision*, *16(2):3*, 1–16. https://doi.org/10.1167/16.2.3

Oliva, A. (2005). Gist of the scene. In L. Itti, G. Rees, & J. K. Tsotsos (Eds.), *Neurobiology of attention* (pp. 251–256). San Diego, CA: Elsevier.

Pajak, M., & Nuthmann, A. (2013). Object-based saccadic selection during scene perception: Evidence from viewing position effects. *Journal of Vision*, *13(5):2*, 1–21. https://doi.org/10.1167/13.5.2

Palmer, S. E. (1999). *Vision science: Photons to Phenomenology*. *System*. MIT Press. Retrieved from https://mitpress.mit.edu/books/vision-science

Pannasch, S., Helmert, J. R., Roth, K., Herbold, A.-K., & Walter, H. (2008). Visual fixation durations and saccade amplitudes: Shifting relationship in a variety of conditions. *Journal of Eye Movement Research*, *2(2):4*, 1–19. https://doi.org/10.16910/jemr.2.2.4

Pepping, G.-J., & Grealy, M. A. (2007). *Closing the gap : the scientific writings of David N. Lee*. Lawrence Erlbaum.

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

Pomplun, M., Garaas, T. W., & Carrasco, M. (2013). The effects of task difficulty on visual search strategy in virtual 3D displays. *Journal of Vision*, *13*(3), 24–24. https://doi.org/10.1167/13.3.24

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

Potter, M. C. (1975). Meaning in visual search. *Science*, *187*(4180), 965–966. https://doi.org/10.1126/science.1145183

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372–422. https://doi.org/10.1037//0033-2909.124.3.372

Rayner, K., & Bertera, J. (1979). Reading without a fovea. *Science*, *206*(4417), 468–469. https://doi.org/10.1126/science.504987

Reichle, E. D., Vanyukov, P. M., Laurent, P. A., & Warren, T. (2008). Serial or parallel? Using depth-of-processing to examine attention allocation during reading. *Vision Research*, *48*(17), 1831–1836. https://doi.org/10.1016/j.visres.2008.05.007

Reinagel, P., & Zador, A. M. (1999). Natural scene statistics at the centre of gaze. *Network: Computation in Neural Systems*, *10*(4), 341–350. https://doi.org/10.1088/0954-898X/10/4/304

Reinhardt-Rutland, A. H. (1988). Induced movement in the visual modality: An overview. *Psychological Bulletin*. US: American Psychological Association. https://doi.org/10.1037/0033-2909.103.1.57

Rensink, R. A. (2000). Seeing, sensing, and scrutinizing. In *Vision Research* (Vol. 40, pp. 1469–1487). https://doi.org/10.1016/S0042-6989(00)00003-1

Risko, E. F., Laidlaw, K., Freeth, M., Foulsham, T., & Kingstone, A. (2012). Social attention with real versus reel stimuli: toward an empirical approach to concerns about ecological validity. *Frontiers in Human Neuroscience*, *6*, 143. https://doi.org/10.3389/fnhum.2012.00143

Risko, E. F., Richardson, D. C., & Kingstone, A. (2016). Breaking the Fourth Wall of Cognitive Science. *Current Directions in Psychological Science*, *25*(1), 70–74. https://doi.org/10.1177/0963721415617806

Rizzolatti, G., Riggio, L., Dascola, I., & Umiltá, C. (1987). Reorienting attention across the horizontal and vertical meridians: evidence in favor of a premotor theory of attention. *Neuropsychologia*, *25*(1A), 31–40. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3574648

Rovamo, J., & Virsu, V. (1979). An estimation and application of the human cortical magnification factor. *Experimental Brain Research*, *37*(3), 495–510.

Royden, C. S., & Hildreth, E. C. (1999). Differential effects of shared attention on perception of heading and 3-D object motion. *Perception & Psychophysics*, *61*(1), 120–133. https://doi.org/10.3758/BF03211953

Royden, C. S., Wolfe, J. M., & Klempen, N. (2001). Visual search asymmetries in motion and optic flow fields. *Perception & Psychophysics*, *63*(3), 436–444. https://doi.org/10.3758/BF03194410

Rushton, S. K., Bradshaw, M. F., & Warren, P. A. (2007). The pop out of scene-relative object movement against retinal motion due to self-movement. *Cognition*, *105*(1), 237–245. https://doi.org/10.1016/j.cognition.2006.09.004

Rushton, S. K., Harris, J. M., Lloyd, M. R., & Wann, J. P. (1998). Guidance of locomotion on foot uses perceived target location rather than optic flow. *Current Biology*, *8*(21), 1191–1194. https://doi.org/10.1016/S0960-9822(07)00492-7

Rushton, S. K., Niehorster, D. C., Warren, P. A., & Li, L. (2018). The Primary Role of Flow Processing in the Identification of Scene-Relative Object Movement. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *38*(7), 1737–1743. https://doi.org/10.1523/JNEUROSCI.3530-16.2017

Rushton, S. K., & Warren, P. A. (2005). Moving observers, relative retinal motion and the detection of object movement. *Current Biology*, *15*(14), R542–R543. https://doi.org/10.1016/J.CUB.2005.07.020

Scialfa, C. T., & Joffe, K. M. (1998). Response times and eye movements in feature and conjunction search as a function of target eccentricity. *Perception & Psychophysics*, *60*(6), 1067–1082. https://doi.org/10.3758/BF03211940

Scialfa, C. T., & Kline, D. W. (1988). Effects of Noise Type and Retinal Eccentricity on Age Differences in Identification and Localization. *Journal of Gerontology*, *43*(4), P91–P99. https://doi.org/10.1093/geronj/43.4.P91

Scialfa, C. T., Kline, D. W., & Lyman, B. J. (1987). Age differences in target identification as a function of retinal location and noise level: Examination of the useful field of view. *Psychology and Aging*, *2*(1), 14–19. https://doi.org/10.1037/0882-7974.2.1.14

Singmann, H., Bolker, B., Westfall, J., & Aust, F. (2018). afex: Analysis of Factorial Experiments. Retrieved from https://cran.r-project.org/package=afex

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

Smith, T., & Henderson, J. (2008). Attentional synchrony in static and dynamic scenes. *Journal of Vision*, *8*(6), 773. https://doi.org/10.1167/8.6.773

Smith, T. J., & Mital, P. K. (2013). Attentional synchrony and the influence of viewing task on gaze behavior in static and dynamic scenes. *Journal of Vision*, *13*(8), 16–16. https://doi.org/10.1167/13.8.16

Snowden, R., Thompson, P., & Troscianko, T. (2006). *Basic Vision: An Introduction To Visual Perception*. Oxford University Press. Retrieved from https://www.amazon.co.uk/Basic-Vision-Introduction-Visual-Perception/dp/019957202X

Spotorno, S., Malcolm, G. L., & Tatler, B. W. (2015). Disentangling the effects of spatial inconsistency of targets and distractors when searching in realistic scenes. *Journal of Vision*, *15(2):12*, 1–21. https://doi.org/10.1167/15.2.12

Strasburger, H., Rentschler, I., & Juttner, M. (2011). Peripheral vision and pattern recognition: A review. *Journal of Vision*, *11*(5), 1–82. https://doi.org/10.1167/11.5.13

Stuart, G. W., Bossomaier, T. R. J., & Johnson, S. (1993). Preattentive processing of object size: implications for theories of size perception. *Perception*, *22*(10), 1175–1193. https://doi.org/10.1068/p221175

T. Buswell, G. (1935). How people look at pictures: a study of the psychology and perception in art.

Tatler, B. W. (2007). The central fixation bias in scene viewing: Selecting an optimal viewing position independently of motor biases and image feature distributions. *Journal of Vision*, *7*(14), 4. https://doi.org/10.1167/7.14.4

Tatler, B. W., Baddeley, R. J., & Gilchrist, I. D. (2005). Visual correlates of fixation selection: effects of scale and time. *Vision Research*, *45*(5), 643–659. https://doi.org/10.1016/j.visres.2004.09.017

Theeuwes, J., Kramer, A. F., Hahn, S., & Irwin, D. E. (1998). Our Eyes do Not Always Go Where we Want Them to Go: Capture of the Eyes by New Objects. *Psychological Science*, *9*(5), 379–385. https://doi.org/10.1111/1467-9280.00071

Todd, J. T. (1995). The Visual Perception of Three- Dimensional Structure from Motion. *Perception of Space and Motion*, 201–226. https://doi.org/10.1016/B978-012240530-3/50008-0

Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review*, *113*(4), 766–786. https://doi.org/10.1037/0033-295X.113.4.766

Townsend, J. T. (1990). Serial vs. Parallel Processing: Sometimes They Look like Tweedledum and Tweedledee but they can (and Should) be Distinguished. *Psychological Science*, *1*(1), 46–54. https://doi.org/10.1111/j.1467-9280.1990.tb00067.x

Treisman, A. (1988). Features and Objects: The Fourteenth Bartlett Memorial Lecture. *The Quarterly Journal of Experimental Psychology Section A*, *40*(2), 201–237. https://doi.org/10.1080/02724988843000104

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97–136. https://doi.org/10.1016/0010-0285(80)90005-5

Turano, K. A., Geruschat, D. R., & Baker, F. H. (2003). Oculomotor strategies for the direction of gaze tested with a real-world activity. *Vision Research*, *43*(3), 333–346. https://doi.org/10.1016/S0042-6989(02)00498-4

Underwood, G., & Foulsham, T. (2006). Visual saliency and semantic incongruency influence eye movements when inspecting pictures. *Quarterly Journal of Experimental Psychology*, *59*(11), 1931–1949. https://doi.org/10.1080/17470210500416342

Underwood, G., Foulsham, T., van Loon, E., Humphreys, L., & Bloyce, J. (2006). Eye movements during scene inspection: A test of the saliency map hypothesis. *European Journal of Cognitive Psychology*, *18*(3), 321–342. https://doi.org/10.1080/09541440500236661

Unema, P. J. A., Pannasch, S., Joos, M., & Velichkovsky, B. M. (2005). Time course of information processing during scene perception: The relationship between saccade amplitude and fixation duration. *Visual Cognition*, *12*(3), 473–494. https://doi.org/10.1080/13506280444000409

Van der Stigchel, S., Bethlehem, R. a I., Klein, B. P., Berendschot, T. T. J. M., Nijboer, T. C. W., & Dumoulin, S. O. (2013). Macular degeneration affects eye movement behavior during visual search. *Frontiers in Psychology*, *4*(September), 579. https://doi.org/10.3389/fpsyg.2013.00579

Van Diepen, P. M. J., De Graef, P., & Van Rensbergen, J. (1994). On-line control of moving masks and windows on a complex background using the ATVista videographics adapter. *Behavior Research Methods, Instruments, & Computers*, *26*(4), 454–460. https://doi.org/10.3758/BF03204665

van Diepen, P. M., Ruelens, L., & d'Ydewalle, G. (1999). Brief foveal masking during scene perception. *Acta Psychologica*, *101*(1), 91–103. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10100455

Viviani, P. (1990). Eye movements in visual search: cognitive, perceptual and motor control aspects. *Reviews of Oculomotor Research*, *4*, 353–393. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7492533

Võ, M. L. H., & Wolfe, J. M. (2013). Differential Electrophysiological Signatures of Semantic and Syntactic Scene Processing. *Psychological Science*, *24*(9), 1816–1823. https://doi.org/10.1177/0956797613476955

Warren, P. A., & Rushton, S. K. (2007). Perception of object trajectory: Parsing retinal motion into self and object movement components. *Journal of Vision*, *7*(11), 2. https://doi.org/10.1167/7.11.2

Warren, P. A., & Rushton, S. K. (2008a). Evidence for flow-parsing in radial flow displays. *Vision Research*, *48*(5), 655–663. https://doi.org/10.1016/j.visres.2007.10.023

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

Warren, W. H., & Fajen, B. R. (2004). Behavioral dynamics of human locomotion. *Ecological Psychology*, *16*(1), 61–66. https://doi.org/10.1207/s15326969eco1601_8

Wolfe, J., Cain, M., Ehinger, K., & Drew, T. (2015). Guided Search 5.0: Meeting the challenge of hybrid search and multiple-target foraging. Journal of vision (Vol. 15). https://doi.org/10.1167/15.12.1106

Wolfe, J. M. (1992). The Parallel Guidance of Visual Attention. *Current Directions in Psychological Science*, *1*(4), 124–128. https://doi.org/10.1111/1467-8721.ep10769733

Wolfe, J. M. (1994). Guided Search 2.0 A revised model of visual search. *Psychonomic Bulletin & Review*, *1*(2), 202–238. https://doi.org/10.3758/BF03200774

Wolfe, J. M. (1998). What Can 1 Million Trials Tell Us About Visual Search? *Psychological Science*, *9*(1), 33–39. https://doi.org/10.1111/1467-9280.00006

Wolfe, J. M. (2010). Guided Search 4.0: A guided search model that does not require memory for rejected distractors. *Journal of Vision*, *1*(3), 349–349. https://doi.org/10.1167/1.3.349

Wolfe, J. M., Alvarez, G. A., Rosenholtz, R., Kuzmova, Y. I., & Sherman, A. M. (2011). Visual search for arbitrary objects in real scenes. *Attention Perception & Psychophysics*, *73*(6), 1650–1671. https://doi.org/10.3758/s13414-011-0153-3

Wolfe, J. M., & Gancarz, G. (1997). Guided Search 3.0 (pp. 189–192). Springer, Dordrecht. https://doi.org/10.1007/978-94-011-5698-1_30

Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nature Human Behaviour*, *1*, article number 0058. https://doi.org/10.1038/s41562-017-0058

Wolfe, J. M., O'Neill, P., & Bennett, S. C. (1998). Why are there eccentricity effects in visual search? Visual and attentional hypotheses. *Perception & Psychophysics*, *60*(1), 140–156. https://doi.org/10.3758/BF03211924

Wyszecki, G., & Stiles, W. S. (Walter S. (2000). *Color science : concepts and methods, quantitative data, and formulae*. John Wiley & Sons.

Yarbus, A. L. (1967). Eye movements and vision. *Neuropsychologia*, *6*(4), 222. https://doi.org/10.1016/0028-3932(68)90012-2

Yeshurun, Y., & Carrasco, M. (1998). Attention improves or impairs visual performance by enhancing spatial resolution. *Nature*, *396*(6706), 72–75. https://doi.org/10.1038/23936

Yiğit, S., & Mendes, M. (2018). Which Effect Size Measure is Appropriate for One-Way and Two-Way ANOVA Models? A Monte Carlo Simulation Study. *Revstat - Statistical Journal*, *16*, 295–313.

Zelinsky, G. J. (2008). A theory of eye movements during target acquisition. *Psychological Review*, *115*(4), 787–835. https://doi.org/10.1037/a0013118

Zelinsky, G., Zhang, W., Yu, B., Chen, X., & Samaras, D. (2006). The role of top-down and bottom-up processes in guiding eye movements during visual search. *Advances in Neural Information Processing Systems 18*, 1569–1576. Retrieved from https://pdfs.semanticscholar.org/f2ac/99372bdebe08e4aebc73c6d35580e9a5117b.pdf

Zetzsche, C. (2005). Natural Scene Statistics and Salient Visual Features. *Neurobiology of Attention*, 226–232. https://doi.org/10.1016/B978-012375731-9/50041-0

# 10 Appendices

# Appendix A – Experiment 3 post-hoc pairwise comparisons

Reaction Time

| | contrast | estimate | SE | df | t.ratio | p.value |
|---|---|---|---|---|---|---|
| 1 | Small,Low - Large,Low | 1203.98 | 105.813 | 60.9645 1 | 11.3783 8 | 2.04E-11 |
| 2 | Small,Low - Small,High | 1566.155 | 103.6075 | 61.4694 | 15.1162 2 | 1.99E-11 |
| 3 | Small,Low - Large,High | 1988.547 | 110.4252 | 61.9111 7 | 18.0081 | 1.91E-11 |
| 4 | Large,Low - Small,High | 362.1741 | 110.4252 | 61.9111 7 | 3.27981 5 | 0.00903 |
| 5 | Large,Low - Large,High | 784.5663 | 103.6075 | 61.4694 | 7.57248 3 | 1.39E-09 |
| 6 | Small,High - Large,High | 422.3922 | 105.813 | 60.9645 1 | 3.99187 5 | 0.00100 3 |

Initiation Time

| | contrast | estimate | SE | df | t.ratio | p.value |
|---|---|---|---|---|---|---|
| 1 | Small,Low - Large,Low | 1.943548 | 4.590598 | 56.4880 1 | 0.42337 6 | 0.97425 4 |
| 2 | Small,Low - Small,High | 12.26241 | 4.628384 | 56.1244 | 2.64939 3 | 0.04979 7 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 3 | Small,Low - Large,High | 27.10435 | 5.291956 | 61.99041 | 5.121802 | 1.86E-05 |
| 4 | Large,Low - Small,High | 10.31886 | 5.291956 | 61.99041 | 1.949915 | 0.218289 |
| 5 | Large,Low - Large,High | 25.1608 | 4.628384 | 56.1244 | 5.436196 | 7.19E-06 |
| 6 | Small,High - Large,High | 14.84194 | 4.590598 | 56.48801 | 3.233117 | 0.010715 |

Scanning Time

| | contrast | estimate | SE | df | t.ratio | p.value |
|---|---|---|---|---|---|---|
| 1 | Small,Low - Large,Low | 973.7776 | 95.55778 | 61.90166 | 10.19046 | 1.92E-11 |
| 2 | Small,Low - Small,High | 1262.703 | 91.74221 | 61.89255 | 13.76368 | 1.91E-11 |
| 3 | Small,Low - Large,High | 1582.635 | 93.70482 | 61.59199 | 16.88958 | 1.98E-11 |
| 4 | Large,Low - Small,High | 288.9256 | 93.70482 | 61.59199 | 3.083359 | 0.01578 |
| 5 | Large,Low - Large,High | 608.8573 | 91.74221 | 61.89255 | 6.636613 | 5.47E-08 |
| 6 | Small,High - Large,High | 319.9317 | 95.55778 | 61.90166 | 3.348044 | 0.0074 |

Verification Time

| | contrast | estimate | SE | df | t.ratio | p.value |
|---|---|---|---|---|---|---|

| | | | | | | |
|---|---|---|---|---|---|---|
| 1 | Small,Low - Large,Low | 228.2593 | 34.7866 | 52.37722 | 6.561703 | 1.42E-07 |
| 2 | Small,Low - Small,High | 291.1889 | 34.03093 | 53.33864 | 8.556596 | 8.17E-11 |
| 3 | Small,Low - Large,High | 378.8075 | 40.9486 | 61.94042 | 9.250804 | 2.08E-11 |
| 4 | Large,Low - Small,High | 62.92958 | 40.9486 | 61.94042 | 1.536794 | 0.422132 |
| 5 | Large,Low - Large,High | 150.5482 | 34.03093 | 53.33864 | 4.423864 | 0.000275 |
| 6 | Small,High - Large,High | 87.61861 | 34.7866 | 52.37722 | 2.518746 | 0.068596 |

Saccade Amplitude

| | contrast | estimate | SE | df | t.ratio | p.value |
|---|---|---|---|---|---|---|
| 1 | Small,Low - Large,Low | 0.300868 | 0.091642 | 52.9394 | 3.283081 | 0.009571 |
| 2 | Small,Low - Small,High | 0.448481 | 0.082354 | 57.67055 | 5.445789 | 6.52E-06 |
| 3 | Small,Low - Large,High | 0.955184 | 0.101275 | 60.49791 | 9.431628 | 2.14E-11 |
| 4 | Large,Low - Small,High | 0.147613 | 0.101275 | 60.49791 | 1.457548 | 0.469197 |
| 5 | Large,Low - Large,High | 0.654315 | 0.082354 | 57.67055 | 7.945182 | 4.83E-10 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 6 | Small,High - Large,High | 0.50670 3 | 0.09164 2 | 52.9394 | 5.52914 9 | 5.92E-06 |

# Appendix B – Experiment 4 post-hoc pairwise comparisons

Reaction Time

| | contrast | estimate | SE | df | t.ratio | p.value |
|---|---|---|---|---|---|---|
| 1 | Control,low - Central,low | -518.495 | 195.9637 | 132.0569 | -2.64587 | 0.093638 |
| 2 | Control,low - Window,low | -3309.38 | 195.9637 | 132.0569 | -16.8877 | 0 |
| 3 | Control,low - Control,high | 1561.933 | 171.8553 | 104.9271 | 9.088655 | 1.64E-13 |
| 4 | Control,low - Central,high | 1222.267 | 197.3636 | 132.1707 | 6.192968 | 1.04E-07 |
| 5 | Control,low - Window,high | -2720.67 | 197.3636 | 132.1707 | -13.785 | 0 |
| 6 | Central,low - Window,low | -2790.89 | 195.9637 | 132.0569 | -14.2418 | 0 |
| 7 | Central,low - Control,high | 2080.429 | 197.3636 | 132.1707 | 10.54109 | 2.54E-14 |
| 8 | Central,low - Central,high | 1740.762 | 171.8553 | 104.9271 | 10.12923 | 5.72E-14 |
| 9 | Central,low - Window,high | -2202.17 | 197.3636 | 132.1707 | -11.1579 | 8.44E-15 |
| 10 | Window,low - Control,high | 4871.314 | 197.3636 | 132.1707 | 24.68192 | 0 |
| 11 | Window,low - Central,high | 4531.647 | 197.3636 | 132.1707 | 22.9609 | 0 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 12 | Window,low - Window,high | 588.7151 | 171.8553 | 104.9271 | 3.425644 | 0.011037 |
| 13 | Control,high - Central,high | -339.667 | 195.9637 | 132.0569 | -1.73331 | 0.512451 |
| 14 | Control,high - Window,high | -4282.6 | 195.9637 | 132.0569 | -21.854 | 0 |
| 15 | Central,high - Window,high | -3942.93 | 195.9637 | 132.0569 | -20.1207 | 0 |

Timeout

| | contrast | estimate | SE | df | t.ratio | p.value |
|---|---|---|---|---|---|---|
| 1 | Control,low - Central,low | -2240.69 | 314.5806 | 123.5899 | -7.1228 | 1.15E-09 |
| 2 | Control,low - Window,low | -4371.1 | 314.5806 | 123.5899 | -13.895 | 1.98E-14 |
| 3 | Control,low - Control,high | 2816.01 | 272.9279 | 100.1519 | 10.31778 | 1.07E-13 |
| 4 | Control,low - Central,high | 2330.576 | 332.4903 | 126.0418 | 7.009456 | 1.92E-09 |
| 5 | Control,low - Window,high | -2650.83 | 332.4903 | 126.0418 | -7.97265 | 1.21E-11 |
| 6 | Central,low - Window,low | -2130.4 | 314.5806 | 123.5899 | -6.7722 | 6.83E-09 |
| 7 | Central,low - Control,high | 5056.703 | 332.4903 | 126.0418 | 15.20858 | 0 |
| 8 | Central,low - Central,high | 4571.269 | 272.9279 | 100.1519 | 16.749 | 5.86E-14 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 9 | Central,low - Window,high | -410.134 | 332.4903 | 126.0418 | -1.23352 | 0.819589 |
| 10 | Window,low - Control,high | 7187.106 | 332.4903 | 126.0418 | 21.61599 | 0 |
| 11 | Window,low - Central,high | 6701.672 | 332.4903 | 126.0418 | 20.156 | 0 |
| 12 | Window,low - Window,high | 1720.268 | 272.9279 | 100.1519 | 6.303014 | 1.18E-07 |
| 13 | Control,high - Central,high | -485.434 | 314.5806 | 123.5899 | -1.54312 | 0.637213 |
| 14 | Control,high - Window,high | -5466.84 | 314.5806 | 123.5899 | -17.3782 | 1.98E-14 |
| 15 | Central,high - Window,high | -4981.4 | 314.5806 | 123.5899 | -15.8351 | 1.98E-14 |

Total number of fixations

| | contrast | estimate | SE | df | t.ratio | p.value |
|---|---|---|---|---|---|---|
| 1 | Control,low - Central,low | -1.53241 | 0.760458 | 128.8805 | -2.01512 | 0.339544 |
| 2 | Control,low - Window,low | -12.3317 | 0.760458 | 128.8805 | -16.2161 | 5.28E-14 |
| 3 | Control,low - Control,high | 5.730937 | 0.65328 | 104.6135 | 8.772554 | 6.22E-13 |
| 4 | Control,low - Central,high | 4.924635 | 0.772424 | 129.5799 | 6.375561 | 4.39E-08 |
| 5 | Control,low - Window,high | -10.1688 | 0.772424 | 129.5799 | -13.1648 | 7.66E-15 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 6 | Central,low - Window,low | -10.7993 | 0.760458 | 128.8805 | -14.201 | 5.28E-14 |
| 7 | Central,low - Control,high | 7.263351 | 0.772424 | 129.5799 | 9.403324 | 6.82E-14 |
| 8 | Central,low - Central,high | 6.457049 | 0.65328 | 104.6135 | 9.88404 | 9.50E-14 |
| 9 | Central,low - Window,high | -8.63642 | 0.772424 | 129.5799 | -11.1809 | 3.44E-14 |
| 10 | Window,low - Control,high | 18.06262 | 0.772424 | 129.5799 | 23.38434 | 7.33E-15 |
| 11 | Window,low - Central,high | 17.25631 | 0.772424 | 129.5799 | 22.34048 | 7.33E-15 |
| 12 | Window,low - Window,high | 2.162842 | 0.65328 | 104.6135 | 3.31074 | 0.015724 |
| 13 | Control,high - Central,high | -0.8063 | 0.760458 | 128.8805 | -1.06028 | 0.896063 |
| 14 | Control,high - Window,high | -15.8998 | 0.760458 | 128.8805 | -20.9082 | 5.28E-14 |
| 15 | Central,high - Window,high | -15.0935 | 0.760458 | 128.8805 | -19.8479 | 5.28E-14 |

Saccade Amplitude

| | contrast | estimate | SE | df | t.ratio | p.value |
|---|---|---|---|---|---|---|
| 1 | Control,low - Central,low | -1.73257 | 0.15623 | 108.6358 | -11.0898 | 3.45E-14 |
| 2 | Control,low - Window,low | 2.452518 | 0.15623 | 108.6358 | 15.69812 | 3.44E-15 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 3 | Control,low - Control,high | 0.485486 | 0.105706 | 104.976 | 4.59281 | 0.000174 |
| 4 | Control,low - Central,high | -1.14398 | 0.155847 | 108.0737 | -7.34039 | 6.28E-10 |
| 5 | Control,low - Window,high | 2.490575 | 0.155847 | 108.0737 | 15.98085 | 1.41E-14 |
| 6 | Central,low - Window,low | 4.185083 | 0.15623 | 108.6358 | 26.78796 | 3.44E-15 |
| 7 | Central,low - Control,high | 2.218051 | 0.155847 | 108.0737 | 14.23219 | 1.41E-14 |
| 8 | Central,low - Central,high | 0.588584 | 0.105706 | 104.976 | 5.568139 | 2.94E-06 |
| 9 | Central,low - Window,high | 4.22314 | 0.155847 | 108.0737 | 27.0979 | 1.41E-14 |
| 10 | Window,low - Control,high | -1.96703 | 0.155847 | 108.0737 | -12.6215 | 1.61E-14 |
| 11 | Window,low - Central,high | -3.5965 | 0.155847 | 108.0737 | -23.077 | 1.41E-14 |
| 12 | Window,low - Window,high | 0.038056 | 0.105706 | 104.976 | 0.360021 | 0.999186 |
| 13 | Control,high - Central,high | -1.62947 | 0.15623 | 108.6358 | -10.4299 | 5.23E-14 |
| 14 | Control,high - Window,high | 2.005089 | 0.15623 | 108.6358 | 12.83421 | 5.22E-15 |
| 15 | Central,high - Window,high | 3.634556 | 0.15623 | 108.6358 | 23.26413 | 3.44E-15 |

Fixation Duration

| | contrast | estimate | SE | df | t.ratio | p.value |
|---|---|---|---|---|---|---|
| 1 | Control,low - Central,low | -4.16547 | 5.101888 | 92.9241 | -0.81646 | 0.963819 |
| 2 | Control,low - Window,low | -19.5562 | 5.101888 | 92.9241 | -3.83313 | 0.003064 |
| 3 | Control,low - Control,high | 27.55031 | 2.870366 | 103.3488 | 9.598185 | 1.10E-13 |
| 4 | Control,low - Central,high | 4.447106 | 5.173554 | 96.65585 | 0.859584 | 0.95506 |
| 5 | Control,low - Window,high | -18.9608 | 5.173554 | 96.65585 | -3.66495 | 0.005274 |
| 6 | Central,low - Window,low | -15.3907 | 5.101888 | 92.9241 | -3.01667 | 0.037591 |
| 7 | Central,low - Control,high | 31.71578 | 5.173554 | 96.65585 | 6.130367 | 2.83E-07 |
| 8 | Central,low - Central,high | 8.612578 | 2.870366 | 103.3488 | 3.000515 | 0.038554 |
| 9 | Central,low - Window,high | -14.7953 | 5.173554 | 96.65585 | -2.8598 | 0.056651 |
| 10 | Window,low - Control,high | 47.10652 | 5.173554 | 96.65585 | 9.105253 | 3.96E-10 |
| 11 | Window,low - Central,high | 24.00332 | 5.173554 | 96.65585 | 4.639619 | 0.000157 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 12 | Window,low - Window,high | 0.595388 | 2.870366 | 103.3488 | 0.207426 | 0.999946 |
| 13 | Control,high - Central,high | -23.1032 | 5.101888 | 92.9241 | -4.52836 | 0.000249 |
| 14 | Control,high - Window,high | -46.5111 | 5.101888 | 92.9241 | -9.11645 | 4.90E-10 |
| 15 | Central,high - Window,high | -23.4079 | 5.101888 | 92.9241 | -4.58809 | 0.000198 |

Initiation Time

| | contrast | estimate | SE | df | t.ratio | p.value |
|---|---|---|---|---|---|---|
| 1 | Control,low - Central,low | -17.244 | 10.27429 | 106.3414 | -1.67836 | 0.549022 |
| 2 | Control,low - Window,low | -77.6228 | 10.27429 | 106.3414 | -7.55505 | 2.30E-10 |
| 3 | Control,low - Control,high | 13.38449 | 6.707557 | 104.8554 | 1.995434 | 0.35195 |
| 4 | Control,low - Central,high | -5.13237 | 10.21663 | 104.9753 | -0.50235 | 0.995987 |
| 5 | Control,low - Window,high | -70.9041 | 10.21663 | 104.9753 | -6.94006 | 5.01E-09 |
| 6 | Central,low - Window,low | -60.3788 | 10.27429 | 106.3414 | -5.87669 | 7.21E-07 |
| 7 | Central,low - Control,high | 30.6285 | 10.21663 | 104.9753 | 2.997906 | 0.03873 |
| 8 | Central,low - Central,high | 12.11164 | 6.707557 | 104.8554 | 1.805671 | 0.466637 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 9 | Central,low - Window,high | -53.6601 | 10.21663 | 104.9753 | -5.25223 | 1.16E-05 |
| 10 | Window,low - Control,high | 91.00726 | 10.21663 | 104.9753 | 8.907757 | 2.99E-13 |
| 11 | Window,low - Central,high | 72.49041 | 10.21663 | 104.9753 | 7.095334 | 2.35E-09 |
| 12 | Window,low - Window,high | 6.718706 | 6.707557 | 104.8554 | 1.001662 | 0.916424 |
| 13 | Control,high - Central,high | -18.5169 | 10.27429 | 106.3414 | -1.80225 | 0.468748 |
| 14 | Control,high - Window,high | -84.2886 | 10.27429 | 106.3414 | -8.20383 | 8.73E-12 |
| 15 | Central,high - Window,high | -65.7717 | 10.27429 | 106.3414 | -6.40158 | 6.33E-08 |

Scanning Time

| | contrast | estimate | SE | df | t.ratio | p.value |
|---|---|---|---|---|---|---|
| 1 | Control,low - Central,low | 240.6393 | 166.7072 | 130.5344 | 1.443485 | 0.70043 |
| 2 | Control,low - Window,low | -3273.19 | 166.7072 | 130.5344 | -19.6344 | 0 |
| 3 | Control,low - Control,high | 1333.5 | 145.2567 | 104.7049 | 9.180298 | 1.36E-13 |
| 4 | Control,low - Central,high | 1278.94 | 169.066 | 130.906 | 7.56474 | 9.13E-11 |
| 5 | Control,low - Window,high | -2833.95 | 169.066 | 130.906 | -16.7624 | 0 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 6 | Central,low - Window,low | -3513.83 | 166.7072 | 130.5344 | -21.0779 | 0 |
| 7 | Central,low - Control,high | 1092.861 | 169.066 | 130.906 | 6.464109 | 2.76E-08 |
| 8 | Central,low - Central,high | 1038.301 | 145.2567 | 104.7049 | 7.148038 | 1.84E-09 |
| 9 | Central,low - Window,high | -3074.59 | 169.066 | 130.906 | -18.1857 | 0 |
| 10 | Window,low - Control,high | 4606.692 | 169.066 | 130.906 | 27.2479 | 0 |
| 11 | Window,low - Central,high | 4552.132 | 169.066 | 130.906 | 26.92518 | 0 |
| 12 | Window,low - Window,high | 439.2454 | 145.2567 | 104.7049 | 3.023924 | 0.03607 |
| 13 | Control,high - Central,high | -54.5602 | 166.7072 | 130.5344 | -0.32728 | 0.999491 |
| 14 | Control,high - Window,high | -4167.45 | 166.7072 | 130.5344 | -24.9986 | 0 |
| 15 | Central,high - Window,high | -4112.89 | 166.7072 | 130.5344 | -24.6713 | 0 |

Verification Time

| | contrast | estimate | SE | df | t.ratio | p.value |
|---|---|---|---|---|---|---|
| 1 | Control,low - Central,low | -741.891 | 106.7728 | 129.5415 | -6.94831 | 2.42E-09 |
| 2 | Control,low - Window,low | 41.43446 | 106.7728 | 129.5415 | 0.388062 | 0.998837 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 3 | Control,low - Control,high | 215.0488 | 90.4381 | 104.999 | 2.377856 | 0.1736 |
| 4 | Control,low - Central,high | -51.5407 | 106.8567 | 129.5881 | -0.48233 | 0.996707 |
| 5 | Control,low - Window,high | 184.1855 | 106.8567 | 129.5881 | 1.723668 | 0.518804 |
| 6 | Central,low - Window,low | 783.325 | 106.7728 | 129.5415 | 7.336373 | 3.20E-10 |
| 7 | Central,low - Control,high | 956.9393 | 106.8567 | 129.5881 | 8.955351 | 1.08E-13 |
| 8 | Central,low - Central,high | 690.3498 | 90.4381 | 104.999 | 7.633396 | 1.65E-10 |
| 9 | Central,low - Window,high | 926.076 | 106.8567 | 129.5881 | 8.666522 | 2.97E-13 |
| 10 | Window,low - Control,high | 173.6143 | 106.8567 | 129.5881 | 1.62474 | 0.583806 |
| 11 | Window,low - Central,high | -92.9752 | 106.8567 | 129.5881 | -0.87009 | 0.95295 |
| 12 | Window,low - Window,high | 142.751 | 90.4381 | 104.999 | 1.578439 | 0.614477 |
| 13 | Control,high - Central,high | -266.59 | 106.7728 | 129.5415 | -2.49679 | 0.13268 |
| 14 | Control,high - Window,high | -30.8633 | 106.7728 | 129.5415 | -0.28906 | 0.999723 |
| 15 | Central,high - Window,high | 235.7262 | 106.7728 | 129.5415 | 2.207737 | 0.241369 |

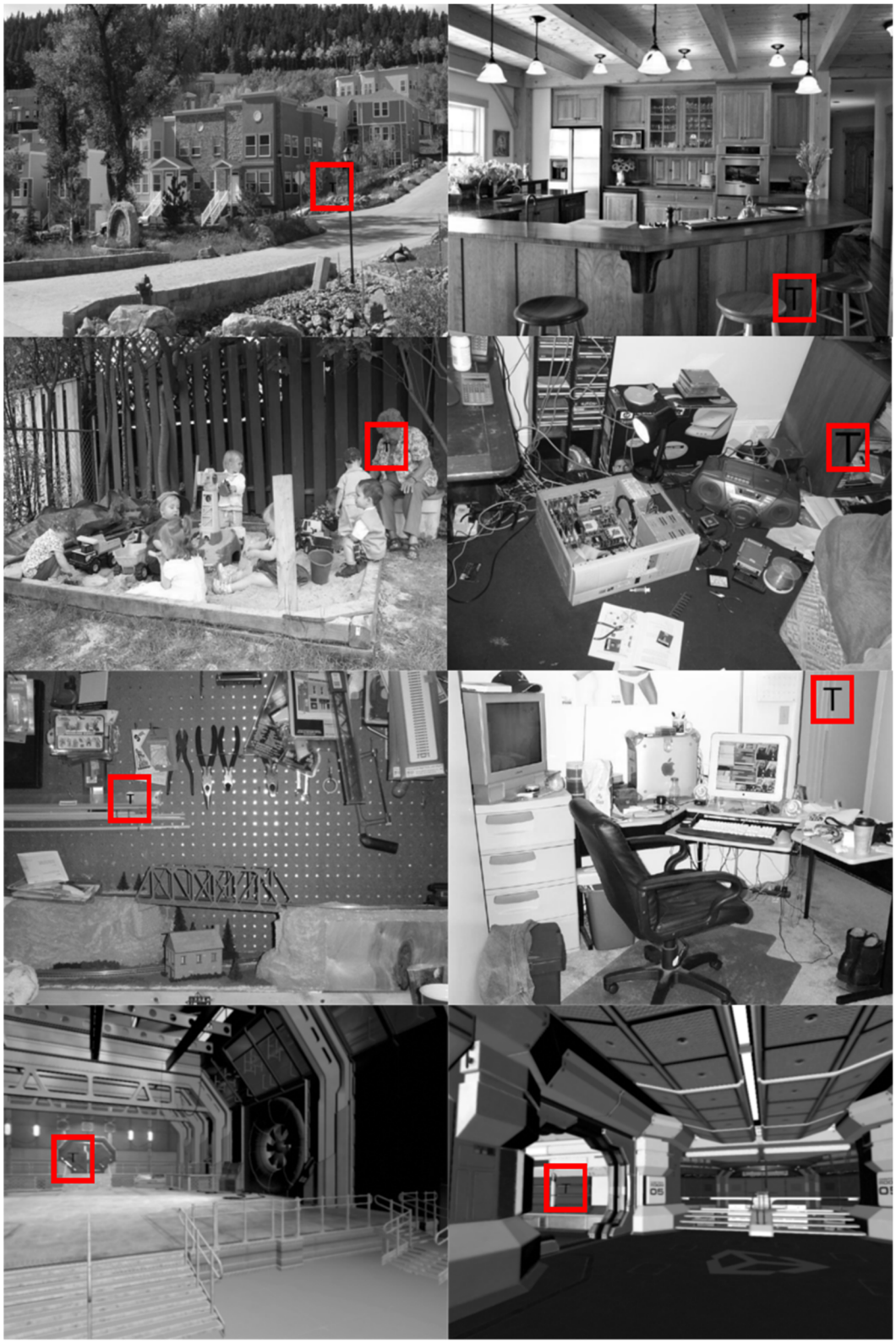
# Appendix C – URLs for all stimuli

Experiments 5, 6 and 7 stimuli

https://mega.nz/#F!sYkzVKqA!Q6fwNaGdcOE0dWFZYnGBFQ

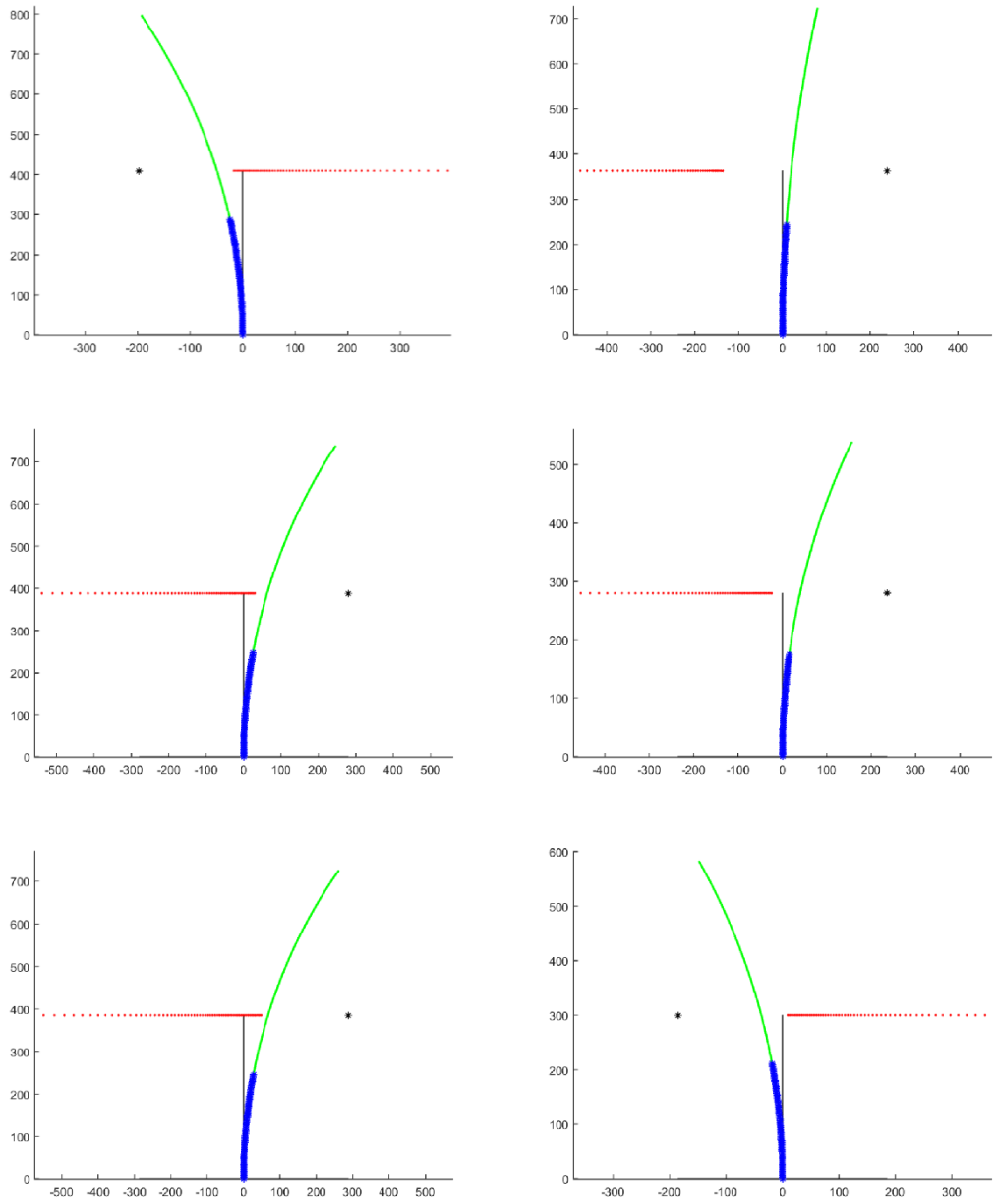
Experiments 1-4 stimuli

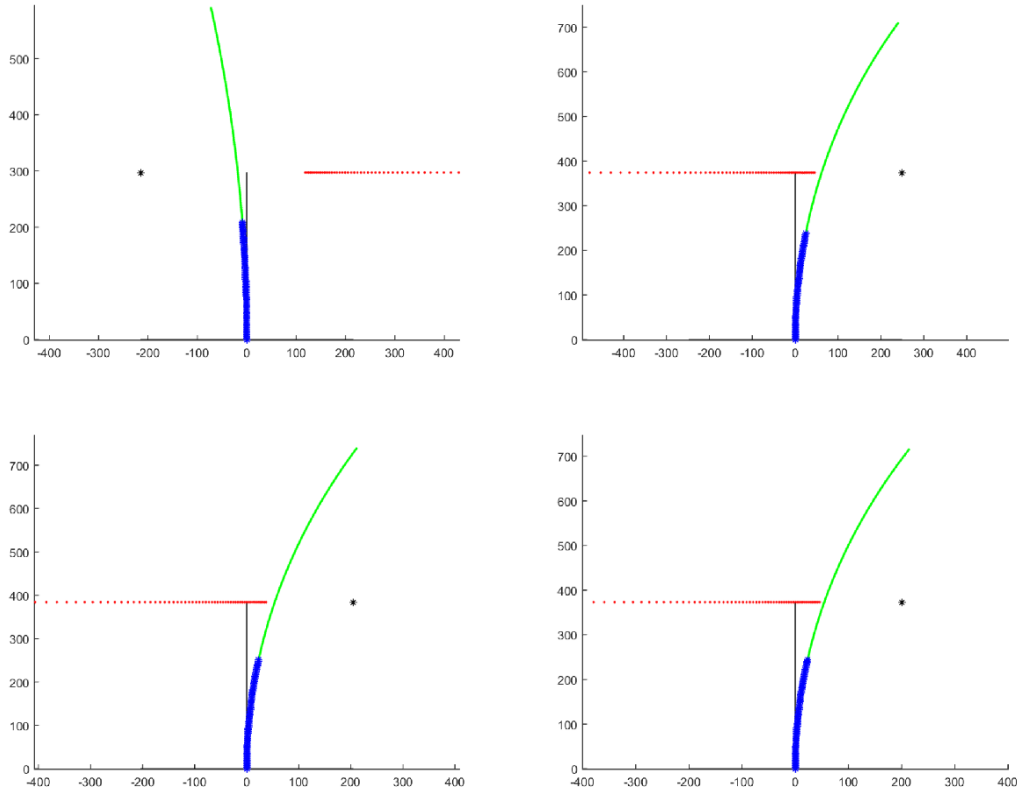
https://mega.nz/#F!cdFkBQIQ!5wL24yffNrvF742VFOPXkw

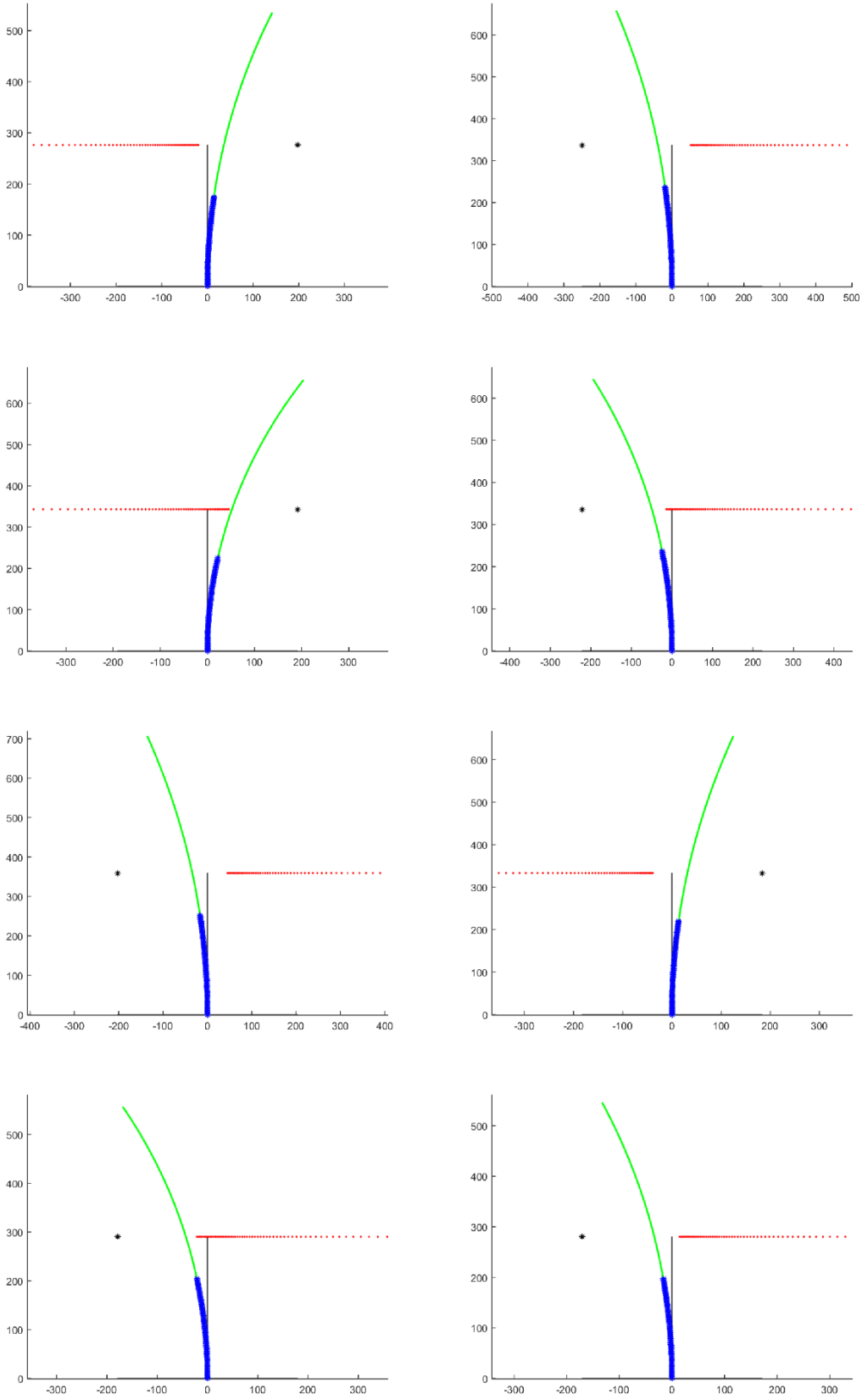


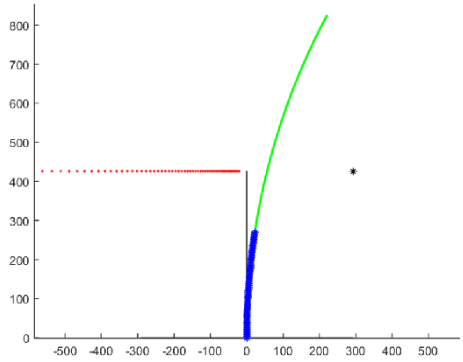
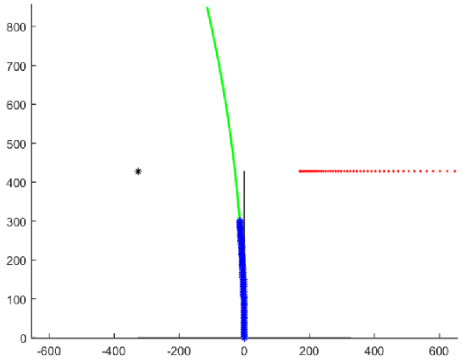
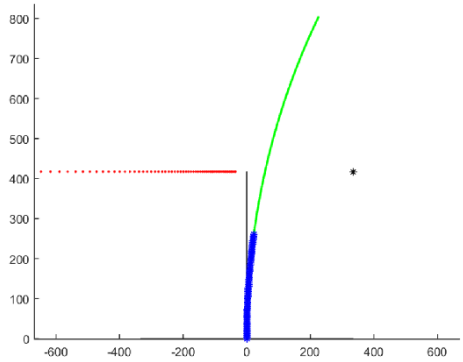
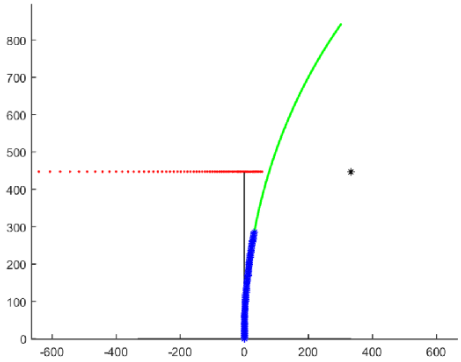
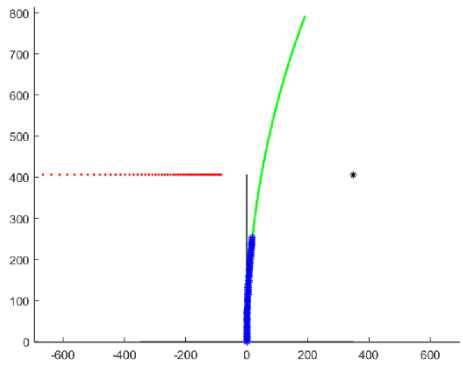
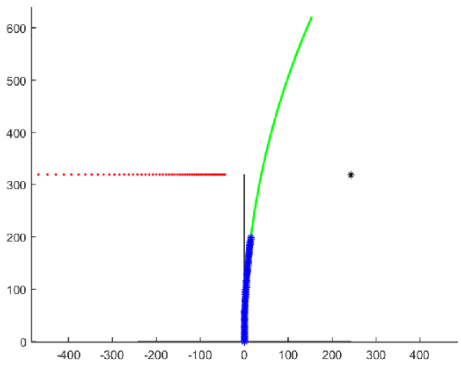
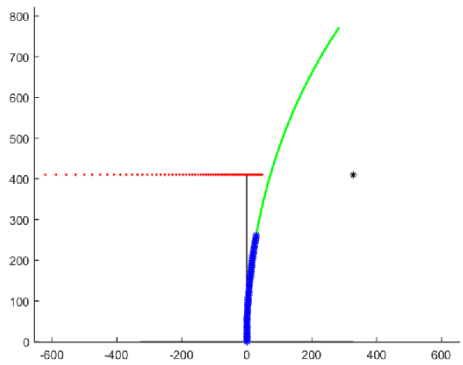
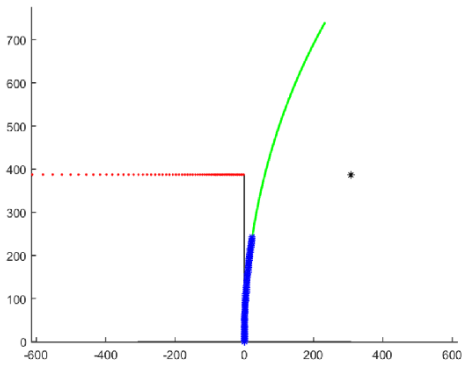
Examples of stimuli ranging from 2D to 3D. Rows are in order from lower quartile, to median, to upper quartile, to 3D.

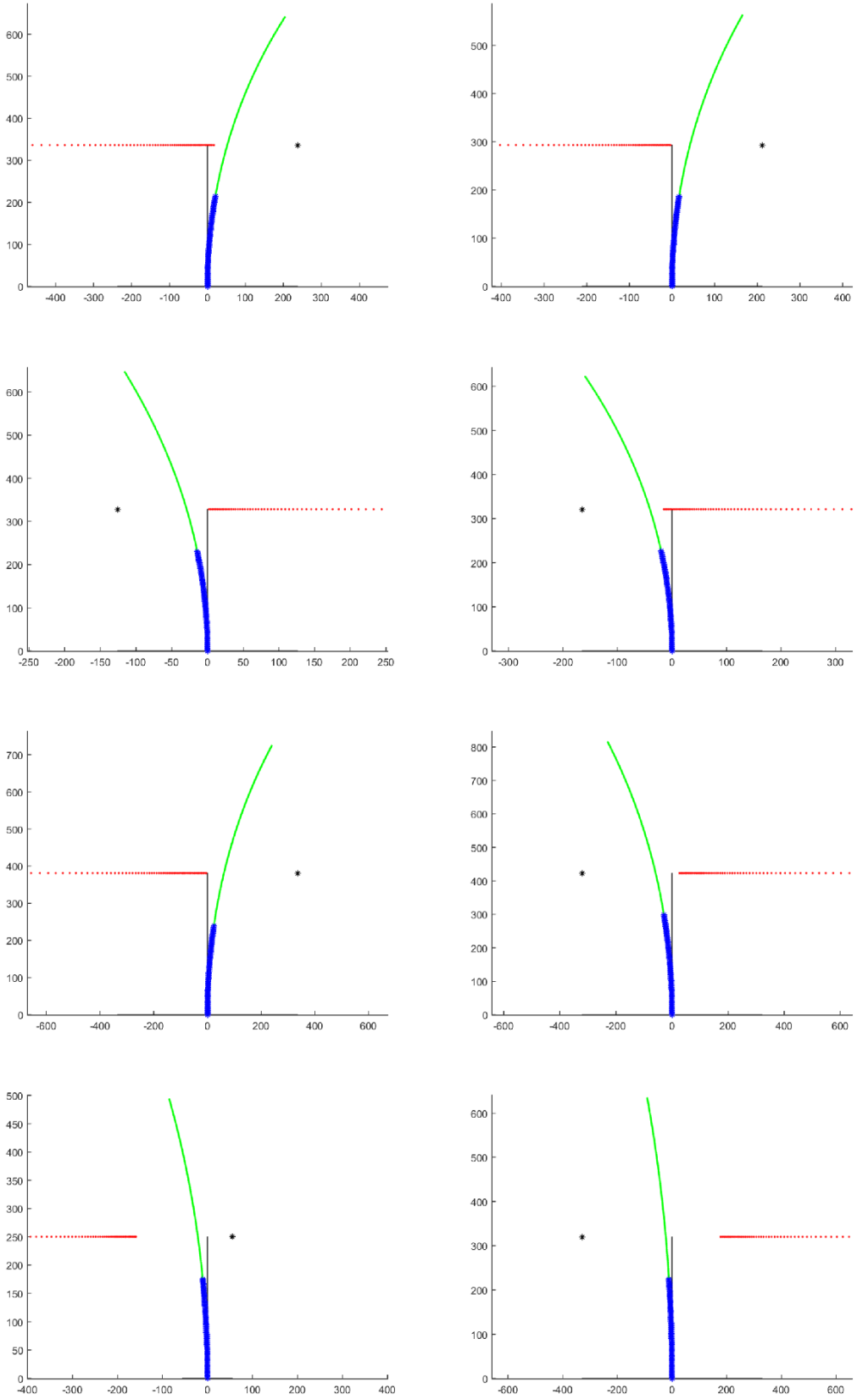
# Appendix D - Calculated target positions along a trajectory for Experiment 7

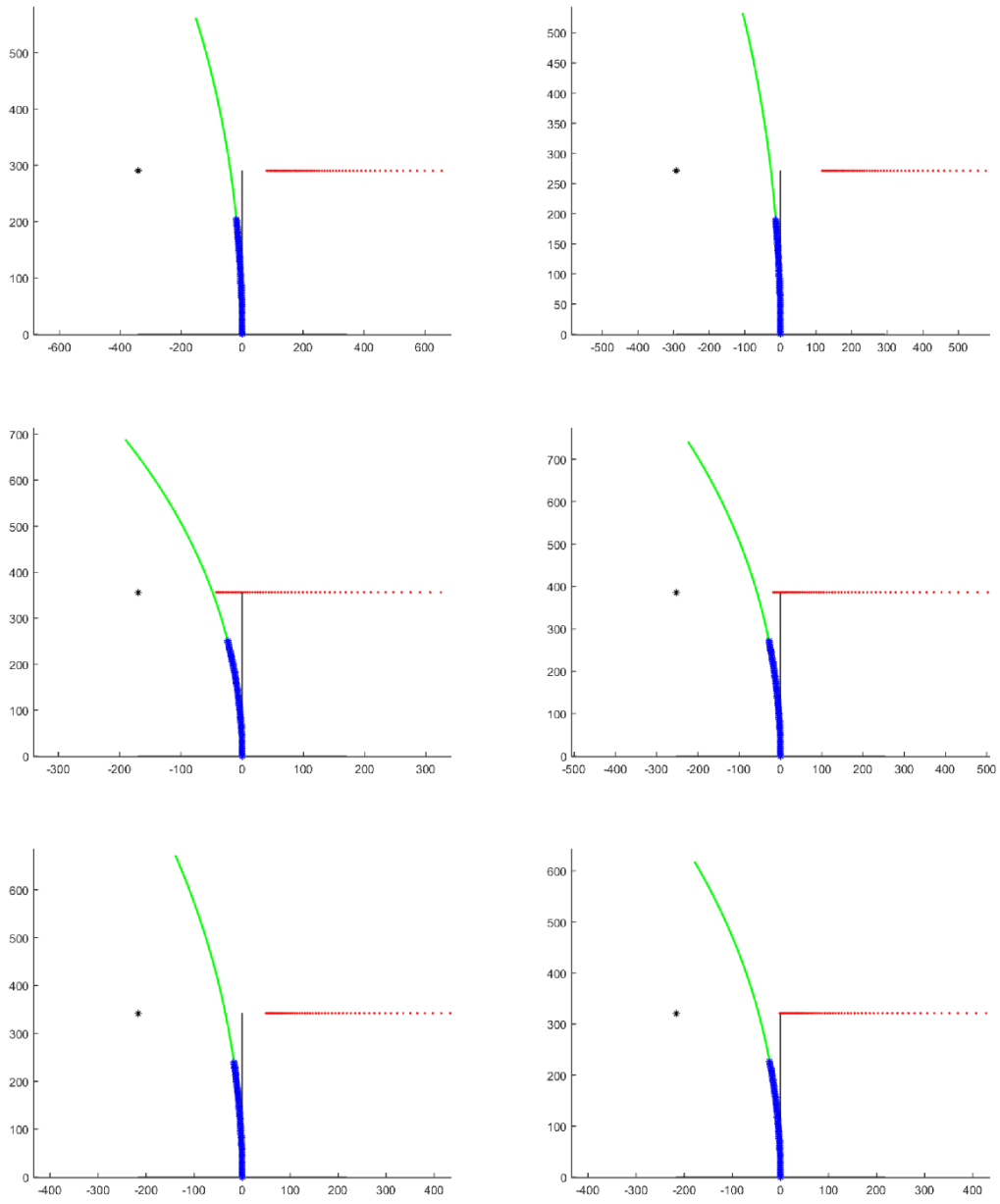



500
400
300
200
100
0
-400 -300 -200 -100 0 100 200 300 400
700
600
500
400
300
200
100
0
-400 -300 -200 -100 0 100 200 300 400
700
600
500
400
300
200
100
0
-400 -300 -200 -100 0 100 200 300 400
700
600
500
400
300
200
100
0
-400 -300 -200 -100 0 100 200 300 400






500
400
300
200
100
0
-600 -400 -200 0 200 400 600
500
450
400
350
300
250
200
150
100
50
0
-500 -400 -300 -200 -100 0 100 200 300 400 500
700
600
500
400
300
200
100
0
-300 -200 -100 0 100 200 300
700
600
500
400
300
200
100
0
-500 -400 -300 -200 -100 0 100 200 300 400 500
600
500
400
300
200
100
0
-400 -300 -200 -100 0 100 200 300 400
600
500
400
300
200
100
0
-400 -300 -200 -100 0 100 200 300 400

# APPENDIX E – T.E.A. ALGORITHM

```
function varargout = teaAlgorithm(varargin)
% TEAALGORITHM MATLAB code for teaAlgorithm.fig
%      TEAALGORITHM, by itself, creates a new TEAALGORITHM or raises
the existing
%      singleton*.
%
%      H = TEAALGORITHM returns the handle to a new TEAALGORITHM or
the handle to
%      the existing singleton*.
%
%      TEAALGORITHM('CALLBACK',hObject,eventData,handles,...) calls
the local
%      function named CALLBACK in TEAALGORITHM.M with the given
input arguments.
%
%      TEAALGORITHM('Property','Value',...) creates a new
TEAALGORITHM or raises the
%      existing singleton*.  Starting from the left, property value
pairs are
%      applied to the GUI before teaAlgorithm_OpeningFcn gets
called.  An
%      unrecognized property name or invalid value makes property
application
%      stop.  All inputs are passed to teaAlgorithm_OpeningFcn via
varargin.
%
%      *See GUI Options on GUIDE's Tools menu.  Choose "GUI allows
only one
%      instance to run (singleton)".
%
% See also: GUIDE, GUIDATA, GUIHANDLES

% Edit the above text to modify the response to help teaAlgorithm

% Last Modified by GUIDE v2.5 23-Apr-2017 00:00:29

% Begin initialization code - DO NOT EDIT
gui_Singleton = 1;
gui_State = struct('gui_Name',       mfilename, ...
                   'gui_Singleton',  gui_Singleton, ...
                   'gui_OpeningFcn', @teaAlgorithm_OpeningFcn, ...
                   'gui_OutputFcn',  @teaAlgorithm_OutputFcn, ...
                   'gui_LayoutFcn',  [] , ...
                   'gui_Callback',   []);
if nargin && ischar(varargin{1})
    gui_State.gui_Callback = str2func(varargin{1});
end

if nargout
    [varargout{1:nargout}] = gui_mainfcn(gui_State, varargin{:});
```

```
else
    gui_mainfcn(gui_State, varargin{:});
end
% End initialization code - DO NOT EDIT


% --- Executes just before teaAlgorithm is made visible.
function teaAlgorithm_OpeningFcn(hObject, eventdata, handles,
varargin)
% This function has no output args, see OutputFcn.
% hObject    handle to figure
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)
% varargin   command line arguments to teaAlgorithm (see VARARGIN)

% Choose default command line output for teaAlgorithm
handles.output = hObject;

% Update handles structure
guidata(hObject, handles);

% UIWAIT makes teaAlgorithm wait for user response (see UIRESUME)
% uiwait(handles.figure1);

axes(handles.axes1)
matlabImage = imread('TEALOGO.png');
image(matlabImage)
axis off
axis image


% --- Outputs from this function are returned to the command line.
function varargout = teaAlgorithm_OutputFcn(hObject, eventdata,
handles)
% varargout  cell array for returning output args (see VARARGOUT);
% hObject    handle to figure
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)

% Get default command line output from handles structure
varargout{1} = handles.output;


% --- Executes on button press in imagePath.
function imagePath_Callback(hObject, eventdata, handles)
% hObject    handle to imagePath (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)
global PROJECTPATH
PROJECTPATH = uigetdir();
addpath(PROJECTPATH)
set(handles.imgpath, 'String', PROJECTPATH);
```

```
% --- Executes on button press in exportPath.
function exportPath_Callback(hObject, eventdata, handles)
% hObject    handle to exportPath (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)
global EXPPATH
EXPPATH = uigetdir();
addpath(EXPPATH)
set(handles.exppath, 'String', EXPPATH);


% --- Executes on button press in addFunctionPath.
function addFunctionPath_Callback(hObject, eventdata, handles)
% hObject    handle to addFunctionPath (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)
global FUNCTIONPATH
FUNCTIONPATH = uigetdir();
addpath(FUNCTIONPATH)
set(handles.functionpaths, 'String', FUNCTIONPATH);

% --- Executes on button press in tCheck.
function tCheck_Callback(hObject, eventdata, handles)
% hObject    handle to tCheck (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)

% Hint: get(hObject,'Value') returns toggle state of tCheck
set(handles.tCheck,'value',1)
set(handles.lCheck,'value',0)
global LETTER
LETTER = 1;

% --- Executes on button press in lCheck.
function lCheck_Callback(hObject, eventdata, handles)
% hObject    handle to lCheck (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)

% Hint: get(hObject,'Value') returns toggle state of lCheck
set(handles.lCheck,'value',1)
set(handles.tCheck,'value',0)
global LETTER
LETTER = 2;


function width_Callback(hObject, eventdata, handles)
% hObject    handle to width (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)

% Hints: get(hObject,'String') returns contents of width as text
%        str2double(get(hObject,'String')) returns contents of width
as a double
```

```
width = str2double(get(hObject,'String'));
set(handles.widthValue, 'String', width);
global WIDTH
WIDTH = width;

% --- Executes during object creation, after setting all properties.
function width_CreateFcn(hObject, eventdata, handles)
% hObject    handle to width (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    empty - handles not created until after all CreateFcns
called

% Hint: edit controls usually have a white background on Windows.
%       See ISPC and COMPUTER.
if ispc && isequal(get(hObject,'BackgroundColor'),
get(0,'defaultUicontrolBackgroundColor'))
    set(hObject,'BackgroundColor','white');
end

function height_Callback(hObject, eventdata, handles)
% hObject    handle to height (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)

% Hints: get(hObject,'String') returns contents of height as text
%        str2double(get(hObject,'String')) returns contents of
height as a double
height = str2double(get(hObject,'String'));
set(handles.heightValue, 'String', height);
global HEIGHT
HEIGHT = height;

% --- Executes during object creation, after setting all properties.
function height_CreateFcn(hObject, eventdata, handles)
% hObject    handle to height (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    empty - handles not created until after all CreateFcns
called

% Hint: edit controls usually have a white background on Windows.
%       See ISPC and COMPUTER.
if ispc && isequal(get(hObject,'BackgroundColor'),
get(0,'defaultUicontrolBackgroundColor'))
    set(hObject,'BackgroundColor','white');
end



% --- Executes on button press in pracTrial.
function pracTrial_Callback(hObject, eventdata, handles)
% hObject    handle to pracTrial (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)

% Hint: get(hObject,'Value') returns toggle state of pracTrial
```

```
set(handles.pracTrial,'value',1)
set(handles.mainTrial,'value',0)
global MAP
MAP = 1;

% --- Executes on button press in mainTrial.
function mainTrial_Callback(hObject, eventdata, handles)
% hObject    handle to mainTrial (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)

% Hint: get(hObject,'Value') returns toggle state of mainTrial
set(handles.mainTrial,'value',1)
set(handles.pracTrial,'value',0)
global MAP
MAP = 2;

% --- Executes on button press in readImages.
function readImages_Callback(hObject, eventdata, handles)
% hObject    handle to readImages (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)
global PROJECTPATH EXPPATH WIDTH HEIGHT imageList
contrastMatrixFolder suffix MAP LETTER nrImages imageFolderName
if ispc %MATLAB function 'ispc' to check if the current computer is
windows
    str = '\';
elseif ismac %MATLAB function 'ismac' to check if the current
computer is a mac
    str = '/';
end

TorL = {'T' 'L'};
myLet = TorL{LETTER};
region_buffer = 3;
letterMatrix = [1 2 3 4];
mapVector = {'PRACTICEMAPS','EXPMAPS'};
[baseName, folder] = uigetfile([PROJECTPATH str '*.stimuli']);
stimuliFile = fullfile(folder, baseName);
%stimuliFile = uigetfile('*.stimuli'); % SCOT.stimuli
fileNames_sorted = textread(stimuliFile,'%s');
sortedImages = cell(size(fileNames_sorted));
w_pivot = WIDTH; % 800
h_pivot = HEIGHT; % 600
%unusedCheck = 0;
newIndex = 0;
%unusedIndex = 0;
maxIndex = length(fileNames_sorted);
set(handles.imageOutput, 'String', 'Checking image dimensions...');
% remove images with deviating resolution
for resIndex = 1:maxIndex
    currentInfo = imfinfo([fileNames_sorted{resIndex}]);
    if(currentInfo.Width == w_pivot && currentInfo.Height ==
h_pivot)
        newIndex = newIndex + 1;
```

```
        sortedImages(newIndex) = fileNames_sorted(resIndex);
    end
end
if(maxIndex-newIndex ~= maxIndex)
    set(handles.imageOutput, 'String', ['Total Images Usable: '
num2str(newIndex), ' Total Images Unusable: ', num2str(maxIndex-
newIndex)]);
else
    set(handles.imageOutput, 'String', 'None of the images in this
set can be used, please enter a new list and try again');
    return;
end
finalSet = cell(newIndex,1);
finalSet = sortedImages(~cellfun(@isempty,sortedImages));
nrImages = length(finalSet);
imageList = cell(size(nrImages));
for imIndex = 1:nrImages
    imgfile = finalSet{imIndex};
    imageList{imIndex}.image = imread([PROJECTPATH str imgfile]);
    imageList{imIndex}.name = imgfile;
    %fprintf(imageList{imIndex})
end
imageFolderName = [mapVector{MAP} num2str(nrImages)];

contrastMatrixFolder = {'contrast_matrices_without'
'contrast_matrices_with' 'contrast_matrices_diff'};
for i = 1:length(letterMatrix)
    rs = num2str(region_buffer);
    sol = num2str(letterMatrix(i));
    contrastFolderName = ['contrast_region_' rs '_letter' myLet '_'
sol];
    mkdir([EXPPATH str 'TStimuli' str imageFolderName
str],contrastFolderName);
    mkdir([EXPPATH str 'TStimuli' str imageFolderName str
contrastFolderName str],contrastMatrixFolder{1});
    mkdir([EXPPATH str 'TStimuli' str imageFolderName str
contrastFolderName str],contrastMatrixFolder{2});
    mkdir([EXPPATH str 'TStimuli' str imageFolderName str
contrastFolderName str],contrastMatrixFolder{3});
end
% contrast matrices: suffix for file names
suffix = {'_woLet_' '_wLet_' '_diff_'}; % without letter, with
letter, difference map


% --- Executes on button press in createMaps.
function createMaps_Callback(hObject, eventdata, handles)
% hObject    handle to createMaps (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)
global EXPPATH imageList contrastMatrixFolder suffix LETTER nrImages
imageFolderName
if ispc %MATLAB function 'ispc' to check if the current computer is
windows
    str = '\';
```

```
elseif ismac %MATLAB function 'ismac' to check if the current
computer is a mac
    str = '/';
end
currentImage = imageList{1}.image;
res_x = size(currentImage,2);
res_y = size(currentImage,1);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%
% 7. Loop across images
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%
TorL = {'T' 'L'};
myLet = TorL{LETTER};
region_buffer = 3;
letterMatrix = [1 2 3 4];
rs = num2str(region_buffer);
noOfLetters = length(letterMatrix);
test = 0; % set test to 1 if you want to show letter running through
image
num_rows = size(currentImage,1);
num_cols = size(currentImage,2);
%thickness of the letter
horizontalWidth = [1 2 2 4]; % 6
verticalWidth = [1 3 3 5]; % 7
%actual width and height
widthVector = [7 13 19 33];
heightVector = [9 16 23 40];
for iIndex = 1:nrImages
    for fontIndex = 1:noOfLetters
        size_of_letter = letterMatrix(fontIndex);
        sol = num2str(size_of_letter);
        contrastFolderName = ['contrast_region_' rs '_letter' myLet
'_' sol];
        hWidth = horizontalWidth(fontIndex);
        vWidth = verticalWidth(fontIndex);
        myHeight = heightVector(fontIndex);
        myWidth = widthVector(fontIndex);
        mySize = max([myHeight myWidth]);
        if mod(mySize,2) ~= 0 % if odd number
            mySize = mySize + 1; % add 1 to make it even number
        end
        centraliseLetter = floor(hWidth/2);
        region_size = mySize + 2*region_buffer;
        colStart = 1 + mySize/2+region_buffer;
        rowStart = 1 + mySize/2+region_buffer;
        colabsEnd = num_cols - region_size/2;
        rowabsEnd = num_rows - region_size/2;
        % initialize output matrices for contrast values
        contrast_map_withLet = ones(res_y,res_x)*NaN;
        contrast_map_withoutLet = ones(res_y,res_x)*NaN;

        currentImage =
imageList{iIndex}.image;%rgb2gray(imageList{iIndex}.image); % get
current image
```

```
        currentName  = imageList{iIndex}.name; % get current image
name
        meanCimg     = mean(mean(currentImage));
        set(handles.mapProgress, 'String', ['processing image
',num2str(iIndex),currentName]);
        drawnow
        % move letter pixelwise and calculate local contrast
        for rowIndex = rowStart:rowabsEnd % rowStart:num_rows % row
            if mod(rowIndex,res_y/10) == 0
                set(handles.mapProgress, 'String',
[num2str(rowIndex/(res_y/10)*10) '% done.']);
                drawnow
            end
            for colIndex = colStart:colabsEnd
                if(myLet == 'L')
                    newImage = currentImage;
                    newImage(rowIndex-floor(mySize/2):rowIndex-
floor(mySize/2) + mySize,colIndex-floor(myWidth/2):colIndex-
floor(myWidth/2) + (vWidth-1)) = 0;
                    newImage((rowIndex+floor(mySize/2)-(hWidth-
1)):(rowIndex+floor(mySize/2)-(hWidth-1))+(hWidth-1),colIndex-
floor(myWidth/2):colIndex-floor(myWidth/2) + myWidth-1) = 0;
                else
                    newImage = currentImage;
                    newImage(rowIndex-floor(mySize/2):rowIndex-
floor(mySize/2) + mySize,colIndex-centraliseLetter:colIndex-
centraliseLetter + (vWidth-1)) = 0;
                    newImage((rowIndex-floor(mySize/2)):(rowIndex-
floor(mySize/2)) + (hWidth-1),colIndex-floor(myWidth/2):colIndex-
floor(myWidth/2) + myWidth-1) = 0;
                end
                % cut out evaluation region = box around the letter
                newSection = newImage(rowIndex-
region_size/2:rowIndex+region_size/2,colIndex-
region_size/2:colIndex+region_size/2);
                % compute local contrast for box *with* letter
                vec =
reshape(newSection,size(newSection,1)*size(newSection,2),1);
                dev = std(double(vec));
                lcontrast    = dev/mean(mean(newImage));
                contrast_map_withLet(rowIndex,colIndex) = lcontrast;
                % compute local contrast for box *without* letter
                oldSection = currentImage(rowIndex-
region_size/2:rowIndex+region_size/2,colIndex-
region_size/2:colIndex+region_size/2);
                vec2 =
reshape(oldSection,size(oldSection,1)*size(oldSection,2),1);
                dev2 = std(double(vec2));
                lcontrast2   = dev2/meanCimg;
                contrast_map_withoutLet(rowIndex,colIndex) =
lcontrast2;
                % visualization
                if test == 1
                    imshow(newSection);
                    %imshow(newImage);
                end
```

```
            end
        end
        % save and plot contrast matrices
        contrast_map_difference = contrast_map_withLet -
contrast_map_withoutLet;
        for i=1:3
            % save matrix
            fout = fopen([EXPPATH str 'TStimuli' str imageFolderName
str contrastFolderName str contrastMatrixFolder{i} str currentName
suffix{i} myLet '.dat'],'w');
            switch i
                case 1
                    contrast_map = contrast_map_withoutLet;
                case 2
                    contrast_map = contrast_map_withLet;
                case 3
                    contrast_map = contrast_map_difference;
            end
            fprintf(fout,[repmat('%f\t',1,res_x-1)
'%f\n'],contrast_map');
            fclose(fout);
        end
    end
end


% --- Executes on button press in minVals.
function minVals_Callback(hObject, eventdata, handles)
% hObject    handle to minVals (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)


% --- Executes on button press in lqVals.
function lqVals_Callback(hObject, eventdata, handles)
% hObject    handle to lqVals (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)


% --- Executes on button press in medVals.
function medVals_Callback(hObject, eventdata, handles)
% hObject    handle to medVals (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)


% --- Executes on button press in uqVals.
function uqVals_Callback(hObject, eventdata, handles)
% hObject    handle to uqVals (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)
```

```
% --- Executes on button press in maxVals.
function maxVals_Callback(hObject, eventdata, handles)
% hObject    handle to maxVals (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)


% --- Executes on button press in writeIms.
function writeIms_Callback(hObject, eventdata, handles)
% hObject    handle to writeIms (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)
global selectedValue EXPPATH MAP LETTER PROJECTPATH FUNCTIONPATH
numDats summedMaps
global newMedians newMaxs newLQs newUQs newMins folderName
fid = 0;
addpath(FUNCTIONPATH)
mapVector = {'PRACTICEMAPS','EXPMAPS'};

if ispc %MATLAB function 'ispc' to check if the current computer is
windows
    str = '\';
elseif ismac %MATLAB function 'ismac' to check if the current
computer is a mac
    str = '/';
end

TorL = {'T' 'L'};
cLetter = TorL{LETTER};

switch selectedValue
    case 1
        choiceStr = 'min';
        currentValue = newMins;
        resultFolder = [EXPPATH str 'sizeImages' str mapVector{MAP}
str 'Minresults_' cLetter'];
        mkdir(resultFolder);
    case 2
        choiceStr = 'lowerQ';
        currentValue = newLQs;
        resultFolder = [EXPPATH str 'sizeImages' str mapVector{MAP}
str 'LQresults_' cLetter'];
        mkdir(resultFolder);
    case 3
        choiceStr = 'median';
        currentValue = newMedians;
        resultFolder = [EXPPATH str 'sizeImages' str mapVector{MAP}
str 'Medianresults_' cLetter'];
        mkdir(resultFolder);
    case 4
        choiceStr = 'upperQ';
        currentValue = newUQs;
        resultFolder = [EXPPATH str 'sizeImages' str mapVector{MAP}
str 'UQresults_' cLetter'];
        mkdir(resultFolder);
```

```
    case 5
        choiceStr = 'max';
        currentValue = newMaxs;
        resultFolder = [EXPPATH str 'sizeImages' str mapVector{MAP}
str 'Maxresults_' cLetter'];
        mkdir(resultFolder);
end

final1folder = [EXPPATH str folderName str mapVector{MAP} str
choiceStr '_FS1_' cLetter];
final2folder = [EXPPATH str folderName str mapVector{MAP} str
choiceStr '_FS2_' cLetter];
final3folder = [EXPPATH str folderName str mapVector{MAP} str
choiceStr '_FS3_' cLetter];
final4folder = [EXPPATH str folderName str mapVector{MAP} str
choiceStr '_FS4_' cLetter];

mkdir(final1folder);
mkdir(final2folder);
mkdir(final3folder);
mkdir(final4folder);

filename = dir([PROJECTPATH str '*.jpg']);

%-----------------------------------------------------%
%   1. General settings for the experiment
%       customise accordingly
% Mimics:
% [~, res_x, res_y, DEG_PER_PIX, ~, ~, ~] = corpus_scenes_setup;
%-----------------------------------------------------%

SAMPLING = 1000;    % sampling rate of the eyetracker - EyeLink 2K
binoculuar

% monitor resolution: 800 x 600
% im = imread('blank.jpg');
x_res = 800;
y_res = 600;
% monitor: 21inch (best measure the actual distance yourself)
x_monitor = 40.5; % visible monitor width (cm)
y_monitor = 30.5; % visible monitor height (cm)
% viewing distance: 90cm ?
MonitorSubjectDistance = 90;    % distance between monitor and
subject (cm)

% compute degree-per-pixel value
ang = atan2(1,MonitorSubjectDistance)*180/pi; % how many degrees
correspond to 1 cm on the subject screen?
r = x_res/x_monitor;     % 1280 pixel (x axis) equal 40.5 cm => 'r'
carries the ratio between pixel and cm
DEG_PER_PIX = ang / r;   % degree per pixel

%-----------------------------------------------------%
%       END of general settings customisation
```

```
%-------------------------------------------------%

res_x = x_res;
res_y = y_res;

PIX_PER_DEG = 1/DEG_PER_PIX;

radius_deg = 3;
region_buffer = 3;
radius_pix = radius_deg * PIX_PER_DEG;

N = 256;
t = (0:N)*2*pi/N;
x = radius_pix*cos(t)+res_x/2;
y = radius_pix*sin(t)+res_y/2;
CircleMask = poly2mask(x,y,res_y,res_x);

horizontalWidth = [1 2 2 4]; % 6
verticalWidth = [1 3 3 5]; % 7
widthVector = [7 13 19 33];
heightVector = [9 16 23 40];
total_images = numDats;

letterMatrix = [1 2 3 4];
fontVector = letterMatrix;
for imageIndex = 1:total_images
    for fontIndex = 1:length(fontVector)
        myHeight = heightVector(fontIndex);
        myWidth = widthVector(fontIndex);
        mySize = max([myHeight myWidth]);
        if mod(mySize,2) ~= 0 % if odd number
            mySize = mySize + 1; % add 1 to make it even number
        end
        switch fontIndex
            case 1
                final1Image = imread(filename(imageIndex).name);
                hWidth1 = horizontalWidth(fontIndex);
                vWidth1 = verticalWidth(fontIndex);
                centraliseLetter1 = floor(vWidth1/2);
                myHeight1 = myHeight;
                myWidth1 = myWidth;
                mySize1 = mySize;
                region_size1 = mySize1+2*region_buffer;
            case 2
                final2Image = imread(filename(imageIndex).name);
                hWidth2 = horizontalWidth(fontIndex);
                vWidth2 = verticalWidth(fontIndex);
                centraliseLetter2 = floor(vWidth2/2);
                myHeight2 = myHeight;
                myWidth2 = myWidth;
                mySize2 = mySize;
                region_size2 = mySize2+2*region_buffer;
            case 3
                final3Image = imread(filename(imageIndex).name);
                hWidth3 = horizontalWidth(fontIndex);
```

```
                vWidth3 = verticalWidth(fontIndex);
                centraliseLetter3 = floor(vWidth3/2);
                myHeight3 = myHeight;
                myWidth3 = myWidth;
                mySize3 = mySize;
                region_size3 = mySize3+2*region_buffer;
            case 4
                final4Image = imread(filename(imageIndex).name);
                hWidth4 = horizontalWidth(fontIndex);
                vWidth4 = verticalWidth(fontIndex);
                centraliseLetter4 = floor(vWidth4/2);
                myHeight4 = myHeight;
                myWidth4 = myWidth;
                mySize4 = mySize;
                [fMEX75,fMEY75,~,~,~,~,~,~,~] =
placeWithType(currentValue{imageIndex},myWidth4,myHeight4,summedMaps
{imageIndex},fid,choiceStr,CircleMask,res_x,res_y,mySize4,radius_pix
);
                imageName = filename(imageIndex).name;

                if(cLetter == 'T')
                    final4Image(fMEY75-floor(mySize4/2):fMEY75-
floor(mySize4/2) + mySize4,fMEX75-centraliseLetter4:fMEX75-
centraliseLetter4 + (vWidth4-1)) = 0;
                    final4Image((fMEY75-floor(mySize4/2)):(fMEY75-
floor(mySize4/2)) + (hWidth4-1),fMEX75-floor(myWidth4/2):fMEX75-
floor(myWidth4/2) + myWidth4-1) = 0;
                    final3Image(fMEY75-floor(mySize3/2):fMEY75-
floor(mySize3/2) + mySize3,fMEX75-centraliseLetter3:fMEX75-
centraliseLetter3 + (vWidth3-1)) = 0;
                    final3Image((fMEY75-floor(mySize3/2)):(fMEY75-
floor(mySize3/2)) + (hWidth3-1),fMEX75-floor(myWidth3/2):fMEX75-
floor(myWidth3/2) + myWidth3-1) = 0;
                    final2Image(fMEY75-floor(mySize2/2):fMEY75-
floor(mySize2/2) + mySize2,fMEX75-centraliseLetter2:fMEX75-
centraliseLetter2 + (vWidth2-1)) = 0;
                    final2Image((fMEY75-floor(mySize2/2)):(fMEY75-
floor(mySize2/2)) + (hWidth2-1),fMEX75-floor(myWidth2/2):fMEX75-
floor(myWidth2/2) + myWidth2-1) = 0;
                    final1Image(fMEY75-floor(mySize1/2):fMEY75-
floor(mySize1/2) + mySize1,fMEX75-centraliseLetter1:fMEX75-
centraliseLetter1 + (vWidth1-1)) = 0;
                    final1Image((fMEY75-floor(mySize1/2)):(fMEY75-
floor(mySize1/2)) + (hWidth1-1),fMEX75-floor(myWidth1/2):fMEX75-
floor(myWidth1/2) + myWidth1-1) = 0;
                else
                    final4Image(fMEY75-floor(mySize4/2):fMEY75-
floor(mySize4/2) + mySize4,fMEX75-floor(myWidth4/2):fMEX75-
floor(myWidth4/2) + (vWidth4-1)) = 0;
                    final4Image((fMEY75+floor(mySize4/2)-(hWidth4-
1)):(fMEY75+floor(mySize4/2)-(hWidth4-1))+(hWidth4-1),fMEX75-
floor(myWidth4/2):fMEX75-floor(myWidth4/2) + myWidth4-1) = 0;
                    final3Image(fMEY75-floor(mySize3/2):fMEY75-
floor(mySize3/2) + mySize3,fMEX75-floor(myWidth3/2):fMEX75-
floor(myWidth3/2) + (vWidth3-1)) = 0;
```

```
                    final3Image((fMEY75+floor(mySize3/2)-(hWidth4-
1)):(fMEY75+floor(mySize3/2)-(hWidth3-1))+(hWidth3-1),fMEX75-
floor(myWidth3/2):fMEX75-floor(myWidth3/2) + myWidth3-1) = 0;
                    final2Image(fMEY75-floor(mySize2/2):fMEY75-
floor(mySize2/2) + mySize2,fMEX75-floor(myWidth2/2):fMEX75-
floor(myWidth2/2) + (vWidth2-1)) = 0;
                    final2Image((fMEY75+floor(mySize2/2)-(hWidth2-
1)):(fMEY75+floor(mySize2/2)-(hWidth2-1))+(hWidth2-1),fMEX75-
floor(myWidth2/2):fMEX75-floor(myWidth2/2) + myWidth2-1) = 0;
                    final1Image(fMEY75-floor(mySize1/2):fMEY75-
floor(mySize1/2) + mySize1,fMEX75-floor(myWidth1/2):fMEX75-
floor(myWidth1/2) + (vWidth1-1)) = 0;
                    final1Image((fMEY75+floor(mySize1/2)-(hWidth1-
1)):(fMEY75+floor(mySize1/2)-(hWidth1-1))+(hWidth1-1),fMEX75-
floor(myWidth1/2):fMEX75-floor(myWidth1/2) + myWidth1-1) = 0;
                end
                % Write images if user chose this input
                imwrite(final1Image,[final1folder str
imageName(1:end-4) '.jpg']);
                imwrite(final2Image,[final2folder str
imageName(1:end-4) '.jpg']);
                imwrite(final3Image,[final3folder str
imageName(1:end-4) '.jpg']);
                imwrite(final4Image,[final4folder str
imageName(1:end-4) '.jpg']);
                region_size4 = mySize4+2*region_buffer;
        end
    end
end
fprintf('finishedFunction')
% --- Executes on button press in tlMin.
function tlMin_Callback(hObject, eventdata, handles)
% hObject    handle to tlMin (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)
global numDats summedMaps
newMins = cell(1,numDats);
for sumIndex = 1:numDats
    newMins{sumIndex} = nanmin(nanmin(summedMaps{sumIndex}));
end

% --- Executes on button press in tlLQ.
function tlLQ_Callback(hObject, eventdata, handles)
% hObject    handle to tlLQ (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)
global numDats summedMaps
newLQs = cell(1,numDats);
for sumIndex = 1:numDats
    sortedSums = sort(summedMaps{sumIndex});
    newLQs{sumIndex} =
nanmedian(sortedSums(find(sortedSums<nanmedian(nanmedian(sortedSums)
))));
end
```

```
% --- Executes on button press in tlMed.
function tlMed_Callback(hObject, eventdata, handles)
% hObject    handle to tlMed (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)
global numDats summedMaps
newMedians = cell(1,numDats);
for sumIndex = 1:numDats
    newMedians{sumIndex} =
nanmedian(nanmedian(summedMaps{sumIndex}));
end

% --- Executes on button press in tlUQ.
function tlUQ_Callback(hObject, eventdata, handles)
% hObject    handle to tlUQ (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)
global numDats summedMaps
newUQs = cell(1,numDats);
for sumIndex = 1:numDats
    sortedSumsUP = sort(summedMaps{sumIndex});
    newUQs{sumIndex} =
nanmedian(sortedSumsUP(find(sortedSumsUP>nanmedian(nanmedian(sortedS
umsUP)))));
end

% --- Executes on button press in tlMax.
function tlMax_Callback(hObject, eventdata, handles)
% hObject    handle to tlMax (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)
global numDats summedMaps
newMaxs = cell(1,numDats);
for sumIndex = 1:numDats
    newMaxs{sumIndex} = nanmax(nanmax(summedMaps{sumIndex}));
end

% --- Executes on button press in summedMaps.
function summedMaps_Callback(hObject, eventdata, handles)
% hObject    handle to summedMaps (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)
global EXPPATH imageFolderName LETTER numDats summedMaps
selectedValue folderName
global newMedians newMins newLQs newUQs newMaxs
contrastMatrixFolder = 'contrast_matrices_diff';
if ispc %MATLAB function 'ispc' to check if the current computer is
windows
    str = '\';
elseif ismac %MATLAB function 'ismac' to check if the current
computer is a mac
    str = '/';
end
TorL = {'T' 'L'};
myLet = TorL{LETTER};
```

```
letterMatrix = [1 2 3 4];
folderName = [myLet, '_Maps'];
for i = 1:length(letterMatrix)
    %Generate values for experiment trials
    mapPath{i} = [EXPPATH str 'TStimuli' str imageFolderName str
'contrast_region_3_letter' myLet '_'];
end
newMins                         = cell(1,numDats);
newLQs                          = cell(1,numDats);
newMedians                      = cell(1,numDats);
newUQs                          = cell(1,numDats);
newMaxs                         = cell(1,numDats);
fontVector                      = letterMatrix;
STIMULUS_PATH = [mapPath{1} num2str(fontVector(1)) str
contrastMatrixFolder str];
addpath([mapPath{1} num2str(fontVector(1)) str contrastMatrixFolder
str])
%Obtain new set of .dat files
filename = dir([STIMULUS_PATH '*.dat']);
numDats = length(filename);
%Preallocate memory
summedMaps                      = cell(1,numDats);
diffMatrices = cell(length(fontVector),numDats);

for fontSize = 1:length(fontVector)
    if(fontSize ~= 1)
        %Update path on each iteration to loop through font sizes
        STIMULUS_PATH = [mapPath{fontSize}
num2str(fontVector(fontSize)) str contrastMatrixFolder str];
        addpath([mapPath{fontSize} num2str(fontVector(fontSize)) str
contrastMatrixFolder str]);
        %Obtain new set of .dat files
        filename = dir([STIMULUS_PATH '*.dat']);
    end
    matrixIndex = 1;
    for currentDat = 1:numDats
        diffMatrices{fontSize,matrixIndex} =
load(filename(currentDat).name);
        matrixIndex = matrixIndex+1;
    end
end

for sumIndex = 1:numDats
    currentMap = diffMatrices{1,sumIndex};
    for cFont = 2:length(fontVector)
        currentMap = abs(currentMap + diffMatrices{cFont,sumIndex});
        if(cFont == length(fontVector))
            summedMaps{sumIndex} = currentMap;
        end
    end
end

switch selectedValue
    case 1
        for sumIndex = 1:numDats
```

```
            newMins{sumIndex} =
nanmin(nanmin(summedMaps{sumIndex}));
        end
    case 2
        for sumIndex = 1:numDats
            sortedSums = sort(summedMaps{sumIndex});
            newLQs{sumIndex} =
nanmedian(sortedSums(find(sortedSums<nanmedian(nanmedian(sortedSums)
))));
        end
    case 3
        for sumIndex = 1:numDats
            newMedians{sumIndex} =
nanmedian(nanmedian(summedMaps{sumIndex}));
        end
    case 4
        for sumIndex = 1:numDats
            sortedSumsUP = sort(summedMaps{sumIndex});
            newUQs{sumIndex} =
nanmedian(sortedSumsUP(find(sortedSumsUP>nanmedian(nanmedian(sortedS
umsUP)))));
        end
    case 5
        for sumIndex = 1:numDats
            newMaxs{sumIndex} =
nanmax(nanmax(summedMaps{sumIndex}));
        end
end
fprintf('Done!')
% --- Executes on button press in tlSummedMaps.
function tlSummedMaps_Callback(hObject, eventdata, handles)
% hObject    handle to tlSummedMaps (see GCBO)
% handles    structure with handles and user data (see GUIDATA)
% eventdata  reserved - to be defined in a future version of MATLAB
global EXPPATH imageFolderName LETTER numDats summedMaps
selectedTLValue folderName

contrastMatrixFolder = 'contrast_matrices_diff';
if ispc %MATLAB function 'ispc' to check if the current computer is
windows
    str = '\';
elseif ismac %MATLAB function 'ismac' to check if the current
computer is a mac
    str = '/';
end
folderName = 'TL_Maps';
TorL = {'T' 'L'};
myLet = TorL{LETTER};
letterMatrix = [1 2 3 4];

for i = 1:length(letterMatrix)
    %Generate values for experiment trials
    mapPath{i} = [EXPPATH str 'TStimuli' str imageFolderName str
'contrast_region_3_letter' myLet '_'];
end
```

```
fontVector                          = letterMatrix;
STIMULUS_PATH = [mapPath num2str(fontVector(1)) str
contrastMatrixFolder str];
addpath([mapPath num2str(fontVector(1)) str contrastMatrixFolder
str]);
%Obtain new set of .dat files
filename = dir([STIMULUS_PATH '*.dat']);
numDats = length(filename);
%Preallocate memory
summedMaps                          = cell(1,numDats);
diffMatrices = cell(length(fontVector),numDats);
valueMatrix = cell(size(diffMatrices1));

for fontSize = 1:length(fontVector)
    if(fontSize ~= 1)
        %Update path on each iteration to loop through font sizes
        STIMULUS_PATH = [mapPath{fontSize}
num2str(fontVector(fontSize)) str contrastMatrixFolder str];
        addpath([mapPath{fontSize} num2str(fontVector(fontSize)) str
contrastMatrixFolder str]);
        %Obtain new set of .dat files
        filename = dir([STIMULUS_PATH '*.dat']);
    end
    matrixIndex = 1;
    for currentDat = 1:numDats
        diffMatrices{fontSize,matrixIndex} =
load(filename(currentDat).name);
        switch selectedTLValue
            case 1
                valueMatrix{fontSize,matrixIndex} =
nanmin(nanmin(diffMatrices{matrixIndex}));
            case 2
                sortedSums = sort(diffMatrices{matrixIndex});
                valueMatrix{fontSize,matrixIndex} =
nanmedian(sortedSums(find(sortedSums<nanmedian(nanmedian(sortedSums)
))));
            case 3
                valueMatrix{fontSize,matrixIndex} =
nanmedian(nanmedian(diffMatrices{matrixIndex}));
            case 4
                sortedSumsUP = sort(diffMatrices{matrixIndex});
                valueMatrix{fontSize,matrixIndex} =
nanmedian(sortedSumsUP(find(sortedSumsUP>nanmedian(nanmedian(sortedS
umsUP)))));
            case 5
                valueMatrix{fontSize,matrixIndex} =
nanmax(nanmax(diffMatrices{matrixIndex}));
        end
        matrixIndex = matrixIndex+1;
    end
end

for sumIndex = 1:numDats
    currentMap = diffMatrices{1,sumIndex} - valueMatrix{1,sumIndex};
```

```
    for cFont = 2:length(fontVector)
        currentMap = abs(currentMap + (diffMatrices{cFont,sumIndex}-
valueMatrix{cFont,sumIndex}));
        if(cFont == length(fontVector))
            if(myLet == 'T') %Update to find if variable name exists
and set values that way
                summedMapsT{sumIndex} = currentMap;
            else
                summedMapsL{sumIndex} = currentMap;
            end
        end
    end
end

for sumIndex = 1:numDats
    summedMaps{sumIndex} = summedMapsT{sumIndex} +
summedMapsL{sumIndex};
end

switch{selectedTLValue}
    case 1
        for sumIndex = 1:numDats
            newMins{sumIndex} =
nanmin(nanmin(summedMaps{sumIndex}));
        end
    case 2
        for sumIndex = 1:numDats
            sortedSums = sort(summedMaps{sumIndex});
            newLQs{sumIndex} =
nanmedian(sortedSums(find(sortedSums<nanmedian(nanmedian(sortedSums)
))));
        end
    case 3
        for sumIndex = 1:numDats
            newMedians{sumIndex} =
nanmedian(nanmedian(summedMaps{sumIndex}));
        end
    case 4
        for sumIndex = 1:numDats
            sortedSumsUP = sort(summedMaps{sumIndex});
            newUQs{sumIndex} =
nanmedian(sortedSumsUP(find(sortedSumsUP>nanmedian(nanmedian(sortedS
umsUP)))));
        end
    case 5
        for sumIndex = 1:numDats
            newMaxs{sumIndex} =
nanmax(nanmax(summedMaps{sumIndex}));
        end
end

% --- Executes on selection change in TLdifference.
function TLdifference_Callback(hObject, eventdata, handles)
% hObject    handle to TLdifference (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
```

```
% handles    structure with handles and user data (see GUIDATA)

% Hints: contents = cellstr(get(hObject,'String')) returns
TLdifference contents as cell array
%        contents{get(hObject,'Value')} returns selected item from
TLdifference
global selectedTLValue
str = get(hObject, 'String');
val = get(hObject,'Value');
% Set current data to the selected data set.
switch str{val}
case 'Please select a value...' % User selects peaks.
   selectedTLValue = 0;
case 'Generate Mins'
    selectedTLValue = 1;
case 'Generate Lower Quartiles'
    selectedTLValue = 2;
case 'Generate Medians'
    selectedTLValue = 3;
case 'Generate Upper Quartiles'
    selectedTLValue = 4;
case 'Generate Maxes'
    selectedTLValue = 5;
end


% --- Executes during object creation, after setting all properties.
function TLdifference_CreateFcn(hObject, eventdata, handles)
% hObject    handle to TLdifference (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    empty - handles not created until after all CreateFcns
called

% Hint: popupmenu controls usually have a white background on
Windows.
%       See ISPC and COMPUTER.
if ispc && isequal(get(hObject,'BackgroundColor'),
get(0,'defaultUicontrolBackgroundColor'))
    set(hObject,'BackgroundColor','white');
end


% --- Executes on selection change in TorLdifference.
function TorLdifference_Callback(hObject, eventdata, handles)
% hObject    handle to TorLdifference (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)

% Hints: contents = cellstr(get(hObject,'String')) returns
TorLdifference contents as cell array
%        contents{get(hObject,'Value')} returns selected item from
TorLdifference
% Determine the selected data set.
global selectedValue
```

```
str = get(hObject, 'String');
val = get(hObject,'Value');
% Set current data to the selected data set.
switch str{val}
case 'Please select a value...' % User selects peaks.
   selectedValue = 0;
case 'Generate Mins'
    selectedValue = 1;
case 'Generate Lower Quartiles'
    selectedValue = 2;
case 'Generate Medians'
    selectedValue = 3;
case 'Generate Upper Quartiles'
    selectedValue = 4;
case 'Generate Maxes'
    selectedValue = 5;
end

fprintf('selectedValue: %i',selectedValue)

% --- Executes during object creation, after setting all properties.
function TorLdifference_CreateFcn(hObject, eventdata, handles)
% hObject    handle to TorLdifference (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    empty - handles not created until after all CreateFcns
called

% Hint: popupmenu controls usually have a white background on
Windows.
%       See ISPC and COMPUTER.
if ispc && isequal(get(hObject,'BackgroundColor'),
get(0,'defaultUicontrolBackgroundColor'))
    set(hObject,'BackgroundColor','white');
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Title:    Single Set Algorithm (Unbiased Selection) FUNCTION
% Author:   Adam Clayden and Antje Nuthmann
% Date:     29/01/2015
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function
[fxCoor,fyCoor,xArray,yArray,xCri,yCri,xallPos,yallPos,trueValue] =
placeWithType(valueType,fontWidth,fontHeight,csumMap,fid,choiceStr,C
ircleMask,res_x, res_y,mySize75,radiusC)

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%Input:
%   valueType: target value which can be the min, lower quartile,
median,
%   upper quartile or the max within each difference map.
%   fontWidth, fontHeight: target object's width and height
%   csumMap: current summed difference map
%   fid: file identifier (1 if writing images, otherwise, 0)
%   choiceStr: name of the valueType eg: "median"
```

```
%   CircleMask: circular section in the centre to exclude
%   res_x, res_y: 600x800
%   mySize75: max[width, height]
%   radiusC: radius of the CircleMask
%
%Output:
%   fxCoor, fyCoor: final x,y positions to use for object insertion
%   xArray, yArray: x,y arrays of all invalid coordinates
%   xCri, yCri: all locations that passed first criterion but not
second
%   xallPos,yallPos: all locations of the same value within a matrix
%   trueValue: returns either "True median" or "Close median"
depending on
%   the outcome. Or, whichever valueType was chosen.
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Instantiate Variables
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
breakloop = 0; %Break condition for while loop
xallPos = []; %All x positions if there are duplicate values
yallPos = []; %All y positions if there are duplicate values
xArray = []; %Invalid x location
yArray = []; %Invalid y location
xCri = []; %All x positions that passed both criterion
yCri = []; %All y positions that passed both criterion
fxCoor = []; %Final x coordinate chosen
fyCoor = []; %Final y coordinate chosen
size75 = mySize75; %Size of letter (max between width and height)
letterRad = size75/2;
trueValue = ''; %Blank string for visualisation to tell user if the
value chosen was a true value or a value close to it
rcidx = 0; %row col index for passed criterion
cRow = res_y/2; %centre of circleMask
cCol = res_x/2; %centre of circleMask

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Find Median in Image
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

[sumedY12, sumedX12] = find(csumMap == valueType);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Begin Loop
% Finds median value in summed difference map or a value close to
said
% value.
% The value (median, max etc) could be located in multiple positions
within
% a given image, and so each one is analysed before 1 being chosen
at
% random to yield our final x,y coordinates.
% Equation that the loop represents: SUMi(SUMj(1< Xij < WH : Xij ~=
C[ij]))
% x,y coordinates of value is checked against two criteria:
```

```
% 1. Boundaries of image to prevent letter from truncating off
% 2. Inner circle to imitate foveal vision as letter should be
placed
%    outside of this region (either parafoveal or peripheral region)
% If value passes both checks then we choose this value.
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
while(breakloop == 0)
    % If empty
    if(isempty(sumedX12))
        % Create temporary matrix of absolute values
        tmpsumlq = abs(csumMap - valueType);
        % Reshape into vector
        tmpVector = reshape(tmpsumlq,1,numel(tmpsumlq));
        % Transpose vector and place values in first column
        mat3(:,1) = tmpVector';
        % Place rows in second column
        mat3(:,2) = repmat(1:size(csumMap,1),1,size(csumMap,2))';
        tmp = [];
        % Place columns in third column
        for i=1:size(csumMap,2)
            tmp = [tmp; repmat(ones(1,size(csumMap,1))*i,1)'];
        end
        mat3(:,3) = tmp;
        % Sort matrix rows by the first column to place coordinates
in
        % order
        mat3 = sortrows(mat3,1);
        % Find all unique values in the first column
        tmpvals = mat3(:,1);
        tmpvals_unique = unique(tmpvals);
        % Open for loop 1 - number of unique values
        for nextmin = 1:length(tmpvals_unique)
            % Find current unique value in matrix and store how many
times
            % it occurs
            idx = find(mat3(:,1)==tmpvals_unique(nextmin));
            for posIndex = 1:length(idx)
                xallPos(end+1) = mat3(idx(posIndex),3);
                yallPos(end+1) = mat3(idx(posIndex),2);
            end
            %fprintf(fid,'\nunique value %i :
%.15f\n',nextmin,tmpvals_unique(nextmin));
            %fprintf(fid,'\nTimes Occurred: %i\n',length(idx));
            % Open for loop to iterate through these multiple values
(if
            % any)
            for rpInd = 1:length(idx)
                % Store current value and coordinates
                mypick = idx(rpInd);
                row = mat3(mypick,2); % X is 2 or 3
                col = mat3(mypick,3);
                % Check against criterion (boundary + centre)
                if(row > (1 + fontHeight) && row < (res_y -
fontHeight) && col > (1 + fontWidth) && col < (res_x - fontWidth))
                    d = sqrt((cRow - row).^2 + (cCol - col).^2);
                    if(d>radiusC+letterRad)
```

```
                    %if(CircleMask((row+size75/2),(col-size75/2)) ==
0 && CircleMask((row-size75/2),(col+size75/2)) == 0 &&
CircleMask((row-size75/2),(col-size75/2))==0 &&
CircleMask((row+size75/2),(col+size75/2))==0)
                        % add row and column values into array
                        rcidx = rcidx + 1;
                        rowcolArray(rcidx,2) = col;
                        rowcolArray(rcidx,1) = row;
                        if(rpInd == length(idx))
                            breakloop = 1;
                        end
                    else
                        xCri(end+1) = col;
                        yCri(end+1) = row;
                    end
                else
                    % Else = invalid location
                    xArray(end+1) = col;
                    yArray(end+1) = row;
                end
                % If all locations don't pass, loop to the next
uniqe value
                % and store how many times it occurs
            end
            if(breakloop == 1)
                % Select one at random and make this our final
                % position
                ranPerm = randperm(numel(rowcolArray(1:rcidx,1)));
                fxCoor = rowcolArray(ranPerm(1),2);
                fyCoor = rowcolArray(ranPerm(1),1);
                break;
            end
        end
        trueValue = ['Close ' choiceStr];
    else
        % If not empty, check each value against criteria (if there
are
        % multiple)
        rcidx = 0;
        trueValue = ['True ' choiceStr];
        row = sumedY12;
        col = sumedX12;
        for tmidx = 1:numel(row)
            if(row(tmidx) > (1 + fontHeight) && row(tmidx) < (res_y
- fontHeight) && col(tmidx) > (1 + fontWidth) && col(tmidx) < (res_x
- fontWidth))
                d = sqrt((cRow - row).^2 + (cCol - col).^2);
                if(d>radiusC+letterRad)
                %if(CircleMask((row(tmidx)+size75/2),(col(tmidx)-
size75/2)) == 0 && CircleMask((row(tmidx)-
size75/2),(col(tmidx)+size75/2)) == 0 && CircleMask((row(tmidx)-
size75/2),(col(tmidx)-size75/2))==0 &&
CircleMask((row(tmidx)+size75/2),(col(tmidx)+size75/2))==0)
                    % Add row and column values into array
                    rcidx = rcidx + 1;
                    rowcolArray(rcidx,2) = col(tmidx);
```

```
                    rowcolArray(rcidx,1) = row(tmidx);
                    % Select one at random and make this our final
position
                    ranPerm = randperm(numel(rowcolArray(:,1)));
                    fxCoor = rowcolArray(ranPerm(1),2);
                    fyCoor = rowcolArray(ranPerm(1),1);
                    if(tmidx == numel(row))
                        breakloop = 1;
                        if(fid ~= 0)
                            fprintf(fid,['True ' choiceStr]);
                        end
                    end
                else
                    xCri(end+1) = col(tmidx);
                    yCri(end+1) = row(tmidx);
                end
            else
                xArray(end+1) = col(tmidx);
                yArray(end+1) = row(tmidx);
            end
        end
        % If all fail, loop back up to top and find next closest
        if(breakloop == 0)
            sumedX12 = [];
        end
    end
    if(breakloop == 1)
        % Select one at random and make this our final position
        ranPerm = randperm(numel(rowcolArray(1:rcidx,1)));
        fxCoor = rowcolArray(ranPerm(1),2);
        fyCoor = rowcolArray(ranPerm(1),1);
        break;
    end
end
```

# Appendix F – Dead-Zone Extension

```
for (i in 1:nrow(b)) { # for a given trial
  fixationKey <- 1
  # fixation index (ordinal number of fixation)
  fixRep$FIX_INDEX[b[i,1]:b[i,2]] <- seq(1,b[i,3])
# code number of fixations on the target object
  tmp <- b[i,1]:b[i,2]
  # idxx <- which(a$HIT_FIX[tmp]==T) # my IA hit check
  idxx <-
which(!is.na(fixRep$CURRENT_FIX_INTEREST_AREA_FIX_COUNT[tmp])) # DV
IA hit check

  if (length(idxx>0)) {
    fixRep$FIX_IDX_TAR[tmp[idxx]] <- 1:length(idxx)
    # beginning of the first fixation on the target object
    # copy to 'tab' for calculation of 'latency to first target
fixation'
    if(refixCheck == 1)
    {
      if(length(idxx) > 1)
      {
        for(j in 1:length(idxx))
        {
          key = j + 1
        }
      }
    }
    if(extendVeri == 1)
    {
      if(length(idxx) > 1)
      {
        #d = sqrt((cRow - a$).^2 + (cCol - col).^2);
        #(x - center_x)^2 + (y - center_y)^2 < radius^2
        #avSacLength <-
sum(a[tmp[2]:(tmp[idxx[1]]),"PREVIOUS_SAC_AMPLITUDE"],na.rm=T)/lengt
h(a[tmp[2]:(tmp[idxx[1]]),"PREVIOUS_SAC_AMPLITUDE"])
        for(j in 1:length(idxx))
        {
          key = j+1
          if(key <= length(idxx) & j != length(idxx))
          {

if(!is.na(fixRep$PREVIOUS_SAC_DIRECTION[tmp[idxx[j]]+1]))
            {

if(!is.na(fixRep$PREVIOUS_SAC_DIRECTION[tmp[idxx[j]]+2]))
              {
                if(fixRep$FIX_INDEX[tmp[idxx[key]]] -
fixRep$FIX_INDEX[tmp[idxx[j]]] != 1)
                {
```

```
                    if(fixRep$PREVIOUS_SAC_DIRECTION[tmp[idxx[j]]+1]
== "RIGHT")
                    {
                      oppositePlus1 <- "LEFT"
                    } else
if(fixRep$PREVIOUS_SAC_DIRECTION[tmp[idxx[j]]+1] == "LEFT")
                    {
                      oppositePlus1 <- "RIGHT"
                    } else
if(fixRep$PREVIOUS_SAC_DIRECTION[tmp[idxx[j]]+1] == "UP")
                    {
                      oppositePlus1 <- "DOWN"
                    } else
if(fixRep$PREVIOUS_SAC_DIRECTION[tmp[idxx[j]]+1] == "DOWN")
                    {
                      oppositePlus1 <- "UP"
                    }
                    if(!is.na(fixRep$CURRENT_FIX_X_OTHER[b[i,1]]) &
!is.na(fixRep$CURRENT_FIX_X[b[i,1]]))
                    {
                      finalX <- (fixRep$CURRENT_FIX_X[b[i,1]] +
fixRep$CURRENT_FIX_X_OTHER[b[i,1]])/2
                      finalY <- (fixRep$CURRENT_FIX_Y[b[i,1]] +
fixRep$CURRENT_FIX_Y_OTHER[b[i,1]])/2
                    } else {
                      finalX <- fixRep$CURRENT_FIX_X[b[i,1]]
                      finalY <- fixRep$CURRENT_FIX_Y[b[i,1]]
                    }
                    if(!is.na(fixRep$CURRENT_FIX_X_OTHER[tmp[idxx[j]
]+2]) & !is.na(fixRep$CURRENT_FIX_X[tmp[idxx[j]]+2]))
                    {
                      currentX <- (fixRep$CURRENT_FIX_X[tmp[idxx[j]
]+2] + fixRep$CURRENT_FIX_X_OTHER[tmp[idxx[j]]+2])/2
                      currentY <- (fixRep$CURRENT_FIX_Y[tmp[idxx[j]
]+2] + fixRep$CURRENT_FIX_Y_OTHER[tmp[idxx[j]]+2])/2
                    } else {
                      currentX <- fixRep$CURRENT_FIX_X[tmp[idxx[j]
]+2]
                      currentY <- fixRep$CURRENT_FIX_Y[tmp[idxx[j]
]+2]
                    }
                    if(fixRep$PREVIOUS_SAC_DIRECTION[tmp[idxx[j]
]+2] != oppositePlus1 & (sqrt((currentY - finalY)^2 + (currentX -
finalX)^2) > fixRep$ia_w[1]*3))
                    {
                      tab[i,"SCANADD"] <- fixRep[tmp[idxx[j]
],"CURRENT_FIX_START"]
                    } else
                    { #otherwise they are verifying
                      fixationKey <- j
                    }
                  }
                }
              }
            }
          }
```

```
      }
    }
    tab[i,"FFIX_TAR_START"] <-
fixRep[tmp[idxx[fixationKey]],"CURRENT_FIX_START"]
    # summed distance (saccade amplitudes) between all fixations
from scene onset to the first fixation on the target
  }
}
tab$SEARCH_INIT <- fixRep[fixRep$FIX_INDEX==1,"CURRENT_FIX_END"]

# LAT latency to first target fixation
tab$LAT <- tab$FFIX_TAR_START # LAT latency to first target fixation
tab$SCAN <- tab$LAT - tab$SEARCH_INIT # scanning time

if(extendVeri == 1)
{
  SCANNEWidx <- which(!is.na(tab$SCANADD))
  checkScan <- tab$SCANADD[SCANNEWidx] - tab$SCAN[SCANNEWidx]
  tab$SCAN[SCANNEWidx] <- tab$SCAN[SCANNEWidx] + checkScan
  newVeri = (tab$FFIX_TAR_START[SCANNEWidx] + checkScan)
  tab$VERI <- tab$RT - tab$FFIX_TAR_START
  tab$VERI[SCANNEWidx] <- tab$RT[SCANNEWidx] - newVeri
# verification time new
}
```

# Appendix G – Partial omega squared effect sizes

Below are the $\omega_p^2$ effect sizes for all experiments:

Experiment 1

| Reaction Time | |
|---|---|
| scotoma | font size |
| 0.020544 | 0.851923 |

| Initiation Time | |
|---|---|
| scotoma | font size |
| 0.055781 | 0.043031 |

| Scanning Time | |
|---|---|
| scotoma | font size |
| - 0.02508 | 0.796687 |

| Verification Time | |
|---|---|
| scotoma | font size |
| 0.157026 | 0.550414 |

| Saccade Amplitude | |
|---|---|
| scotoma | font size |
| 0.633615 | 0.262474 |

| Fixation Duration | |
|---|---|
| scotoma | font size |
| 0.294731 | 0.384966 |

Experiment 2

| Reaction Time | |
|---|---|
| scotoma | font size |
| 0.041824 | 0.72568 |

| Initiation Time | |
|---|---|
| scotoma | font size |
| 0.089337 | 0.056834 |

| Scanning Time | |
|---|---|
| scotoma | font size |
| 0.068287 | 0.747616 |

| Verification Time | |
|---|---|
| scotoma | font size |
| 0.441757 | 0.342049 |

| Saccade Amplitude | |
|---|---|
| scotoma | font size |
| 0.611535 | 0.230222 |

| Fixation Duration | |
|---|---|
| scotoma | font size |

| | |
|---|---|
| 0.259903 | 0.149764 |

Experiment 3

| Reaction Time | | | | | |
|---|---|---|---|---|---|
| scotoma | salience | target_size | scotoma:salience | scotoma:targetsize | salience:targetsize |
| 0.011385 | 0.876639 | 0.758246713 | -0.012910643 | -0.031196368 | 0.479235538 |

| Initiation Time | | | | | |
|---|---|---|---|---|---|
| scotoma | salience | target_size | scotoma:salience | scotoma:targetsize | salience:targetsize |
| 0.211066 | 0.417963 | 0.110363 | 0.012751 | -0.01915 | 0.125608 |

| Scanning Time | | | | | |
|---|---|---|---|---|---|
| Scotoma | salience | target_size | scotoma:salience | scotoma:targetsize | salience:targetsize |
| -0.01077 | 0.86754 | 0.725304 | 0.050886 | -0.02992 | 0.414698 |

| Verification Time | | | | | |
|---|---|---|---|---|---|
| scotoma | salience | target_size | scotoma:salience | scotoma:targetsize | salience:targetsize |
| 0.105144 | 0.641493 | 0.457754 | 0.018208 | -0.02557 | 0.287314 |

| Saccade Amplitude | | | | | |
|---|---|---|---|---|---|
| scotoma | salience | target_size | scotoma:salience | scotoma:targetsize | salience:targetsize |
| 0.428779 | 0.67766 | 0.445054 | -0.03114 | 0.097165 | 0.090966 |

| Fixation Duration | | | | | |
|---|---|---|---|---|---|
| scotoma | salience | target_size | scotoma:salience | scotoma:targetsize | salience:targetsize |
| 0.262961 | 0.641058 | 0.297418 | -0.0111 | 0.16703 | 0.004952 |

Experiment 4

| Reaction Time | | |
|---|---|---|
| scotoma | salience | scotoma:salience |
| 0.916135 | 0.832288 | 0.270917259 |

| Initiation Time | | |
|---|---|---|
| scotoma | salience | scotoma:salience |
| 0.564988 | 0.177366 | -0.02258 |

| Scanning Time | | |
|---|---|---|
| scotoma | salience | scotoma:salience |
| 0.941959 | 0.777206 | 0.218448 |

| Verification Time | | |
|---|---|---|
| scotoma | salience | scotoma:salience |
| 0.404355 | 0.569068 | 0.229766 |

| Saccade Amplitude | | |
|---|---|---|
| scotoma | salience | scotoma:salience |
| 0.92535 | 0.52656 | 0.166337 |

| Fixation Duration | | |
|---|---|---|
| scotoma | salience | scotoma:salience |
| 0.418397 | 0.578951 | 0.426437 |

Experiment 5

| Reaction Time | |
|---|---|
| scotoma | |
| 0.347554 | |

| Initiation Time | |
|---|---|
| scotoma | |
| 0.078402 | |

| Scanning Time | |
|---|---|
| scotoma | |
| 0.294207 | |

| Verification Time | |
|---|---|
| scotoma | |
| 0.225996 | |

| Fixation Duration | |
|---|---|
| scotoma | |
| -0.03069 | |

| Saccade Amplitude | |
|---|---|
| scotoma | |
| 0.676181 | |

Experiment 6

| Reaction Time | |
|---|---|
| scotoma | |
| 0.339953 | |

| Initiation Time | |
|---|---|
| scotoma | |
| 0.41162 | |

| Scanning Time | |
|---|---|
| scotoma | |

| 0.005936 |
|---|

| Verification Time |
|---|
| scotoma |
| 0.386576 |

| Fixation Duration |
|---|
| scotoma |
| 0.025788 |

| Saccade Amplitude |
|---|
| scotoma |
| 0.72441 |

Experiment 7

| Reaction Time | |
|---|---|
| scotoma | optimalUnoptimal |
| 0.341018 | 0.251417654 |

| Initiation Time | |
|---|---|
| scotoma | optimalUnoptimal |
| 0.03636 | 0.0598 |

| Scanning Time | |
|---|---|
| scotoma | optimalUnoptimal |
| 0.252879 | 0.264272 |

| Verification Time | |
|---|---|
| scotoma | optimalUnoptimal |
| 0.120461 | -0.02139 |

| Saccade Amplitude | |
|---|---|
| scotoma | optimalUnoptimal |
| 0.818863 | 0.01167 |

| Fixation Duration | |
|---|---|
| scotoma | optimalUnoptimal |
| 0.03636 | 0.0598 |