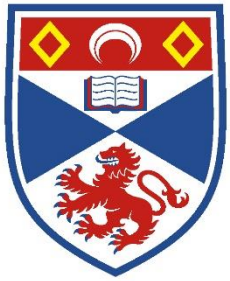
Exploring how object shape and binocular vision interact to make or break camouflage

Philip Cammack



University of
St Andrews

This thesis is submitted in partial fulfilment for the degree of PhD
at the
University of St Andrews

23rd September 2016

Abstract

Depth perception is a major component of 3D vision. There are many cues to depth; one particularly sensitive aspect is the vivid perception of depth created from having eyes with overlapping visual fields (binocular vision). As the eyes are located at different points in space, they see different views of the scene – these slight differences (called binocular disparity) can be used to obtain depth information. However, extracting depth from disparity requires complex visual processing. So why use binocular vision?

Julesz (1971) proposed an explanation – camouflaged animals can fool the perception of some cues to 3D shape, but camouflage is ineffective against binocular vision. We would expect that animals with binocular vision could see the 3D shape of animals, despite their camouflage. Whilst commonly accepted, this hypothesis has not been tested in detail. In this thesis, we present experiments designed to establish how depth from binocular vision interacts with camouflage and object shape. Two main questions were addressed:

First, we explored how the visual system represented depth information about 3D objects from binocular disparity. Objects with smooth depth edges (hill-shaped) were perceived with less depth than sharper edged objects. A computational model that segregated the object, then averaged the disparity over the segregated region emulated human performance. Finally, we found that disparity and luminance cues interacted to alter perceived depth.

Secondly, we investigated if binocular vision could overcome camouflage. We found that camouflaged objects defined by luminance were detected faster when also defined by depth from disparity, thus reduces the effect of camouflage. Smooth objects were detected slower than sharp objects: an effect that was replicated in the real world, suggesting a camouflage technique to counter binocular vision.

In summary, binocular vision is useful because it can detect camouflaged objects. However, smoother shapes take longer to spot, forming binocular (or stereoscopic) camouflage.

Contents

Table of Figures

1 Introduction

We exist in a rich and complex environment, with a wide variety of complex visual information. To exist and navigate in our three-dimensional world, we must take the visual input from the eye and reconstruct the world which in we live in to enough detail and accuracy to enable us to live, eat, catch prey, navigate and avoid predators. As the world is three dimensional, it is not enough to only identify and know which direction objects are in – we must know how far away they are. This is essential to many of the uses of vision: to navigate, we must know the 3D arrangement of objects, and their shapes; to know if we are in danger we need to know if a predator is close enough to strike.

There are many ways in which the human visual system can extract information about depth and shape in a scene, from head movements to shape from shading. However, perhaps the most intriguing is binocular vision – the ability to extract depth information purely from the difference in the images between the two eyes. Binocular vision creates a particularly vivid impression of depth, and is excellent at distinguishing very small depth differences e.g. (McKee, 1983; Westheimer, 1975, 1979; Westheimer & McKee, 1977), indicating that it may be an important cue to extracting 3D object shape (Anzai & DeAngelis, 2010). Overall, binocular vision provides an exquisitely sensitive cue to depth without requiring movement that might give away the predator's presence.

Despite the advantages, extracting depth from binocular vision turns out to require a complex and metabolically costly processing system (Marr, 1982). This is because in order to extract the depth of a feature in the scene using two eyes, a feature must be identified as belonging to the same object in both eyes. This is not a trivial problem to solve – depending on the depth of the object a single feature could be located in a large range of different positions in each eye, requiring a huge number of comparisons between small regions in each eye. And this is only a tiny fraction of the way to even identifying the existence of a feature, let alone matching it to the depth of an object – the visual system must eliminate false matches, compensate for regions with no matches, detect edges, segregate out objects and many more steps besides (see Section 2.2 for an in-depth discussion of binocular vision).

Given all the neural apparatus used to extract depth from two eyes, perhaps it is surprising that the visual system does not rely entirely on other cues to depth. There are several other cues that can provide an impression of depth, and for a lot of animals these cues are their sole source of depth information e.g. (Heesy, 2009; Isbell, 2006; Srinivasan, 1995), this can be the case even when the animals have eyes with overlapping visual fields e.g. (Martin, 2015; Ott, Schaeffel, & Kirmse, 1998). For example, shape from shading can provide a vivid impression of depth of the shape of a surface by the luminance of the surface e.g. (Norman, Todd, Norman, Clayton, & McBride, 2006; Todd, 2004)- this cue can even override depth from binocular disparity in certain circumstances e.g.(Chen & Tyler, 2015; Lovell, Bloj, &

Harris, 2012). Other cues are also available on a comparable scale: self-motion causes closer objects to move across objects further away, enabling depth calculations e.g. (Heesy, 2009; Snowden & Freeman, 2004; Srinivasan, 1995). Artists commonly use other cues to depth in their work – for example objects further away are smaller and higher up in the visual scene.

(a) (b)

Figure 1.1: A demonstration of countershading: this is the same caterpillar with different orientations. Notice how in (a) the 3D shape is clear from the shading, where in (b) it has been obscured. Images reproduced by kind permission of Olivier Penacchio.

As there is an abundance of other reliable cues to depth, then why use binocular vision? It clearly has some advantages as it has evolved independently in several animals (Iwaniuk & Wylie, 2006; Stevens, 2006; Willigen, Frost, & Wagner, 1998). One of the most commonly accepted reasons is that binocular vision is likely to be used to spot well-camouflaged prey in the environment. Often, animals create exquisite camouflage through a variety of methods, such as eliminating shape-from-shading cues via countershading (Stevens & Merilaita, 2009) as in Figure 1.1, or creating disruptive camouflage patterns that break up the shape of the animal. However, with the ability to see in 3D using two eyes, then a binocular observer can perceive the shape of the camouflaged object, no matter how well their patterning matches their background. This is hypothesized to create a pop out or jump out effect, making the animal easy to detect due to the additional depth information.

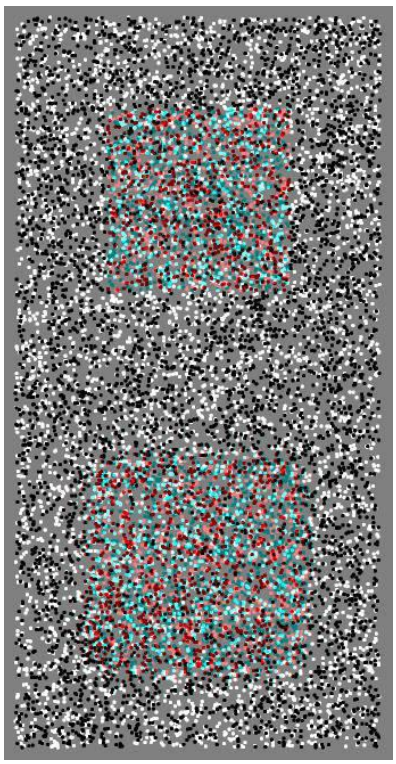


Figure 1.2: Abstract display of random dots similar to that used by Julesz (Burt & Julesz, 1980). This is the stimulus from experiment 1. These images can be viewed with red/blue anaglyph glasses to get an impression of depth.

This idea, that “even under ideal monocular camouflage, the hidden objects jump out in depth when stereoscopically fused” (Julesz (1971), p145), first stated by Julesz in 1971 has often been assumed to be correct e.g.(Heesy, 2009; Isbell, 2006). However, this observation

was made from a highly abstract display of random dots such as that in Figure 1.2 using objects with sharp edges – a far reach from realistic environments where animals have evolved to be as hard to detect as possible. In the forty years since Julesz made this proposal, this theory has been widely accepted but has not been subjected to rigorous scientific enquiry. In this thesis, I break this observation down into two halves for in-depth study:

1. What is a depth defined object? What happens when the boundary between the object and the background is smoothed?
2. Does depth perception break camouflage? Is there a way in which prey animals can break stereoscopic depth perception from binocular vision, thus reforming their camouflage?

1.1 Roadmap

In Chapter 2 I consider the background behind the current research, with two primary focuses: in Section 2.1 we present an overview of the biological background of animal camouflage, and in Section 2.2 we conduct an in depth discussion of stereopsis. In Section 2.3 we review visual search tasks relevant to stereopsis and camouflage. Finally, in Chapter 3 we discuss the methodologies used in this thesis for data extraction and analysis.

From this point, the thesis splits into two main strands as shown in the flowchart in Figure 1.3: those considering object perception in Chapters 4, 5 and 6; and those considering camouflage in Chapters 7, 8 and 9. We then briefly overview all the results together in the summary in Chapter 10. At the end of the thesis there is a glossary and references.

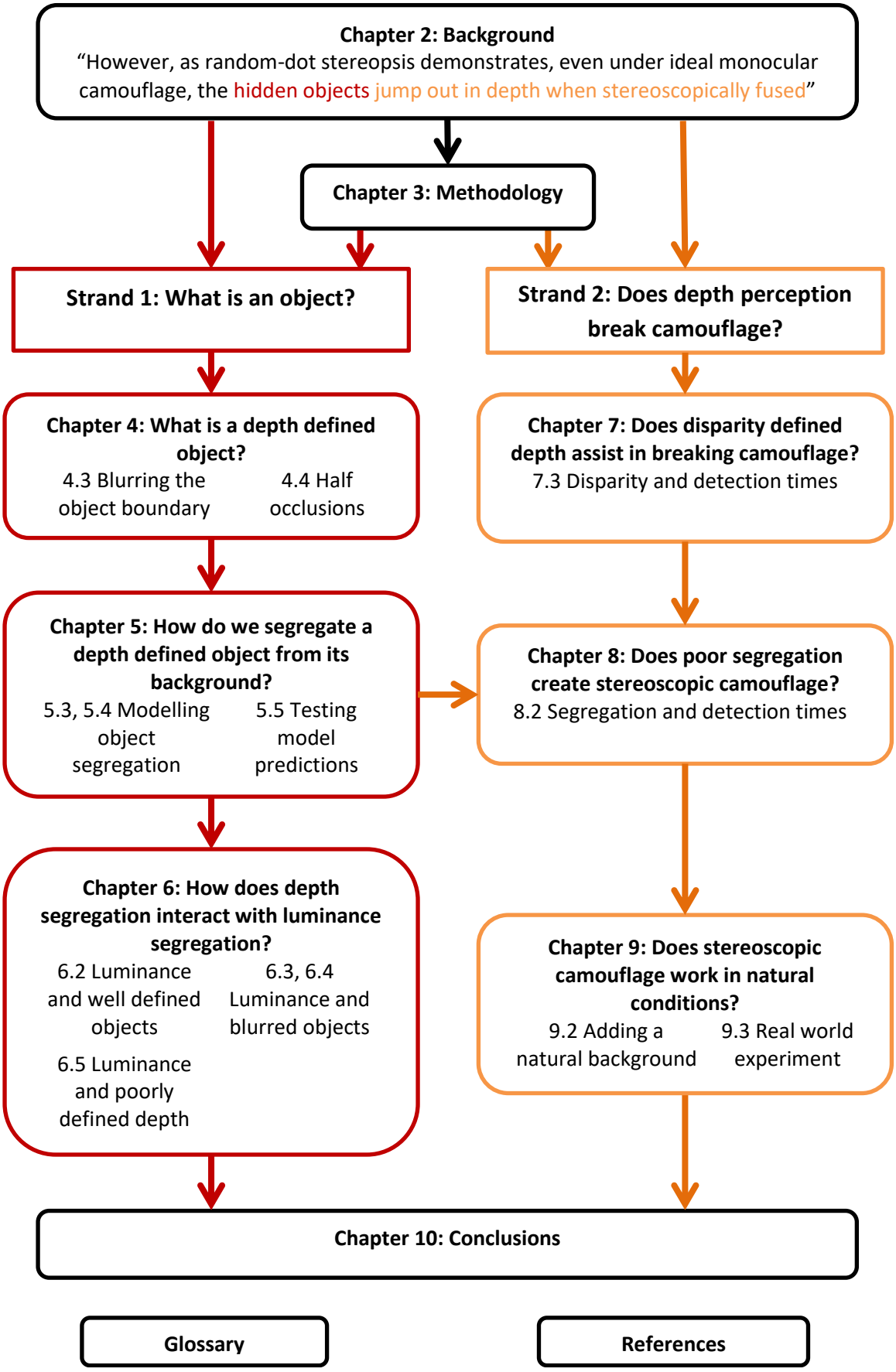


Figure 1.3: Flowchart of the thesis structure.

2 Background

- **Forms of camouflage such as crypsis and background matching.**
- **Detection time.**
- **Overview of binocular vision.**
- **Description of binocular disparity and other cues to depth.**
- **Modelling of disparity extraction.**
- **Overview of visual search.**

Figure 2.1: An octopus (Octopus rubescens, the central rock), displaying camouflage via background matching, disruptive colouration and mimicry.
Image reproduced with permission, (Rob, 2011).

2.1 Camouflage

We split our discussion of camouflage techniques into three major Sections: First, we discuss the different ways in which an animals can avoid detection. Second, we consider how predators react to these different methods of camouflage, and how they detect their prey. Finally, we review the current literature relevant to the interaction of depth perception and camouflage to see how binocular depth perception may aid in detecting camouflaged prey.

2.1.1 Camouflage techniques and depth perception

Camouflage takes many forms, but in the most generic it is the ability of an animal to fool perception of an observer so that it thinks the animal is either not present, or is not of interest, thus increasing its chance of survival (Darwin, 1859). Here we only consider the camouflage techniques in the presence of a surface (for example the soil, or a tree trunk) as in the open ocean the challenges of camouflage are different enough to require a different approach (Johnsen, 2002, 2014; Johnsen & Sosik, 2003). The methods of camouflage in land and most aquatic environments reflect a major split in camouflage techniques (Endler, 2006):

1. Mimicry and masquerade, which is attempting to be perceived as something else – either not of interest to the observer (typically called masquerade e.g. (Ruxton & Sherratt, 2004; Stevens & Merilaita, 2009)), or as a well defended object that is not worth predating (typically called mimicry e.g. (Edmunds & Edmunds, 1974; Ruxton & Sherratt, 2004)).
2. Crypsis and background matching – this is appearing not to be present, but rather perceived as part of the background and thus avoiding detection and reducing the chances of predation (Cooper & Allen, 1994; Cuthill et al., 2005; Edmunds & Edmunds, 1974; Feltmate & Williams, 1989; Johannesson & Ekendahl, 2002; Ruxton & Sherratt, 2004).
3. Disruptive colouration, is the use of bold colours or contours (such as in Figure 2.2) to break up an animals' outline. The observer then sees the animal as many different elements that are not of interest (Cott, 1940; Cuthill et al., 2005; Lovell, Egan, Scott-Brown, & Sharman, 2015).
4. Self-shadow concealment is when an animal changes shape or colouration to hide shadows that might give away the location of an animal that is otherwise well camouflaged. This comes in two forms: hiding shadows on the animal, normally by colouration called countershading e.g.(Penacchio, Lovell, Cuthill, Ruxton, & Harris, 2015; Rowland, 2009; Ruxton, Speed, & Kelly, 2004), or by changing shape to minimise shadows on the animal or background e.g.(Penacchio, Lovell, Cuthill, et al., 2015).

Figure 2.2: A herd of zebras displaying disruptive colouration, making it hard to distinguish one animal from another. the primary purpose of their stripes is debated (Caro, Izzo, Reiner, Walker, & Stankowich, 2014; Egri et al., 2012; How & Zanker, 2013). Image reproduced with permission, Kaisl (2007).

Many animals employ a mix of these tactics, for example in Figure 2.1, where the Octopus is mimicking seaweed in order to match its background – this specific combination is often known as masquerade. These mixings of camouflaging techniques suggest there is no clear boundary between the different techniques e.g. (Cuthill et al., 2005; Endler, 1981). For clarity in this thesis we use the definitions listed above.

We are interested in the interaction of camouflage with binocular depth perception. A binocular predator has eyes placed such that an area of the environment is visible in both eyes. Due to the different positions in space of the two eyes, each eye sees a slightly different image of the scene. It is possible to use these small differences to extract the depth of the scene (discussed in detail in Section 2.2). It is commonly thought that the ability to extract depth causes camouflaged objects in the scene to become obvious, making them easy to detect e.g. (Heesy, 2009; Isbell, 2006; Julesz, 1971). To investigate this claim, we consider how well these three camouflage mechanisms may perform when viewed under the effects of binocular depth perception.

Mimicry and masquerade are related techniques, where an animal imitates either another animal or an aspect of its environment, such as a leaf. In the majority of cases a mimic must attempt to not only match the colouration and patterning of the mimicked animal, but in order to appear to be the mimicked object from any angle, and so will match the overall

shape. This makes depth information of little use in the detection of the majority of animals employing mimicry. Due to this, we do not spend much time considering the interaction of mimicry with depth perception.

It is less clear how disruptive colouration may interact with depth perception – for some patterns and shapes it may be that the breaking up of the prey's shape into many elements would cause the perception of several different objects with slightly different depths. However, it is also possible that the bold edges used in disruptive colouration would be discounted due to the continuous change in depth, thus enabling depth perception to detect the form of the animal (Lovell, Scott-Brown, Egan, & Sharman, 2016). While of interest (we briefly consider the interaction of disruptive colouration and depth perception in Chapter 6), we are mainly concerned in investigating Julesz's claims, which relate to objects camouflaged with no monocular cues to shape – such as in background matching.

It is thought that background matching can be 'broken' (i.e. it will no longer have a deceptive effect on the observer) using depth perception (Julesz, 1971; Troscianko, Benton, Lovell, Tolhurst, & Pizlo, 2009a). Typically, an animal matching its background will match, as closely as possible (from the point of view of the viewer), the surface properties of the background using pigmentation and patterning (Endler, 1981, 1984; Zylinski, Osorio, & Shohet, 2009). Typically, an animal that matches a background well enough to be hard to detect is described as being cryptic. We do not go into the intricacies of the difference between crypsis and background matching here, for more detail see Ruxton & Sherratt (2004). However, even if the animal matches the colour, luminance and texture of the background, it will still have a 3D shape that is different to the background, particularly in cases where the background is relatively flat. This means that to an observer with depth perception, the animal will still stand out from the background, thus breaking camouflage (Isbell, 2006). We discuss this in detail in Section 2.1.3, after considering how camouflaged animals are detected in the next Section.

2.1.2 Detection of camouflaged animals

In order to study the interaction of camouflage with processes that cause the camouflage to no longer be effective, called the camouflage 'breaking' processes, we must understand how a predator detects a camouflaged prey item. Note that prey will also be attempting to detect camouflaged predators, but we use a 'predator searching for prey' description for convenience of linguistics.

To detect camouflaged prey, the predator must search an area of the environment for the prey item. The predator must decide how long to search the environment – if it spends a long time searching an area without prey in or attacks a feature of the environment that is not prey, then it has wasted time and energy (Geisler & Diehl, 2002, 2003; Maloney, 2003; Smith, 2009). However, if it moves on too fast and misses potential prey then it has lost a potential opportunity to feed (Geisler & Diehl, 2002, 2003; Maloney, 2003). This leads us to one of the major attributes describing predator activity – the length of time that the

predator spends searching for a prey item before moving on, sometimes called the quitting threshold (Godwin, Menneer, Riggs, Cave, & Donnelly, 2014; Schwark, MacDonald, Sandry, & Dolgov, 2013). The optimal length of time spent searching a region of the environment is determined by the overall abundance of prey, and how well it is camouflaged – typically a predator will choose how fast it moves on to account for these factors (Gendron & Staddon, 1983).

The total chance of detection of an animal when encountering a predator is therefore a function of the length of time a predator searches for and the probability of the animal being detected per unit time. If the animal takes a long time to detect (i.e. is very well camouflaged) and the predator spends a long time searching, then the overall probability of survival could be the same as an if a predator spends a short time searching for an animal that is easy to detect.

In the experiments presented in this thesis, the participants know that their 'prey' is always present, therefore are incentivised to search for a long time before giving up. The longer it takes the participant to detect the 'prey' the better camouflaged it is, as they will only infrequently quit the search. We therefore we use detection time as a measure of how well camouflaged our token prey item is e.g. (Cuthill et al., 2005; Lovell et al., 2015; Penacchio et al., 2016; Penacchio, Lovell, Sanghera, et al., 2015). In an environment without the knowledge of the target being present, then if an animal takes longer to detect, then it will be more likely to go unnoticed until the predator moves onto the next area. As participants will sometimes get bored and quit their search, we also look at number of trials where they correctly detect the prey – called the accuracy.

The theory that predators modify the length of time they search the environment for prey dependent on prey abundance and camouflage is called the search rate theory (Ruxton & Sherratt, 2004). There are other competing theories, most notably search image theory – this suggests that once one prey item has been detected, the predator will be better at detecting a second similar prey item for a short time (Dukas & Kamil, 2001; Pietrewicz & Kamil, 1979; Plaisted & Mackintosh, 1995; Tinbergen, 1960). This mechanism leads to similar effects to the search rate theory (Ruxton & Sherratt, 2004), and there are some key differences these affect situations with a mix of prey types and prey frequencies (Dawkins, 1971; Plaisted & Mackintosh, 1995; Reid & Shettleworth, 1992). While important to understanding predator-prey dynamics in a mixed environment, this does not assist us in evaluating how well cryptically camouflaged (hard to detect) a single prey item is when presented in isolation – for both theories a prey item that takes longer to find will have a greater survival rate. Hence my studies will focus on the detection of single items.

Using the detection time as a measure of camouflage is a well-established technique (an extensive number of examples can be found in *Avoiding Attack*, by Ruxton & Sherratt (2004)).In biological contexts, a set of artificial prey items are displayed to the predators. As the presence of the prey cannot be continuously monitored, the experimenter instead

measures proportion eaten at multiple time points, with the lowest proportion eaten considered to be the best camouflaged e.g.(Bond & Kamil, 1998; Cuthill et al., 2005; Stevens et al., 2007). This is the equivalent to measuring detection time – as those objects/prey items that take longer to detect will have the fewest casualties at a given time point. In a lab setting, we can measure how well camouflaged a prey item is by directly measuring both accuracy and time to detection (Bond & Kamil, 2002; Cuthill et al., 2005; Lovell et al., 2015; Penacchio, Lovell, Sanghera, et al., 2015).

2.1.3 Binocular vision and camouflage

An essential component to discuss when considering the evolution of camouflage and camouflage breaking techniques is the evolutionary arms race. The arms race is essentially a competition between predator and prey, where an adaptation on one side is countered by an adaptation on the other. An example is the use of specific behaviours to manipulate prey into startling towards predators, rather than away from them (Catania, 2009). However, it appears that these alterations typically involve morphological changes, rather than changes to the sensory systems (Abrams, 2000). There are several reasons for this: Firstly, in the arms race prey animals are thought to have the advantage – if the predators lose they only miss a meal, whereas if prey fails it loses its life, thus there is a greater evolutionary pressure on the prey item (Dawkins & Krebs, 1979). Additionally, sensory systems are not just for one function, as they are required for the detection of several prey types, communication, navigation etc. Therefore, alterations to a sensory system must not impair other functions (Stevens, 2013), enabling some animals to exploit the properties of their predator's visual systems (Merilaita, 2003; Osorio & Srinivasan, 1991).

Figure 2.3: A mossy gecko (*Rhacodactylus chahoua)* on a tree, flattened against the left hand side of the trunk. Reproduced by kind permission of Dr Hobaiter.

As discussed in the introduction (Chapter 1), a suggested counter adaptation to camouflage is binocular vision (Heesy, 2009; Julesz, 1971; Mckee, Watamaniuk, Harris, Smallman, & Taylor, 1997; Wardle, Cass, Brooks, & Alais, 2010) – this is the ability to extract depth from the environment using two eyes. Being able to see the environment in 3D is thought to allow the shape and therefore presence of an animal to be detected easily (Heesy, 2009; Isbell, 2006; Julesz, 1971), despite any background matching (see point 2 in Section 2.1.1). As required above, the adaptation to use depth from binocular vision does not inhibit other functions of the visual system, and indeed may assist with some tasks such as nocturnal navigation and manoeuvring in trees (see Heesy (2009) for a good overview). If binocular vision is used to break camouflage, then as outlined above we would expect prey animals to be ahead in the evolutionary arms race and possess counter adaptations that would inhibit depth from binocular vision. In the second strand of this thesis, we establish if depth from binocular vision assists with detecting a monocularly camouflaged object. We then move onto investigating if any small alteration to shape or colouring could inhibit depth perception (Merilaita, 2003), thus creating a counter adaptation which we dub 'stereoscopic camouflage'.

Interestingly for the implications of the camouflage-breaking properties of binocular depth perception, in certain circumstances binocular depth is only dominant over other cues to object segregation when these other cues are weak (Dobias & Papathomas, 2013; Pizlo, Li, & Steinman, 2008). This may imply that binocular vision is suited to break camouflage: a well camouflaged animal will cause all the cues to depth and object segregation to fail, leaving binocular disparity as the only remaining cue to the presence of an object.

In order to fully understand the interaction of camouflage and depth perception, we need an in-depth understanding of the how the visual system extracts depth from the scene using two eyes. In this thesis we use the human visual system as an overall model of generic binocular vision. While it is reasonable to assume that human binocular vision is similar to

other vertebrates (Kiltie & Laine, 1992), it is worth bearing in mind that different animals do perceive their environment in different ways (Briscoe & Chittka, 2001; Heiling, Herberstein, & Chittka, 2003; Land, 1997; Thoen, How, Chiou, & Marshall, 2014). In the next Section, we cover the relevant sections of disparity extraction and processing that may help inform us of the interaction of camouflage and stereopsis.

2.2 Binocular vision

2.2.1 What is binocular vision?

Binocular vision is the ability to extract depth information using two eyes – there are many different cues created by having the eyes located in different points in space. When both eyes' views overlap, then there is a region in which the two eyes see slightly different views of the same object, with the images of the objects in the scene falling on different points on the retinae of each eye. By processing the differences in retinal position between the two eyes (called binocular disparity), it is possible to calculate the relative 3D locations of the objects in the scene e.g. (Howard, 2002; Howard & Rogers, 2002). There are other cues to 3D location possible using binocular vision – for example, when an object occludes the background, one eye will be able to see more of the background region than the other. The region visible to only one eye is called a half occlusion or monocular zone (see Harris & Wilcox (2009) for a good overview). This wealth of cues has been well-explored, but in this Section we primarily consider disparity and a discussion on half occlusions.

The impression of depth formed from binocular vision and other sources is called stereopsis: when Julesz wrote about stereopsis and camouflage he was referring to depth caused by binocular vision (Julesz, 1971). However, it is possible to have a vivid impression of depth from just one eye (Vishwanath & Hibbard, 2013). For clarity in this thesis we refer to the impression of depth from two eyes as binocular vision (and avoid the word 'stereopsis' when possible), as this is the source of depth information that has been hypothesized to break camouflage.

An in-depth knowledge of depth from disparity is essential to both strands of this thesis: in the first strand we need to understand what binocular information we know the visual system has available for object perception. In the second strand we need to understand what extra information binocular vision provides that may assist in breaking camouflage, and in turn how these mechanisms may be exploited to provide prey with a counter adaptation.

2.2.2 Definition of disparity

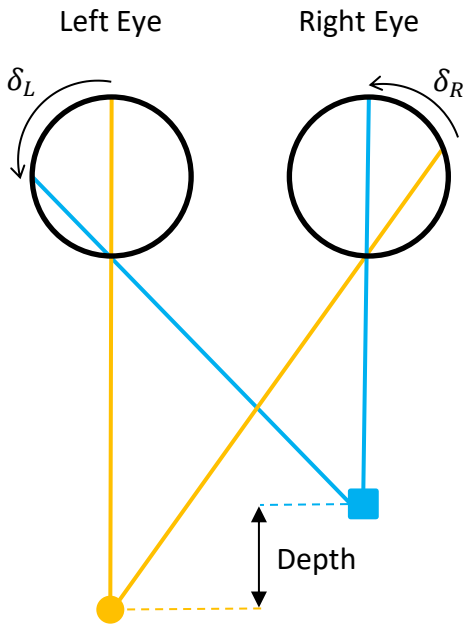


Figure 2.4: Simplified diagram of a pair of eyes (black circles) viewing a scene made of an orange circle and a blue square.

In order to understand how the visual system extracts depth from binocular vision depth, we must first understand what information is available to the binocular visual system. When both eyes of a binocular observer fixate on an object, then the image of the object falls on the fovea of each eye, for example when fixating on the orange circle as in Figure 2.4. To change the fixation from the orange object to the blue object, each eye must move by a different angle. The difference between these angles is called the absolute binocular disparity. The magnitude of this difference is proportional to the depth between the two objects (see Figure 2.4) and is thought to be used very early in the visual system (V1, for reviews of this work, see Deangelis, 2000; Neri, 2004; Parker, 2004).

However, there is a problem with absolute disparity: if the observer changes fixation to the blue square, then the eyes no longer have to move to fixate on the blue square, meaning that the absolute disparity of the blue square is now zero. This means absolute disparity constantly changes as the observer changes fixation, meaning it is not a useful measurement for our experiments.

The use of absolute disparity in experiments is problematic – in order to measure the absolute disparity of an object we have to know the fixation position, and recompute the value every time fixation changes. Instead, we can use a related quantity, called relative binocular disparity. Relative disparity is defined independent of fixation, and instead computes the angular difference between a given object in the scene and all other objects. For example, in Figure 2.4, we can consider disparities relative to our orange circle. The relative disparity between the orange circle and blue square is then $\delta = \delta_L - \delta_R$ no matter where the observer is fixating, and this calculation can be easily extended to many objects. Relative disparity is invariant relative to fixation position and is proportional to the depth difference between the objects. For the rest of this thesis, we measure relative disparity with reference to the plane of the display screen, and call it ‘disparity’ for brevity.

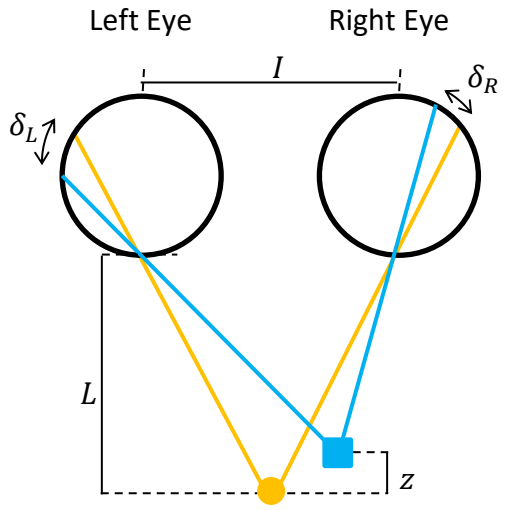


Figure 2.5: Measurements needed to calculate depth from disparity.

Disparity does have another problem, as we are measuring the angular difference between each and every point in the scene. However, there is no way of using these angular dependencies to calculate metric depth between two objects z without first knowing two other quantities: the distance of the object at zero disparity from the observer L, and the intraocular distance I, as shown in Figure 2.5. From these values, it is possible to calculate the approximate difference in depth z (when $I \ll L$ (Harris, 2004)) between the square and circular objects from the disparity δ (rearranged from Howard & Rogers (2002):

$$z = \frac{\delta L^2}{I} = \frac{(\delta_L - \delta_R)L^2}{I} \qquad \text{Eq. 2.1}$$

The dependence on needing a distance measurement to use disparity to calculate the depth difference between two objects presents a conundrum – how do we know the distance to start with? Typically, this distance is estimated using other cues to depth, for example the relative size of familiar objects, or motion information (Brenner & Landy, 1999; Brenner, Smeets, & Landy, 2001; Johnston, Cumming, & Landy, 1994; Landy & Brenner, 2001). This knowledge of distance from the observer is particularly important, as a small change in disparity at a small distance indicates a small difference in depth, while at a long distance it is a much greater difference in depth (as can be seen from the L^2 term in Equation 2.1). While there are normally cues to object distance available in abundance, when they are not the size and scale of an object can become confusing, as shown in Figure 2.6. This effect can create biases in abstract vision experiments and in some situations in the real world, when there are few cues to distance or depth (Cumming,

Figure 2.6: The Pillars of Creation – without familiar scaling cues we have no idea what size these are. The left hand pillar is four light years long. Image reproduced with permission, (NASA, 2014).

Johnston, & Parker, 1991; Glennerster, Rogers, & Bradshaw, 1996; Johnston, 1991). We discuss how we avoid these problems in the methodology Chapter (Chapter 3).

$$\delta = \frac{Iz}{L^2} \qquad \text{Eq. 2.2}$$

It is worth noting that as disparity decreases with the distance squared (see Eq. 2.2), it is typically of most use close to the observer, as this is where disparity can be used for the finest depth discriminations (Ogle, 1958). This means that the camouflage breaking properties of depth perception will be most useful within a short range when the change in disparity δ is significant for the depth of the object z. When the observer is too far away, then monocular camouflage will be sufficient to make the prey item undetectable.

In our previous figures we have considered two distinctively different objects which are easy to distinguish – therefore it is possible to identify which pair of images in the eyes correspond to the same object. However, in a scene with many similar objects, constructing a depth representation suddenly becomes more complex, as shown in Figure 2.7. Here, a pair of identical circular solid objects are placed one in front of another. However, the configuration of images created is entirely consistent with the objects being placed at equal depths as indicated by the hollow circles and dotted lines. From the images alone (if we ignore other cues to depth external to disparity) there is no way to distinguish between these two situations - this is called the correspondence problem (Goutcher & Mamassian, 2005; Marr & Poggio, 1976a; Marr, Poggio, Hildreth, & Grimson, 1991; Seymour & Clifford, 2012). To solve this problem the visual system must consider other cues such as distortion of object surfaces (Vidal-Naquet & Gepshtein, 2012) or nearby distinguishing features (Marr & Poggio, 1976a).

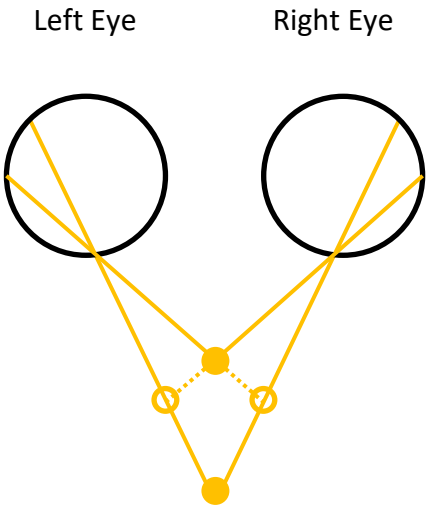


Figure 2.7: The correspondence problem – objects located at the solid circles produce an identical set of images in the eyes as the pair located at the hollow circles.

Another problem with depth judgements from disparity is Panum's fusional area (Panum, 1858). This is the region in which an object appears single. Once an object is placed outside of the fusional area, then it is far away in depth from fixation that the object is no longer fused into appearing as one object, but is rather perceived as two semi-transparent objects, called diplopia. This limit is a limiting factor on the design of experiments, as once the perception of objects in a scene becomes diplopic the judgement of depth is impaired. Typically, the fusional limit lies around 10 to 15arcmin (Ogle, 1952) although reliable disparity discrimination can be made in excess of 1 degree (Blakemore, 1970; Westheimer & Tanzman, 1956). This value is dependent on context (Burt & Julesz, 1980) and is much larger

in the peripheral vision (Blakemore, 1970). Diplopia can also be caused when elements of very different disparities that are placed close to each other – this is called the disparity gradient limit (Burt & Julesz, 1980; McKee & Verghese, 2002). In the experiments presented here we make sure to observe these limits to ensure that it does not cause any problems with the perception of the stimuli.

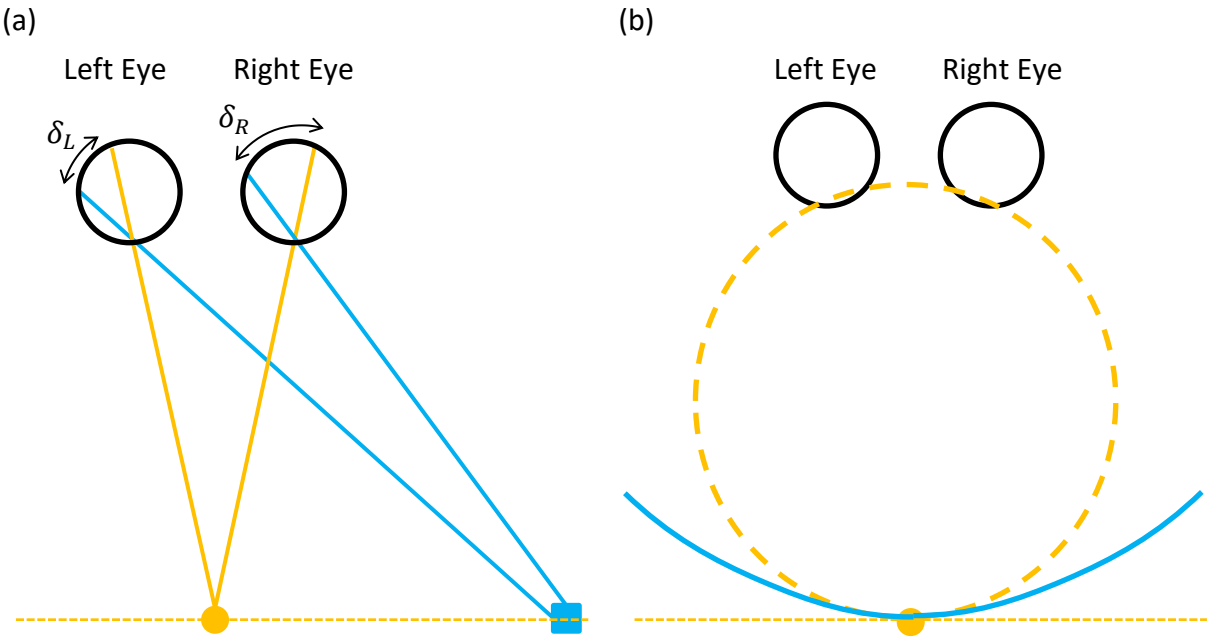


Figure 2.8: a: Despite being on the same depth plane as indicated by the dotted orange line, the disparity between the orange circle and blue square is non-zero –the angles δ_L and δ_R are different. b: The horopter is theoretically a circle centred on the object and the focal points of the eyes, as shown by the dashed orange circle. Experimentally it has found to be a shallower arc similar to the solid blue line.

An additional simplification we have been using in discussing disparity is that all the situations presented so far have considered objects located in front, or almost directly in front of the viewer. When considering disparity relative to a point, any object located at that point has zero disparity. However, if we translate an object laterally as in Figure 2.8a from the location of the orange circle to the blue square, we can see that once the object has been moved sufficiently far laterally along the depth plane (called a front-to-parallel plane, thin orange dotted line), the object's disparity begins to change. We therefore introduce the concept of the horopter – this is the line along which all objects will be of zero disparity. Theoretically, the horopter is a circle, centred at fixation and the focal point of the lenses in each eye (orange dashed line in Figure 2.8b), however in experiments it is found that it is perceptually a much shallower arc (shown as blue line Figure 2.8b)(Blakemore, 1970; Tyler, 1991). In the experiments in this thesis, we present objects directly in front of the observers to avoid problems with the shape of the horopter. This also avoids problems with vertical disparity, where the vertical extent of an object is greater when closer to one eye but not the other, e.g. (Backus, Banks, Van Ee, & Crowell, 1999; Brenner et al., 2001; Cumming et al., 1991; Garding, Porrill, Mayhew, & Frisby, 1995).

2.2.3 Half occlusions

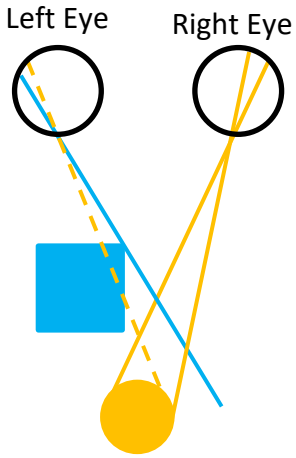


Figure 2.9: The origin of half occlusions: the left eye's view of the orange circle is blocked by the opaque blue square.

The most basic form of half occlusion or monocular zone occurs when an opaque object blocks the view of an object for one eye only. For example, in Figure 2.9, the right eye can see the entirety of the circle. However, the left eye is unable to view the circle as the square blocks the view – in order to see the edge of the circle, the light would have to follow the dashed line, passing straight through the square object. Indeed, the eye cannot see anything below the blue line. This form of half occlusion where one object completely obscures the view of another is fairly common but cannot give much information about a scene; typically the circle is perceived to be as close as possible in depth to the blue square (Nakayama & Shimojo, 1990).

There are many other forms of half occlusions (see Harris & Wilcox (2009) for an in-depth discussion) some of which have been found to be able to cause a percept of depth e.g. (Anderson & Nakayama, 1994; Julesz, 1971; Nakayama & Shimojo, 1990; Tsirlin, Wilcox, & Allison, 2010). Here, however we restrict discussion primarily to half occlusions of a textured background by an opaque binocularly visible foreground, such as in Figure 2.10(a, b), as these are the type that occurs in the stimuli we use.

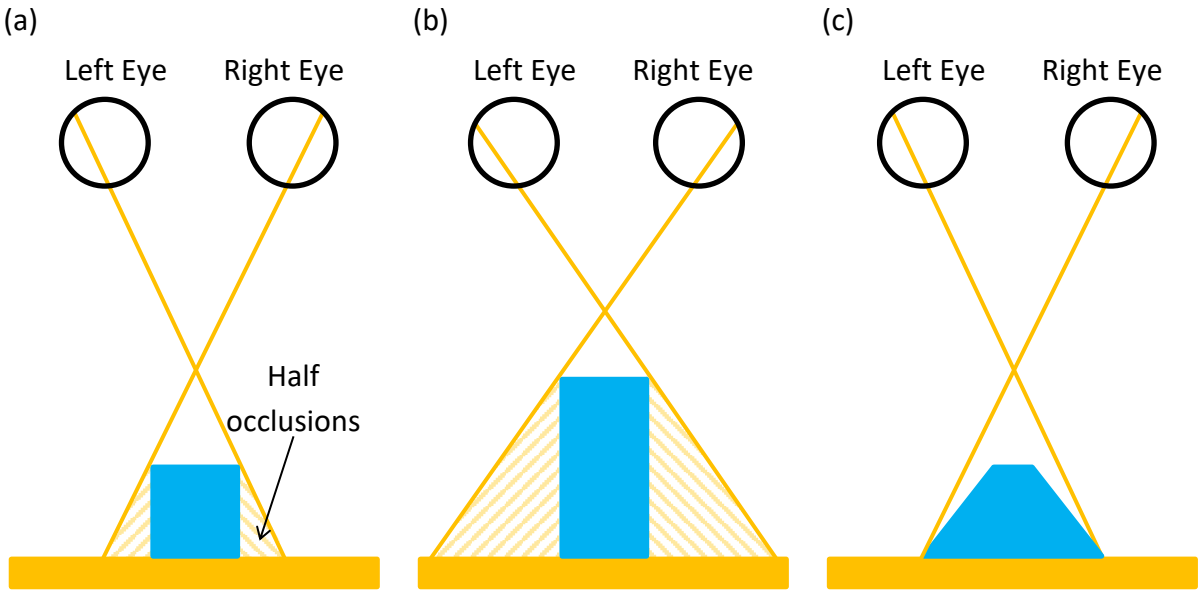


Figure 2.10: Demonstrating half occlusions. (a) An object with less depth has smaller half occlusions (hatched regions) than an object with more depth shown in (b). However, a trapezium with shallow edges (c) never occludes the background to only one eye, thus forming no half occlusions.

When an opaque object is placed on a textured background, the object blocks the view of the background from one eye as in Figure 2.10a. A similar shaped object, but with additional depth will block a greater region of the background Figure 2.10b, thus the size of the half occlusion is proportional to the depth in the object. In these situations, the visual system assumes that the monocularly visible background is of the same depth as the adjacent background, with the occluding object being perceived as having more depth (Anderson & Nakayama, 1994; Collett, 1985; Julesz, 1971; Nakayama & Shimojo, 1990).

Calculating the foreground (the surface of the blue object that faces the view) depth from the size of half occlusions is complicated by the profile of the object. Figure 2.10c shows a simple example where, because of the shape of the object there are no half occlusions present despite the object having the same maximum depth and being viewed at the same distance as in 2.10a. From the current literature, whether an impression of depth is generated appears to be specific to the situation (see Tsirlin et al., 2010 for a discussion). We will therefore design experiments to explore how half occlusions affect depth perception for the kinds of stimuli we use (see Section 4.4).

2.2.4 Early disparity extraction

Having considered the cues available to the visual system to extract depth using binocular vision, we briefly discuss the current knowledge of how the visual system can extract depth. This informs us on what information the visual system can draw on when trying to identify objects in a scene. For a detailed review of the visual system and depth perception, see Parker (2007). In this Section, we present an overview of how disparity is extracted in the early visual system.

The current understanding of disparity extraction is primarily informed though a mix of computational modelling and neurophysiology, and is often concerned with the processing going on in the very early visual system. The first stages of disparity extraction are thought to occur in V1 (Barlow, Blakemore, & Pettigrew, 1967; Nikara, Bishop, & Pettigrew, 1968; Pettigrew, 1980; Poggio, 1995). To understand this process, we consider a basic property of neurons – receptive fields.

A receptive field is, in this case, an area of the visual field in which the chosen neuron responds to changes of stimuli within that area. Stimuli in the receptive field can either excite or inhibit the activity of the neuron, or a mix of the two. For example, a neuron may respond strongly to high luminance in the centre of the receptive field, but the response could be inhibited by the edges of the receptive field also receiving a high luminance stimulus. The output of neurons that respond simply to stimuli can be combined by neurons further up the visual hierarchy to form more complex behaviours, such as edge detectors (Wiesel, 1968). The neurons we are most interested in are those with two receptive fields – one for each eye, as this enables the extraction of disparity. In reality, these receptive fields probably use a mix of inhibitory and excitatory regions that can be described by Gabor filters e.g. (Jones & Palmer, 1987; Qian & Zhu, 1997; Sanger, 1988). Additionally, each

neuron in V1 does not operate independently - the neurons in V1 are arranged in a complex cooperative and competitive set of relationships that helps with the computation of disparity (Samonds, Potetz, & Lee, 2009). However the simplified view is a sufficient to understand basic disparity extraction.

Consider a disparity sensitive neuron with a receptive field centred on the fovea in each eye (i.e. there is zero horizontal offset between the two receptive fields). The neuron will then respond maximally when the stimuli in both receptive fields is the same - for example if both areas are of high luminance. If the receptive field in one eye does not match the receptive field in the other, then the activity of the neuron could be inhibited. As this disparity sensitive neuron has receptive fields placed identically in each eye, then the horizontal offset between the neuron's two receptive fields is zero - therefore it is going to respond maximally when it is presented with a stimulus with zero disparity. In the brain this neuron would be part of a bank of neurons, each with a different horizontal offset between the receptive fields in the two eyes – our exemplar neuron has a zero offset, and detects zero disparity, but it's neighbour might have a small offset and detect a small horizontal disparity. By comparing the responses of a bank of these neurons and selecting the disparity of the neuron with a maximal response, the brain can detect the disparity of a small region of the visual scene. In essence, the visual field is covered by banks of disparity detecting neurons. One can imagine that the extracted disparity at each point on the visual field corresponds to the disparity of the neuron that is responding maximally at that point (DeAngelis, Ohzawa, & Freeman, 1991).

As an aside, there is discussion over whether the pairs of receptive fields belonging to binocular neurons have a lateral separation between the two eyes, or if they are responding to different phases of patterns within each receptive field e.g. (Blake & Wilson, 2011; Chen & Qian, 2004; DeAngelis et al., 1991; Fleet, Wagner, & Heeger, 1996; Read & Cumming, 2007). This detail is not of direct relevance to the thesis, and we do not go into further detail about this discussion.

The arrangements of these neurons are thought to explain some of the basic behaviour of the depth perception. For example, there is a minimum and maximum horizontal offset present between the receptive fields: the minimum offset suggests there is a lower limit for perceiving depth from disparity (Filippini & Banks, 2009; Harris, McKee, & Smallman, 1997; Norcia & Tyler, 1984), and can explain why the visual system is unable to detect depth corrugations of more than 5cpd (cycles per degree) (Tyler & Julesz, 1980). The maximum offset may be responsible for Panum's fusional limit, where features above a certain disparity are perceived as diplopic, as there are no neurons capable of signalling such a high disparity (Joshua & Bishop, 1970; Nikara et al., 1968).

While we have been discussing results primarily found in humans, the methods of disparity extraction seem to be generalizable to many animals: a lot of neurophysiology of the workings of the visual system are informed from studies of primates e.g. (Nassi & Callaway,

2009; Nienborg, Bridge, Parker, & Cumming, 2004; Xiao, Wang, & Felleman, 2003) or cats e.g. (Barlow et al., 1967; Jones & Palmer, 1987; Nikara et al., 1968). The behaviour of the animals in these studies is very similar to that observed in humans, and to our knowledge there are a lot of similarities between the human and other animal's visual systems. This suggests that the results of the experiments we conduct later in the thesis are generalizable to the use of binocular vision in some other animals, particularly our close relatives.

Disparity processing is by no means limited to the V1 and the early visual system (Parker & Cumming, 2001): in fact the responses of V1 cannot explain our sensation of depth from disparity (Cumming, Shapiro, & Parker, 1998). In this thesis we are looking at the relative perception of 3D objects and surfaces – it appears that this occurs much later in the visual system e.g. (Chandrasekaran, Canon, Dahmen, Kourtzi, & Welchman, 2007; Georgieva, Peeters, Kolster, Todd, & Orban, 2009; Welchman, Deubelius, Conrad, Bülthoff, & Kourtzi, 2005).

V2 has particularly interesting interactions for our investigation of depth interaction, as it processes edges and contours (Peterhans & von der Heydt, 1993; Qiu & Von Der Heydt, 2005). Of particular note is that V2 contains neurons that are specialised in detecting depth edges (Von Der Heydt, Zhou, & Friedman, 2000) and calculating the depth ordering of planes (Thomas, Cumming, & Parker, 2002). There is evidence that some cells in V2 use Gestalt cues (see Section 2.2.6) to create representation of local depth order (Bredfeldt & Cumming, 2006; Qiu & Von Der Heydt, 2005). These studies strongly indicate that visual processing after V1 is important in the processing of disparity defined objects, and further indicates that we cannot rely on the models of disparity extraction to explain the perception of 3D objects. Additionally, the processing in V2 is only one step towards segmentation (Bredfeldt, Read, & Cumming, 2008), so we must consider further stages of disparity processing in the visual system.

V3 and V3A have relatively little processing of depth information relative to the earlier areas of the visual system (Anzai, Chowdhury, & DeAngelis, 2011). More disparity processing is performed in V4, where macaque monkeys have been shown to have fine scale disparity discrimination (Shiozaki, Tanabe, Doi, & Fujita, 2012). Of particular interest in V4 are the neurons that are sensitive to the disparity between the centre and surround of a RDS (Umeda et al., 2007). This is relevant to our experiments, as they use RDSs to look at the perceived peak depth of a central depth defined object on a surrounding background – an experiment with very similar characteristics to Umeda et al's. This further confirms that the modelling of V1 cannot fully explain our experimental results, as there is neurophysiological evidence for processing disparity in similar experiments that is completed in V2, V3 and V4.

Finally, there is evidence of complex encoding of 3D surfaces and spatial configurations in V5/IT (Anzai & DeAngelis, 2010; Milner & Goodale, 1998; Orban, Janssen, & Vogels, 2006). There are also neurons in V5/MT that code specifically for the disparity between visual features, as opposed to just disparity present in the scene(Krug & Parker, 2011). The

argument that the early visual system is not the be all and end all of visual processing of depth information also supported by the finding that there is a better link between perception and neuronal activity in MT than in the primary visual cortex (Cumming & DeAngelis, 2001).

In the next Section we review the modelling techniques that have been used to further our understanding of disparity extraction, and consider their successes and limitations.

2.2.5 Disparity modelling

Computational modelling of the early visual system is highly informative about the way in which disparities are processed and perceived. This is of particular relevance to the first strand of the thesis, where we consider what depth-defined objects are, and how they are perceived.

We begin with a short discussion of cognitive neuroscience, and the levels of computational modelling. These were first laid out by Marr in 1976 (Marr, 1976) and have been widely used in the study and interpretation of computational modelling of neural systems e.g. (Churchland & Sejnowski, 1988; Griffiths, Lieder, & Goodman, 2015; Marr & Poggio, 1976b; Poggio, 2012). These levels lay down the different way in which any system that performs a computation can be understood: 1) The basic components, that is how do the neurons or transistors work; 2) The circuitry, how and why are the basic components interconnected; 3) The algorithm, this is the step-by-step processing that is implemented by the hardware; 4) The overall computational goal. While each of these levels are interconnected and knowledge about one can strongly inform another, it is possible to progress understanding at one level without understanding another – for example you can understand how transistors work, but not how they are arranged to calculate a trigonometric function; or know the equation to calculate a point spread function but not know how it is implemented algorithmically.

The most basic models of the early visual system use cross correlational techniques between the two eyes – these models take small areas or windows from each eye's view and perform a cross correlation between the windows. By taking windows with different horizontal separations, then the disparity is calculated as the horizontal offset that creates a maximal response in the cross correlator e.g. (Harris, 2014), similar to the banks of disparity receptive neurons in the early visual system. This simple methodology creates a map of disparities in the visual field, which can then be analysed to consider how the model has performed.

These modelling techniques can recreate many of the fundamental properties of disparity extraction – for example the minimum and maximum limits for perceiving disparity (Banks, Gepshtein, & Landy, 2004; Filippini & Banks, 2009) and can be extended to understand the origins of other features of human stereovision e.g. (Allenmark & Read, 2010, 2011; Anzai & DeAngelis, 2010). The models have been improved until they are highly sophisticated;

adding stages such as spatial frequency filtering (Goutcher & Hibbard, 2014; Kane, Guan, & Banks, 2014) and variable size receptive fields e.g. (Allenmark & Read, 2011). One particularly successful model is the constantly evolving disparity energy model (Read & Cumming, 2003; Tsai & Victor, 2003) which is capable of imitating some very complex behaviour in the binocular visual system, in both human and non-human subjects (e.g. Nienborg et al., 2004).

There are some circumstances in which the current understanding of disparity extraction can cause apparently large scale effects. A prime example of this was discovered by Kaufman et al (Kaufman, Bacon, & Barroso, 1973) when working with random dot stereograms (RDS) – these are a field of randomly located dots displayed separately to each eye, which provide a disparity signal with little to no other cues to form (see Figure 1.2, discussed in detail in Section 3.1). Kaufman et al. used these RDSs to display two overlapping transparent planes of different depth (called stereotransparency). They found that the two planes were perceived as a single plane at the average disparity of the two transparent planes. This effect typically occurs when the disparity between the two planes is smaller than 2-6 arcmin (Parker & Yang, 1989; Stevenson, Cormack, & Schor, 1991; Tsirlin, Allison, & Wilcox, 2008).

The apparent long range interaction between these large overlapping planes can be explained by short range interactions in algorithmic models of the visual system (Harris, 2014; Tsirlin et al., 2008). When the overlapping two planes of random dots are generated, then adjacent elements often have little horizontal separation but are of very different disparities. This means these elements fall within one receptive field, meaning that the neuron with a maximal response (for small separations between the two planes) is then the neuron tuned to detect the disparity that is an average between these two planes as it responds strongly to both elements (Harris, 2014). As would be expected from this result, when the planes are separated into two adjacent opaque planes the effect is no longer observed (Akerstrom & Todd, 1988).

These models imply that the visual system extracts disparity from every point in the scene, and the perceived depth of any given point is then given by the disparity extracted for that exact location. However, this is not the end of the story. The algorithmic models presented here are replicating behaviour in the early visual system, primarily in V1. We know that the behaviour of V1 cannot explain some of the key characteristics of binocular depth perception (Goutcher & Hibbard, 2010; Parker, 2004; Tsai & Victor, 2003). For example, neurons in V1 respond identically to anti-correlated and correlated random dot stereograms, but no depth is perceived in anti-correlated RDS (Cumming & Parker, 1997; Cumming et al., 1998). To fully understand binocular vision, we must therefore consider areas beyond the early visual system.

Of particular interest to us is some of the literature discussed at the end of Section 2.2.4, which shows that the perception of 3D shapes and depth intervals occurs late in the visual

hierarchy. This indicates that in order to fully capture the scope of processing that is involved in detecting and identifying camouflaged objects, we must consider what processing is done after the early visual system, specifically what processing is done to the extracted disparities (Anzai & DeAngelis, 2010). A good way to explore the overall effects of the stages of processing disparity in the visual system is to develop quantitative models to assist in understanding and characterizing the response of the participant to the displayed stimuli. In this thesis we model our results using a general quantitative model with two aims: to have a better understanding of the mechanisms an animal could exploit in for stereoscopic camouflage; and to inform the development of more complex models of the visual system.

In the next Chapter, we review the discoveries that cannot be explained by these models of early disparity extraction. For the first strand of this thesis, we are particularly interested in those relating to the perception of objects, as we wish to investigate what a disparity defined object is, and how it is perceived.

2.2.6 Disparity extraction and object perception

Typically, disparity extraction can account for short range disparity interactions (or apparent long range effects such as stereo transparency, see Section 2.2.5) – normally these effects are attributed to the size of the receptive fields (see Section 2.2.4). Unfortunately, point by point extraction of disparity (which we call a disparity map) does not explain some longer range interactions. One example of this is disparity capture (or interpolation), where areas of ambiguous depth (for example horizontal lines) are displayed flanked by disparity defined surfaces (Georgeson, Yates, & Schofield, 2009; Wilcox, 1999; Wilcox & Duke, 2005). From a disparity extraction point of view, these regions are undefined in depth, and should be perceived as having no depth. However, studies have found that depth is perceived in the ambiguous region and that it is dependent on the disparity of the flankers or surrounding regions (Georgeson et al., 2009; Harris & Gregory, 1973; Ramachandran, 1986; Wilcox & Duke, 2005; Yang & Blake, 1995).

Unexpected biases in the perceived depth of a region do not only occur in poorly defined regions (Yang & Blake, 1995). If two regions of disparity defined depth are placed next to each other, the observers are good at judging the depth difference between the regions. However, if the two regions of disparity are joined with a 'Cornsweet depth profile', as in the solid line in Figure 2.11 (bottom), the perceived depth difference between the two regions is magnified (Anstis, Howard, & Rogers, 1977; Didyk, Ritschel, Eisemann, Myszkowski, & Seidel, 2012; Rogers & Graham, 1983), as shown by the dashed line in Figure 2.11 (bottom). The effect was originally discovered in the luminance domain (Cornsweet, 1970; Craik, 1966; O'Brien, 1958) – the effect can be seen for luminance in Figure 2.11 (top). This depth effect cannot be predicted from simple disparity extraction, as the size of the regions should be sufficient to allow for disparity extraction unaffected by the Cornsweet profile.

Shading and luminance gradients do not only provide analogous effects to the perception of disparity defined depth – as discussed in the introduction, there are many other cues to 3D shape apart from binocular disparity. Some of these cues, such as the size of familiar objects, will not be of much use to detect camouflaged animals, as the camouflaged object has to have already been detected in order to be compared. Others, such as self-motion can cause a perception of depth in an object, but would be self-defeating – the predator may be able to detect the prey item, but the self-motion would give away the predator's presence to its prey (Srinivasan, 1995). The most relevant to our discussion of camouflage of depth defined objects is shape from shading – this is where shadows and shading can give cues to the shape of a 3D object due to the direction of the light source (Norman et al., 2006; Todd, 2004). Indeed, there is an entire camouflage mechanism called self-shadow concealment or Wcountershading (Penacchio, Lovell, Sanghera, et al., 2015; Rowland, 2009; Ruxton et al., 2004; Stevens & Merilaita, 2009) that is thought to have developed in order to reduce or eliminate shading cues (see Section 2.1.1).

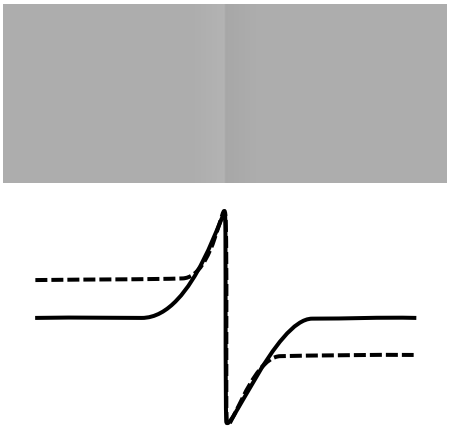


Figure 2.11: Top: Cornsweet illusion in the luminance domain. Both sides have identical luminance. Image reproduced with permission, (Fibonacci, 2005).
Bottom: solid line: cross-section of the Cornsweet profile. Dashed line: representation of the viewer's perception.

Shape from shading has been found to give a vivid impression of depth and surface orientation, due to the direction of the light source (Kleffner & Ramachandran, 1992; Koenderink, Doorn, & Kappers, 1992; Norman & Wiesemann, 2007). We know that the simultaneous presentation of shape from shading and binocular depth perception causes a more vivid impression of depth than either cue presented alone (Bülthoff & Mallot, 1988; Todd, Norman, Koenderink, & Kappers, 1997), implying that the combination may be very effective at breaking camouflage techniques such as background matching. Other studies have looked at situations where shape from shading and binocular depth perception are placed in conflict, and have found that depth from disparity is typically dominant, although in situations where the disparity signal is noisy, shape from shading can be dominant (Chen & Tyler, 2015; Lovell et al., 2012). This is the behaviour we would expect if binocular disparity was used to break monocular camouflage techniques – as discussed above, shape from shading cues can be manipulated by camouflage techniques such as shape from shading or countershading.

As an aside, this combination of shape from shading and binocular depth perception can be looked at from the perspective of cue combination – the brain must take two cues, and

combine them in an optimal way. The exact way in which cues are combined is often complex, and relies on estimates of how reliable each of the two cues are (Hillis, Watt, Landy, & Banks, 2004; Knill & Saunders, 2003). However, a simplified view is that the less reliable a cue, the less weight the visual system places on this cue, resulting in the perception of the object being closer the characteristics of the reliable cue. The estimates of reliability are typically a mix of a judgement of the noise present in the stimulus, and long term judgements of cue reliability (Backus et al., 1999; Gillam, 1968; Knill, 2007; Young, Landy, & Maloney, 1993). If shape from shading is typically judged to be a less reliable cue than disparity, then the depth from disparity would be dominant – as has been observed. However, in this thesis we primarily concentrate on the effects of binocular depth perception and camouflage, so we return to examining the literature on interaction of binocular disparity between different elements.

In the 1980s Mitchison and Mckee did a series of experiments (Mitchison & McKee, 1987a, 1987b) on lines of regularly spaced points and found a the perception of depth in the central dots was dictated by the depth of the dots at the ends of the lines. Under short presentation times the entire line of dots was perceived as having the same disparity as the end points; under longer presentations the depth of the dots in the centre of the line were influenced but not dictated by the disparity of dots at the ends. This goes against the predictions of a point by point depth representation, which would predict the exact estimation of the depth of the central points regardless of the depth of the points at the ends of the lines. Instead, we have a grouping effect, where the perception of entire line is affected by the depth of some of the points. Under short presentations times, the grouping of the points into all having the same depth bears a striking resemblance to the theories of Gestalt grouping, which may provide us with some insight into the origins of these perceptual biases.

Gestalt grouping concerns the collection of elements into objects – an area where disparity maps seem to have considerable trouble. The Gestalt principles of perceptual grouping state that when a set of elements are grouped into single perceived object or collection, they follow certain principles. These ideas were first laid out by Max Wertheimer in 1923 who identified five main aspects listed below (Wertheimer, 1923), each of which we have accompanied with a companion diagram in Figure 2.12.

1. Proximity. Elements located close to each other are grouped together.
2. Similarity. Elements that are similar to each other are grouped together.
3. Good continuity. Elements that are arranged to follow established lines or curves are preferentially grouped over sharp changes.
4. Closure. Elements are preferentially grouped to form closed objects as opposed to ones with gaps or holes.
5. Common fate. Elements that move together are grouped.

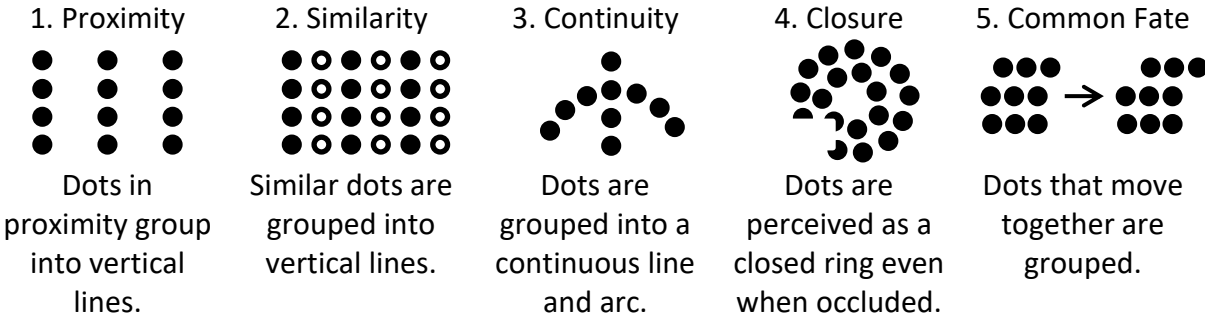


Figure 2.12: Using dots to demonstrate the five main principles of Gestalt Grouping

Several other principles have been added since 1923, for example: symmetry, connectedness (two elements that are connected by a third element) and common regions (elements contained within a common area, e.g. a box) (Wagemans, 2015). Additionally, there appear to be top-down influences in grouping (Beck & Palmer, 2002). As a whole, grouping and subsequent processing on an object level may explain some of the disparity effects discussed here.

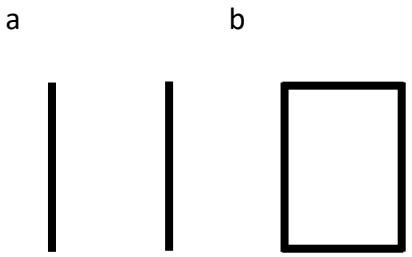


Figure 2.13: Two vertical lines displayed individually (a) are perceived with more depth than when the lines are grouped with horizontal lines (b).

Deas and Wilcox (2014) performed a very interesting study on the effect of grouping and depth perception. Inspired by papers that found horizontal lines or intermediate dots changed the perception of the depth of vertical lines (Fahle & Westheimer, 1988; Mitchison & Westheimer, 1984), they studied the interaction of disparity defined lines and perceptual grouping – participants were asked to judge the depth (defined by disparity) between two vertical lines. The lines were either displayed individually as in Figure 2.13a, and thus not grouped; or displayed with horizontal lines between them as in Figure 2.13b, thus grouping them via closure. Deas and Wilcox found that less depth was perceived when the lines were grouped together – a completely unexpected result, as the addition of horizontal lines does not alter the disparity information present in the vertical lines. This result strongly suggests that disparities are further processed after the elements have been grouped together to from a single object. In a follow-up paper, Deas and Wilcox found objects formed of sets of strongly

grouped elements were faster to detect than objects formed of poorly grouped elements. This was the case despite a decrease in perceived depth in the object (Deas & Wilcox, 2015). Pilzo et al (Pizlo, Li, & Francis, 2005a) also found a similar result: that perceptual grouping has a marked effect on the perception of depth in an object.

Additional processing after early disparity extraction is further indicated by the interaction of depth from binocular disparity and luminance. As discussed earlier, there are strong analogies between several luminance effects and disparity processing effects. This potentially implies that the same mechanisms are being used for processing both disparity and luminance. A few rare studies have even found interactions between luminance and disparity, with luminance edges having been found to effect the perceived depth and depth thresholds (Burge, Peterson, & Palmer, 2005; Didyk, Ritschel, Eisemann, Myszkowski, Seidel, et al., 2012; Peterson & Gibson, 1993). Interactions are not limited to luminance and disparity: depth from motion also affects disparity defined depth (Seymour & Clifford, 2012). If early disparity extraction was the be-all and end- all of disparity processing, then effects such as these would be very unexpected.

These studies give a strong case against a point by point representation of disparity in the visual field being the final stage of disparity processing. We have a variety of different cases in which a simplistic point by point disparity extraction should cause veridical perception. When looked at from the angle of Gestalt grouping, it appears that an additional stage of object processing is dominant over the early levels of disparity extraction. We hypothesize that these biases in perceived depth stem from a mechanism dedicated to the identification of objects. In the first strand of this thesis, we investigate in depth the perception of an object defined solely by disparity to understand how disparity defined objects are perceived, and if poor grouping of the object may cause difficulties in the perception of the object.

2.3 Visual search

Visual search tasks are widely used in the study of visual behaviour, as they give a wealth of data into how different features such as colour and orientation are processed – either independently or in conjunction (McSorley & Findlay, 2001; Wolfe, 1994). In general, a visual search task is when a participant is displayed with a region containing a visual stimulus. The participant is then requested to find a target object within the search area – frequently this is made more complex by the introduction of distractors which may be similar to the target object. By manipulating either the object, or the quantity and similarity of the distractors, it is possible to infer the factors involved in the detection of the target object, and how hard the object is to distinguish from the distractors (Cave & Wolfe, 1990; Treisman, 1982; Treisman & Gelade, 1980; Wolfe, 1994).

In general, observers' performance at a visual search task with distractors is partitioned into two major categories: serial searches, where the time taken to find the target object is affected (typically linearly) by the number of distractors present; and parallel searches,

where the time taken to find the target is unaffected by the number of distractors that are present (Treisman & Gelade, 1980). In parallel searches the target is thought to have a 'pop out' effect, where the target is clearly different to the distractors and therefore comparatively trivial to detect, for example a horizontal line in a field of vertical lines. Serial searches are when an increase in the number of distractors increases the time it takes to detect the target. This behaviour is often caused when the target is defined by a conjunction of two other cues (Treisman & Gelade, 1980), for example when the target is a horizontal red line, not a vertical red line or a horizontal green line (Finlayson, Remington, Retell, & Grove, 2013; Nakayama & Silverman, 1986).

Promisingly for the use of binocular vision to break camouflage, there is evidence that depth information can enable the segregation of the scene into planes that can each be rapidly searched (Nakayama & Silverman, 1986). Segregation in depth also improves the detection of objects superimposed on another scene (Harris & Willis, 2001; Harris & Gregory, 1973; Moraglia & Schneider, 1990; Schneider, Moraglia, & Speranza, 1999; Wardle et al., 2010). However, the use of depth for segregation seems to be limited to large disparities (over 6arcmin (de la Rosa, Moraglia, & Schneider, 2008; Mckee et al., 1997)) and sometimes only in certain circumstances e.g. (Finlayson et al., 2013; Steinman, 1987).

Interestingly, studies looking at disparity judgements in visual search found that people were typically faster when the target was located in front of a background, rather than behind (Becker, Bowd, Shorter, King, & Patterson, 1999; Kim, 2013; O'Toole & Walker, 1997). This is the same geometric arrangement as an object in the environment sitting on an opaque background. This perhaps indicates that binocular vision has evolved to assist in detecting environmental objects, an argument further supported by an increased detection time for convex objects (Bertamini & Lawson, 2008). However, these search tasks all use flat planes in their experiments, rather than three dimensional objects that extend over a range of depths like a real world object. Additionally, the task is frequently to identify a collection of small elements rather than large extended objects.

There is some visual search literature using objects: for example we know it is easier to find a 2D object that is closed than an open one, indicating that the visual system is adapted to detecting entire objects (Bertamini & Lawson, 2008; Elder & Zucker, 1993). The closest object based searches come to considering 3D objects is the work considering shadow perception – for example using abstract 2D objects to cast shadows (Rensink & Cavanagh, 2004). There is also some work looking at the perception of photographs of stones with cast and self-shadows (Lovell, Gilchrist, Tolhurst, & Troscianko, 2009). Unfortunately, these studies investigated if shadow processing uses a separate, faster mechanism than object detection rather than considering the effect of any percept of depth caused by the inclusion of shadows.

Unnatural search tasks are a concern, with the majority of visual search tasks being presented on simple backgrounds (Troscianko et al., 2009a; Wolfe, 1994), or on uniform

grids e.g. (McSorley & Findlay, 2001). Neider et al. did two studies in naturalistic scenes, and found that the effect of the extra information from the naturalistic background could not be solely attributed to the presence of crowding from the extra elements in the scene (Neider, Boot, & Kramer, 2010; Neider, Brotzen, & Zelinsky, 2010). This indicates that it is important to consider the effects of the environment that the target object is located in. Lovell et al. (2015) conducted some naturalistic studies investigating the detection of a snake that was camouflaged via a leaf-like pattern on the snake that matched an artificial leaf background. Interestingly, they managed to make the snake so well camouflaged against an artificial leaf background that they needed no distractors, and could measure how well camouflaged the snake was from detection time alone (Lovell et al., 2015).

Search tasks in the environment using only visual cues are rare, one such study by Foulsham et al (2014) did a search task outside of the lab, using an eye tracker to study the effects of colour on search times (Foulsham et al., 2014). The presence of an eye tracker assisted Foulsham et al. to detect trials when participants looked at the target but did not recognise it.

To us, the most useful naturalistic world search tasks are those that relate to camouflage. Several studies have investigated if visual search can predict participants' performance at a foraging task (active searching of an environment for 'food'), and have found that if the foraging is mainly visually driven, then visual search is a good analogue to foraging for natural objects (Gilchrist, North, & Hood, 2001; Smith, Hood, & Gilchrist, 2008). Naturally, the effect of the environment is important when considering visual search and camouflage, with reaction times increasing with increased similarity of the target to the background (Neider & Zelinsky, 2006). This is caused by background matching – a form of camouflage. Despite these attempts, there is little literature relating camouflage to complex scenes or with discrete objects, with most inferences about visual camouflage being drawn directly from simpler objects and tasks (see Troscianko et al. (2009a) for a discussion).

In developing the visual search tasks presented in this thesis, we heavily drew inspiration from the more complex camouflage tasks, especially those that were constructed with a real environment in mind (Lovell et al., 2015; Neider, Boot, et al., 2010; Neider, Brotzen, et al., 2010). However, perhaps the most influential were the meta analyses discussing the advantages and disadvantages of different visual search paradigms. One common paradigm to use is a search task where distractors are always present but with the target only present in some of the trials e.g. (Gilchrist et al., 2001; Neider & Zelinsky, 2006; O'Toole & Walker, 1997; Rensink & Cavanagh, 2004; Wolfe, 1994). The participant's task is then to identify if the target is present or absent, requiring only a very simple experimental setup. However, this methodology has been found to have serious issues – the prevalence of the target makes a significant difference to both detection rates and reaction times e.g. (Godwin et al., 2014; Schwark et al., 2013) making it hard to analyse the effect of manipulating the target vs the effects of target prevalence.

Additionally, in camouflage tasks we typically wish to compare the participant's performance between several objects to establish if certain attributes cause the target to be better camouflaged. We decided to avoid placing multiple objects in each search task as this introduces problems with interpreting the results. For example, multiple object searches cause an increase in miss rates for the rarer or harder to detect targets – particularly when the participant is tasked to indicate if the object is present or absent (Cain, Adamo, & Mitroff, 2013).

The inclusion of distractors in the visual search task can influence how camouflaged a target object is measured to be. From a study by Neider and Zelinsky (2006) we know that when searching for a camouflaged object, participants will spend time investigating distractors instead of the target similar background. This is a problem, as the location and shape of the distractors will confound our measurements of the target's camouflage. For example, if we have multiple different targets, then the targets that are most similar to the distractors will take longest to be spotted purely due to our choice of distractors, and not due to any inherent properties of the target. For these reasons, we avoid tasks with distractors and tasks where the target is ever absent. We discuss in depth how we combine these studies into a camouflage visual search task in Section 3.2.2.

3 Methodology

- **Outline of stereoscopic stimulus presentation.**
- **Overview of the psychophysical techniques used in the first strand to extract data from the two alternative forced choice experiments.**
- **Overview of the visual search tasks used in the second strand of this thesis.**
- **Participant recruitment.**

Figure 3.1: A Dead Leaf Mantis (*Deroplatys desiccate*, two central brown leaves) displaying mimicry. Animals imitating dead leaves such as this mantis are frequently very thin and flat, making them appear the right shape for a leaf even to a stereoscopic viewer. Image reproduced with permission, (Pingstone, 2005)

In this Chapter, we will review the general methodologies used in the thesis. First, we will describe the setup used to display controlled stereoscopic stimuli on a computer screen. The techniques of stereoscopic display will be used in all but the last Experiment (11). We then go on to discuss the methods of data analysis. Due to using different experimental paradigms in the two strands of the thesis, we discuss the data analysis methods in two separate Sections: For the first strand we describe the psychophysical analysis techniques, for the second strand we discuss the data analysis techniques used for visual search tasks. In the final Section of the Chapter, we discuss participant recruitment techniques, which are identical for both strands of the thesis.

3.1 Display of stereoscopic stimuli

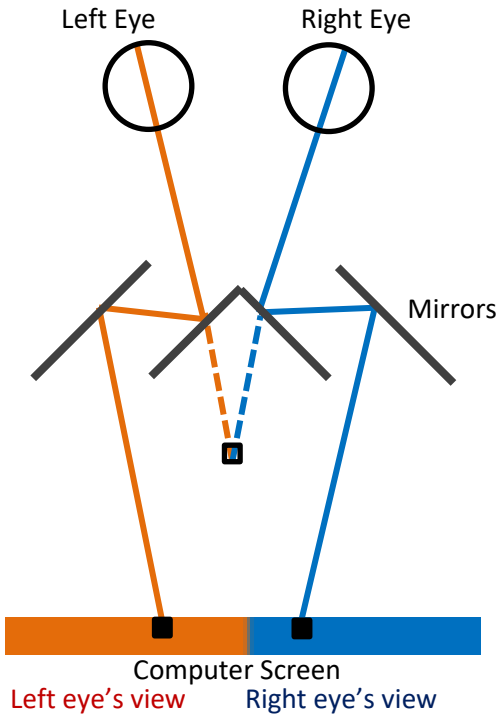


Figure 3.2: Ray diagram of stereoscope: a point in the left half of the screen enters the left eye at an angle that is different to the angle of the ray at the right eye. The visual system, without information about the mirror setup sees an object in depth in front of the computer screen.

The objective of our experiments is to test the visual system's use of depth defined by binocular disparity. To do this, depth must be delivered in a highly controlled manner, or other cues to form and depth may be present. Here, we discuss how this is achieved using random dot stereograms (RDS) (Julesz, 1971) and a Wheatstone Stereoscope (Wheatstone, 1838). Random dot stereograms (for an example, see Figure 1.2) are of particular use as a display method, as they can display binocular disparity in isolation from other cues (Julesz, 1971).

For all computer based experiments, stimuli were displayed using the same apparatus, except that the screen was changed for the visual search tasks in the second strand of the thesis (details in situ). A completely different setup without a computer screen used to create a real world task in Section 9.3 (Experiment 11), this setup is discussed in Chapter 9.

In order to display images with stereoscopic depth on a computer screen, we must present each eye with different but overlapping images – i.e. the images originate from different locations on a computer screen, but are projected into the eyes via mirrors as if they come from the same location (Figure 3.2, detail on exactly how this works is discussed below). To

achieve this, images for the left and right eyes were presented side by side on a luminance calibrated monitor (monitor details are at the beginning of each strand in Sections 4.2.1 and 7.2.1). In order to make the images for each eye appear to overlap as if they originate from the same location in space, they were reflected into stereoscopic presentation using a Wheatstone stereoscope (Wheatstone, 1838)(see Figure 3.2). Viewing position was stabilized with a chin rest placed 1m from the screen. If needed due to differing intraocular distances, the participant adjusted the central stereoscope mirrors to obtain comfortable fusion.

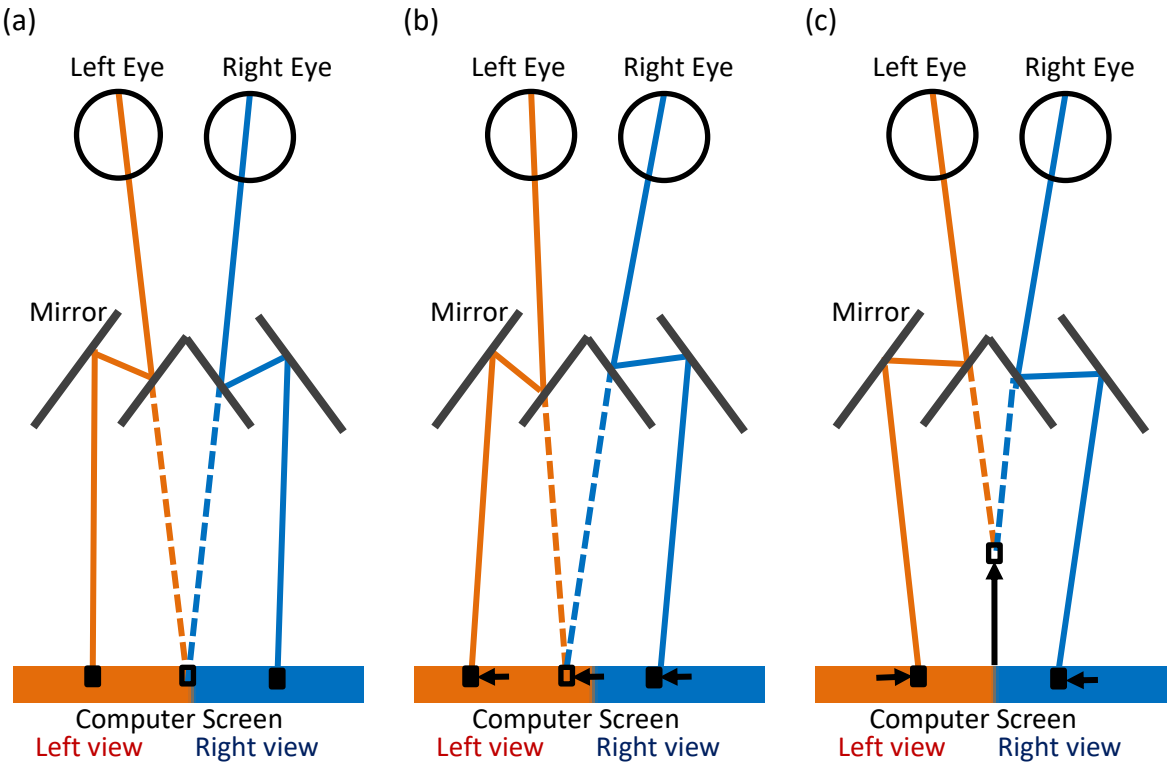


Figure 3.3: Effect of dot movement for a Wheatstone stereoscope. (a) A dot positioned at the centre of each half of the screen (small black squares) is perceived at the centre of the computer screen (hollow black square). (b) Moving dots in each half of the screen to the left moves the perceived dot to the left, as indicated by the arrows. (c) Moving the dots in opposite directions in each eye moves the perception of the dot away from the plane of the screen.

A Wheatstone stereoscope works as follows: When a dot is displayed at the centre of the left half of the screen, and an identical dot is displayed at the centre of the right half of the screen viewer perceives a single dot on the surface of the screen directly ahead of them (Figure 3.3a). If we change the position of the dot, say to the left, in both eyes, then the viewer perceives the single dot at the centre moving towards the left (Figure 3.3b). If we move the dot in different directions in both eyes – for example to the left in the right eye and to the right in the left eye (Figure 3.3c), then the image of the dot falls in different places in the retina in each eye, introducing a disparity cue. As the visual system does not

know about the presence of the mirrors, it assumes that the difference in relative dot position between the two eyes is due to a single dot located in depth between the viewer and the screen (Figure 3.3c). We can use this setup to create the perception of one or multiple disparity defined dots (or objects) in depth relative to the screen. By varying the amount of horizontal displacement we give a dot in each eye, we can vary the apparent position on the screen and in depth, allowing us to accurately control exactly where in depth the viewer perceives the dot.

Many stereoscopic experiments suffer from cross-talk. Cross-talk is most simply described with red/blue anaglyph glasses – these are glasses with a red filter over the one eye and a blue filter over the other (see Figure 3.4). When looking at an anaglyph, such as in Figure 3.5, then the viewer gets a perception of depth in the image. This perception of depth relies on each eye seeing only either the red or the blue of the printed anaglyph. However, frequently the filters on the glasses allow a faint red image though the blue filter or vice versa. This faint image from the other eye can interfere with the perception of depth or cause eye strain and discomfort e.g. (Seuntiëns, Meesters, & IJsselsteijn, 2005; Woods, 2012; Xing, You, Ebrahimi, & Perkis, 2012). Cross talk from overlapping images is not a problem in the Wheatstone stereoscope (Woods, 2010), as the two images originate from separate physical locations.

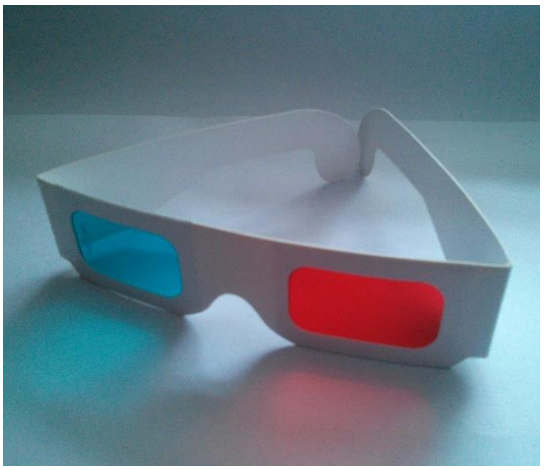


Figure 3.4: Red/blue anaglyph glasses. Image: own work.

Figure 3.5: Anaglyph of an Asiatic hybrid Lilium. To see depth, wear anaglyph glasses with the red filter over the left eye. Image reproduced with permission, (Harrison, 2008)

Wheatstone stereoscopes do have some disadvantages. It is possible for the observer to sometimes see segments of the other eye's view in the edge of the field of view. This cross-viewing can give away information about the properties of the stimulus or distract the participant. In these experiments, we reduce this effect by applying black tape to the two mirrors. However, a small amount of cross-viewing across the central line of the monitor will still have been present due to the individual differences in

participant's intraocular distances. For this reason, the central region of the screen is presented as being blank.

In order to isolate the binocular disparity cue, our experiments used random dot stereograms (RDS (Blake & Wilson, 2011; Gantz & Bedell, 2012; Howard & Rogers, 2002; Julesz, 1971). RDSs (for an example, see Figure 3.6) are very abstract but do deliver disparity defined-depth without any other information to form, either in depth or monocularly. To create a random dot stereogram, a field of random dots at the required density is generated, and then duplicated into left and right displays. Disparity is then added by taking the desired dots and moving them laterally to the left in the right eye's view and vice versa in the left, creating a binocular disparity signal when the two halves of the stereogram are viewed simultaneously through the stereoscope (as in Figure 3.2).

When the change in disparity at the edge of an object is large, this process leaves an area with no dots on one side where the dots were moved from and an area with double density on the other where the dots were moved to. To avoid this, duplicate dots from the background (i.e. those unmoved by adding disparity to the image) are deleted on the high density side and regenerated randomly on the sparse side. This maintains dot density but creates half occlusions (areas of the image only visible to one eye, see Chapter 2.2.3), as would be seen in an opaque real object. With this manipulation, RDSs have the advantage than when under monocular vision, there is no cue to object form: all that can be seen is a field of equal density random dots as can been seen in the experimental screenshot in Figure 3.6. However, it does introduce the disadvantage that we must now control for half occlusions – a potential cue to depth e.g. (Harris & Wilcox, 2009; Tsirlin et al., 2010). Please note that Figure 3.6 is the only true screenshot displayed in the thesis – stimuli in the experimental Chapters are representative of those displayed to the participants but are mock-ups created with exaggerated disparity to ensure visibility on an A4 page.

All experiments were coded in MATLAB® (2013 or 2014) with stimuli being coded and displayed using the psychophysics toolbox (Brainard, 1997; Kleiner, Brainard, & Pelli, 2007; Pelli, 1997). This is a frequently used toolbox within vision science, and allows for accurate and timed presentation of stereoscopic stimuli.

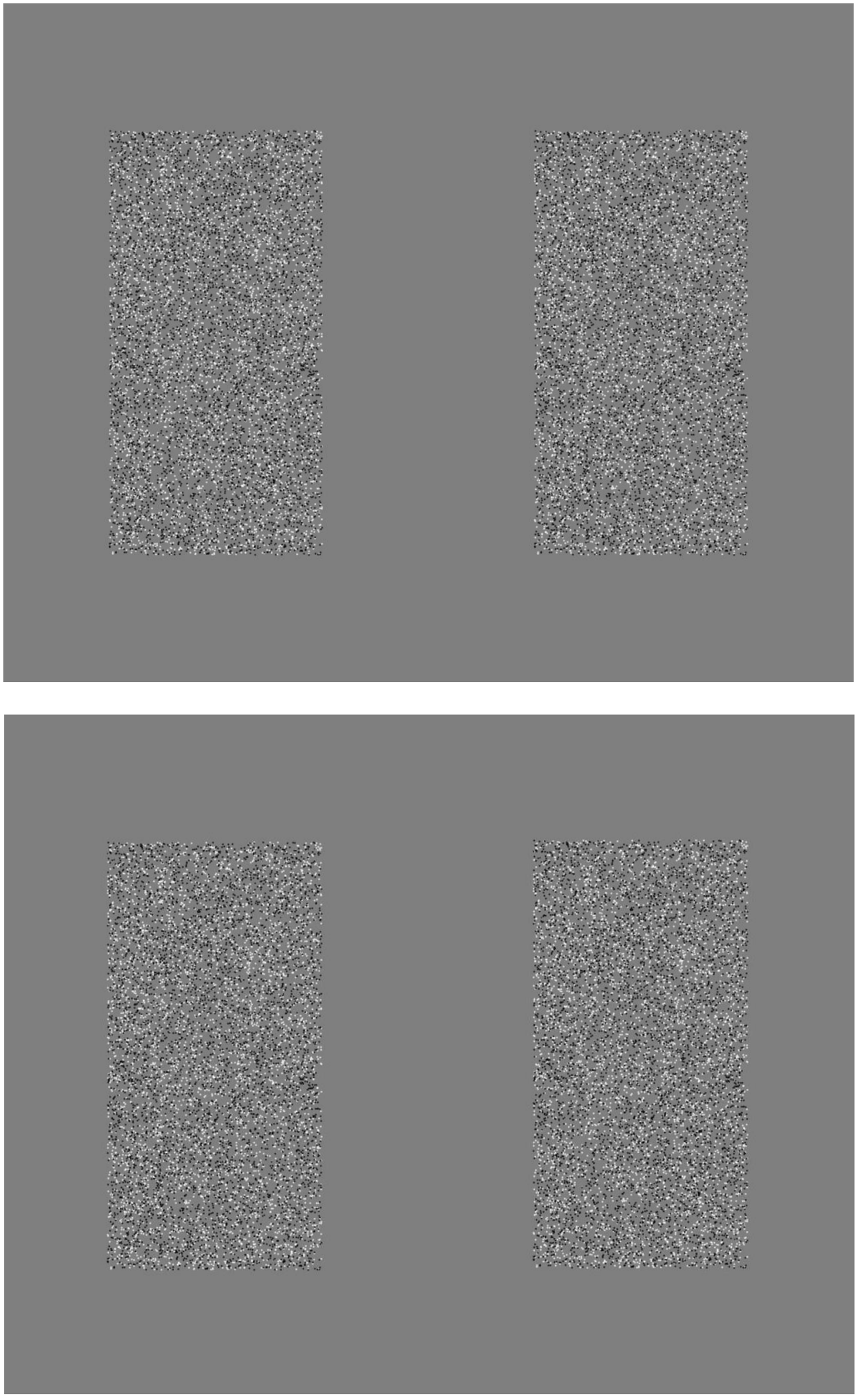


Figure 3.6: Top: A screenshot of the random dot stimulus used in Experiment 1 (Section 4.3). When viewed through the Wheatstone stereoscope, the left half of the image would only be viewed by the left eye, and the right half only by the right. Divergent (top) or cross (bottom) fusion gives a similar percept of two square shapes protruding from the RDS.

3.2 Data analysis

There are two main methods of collecting and analysing data in this thesis: psychophysics, which is used in the first strand of the thesis to characterise the perception of an object; and visual search tasks, which are used in the second strand of the thesis to investigate the interaction of camouflage and depth perception.

3.2.1 Strand 1: Psychophysics

Psychophysics is a paradigm that develops a numerical relationship between a measurable property of a stimulus and the perception of that property (Kingdom & Prins, 2009). This is typically achieved by systematically altering the property of interest in a stimulus, while the participant makes very simple choices such as: can you see the stimulus, or which stimulus is brighter. Here, we primarily consider what is called the two alternative forced choice paradigm (2AFC) using a method of constant stimuli.

The two alternative force choice methodology revolves around displaying two stimuli simultaneously and asking the participant to make a simple choice between the stimuli, such as 'which is brighter'. The participant is forced to choose between the two options, even when the participant does not know which of the two options is correct. The advantage of this is that participants are often well above chance level when uncertain, so forcing a choice allows us to calculate how often they choose each stimulus at the displayed level of the stimulus (Kingdom & Prins, 2009). We can get an idea of the participant's perception by plotting the property of interest (for example in Figure 3.7 we use brightness of stimulus A) against the percentage of times the participant chose one of the two stimuli. The orange line in Figure 3.7 is an example of when the participant is uncertain if stimulus A or B is brighter, but still chooses stimulus A the majority of the time.

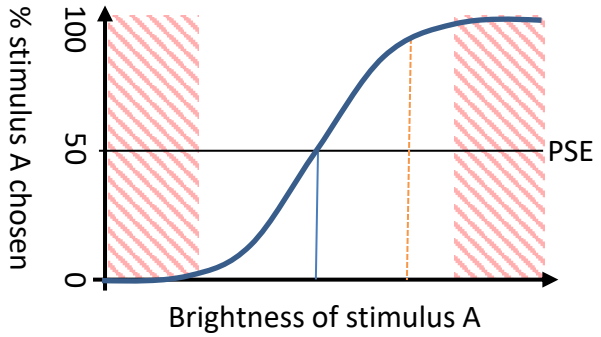


Figure 3.7: Diagram of participant choosing an arbitrary stimulus A as brighter than stimulus B as a function of the brightness of stimulus A. In the red hashed region, the participant is confident and always chooses one of the two stimuli as the brightest. The vertical orange line is an example of when they are uncertain but still above chance level. The blue vertical line indicates the brightness of stimulus A when the participant responds as if they perceive the two stimuli are identical, called the PSE.

We are interested in the point at which participants can no longer tell the difference between the presented stimuli, as indicated by the blue solid vertical line in Figure 3.7. This point, called the PSE (point of subjective equality) indicates when the participant perceives both the presented stimuli to be identical. For a two alternative forced choice experiment,

the PSE is determined by when they choose each stimulus with a frequency of fifty percent. 2AFC methodology also gives a measure of how sensitive the participant is to small changes in the stimulus. This is called threshold, and is calculated by measuring how rapidly the participant changes between choosing one stimulus over another as the stimulus level varies. We go into greater detail on threshold later in this Section.

The 2AFC methodology also helps solve the problem posed in Chapter 2.2.2, that in order to make judgements of the depth from disparity, we need to know how far away the object is – normally estimated by other cues to depth such as the size of familiar objects. This is a problem when displaying abstract stereoscopic stimuli in a dark room there are few cues to distance, therefore making absolute judgements of depth are highly problematic. We avoid this problem by not asking the participants to make an absolute judgement of depth of one object, but rather to distinguish between two objects displayed simultaneously, thus enabling them to make accurate judgements between the stimuli (Glennerster et al., 1996). By joining the background of the two objects, we make it clear they are originating in the same depth plane, therefore eliminating any bias between them that might be caused by incorrect distance judgements.

Although the 2AFC methodology theoretically allows the measurement of point of subjective equality by displaying stimuli at all possible stimulus levels, this is highly impractical. Instead, we need a method of charactering the participant's response over a large range while minimising the number of measurements needed. There are many techniques of doing this, but in this thesis we use one of the simplest, called the method of constant stimuli, rather than the more complex staircase methods. We discuss the reasons for this below.

The method of constant stimuli relies on pre-setting a range of stimulus values to be displayed. The participant is repeatedly presented with each different stimulus value (the value presented is randomised from the pre-set range each trial), with enough repeats that their percentage of each stimulus chosen at each stimulus value can be assessed. From this data, the PSE can be interpolated between the pre-set stimulus values typically by fitting a cumulative normal function e.g. (Glennerster & McKee, 1999; Harris, 2014; Harris, Chopin, Zeiner, & Hibbard, 2012). The diagram in Figure 3.7 is a mock-up of a cumulative normal curve. This method characterises the participant's responses over the tested range and is simple to run, although it results in slow data collection.

In comparison, staircase methods rely on the principle of adjusting the stimuli to make the task easier when the participants make an incorrect judgement, and making the task harder when the participants are correct. Theoretically, in the simplest case the staircase should converge to near the PSE, with the participant switching between being correct and incorrect each time the task is made easier or harder. However, there are many complexities to this method, with different ways of implementing the staircase having different advantages and disadvantages (Kingdom & Prins, 2009). When starting the first

Experiment we therefore used the method of constant stimuli, with the intention to switch to a staircase method if faster results collection was required. However, we found the method of constant stimuli could collect sufficient data in the time allowed, therefore never requiring us to use the staircase alternative.

More specifically, the method of constant stimuli we used presented two stimuli simultaneously on the screen, one placed above the other with the participant asked to make a judgement between them. Unknown to the participant, one of the stimuli (the comparison stimulus) had exactly the same properties in every trial, while the second varied according to pre-set values (the test stimulus). The position of the comparison stimulus was randomised between the upper and lower position (with the test stimulus being placed in the other position) to obscure the constant nature of the comparison stimulus and avoid biases in perception between the upper and lower positions e.g.(Bian & Andersen, 2010; Goutcher & Mamassian, 2002; Harris et al., 2012; Hibbard & Bouzit, 2005). We measured how often the participants chose the comparison stimulus as a function of the varied parameter of the test stimulus.

Once sufficient data is accumulated for each value of the parameter we must then interpolate between the collected data points in order to calculate the point of subjective equality. To achieve this, we use Palamedes toolbox: routines for psychophysical analysis (Prins & Kingdom, 2009), which is an established analysis toolbox for MATLAB®. We first used Palamedes to fit a cumulative normal curve to the participant's data using a maximum likelihood model, the fitted curve was then used to calculate the PSE. Additionally, Palamedes was used to bootstrap error margins on the calculated value PSE. We can measure the bias in perception of the stimulus by measuring the shift in PSE from the objective value at which the test and comparison stimulus are identical, as in Figure 3.8a.

By fitting a cumulative normal we have an additional measure of use – the gradient of the slope near the PSE, called the threshold. Our estimate of threshold, related to the slope of the curve, is calculated as half the difference between the stimulus level at 75% and 25%. Threshold is a measure of how sensitive the participant is to the manipulations made to the test stimulus – a more sensitive participant will have more of a difference in performance between adjacent stimulus level, increasing the slope of the function at the PSE. The difference between a more and a less sensitive participant is shown in Figure 3.8b.

In order to eliminate any trade-off effects where the participant might perform the task faster in order to finish earlier, while another might be slower, we presented each stimulus for a set amount of time. This forced the participants to all spend the same amount of time viewing the stimulus. A 2s stimulus duration was settled on as being sufficient for the participant to view the stimuli and make an informed decision, while being short enough to collect sufficient data in each one-hour session.

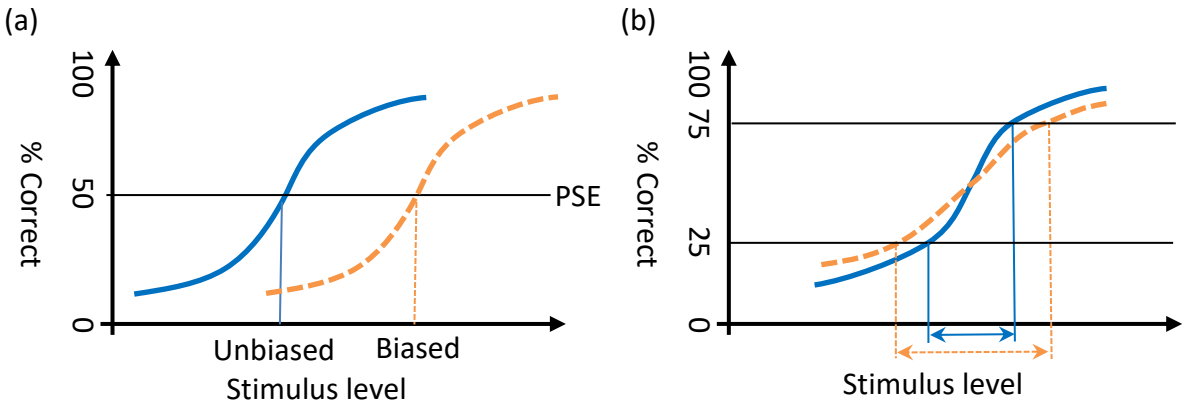


Figure 3.8:Pictorial representation of the alterations made by (a) biased percept (orange dashed) compared to an unbiased (solid blue). Stimulus level at 50% correct is the PSE. (b) Participants that are less sensitive (orange dashed) or more sensitive (blue solid). Arrows indicate the size of the thresholds for each of these lines.

Due to the naïve nature of the participants, we chose not to control where the participants fixated over the 2s stimulus presentation as without including eye tracking we had no way of verifying fixation. This should not be a problem, as our motivation has a real world context where there is no control over fixation, thus controlling fixation would be an unnecessary abstraction from detecting camouflaged objects in the environment.

There are three main disadvantages of the method of constant stimuli we had to overcome:

1. The data needed to be piloted on several participants to determine the range values that need to be displayed before the experiment could be run – however this is not a major disadvantage as all experiments require a certain level of piloting. This means we did require the pilots to either be inexperienced stereoscopic observers or, in some circumstances, a mix of inexperienced and experienced stereoscopic observers.
2. As the depth of the stimuli are pre-set to certain values and constant through the experiment, there are inevitably some stimulus values that are at the ceiling or floor of measurement for some participants. This reduces the amount of useful data collected. In the end, this merely increases the experiment duration, but we have tried to keep ceiling and floor measurements to a minimum by piloting. For the experiments presented here, the variation between participants is small enough that this problem occurs infrequently and does not represent a large portion of the time spent collecting data.
3. Some outlier participants with particularly poor (or good) stereovision had PSEs outside of the measured range. We kept this problem to a minimum by screening participants with the TNO test (see Section 3.3) and rejecting any participants with thresholds above 240 arcmin. This is because our stimuli can only be perceived using stereoscopic vision, therefore someone with poor binocular vision will be unable to see the stimuli or complete the tasks. Secondly, without further testing we do not

know why certain individuals have poor binocular vision, and therefore we would not be able to correctly interpret why their performance differs from the more typical participant. We chose such a large threshold as the performance of a participant changed greatly dependent on the lighting in the room when performing the TNO test. Additionally, the TNO test used stimuli sufficiently different to the experiment that some participants who had trouble passing the TNO test had no difficulty perceiving depth in our stimuli, and vice versa. The TNO test was therefore a high pass filter to ensure consistency and a reference point to other experiments, while we performed further screening in a demo stage of our experiment.

3.2.2 Strand 2: Visual search

Visual search tasks were used in conjunction with the psychophysics used in this thesis. The psychophysical analysis can give us a quantitative measurement of how changing the stimulus changes the perception of the stimulus, it cannot tell us how much harder or easier that alteration makes the stimulus to detect. We therefore need to move to visual search experiments in the second strand of the thesis, where we are looking at the interaction of binocular vision with camouflage.

A visual search experiment is one in which the participant's task is to locate a target in the presented visual scene as quickly as possible. Typically, an object on its own is very easy to spot, so other objects that look like targets are often introduced into the scene – these are called distractors. The participant's task is to find the target amongst the distractors. There is a wide variety of literature using this paradigm, which was discussed in depth in Section 2.3.

The typical visual search task aims to compare how easy it is to detect the target based on certain properties, e.g. if a red or blue target is easier to spot amongst green distractors. There are various ways to evaluate the participant's performance – one technique is to evaluate how much longer it takes to spot a target with an increase in the number of distractors, giving an overall measure of how easy it is to distinguish between the distractors and the target (see Section 2.3) e.g.(Neider & Zelinsky, 2006). In our experiments this presents a problem as we typically need multiple different shaped targets as we are comparing participants' performance across many conditions. Specifying exactly what the target looks like to a naïve participant is therefore extremely hard, and it is known that multiple targets can introduce many extra sources of error (Cain et al., 2013) including biases based on the similarity between the distractors and the different targets (Duncan & Humphreys, 1989; Lovell, Gilchrist, Tolhurst, To, & Troscianko, 2008). Additionally, this regime presents issues: the distractors are examined in preference to areas of the background which may contain an undetected target (Neider & Zelinsky, 2006). Additionally, measuring how much harder distractors make the target to spot is, in camouflage terms,

more akin to the various forms of mimicry and therefore dependent on the exact properties of the distractors as well as the properties of the target (see Chapter 2.1).

Unfortunately, there are few examples of combining binocular vision and camouflage with visual search tasks. The few studies that do consider depth and visual search typically look at the effects of depth on spotting a 2D target amongst 2D distractors, and investigating how the depth of the target relative to the distractors changes the difficulty of detecting the target (e.g. Finlayson et al., 2013; Kim, 2013; O'Toole & Walker, 1997; see Section 2.3 for a discussion of this literature). While of relevance, this is not the regime we are after, as our targets are aiming to resemble prey items, and therefore must have their own distinct 3D shape. We must therefore turn to a different regime for testing how hard it is to spot a disparity defined 3D object on a featureless background.

To achieve this, we altered a much simpler paradigm used by Lovell et al. (2015) – no distractors with the target always present, and looking at the time it takes to detect the single target object. Across trials, we alter the properties of the object that we are interested in: the participant does not need to identify each of these objects as a separate target, as we have only instructed them to find an object in the scene. If the target is indeed camouflaged by our experimental manipulation, then it will be harder to spot and the time it takes for the participant to indicate they have spotted the target will increase.

We refer to the time between stimulus onset and the participant's response as their reaction time and equate this to the detection time (in the context of camouflage, this is the length of time it takes for a predator to spot their prey). Unfortunately, reaction time also includes the length of time it takes for the participant to make a motor response to indicate that they have spotted the target. However, this should be negligible in comparison to the overall reaction time (measured in seconds) and should be similar for all different stimulus conditions.

Some studies use eye tracking to further break the reaction time down into the time until first fixation on the target, and the length of time spent fixating on the target to positively identify it as the target, called the verification time e.g. (Neider, Brotzen, et al., 2010). This can give us some further information about how the participants are detecting their targets – for example Neider et al. found that verification time increased as the scene searched got more cluttered. In the absence of eye-tracking, we cannot distinguish between the time to first fixation of the target and the time spent verifying that the fixated object is the target. We argue that in a camouflage context the distinction between first fixation and verification time is not essential. This is because in order to react to the presence of a prey item, the prey item must be identified as being the predator's target. In order to identify the prey, the predator must fixate and verify the target. We consider the length of time it takes to identify the target as a measure of camouflage - time to identification includes both the length of time till first fixation time and the length of time to verification time. Therefore, how well

camouflaged an object is, is not influenced by the distinction between how much time is spent fixating the target, and how much is spent verifying it.

It is frequent practice in visual search experiments to exclude outliers in reaction time on the grounds that they are typically due to participants not paying attention. However, in our data all the long response times were present in one condition of our experiments (specifically the smoothest objects in Experiments 9, 10 and 11). This indicates that the long response times are due to the experimental stimuli, not participant inattentiveness and any attempt to remove outliers is invalid. To ensure that outliers do not bias our data however, we will compare the trends in both the mean and the median of the data, as the difference between the mean and median results will inform us about the effect of the spread of the data.

3.3 Participants

Participant details presented here are used in all experimental Chapters of the thesis, unless specifically noted otherwise.

Participants were recruited via the University of St Andrews' online recruitment system (SONA) and were recompensed £5 per hour for their time. Due to the nature of the recruitment, the majority of participants were students, although some older participants were recruited thought the study. This should not present a problem to our results for two reasons: Firstly, the binocular visual system is functionally comparable and only deteriorates slowly with age e.g.(Norman et al., 2008). Secondly, we are comparing a human visual system to that of a generic stereoscopic observer – therefore variation within the human population is less important to measure. Participants were only accepted with normal or corrected-to-normal vision.

All participants were tested for stereoactuicity using the TNO stereotest ("TNO Stereotest, Richmond Products," 2014) and any participants who could not correctly identify depths of 240 arcmin were rejected from further study. This threshold was set so high because the TNO test was only a rough approximation to our experiments – each experiment also had a demonstration that was used to exclude participants on the basis of their performance at the demonstration stage. The TNO test is an established test using red/blue anaglyphs and coloured glasses in order to display stereoscopic depth from a booklet. The booklet consists of several simple stereograms to establish that the participant can see stereoscopic depth, then a series of stereograms with decreasing depth to establish an approximate depth threshold. While the TNO test and our experimental setup are different enough that the approximate threshold is not directly relevant to our experiments, they give an indication of how good the participant's stereovision is.

Participation was for a maximum of a single one-hour session each day and for a maximum of eight hours total. Participants could stop at any time without giving a reason. Ethical approval was given by the St Andrews University Teaching and Research Ethics Committee

(UTREC). All research was conducted according to the Declaration of Helsinki and abided by the BPS code of conduct. Participants gave informed consent before the experiment.

For clarity and anonymity, all participants who completed the experiments are labelled with a capital letter in alphabetical order of completion. In each Chapter a single letter will refer to one and only one participant, but labelling is restarted with A for every Chapter, except when explicitly mentioned otherwise.

4 What is a depth defined object?

- **Strand 1, How does depth from binocular vision contribute to object perception?**
- **Investigating: How is a depth defined object perceived with different edge shapes?**
- **Task: Which of two objects has a greater peak depth?**
- **Manipulation: One object (the smooth object) has blurred edges.**
- **Result 1: The smooth object is perceived with less depth than a sharp object.**
- **Result 2: Bias is not due to the absence of half occlusions in the smooth object.**
- **Conclusions: Smooth edge shapes cause a decrease in the perceived peak depth of the smooth object.**

Figure 4.1: A mossy gecko on a tree (*Rhacodactylus chahoua*, flattened against the left hand side of the trunk). Note the ragged edges and flattening into the trunk to disguise the shape in a way that blurs its edge with the background.
Reproduced by kind permission of Dr Hobaiter.

4.1 Introduction

In this thesis we are exploring the interaction of binocular vision with camouflage, particularly with reference to Julesz's assertion in 1971 that "hidden objects jump out in depth when stereoscopically fused" (Julesz, 1971). To do this, we have split the thesis into two main strands: the first explores what a disparity defined object is, and how we can perceive it using depth from two eyes. In the second strand, we investigate how easily spotted objects are when viewed with two eyes, and if depth from two eyes can break camouflage.

This Chapter is the first of three exploring the first strand of this thesis: how does binocular vision contribute to object perception? To explore this, we start by looking at the Gestalt principles of grouping (Wertheimer, 1923) (discussed in depth in Section 2.2.6), which give us an idea of what properties cause a set of elements to be perceptually grouped into a single entity. This approach inspired the work of Deas and Wilcox (2014), who studied the perception of depth separation between two vertical lines. When the two vertical lines were joined with two horizontal lines, grouped into a single rectangular object (see Section 2.2.6), then the depth difference between the vertical lines was perceived as being less than when the objects were not grouped.

Deas and Wilcox's work is one of many examples (see Section 2.2.6 for a full discussion) where the current understanding of disparity processing in the brain cannot explain the observed behaviour. As discussed in Section 2.2.5, the visual system uses small regions of the visual scene to calculate the disparity within that region e.g.(DeAngelis, 2000; Neri, 2004). If there is a depth change on a scale smaller than these regions, then the visual system cannot detect this depth difference and will effectively average over the depths present in that region. This causes measurable effects on our perception: for example a depth corrugation of over 3cpd (cycles per degree) will typically be perceived as a flat plane (Kane et al., 2014; Tyler, 1973, 1975). However, this effect is only on a small scale – on the length scales used by Deas and Wilcox (3.3° by 2.2°) then the disparity should be extracted for each vertical line without any kind of interaction due to the horizontal lines joining them. This illustrates a common problem with the current literature modelling how disparity extraction works: disparity extraction takes place over small length scales, and then conclusions are made about if the model agrees with the observed human behaviour, with little consideration about further processing. However, the extracted disparities from a disparity interaction at small scales cannot explain larger scale grouping effects such as those discovered by Deas and Wilcox.

The majority of objects used in experiments such as Deas and Wilcox are composed of elements that are well defined and on a featureless background. This is rarely the case in the real world, when background features cause visual effects such as crowding (Neider, Boot, et al., 2010). The featureless background is even more abstract when considering camouflage, as the animal is often evolved to match the background as closely as possible

(see Section 2.1.1). In this Chapter, we explore what happens when we smooth (blur) the boundary between the object and the background, making it harder to distinguish the edge between the object and its background.

Perceptually an increase in the smoothness of an object has been described as making the shape appear smoother or more like a 'hill'. The smoother the object, the less abrupt the transition from the object to the background, increasing the continuity between the background and the object. This causes perceptual grouping via good continuity between the object and the background, making it harder to segregate object out from the background. In the case of the sharp object, the edge of the object forms a closed contour, thus grouping the sharp object together via closure and inhibiting grouping with the background (Nakayama, Shimojo, & Silverman, 1989).

We hypothesize that the long range effects that have been found to be dependent on the grouping of the elements are indicative of another stage of object orientated disparity processing that estimates the overall depth of the object. This additional stage could cause a difference in peak depth between sharp and smooth objects. In the case of smooth objects, the boundary between the object and the background is poorly defined compared to the sharp object. If elements of the smooth edge are included in the estimate of the region of smooth object, then the lower non-peak disparities that make up the smooth object may be used in the overall estimation of depth in the smooth object, thus decreasing the perceived depth of it relative to the sharp object.

In this Chapter we discuss (in Section 4.2.1) the objects we used, and how we manipulated their shape to change how well defined the edge of the object was. We then present two experiments:

> Experiment 1 (Section 4.3) tested if an object with a smooth edge is perceived with a different peak depth to an object with a sharp edge. We found that smooth edged objects are perceived with less depth, possibly implying that there are further stages of disparity processing after objects are segregated from their background.
>
> Experiment 2 (Section 4.4) tests if the absence of half occlusions (a potential cue to depth discussed in Section 2.2.3) in the object with a smooth edge caused the difference in perception of their peak depth when compared to the objects with well defined edges that have half occlusions. We found no significant effects of half occlusions on the perception of depth in these objects.

Finally, we sum up the findings of this Chapter in a discussion Section (4.5).

4.2 Methods used in the first strand of the thesis

4.2.1 Stimuli

The overall experimental setup (using a Wheatstone Stereoscope) is discussed in Section 3.1. Stimuli were displayed on a luminance calibrated screen placed 1m from the observer (iiyama HM204DE A Vision Master Pro 514, size 23 by 17 degrees). The stimuli consisted of an RDS (see Chapter), 5.6 degrees wide and 11.2 degrees tall with white (12.24cd/m^2) and black (<0.01cd/m^2) dots of size 2.14 by 2.14 arcmin distributed at 20% density (326 dots per square degree) on a grey (6.10cd/m^2) background.

The RDS (random dot stereogram, see Section 3.1 and Figure 4.4) was comprised of two patches displayed one above the other, each containing disparity defined objects with a side length of 5.6 degrees. A constant comparison object was a depth-defined square with constant crossed disparity of 5.7arcmin and side length 2.8 degrees located centrally (see Figure 4.2a). This object is well defined, and can be easily segregated and the elements perceptually grouped (via Gestalt principles, see Section 2.2.6) as an object distinct from the background. We will call this the sharp object. The sharp object was randomly assigned to either the upper or lower patch. The other object, assigned to the other location was the test object and had a variable depth and shape stimulus defined differently in each experiment (see Figure 4.2b for an example shape). In this experiment, the test object has a smooth edge (as in Figure 4.2b), and we hypothesized that it will be hard for the visual system to determine where the background ends and the object begins. We will call this the smooth object.

(a) (b)

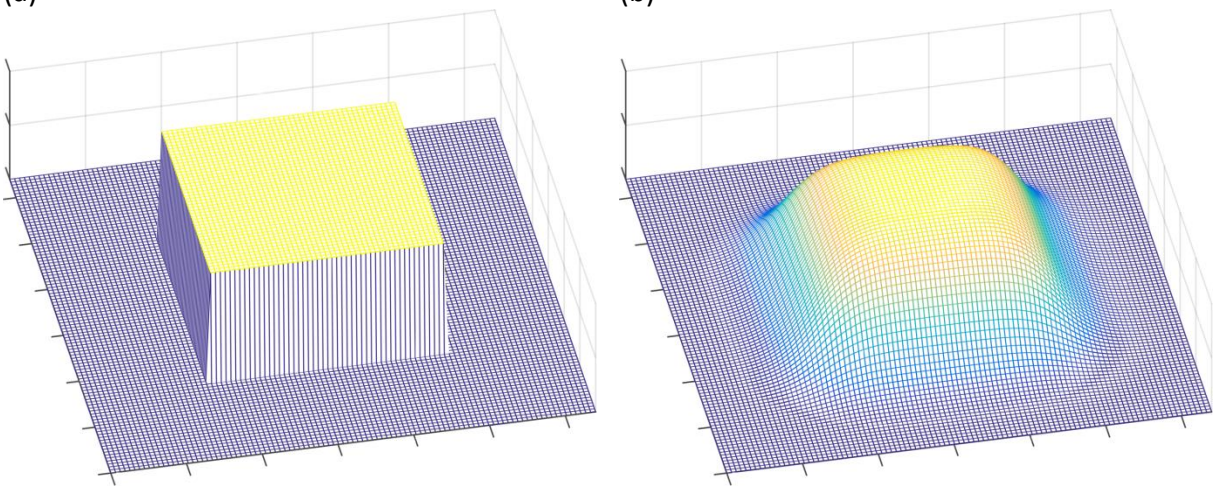


Figure 4.2: A 3D representation of the stimuli used in experiment 1. (a) left: Sharp object with well defined edges, (b) right: smooth object with poorly defined edges.

Throughout the thesis, the profile of the smooth object (the test object) along one dimension x is defined using an underlying hyperbolic tangent function:

$$f(x,p) = \frac{1}{2}\left[\tanh\left(\frac{1}{\sigma}\left(x - \frac{w-p}{2}\right)\right) - \tanh\left(\frac{1}{\sigma}\left(x - \frac{w+p}{2}\right)\right)\right] \quad \text{Eq. 4.1}$$

Where x is the coordinate of the specific point, located between 0 and w, which is the overall width of the stimulus. The two major variables that cause interesting manipulations to the stimulus are σ, the smoothness coefficient and p the plateau size.

The smoothness coefficient, σ is a variable that changes the overall shape of the hyperbolic function without altering the x position of the point corresponding to half the maximum depth, δ_p. The manipulation can be seen in the change of shape of the object in the cross-sections displayed in Figure 4.3(a). Here, an increase in smoothness coefficient decreases the first and second differentials of x (i.e. the gradient and the rate of change of the gradient), thus joining the top of the object more continuously to the background.

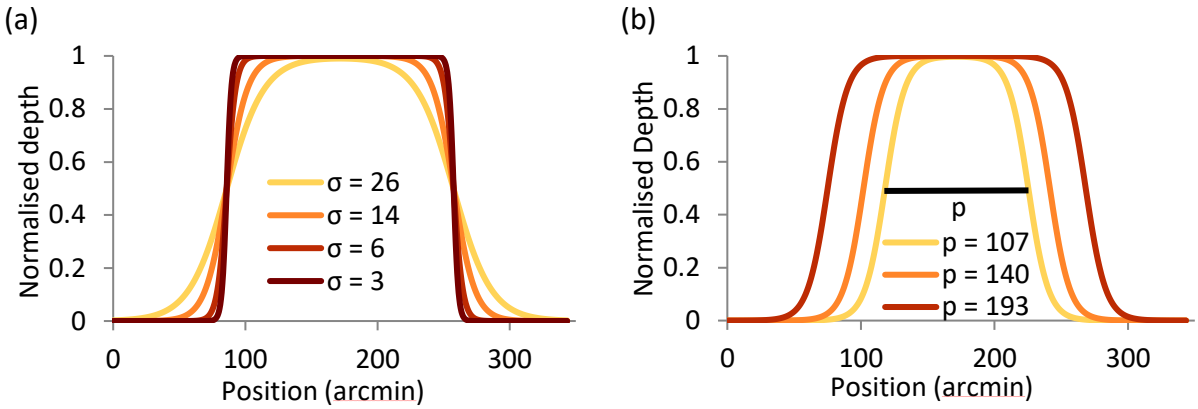


Figure 4.3: A 2D cross-section of the stimulus showing (a) the effect of manipulating smoothness coefficient and (b) plateau size.

Hyperbolic functions were chosen specifically because as long as σ is kept low enough that $f\left(\frac{w}{2}, p\right) \approx \delta_p$ (i.e. the depth of the central point is approximately equal to the input peak disparity) then the average disparity of the entire object is $\delta_p/2$ independent of the smoothness coefficient. If the smoothness coefficient is set too high, then this is invalidated. To avoid this, the code that creates objects to display checks the peak depth at the centre of the object is not less than $0.99\delta_p$.

A very small smoothness coefficient could cause depth edges that are steep enough to cause half occlusions (Harris & Wilcox, 2009) or be beyond the disparity-gradient limit (Burt & Julesz, 1980). In these experiments we manually set smoothness coefficients – so we avoid using any smoothness coefficients that cause these problems. To ensure this is the case the code checks if the edges are steep enough to cause problems and generates an error message if they are. Between the lowest (0) and highest (26) smoothness coefficients used in all experiments in this thesis, the volume, peak depth and average disparity of the

object for any input maximum disparity δ_p is constant to within 1%. For the values used in each experiment, see the stimuli Sections.

The second variable, plateau size, p, has no effect on the first or second derivatives of x. Instead it interacts with the two points of the function that are at $f(x,p) = 0.5\delta_p$, also known as the full width half maxima or the positions at half maximum depth. These also coincide with the inflection points of the function. The plateau size is defined as the distance between the two inflection points, as shown in Figure 4.3(b). Decreasing the plateau size by Δp moves the inflection points towards the centre of the object ($x = w/2$) by $\Delta p/2$ each, and vice versa when p is increased – this manipulation can be seen in Figure 4.3(b). For the first strand of this thesis, plateau size is normally kept constant at 168arcmin (2.8^o), but is manipulated for Experiment 3 between 100 and 200 arcmin. In the second strand, plateau size is reduced to a constant 74.7arcmin except for Experiment 8 where it is 112arcmin.

Overall, the two variables σ and p make for a very useful way of intuitively altering the stimulus shape. The effect of each variable is effectively orthogonal, that is to say that any measurement of the stimulus altered by one variable is left unchanged by the other. This makes our smooth object useful for exploring the effect of the smooth shape on the perception of the object.

4.2.2 Experimental procedure

In the experiments we are trying to establish if and how causing the object to be less well defined causes a bias in the perception of the object. To this end, we asked participants compare the peak depths of the objects. This is because the peak depth is an inherent property of the object this is well defined mathematically, meaning we know the performance of an ideal observer. Additionally, the peak depth is at the central point of the object and this was used to help describe what was meant to the participant using cross sectional line drawings (discussed later in this Section). The alternative would have been to introduce and ask about the depth of a probe at the peak depth e.g. (Koenderink et al., 1992), but then the presence of the probe object would have provided a disparity cue of its own, thus confounding the experiment.

After completing the consent form, participants took the TNO test ("TNO Stereotest, Richmond Products," 2014) for stereoacuity. Those who could not correctly identify depths of 240 arcmin or larger were excluded as discussed in Chapter 3.3.

Prior to the demonstration, we explained the task to the participants using a cross-section of the stimulus, indicating where the peak depth was on the stimulus, stating it was always central and informing them that the rest of the object's shape was irrelevant to the task. This was to ensure that all participants correctly grasped the concept of the peak depth of the object. To ensure participants could correctly view the stimulus, they then viewed a static stereogram screenshot from the experiment (similar to Figure 3.6) and were asked to

describe the shapes present in the stimulus (they did not know prior to this point the 3D shape, only the 2D cross-section) and make a peak depth judgement.

Every participant then completed a short demonstration of the experiment to familiarize them with the procedure. The demonstration consisted of the same stimuli as in the experiments, but with a larger disparity (maximum 9arcmin) and with smoothness coefficients (3 and 26). Additionally, the demonstration trials started at 10s duration, reducing to 2s per trial by the 10th trial. Data from the demonstration was checked to ensure participants understood the task before they continued onto the experiment. We specifically checked that the participant's responses changed with changing disparity, and that the sharp object was indicated as having more depth more often for sharp objects. If either of these was not the case, we checked the participant's understanding of the task by asking them to explain the task to the experimenter. If possible, any participants who misunderstood the task were corrected and re-ran though the demonstration, but this was not always possible and some participants were excluded from further study based on their performance on the demonstration (Heron & Lages, 2012).

Experiments all followed the same procedure: a fixation cross was presented at the centre of the screen (black on mid grey, 69 arcmin high/wide) until the space bar was pressed to initiate the trial. The stimulus was displayed, consisting of a rectangular RDS made up of two adjacent patches, one above the other. The upper patch of the RDS was randomly assigned to be either the smooth or sharp object, and the other object was assigned to the lower patch. This is done to remove any biases the participant might have between the upper and lower patches. The stimulus was presented for 2s, followed by the response prompt screen: black text on mid grey prompted participants to press either the up or down arrow buttons to indicate if the upper or lower stimulus had a greater peak depth. When a response was given, the fixation cross was then redisplayed. Trials were presented in blocks of approximately 300 trials, each block taking 10 – 15min to complete with a forced 60s break between each block. The first experimental session was three blocks preceded by the demonstration, subsequent sessions were four blocks.

4.3 Experiment 1: Is perceived peak depth affected by the shape of an object's edges?

4.3.1 Introduction

In this experiment, we investigate if 3D shape has an effect on object perception. We compare two objects – one with a sharp, well defined edge and the other with a smooth continuous edge that makes it hard to segregate the object from its background (discussed in Section 4.2.1). The objects are large (2.8^{o} across), thus we would expect early disparity extraction mechanisms to correctly obtain the peak depth of the objects (as discussed in the background, Sections 2.2.4 to 2.2.6). If 3D shape does have an effect on perceived depth, this may be due to another stage of object orientated disparity processing after the initial

extraction. This additional stage may be present in order to estimate the overall depth of the object, and therefore may cause a difference in the perceived depth between the well-defined (sharp) object and the poorly-defined (smooth) object. This may be caused by the inclusion of the non-peak disparities in the smooth object being used in the estimation of the overall depth of the smooth object, this decreasing the perceived depth of the smooth object relative to the sharp one.

4.3.2 Stimulus

Here, we discuss the details of the stimulus used for this experiment specifically. General details of the experimental setup, procedure and stimulus parameters are described in Section 4.2. For this experiment, the test stimulus was an object with smooth depth edges as shown in Figure 4.2(b), defined by the function:

$$\delta(x, y) = \delta_p f\left(x, \frac{w}{2}\right), f\left(y, \frac{w}{2}\right) \qquad \text{Eq. 4.2}$$

Where $f(x, p)$ is the function defined in Eq.4.1, and δ_p is the peak depth of the object. A plateau size (p) of 2.8 degrees (w/2) was used for the smooth object, meaning that the distance between the inflection points of this object is identical to the distance between the edges of the sharp object (see Section 4.2.1), which has a plateau size equivalent to 2.8 degrees. This should be a large enough scale such that individual disparity detecting neurons will not encompass both the edge and the centre of the object (see Sections 2.2.4 to 2.2.6).

For this experiment, we ran test stimuli using four different smoothness coefficients (3, 6, 14 and 23) with peak disparities at 5.4 and 8.4arcmin and five disparities equally spaced between these. A higher smoothness coefficient indicates a stimulus with a lower disparity gradient and lower rate of change of disparity gradient in the depth edge, thus forming a 'smoother' shape as shown in Figure 4.3a. Each combination of peak disparity and smoothness coefficient was displayed randomly interleaved in each block. The sharp comparison object was displayed at a peak disparity of 5.8 arcmin – see Figure 4.4 for an example stimulus. We ran a total of seven blocks spread between two sessions, with each participant completing a minimum of 60 trials (maximum 77) for each combination of smoothness coefficient and peak disparity.

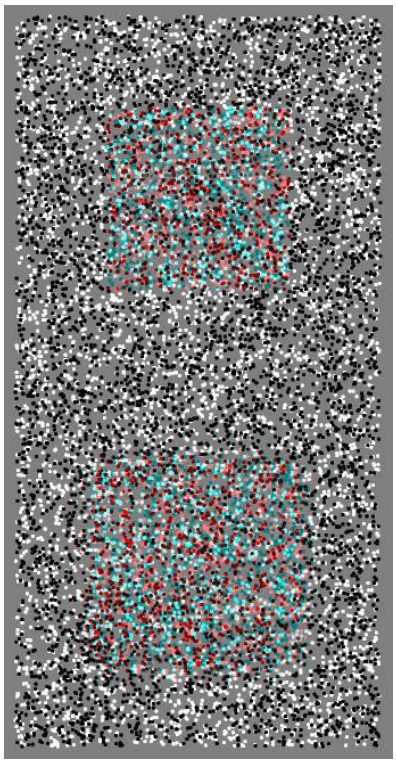


Figure 4.4: Anaglyph of the stimulus used in Experiment 1, with the sharp object top (as shown in Figure 4.2a) and smooth object bottom (see Figure 4.2b).

4.3.3 Results

First, I will discuss the analysis and results for one participant (Participant A or Par A), then show the final results for all participants.

We measured the proportion of occasions the sharp-edged stimulus was judged as having more depth, as a function of peak disparity of the smooth object (Figure 4.5a). For each smoothness coefficient for each participant, we fitted a psychometric function (cumulative normal) to the data. This was done using the Palamedes toolbox (for details see Chapter 3.2.1) and produces a best fit to the data, as shown in Figure 4.5a. The psychometric function gives us two useful quantities – the point of subjective equality (PSE), where the participant chose the sharp stimulus on half the occasions; and the threshold, a measure of the slope of the psychometric function, which relates to the sensitivity of the observer for this task. We define threshold as half the difference in disparity of the smooth object at the 0.75 and 0.25 points on the function fit.

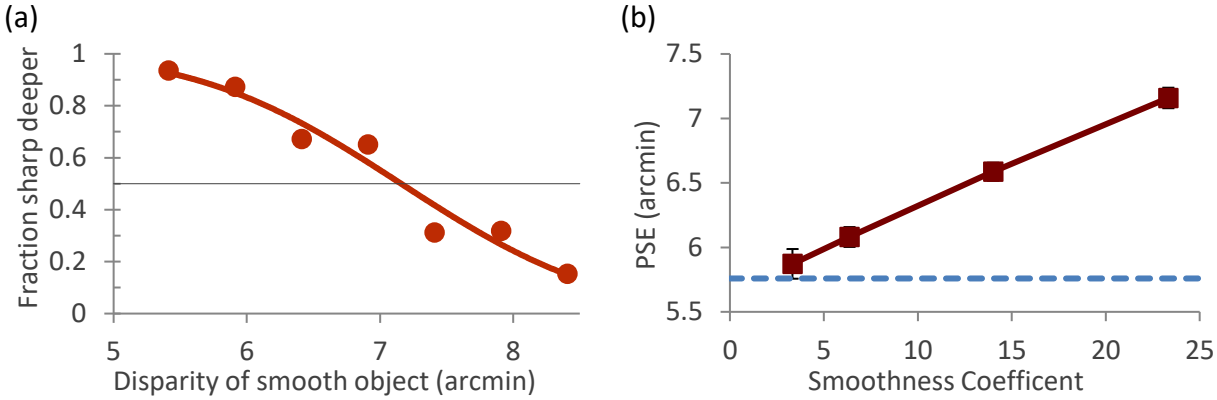


Figure 4.5: Analysis for one participant. (a) The psychometric function, black horizontal line is at 50% . (b) Extracted PSEs. Dashed blue line is the disparity of the sharp stimulus. Error bars are 1SEM.

PSEs are plotted against smoothness coefficient in Figure 4.5b. Error bars on the PSE for each smoothness coefficient were calculated using bootstrapping. A PSE close to 5.8 arcmin (the disparity of the sharp stimulus, indicated by the dashed blue line) indicates that the smooth stimulus was perceived to have the same peak depth as the sharp stimulus – this is the performance we would expect from an unbiased observer. A greater PSE than 5.8 arcmin means the smooth stimulus is displayed with more depth than the sharp stimulus to be perceived as having the same depth, thus meaning that the smooth object is perceived with less depth. The results suggest that the smooth stimulus was perceived as being flatter – it has a lower perceived depth than it was objectively displayed with.

We can see from Figure 4.5b that for this participant the smoothness coefficient makes a difference to the perceived peak depth. When the edges of the smooth object were very steep, there was no difference in the perceived depth. However, once the edges were

smoothed out, as indicated by a higher smoothness coefficient, the smooth object was perceived as progressively flatter.

Looking at the threshold data in Figure 4.7 for Par A, the smoothness coefficient does not appear to have made the task any harder for the participant to perform. This indicates that this participant had a biased but consistent perception of the smooth stimulus.

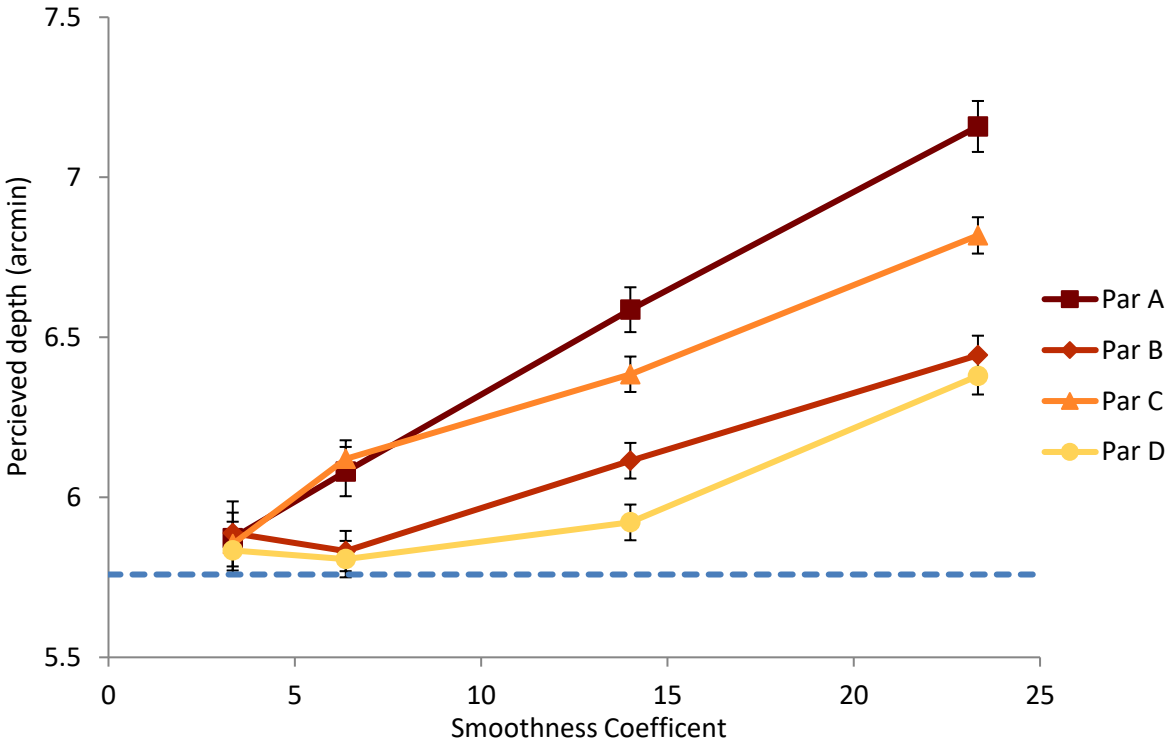


Figure 4.6: Perceived depth (PSE) as a function of smoothness coefficient for all participants in experiment 1, with the ideal observer shown as a horizontal blue dashed line. Error bars are one standard error of the mean (SEM).

Figure 4.6 shows the PSEs for all participants. Although there is individual variation, we can see that the overall trend observed in Par A's data is supported. All participants saw the smoother objects (higher smoothness coefficient) as having less peak depth (i.e. flatter) than the sharp edged object. A repeated measures ANOVA on a group level shows a highly significant effect of smoothness coefficient on the PSE ($F(3,9) = 21.1$, $p<0.0005$).

Figure 4.7 shows the effect of smoothness coefficient on threshold. We find there are no significant effects on a group level between any pair of smoothness coefficients on the thresholds (Bonferroni pairwise comparison, $p>0.1$) indicating that the participants were not finding it any harder to make the depth judgements on the smoother stimuli.

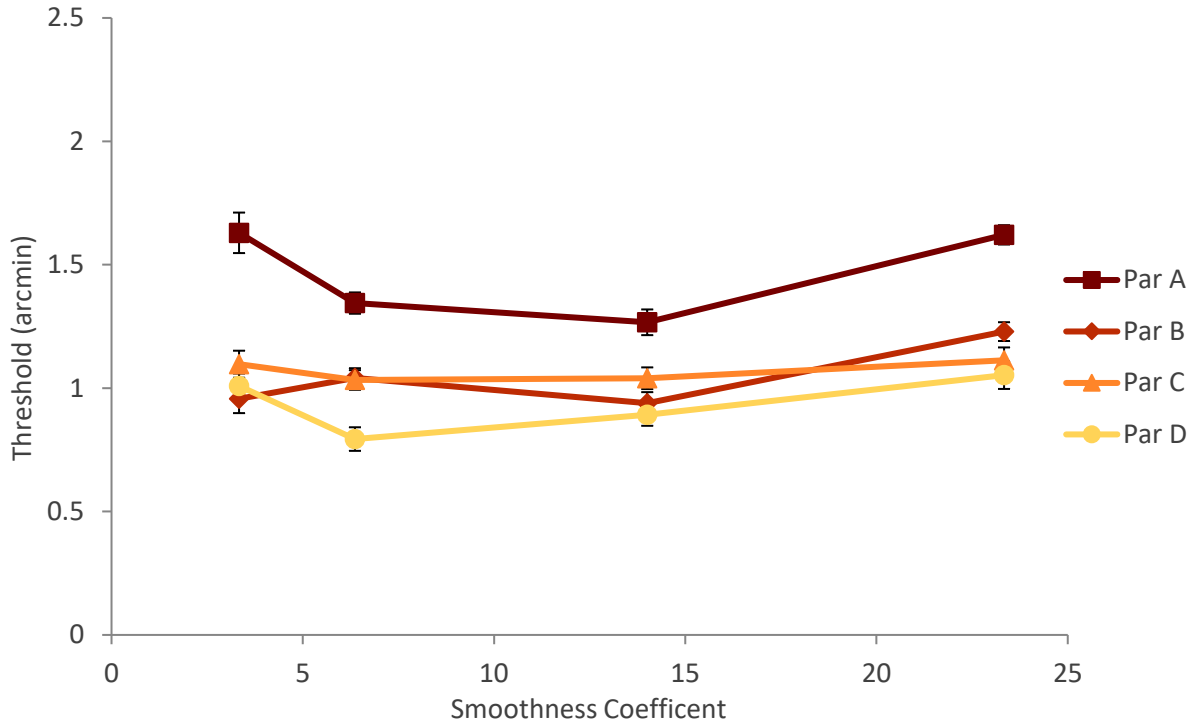


Figure 4.7: The threshold with smoothness coefficient for all participants in experiment 1. Error bars show 1 SEM

4.3.4 Discussion

This Chapter investigated if making an object smoother effects the perception of the object. If so, this effect would likely be caused by an object orientated disparity processing mechanism that takes place after early disparity extraction. To test this, we have created a pair of objects that should be perceived as having identical perceived peak depth due to their size (as discussed in Section 2.2.5). We varied the edge profile of the smooth object to make it harder or easier to segregate from its background (discussed in Section 4.2.1). We found that the smoother (harder to segregate) the object was, then the lower the perceived peak depth in comparison to the sharp object. This indicates that the edge shape of the object has an effect on the perceived peak depth of the object over a long range compared to the size of disparity detectors. We speculate that object segregation and object orientated disparity processing may be having an effect on the perception of depth in the object. Thresholds indicated that the participants did not find the task any harder to complete with the smoother objects.

The key finding of this experiment is that the smoother the depth edge of the displayed object, the lower the perceived peak depth despite having a constant objective peak depth. From a viewpoint of early disparity extraction, these results are highly unexpected as for this size of object we should be able to extract the disparity of the peak of the object without any interaction of with the smooth edges. These results support our hypothesis that the

perceived depth between the two objects is different. We will therefore consider what mechanisms may cause this bias in the perception of depth in the smooth object.

To consider how the disparities of the smooth edges may interact with the peak disparity, we turn to the current models of disparity extraction. These typically involve the comparison of small regions of visual scene between the eyes in order to estimate a disparity for that region, as discussed in depth in Section 2.2.5. Within these regions, the different disparities are effectively averaged, rendering it impossible to extract disparities on a smaller or larger scale than the size of the regions (see Section 4.2.1). To ensure this did not affect the perception of our stimulus, we have deliberately made the object displayed here too large to be affected by these artefacts of disparity extraction (see Section 4.3.2). However, this does not discount the possibility that there is a larger scale mechanism that combines disparities after the currently modelled stages of disparity processing.

There are two major possible mechanisms that could cause this bias in perception of the object, which we discuss in the following subsections.

4.3.4.1 Object based disparity interactions

One possibility is that the decrease in perceived peak depth is due to processing on an object level. The visual system must somehow segregate objects from the background and estimate the depth of these objects – to effectively segregate, there must be a decision made about where the boundary is between the object and the background.

In this experiment, we have been looking at the perceived peak depth of an object which has edges that are grouped with the background via good closure (discussed in Section 4.2.1). As the edge of the object becomes smoother, the grouping between the object and the background should become stronger, making it harder to segregate the object out from the background. This difficulty in segregation may introduce elements of the object's edge into the area that is segregated out as the object. The harder the object is to segregate (caused by an increase in the smoothness of the object), then the more likely the segregated area is to include elements of the edge. These edge elements are lower depth than the peak depth of the object.

According to current models of disparity extraction the inclusion of below-peak edge elements in the segregated object should not influence the perceived peak depth. As discussed earlier, this is because the edge is far enough away it should not interfere with the extraction of disparity at the central peak of the object (see Section 2.2.6 for a discussion of the range of disparity interaction from disparity extraction). We therefore consider object based mechanisms that could cause the edge elements that have been included in the segregated object to decrease the perceived depth of the overall object.

We hypothesize that edge elements may influence the perceived depth of the object. This is may be caused by the visual system averaging over the segregated object to reduce disparity noise in the estimate of the depth of the object introduced by small errors in

disparity extraction. When the segregation between the object and background has been poor, such as in our smooth object, then it could include intermediate disparities present in the smooth edge. Averaging over this segregated area in the smooth stimulus would include the elements of lower than peak disparity, thus averaging across the area would lower the perceived depth of the object. The more intermediate disparities present in the segregated object, then the greater we would expect the bias in perceived depth to be. This prediction lines up with what we have observed (see Figure 4.6), as larger smoothness coefficients have more intermediate disparities present in their edges and result in a larger bias in perceived depth.

In summary, the first hypothesized mechanism is that the decrease in perceived peak depth is due to poor segregation that leads to the inclusion of elements of the smooth edge in the estimated object. In order to decrease disparity noise, the segregated region is averaged over to obtain an estimate of the depth of the object. However, the inclusion of elements of the smooth edge means that the calculated average is lower than the peak depth of the object, causing a decrease in the perceived peak depth. The smoother the object, the more edge elements of lower disparity included in the average, leading to a decrease in perceived peak depth with an increase smoothness coefficient. In Chapter 5 we investigate this hypothesis in detail using a computational model.

4.3.4.2 Half occlusions

Half occlusions (discussed in Section 2.2.3) can under certain circumstances, cause a perception of depth, separate from binocular disparity e.g.(Collett, 1985; Harris, 2010; Tsirlin et al., 2010). These half occlusions are present in the sharp object but are absent in the smooth object, therefore it is possible that the discrepancy in the presence of these cues is causing a bias in the perceived depth that is not specifically due to disparity processing.

Half occlusions in the sharp object are caused by occlusion of the background in only one eye. If there is a greater distance between the occluding surface and the background, then the region visible to only one eye is larger. The size of the half occlusion therefore gives a depth cue to the distance between the occluding surface and background. In our study, the sharp object has half occlusions, as the top of the object occludes one eye's view of the background. However, none of the smooth objects have half occlusions, as all of the continuous edge between the foreground and background can be viewed with both eyes. Indeed, the code is set up so that if the gradient between the foreground and background was high enough to cause half occlusions the code would quit with a terminal error message.

As smaller half occlusions are typically present when there is less depth in an object (and the visual system perceives less depth with smaller half occlusions e.g. (Tsirlin et al., 2010)), it is possible that the visual system is assuming that the complete absence of half occlusions in the smooth stimulus is due to a low peak depth – thus leading to a perception of less depth

in the smooth stimulus. While this hypothesis may account for some of the bias observed in perceived peak depth for the smooth object, it cannot explain why the objects with different smoothness coefficients are perceived as having different perceived peak depths, as half occlusions are absent for all smooth stimuli.

To test if the absence of half occlusions in the smooth object creates any bias in the perception of depth, for Experiment 2 we create a rotatable stimulus that has half occlusions in one orientation and not in the other. If the absence of half occlusions is responsible for some of the bias in perceived peak depth, we would expect there to be less depth perceived in this stimulus when the half occlusions are absent.

4.4 Experiment 2: Do half occlusions have an effect on the perceived peak depth of the sharp and smooth objects?

4.4.1 Stimulus

The sharp comparison object was kept identical to the first experiment (see Chapter 4.2.1), but the test (smooth) object was altered to comprise of a smooth change in depth in one direction, and sharp change in the other, as shown in Figure 4.8. The test object was displayed in one of two orientations: either with the sharp edges displayed orientated left/right as in Figure 4.8a, thus creating half occlusions; or by rotating the stimulus by 90 degrees results in the sharp edges displayed top/bottom as in Figure 4.8, such that there are no half occlusions. The experimental screenshot in Figure 4.9 shows the no half occlusion condition. By rotating the stimulus in this way, we ensure that the overall shape and disparity distribution is constant, with the only difference between the two conditions being the presence or absence of half occlusions – therefore if there is a perceived difference between the two conditions it must be due to the presence or absence of the half occlusions.

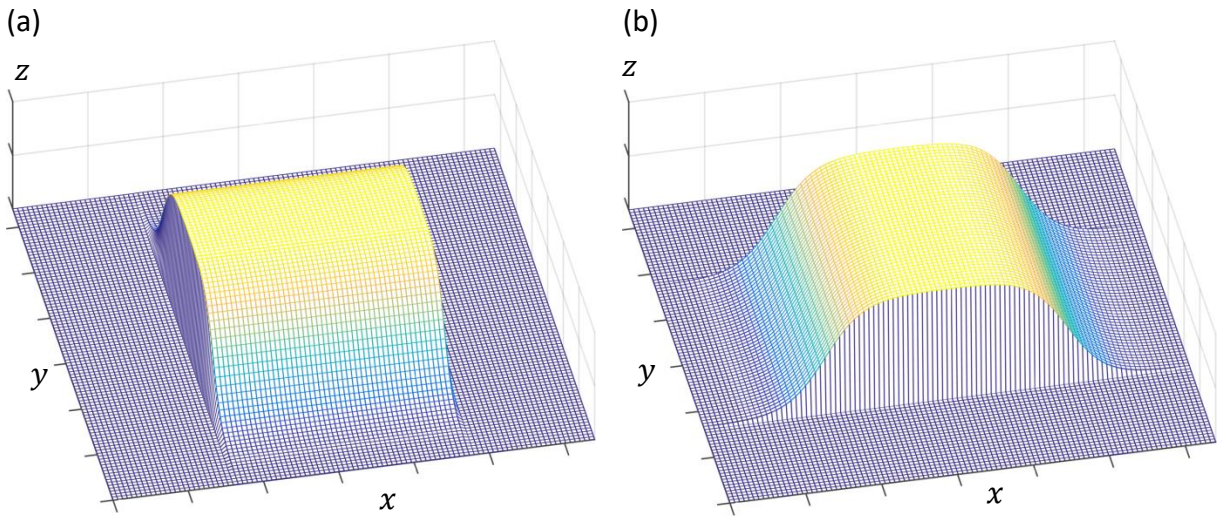


Figure 4.8: A 3D representation of the stimuli used in experiment 2. (a) Stimulus, orientated with half occlusions, and in (b) with no half occlusions.

The smooth edges of the stimulus were the same shape as in the previous experiment, with the disparity of a dot being described by:

$$\delta(x, y) = \delta_p f\left(x, \frac{w}{4}\right) d_v \qquad \text{Eq. 4.3}$$

Were $f\left(x, \frac{w}{4}\right)$ is the equation for the smooth edged object (see Eq. 4.1). The disparity contribution d_v governs the sharp edge, and is calculated by:

$$d_v = \begin{cases} 0 \; if \; y \leq \frac{w}{4} \\ 1 \; if \; \frac{w}{4} < y \leq \frac{3w}{4} \\ 0 \; if \; \frac{3w}{4} < y \end{cases} \qquad \text{Eq. 4.4}$$

These two equations in the current form describe the object with no half occlusions present (Figure 4.8b). By switching x and y we can change this into the second condition with half occlusions (Figure 4.8a).

We displayed the stimuli with a reduced number of smoothness coefficients (SC) compared to experiment 1, removing SC6 in order to bring the total experiment duration down from four hours to three. SC6 was removed as there was very little difference between the perceived peak depth at this value and that at SC3 (see Figure 4.6), meaning these two values are the most redundant. The SC3 tests the behaviour of the perceived peak depth at the smallest smoothness coefficient that does not cause half occlusions, and so is more informative to the overall trend than the intermediary value.

4.4.2 Results

As in Experiment 1, we collected the proportion of occasions the sharp object was judged as having more depth than the smooth, and plotted it as a function of peak depth of the smooth object for each smoothness coefficient. We fitted a cumulative normal curve for every smoothness coefficient, and extracted the PSE (the value when participants chose the sharp object 50% of the time) and threshold (the gradient of the cumulative normal function). The PSE gives us a measure of what the perceived peak depth of the object is, and the threshold is a measure of sensitivity.

Two of six participants had PSEs outside of the measured range or were unable to complete the TNO test, and were so excluded from further study. The PSEs for the remaining 4 participants are displayed on Figure 4.10, plotted against smoothness coefficient. There is one graph for each participant: dotted lines and hollow symbols indicate the condition with half occlusions present; solid lines and solid symbols indicate when half occlusions are absent.

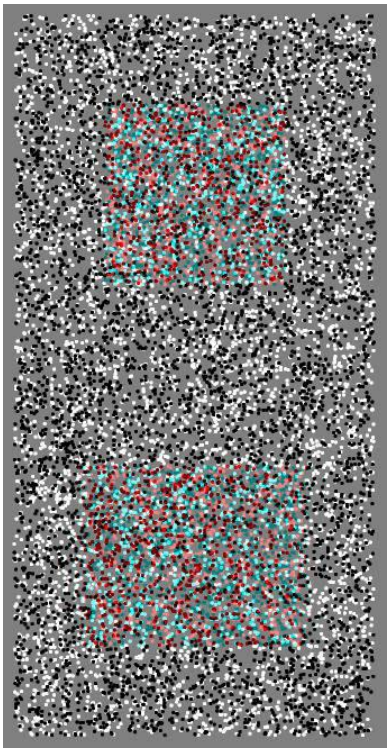
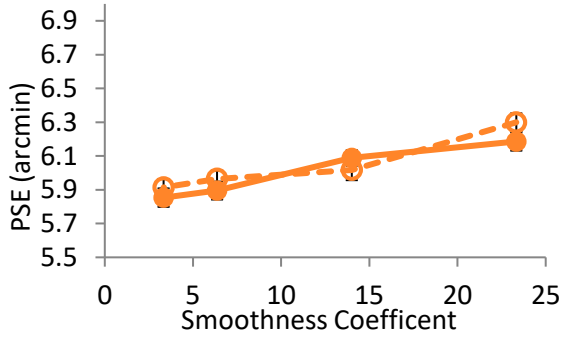


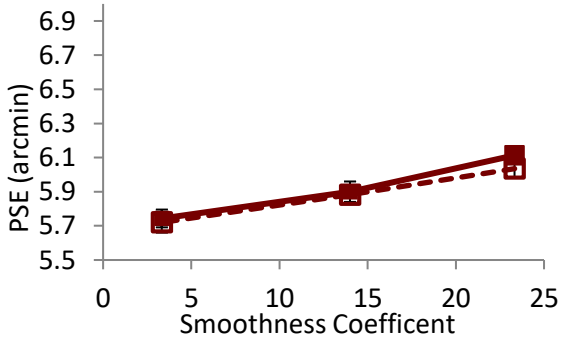
Figure 4.9: Anaglyph of the no half occlusion stimulus used in experiment 2, with the sharp object top and smooth object bottom.

The same overall trend as seen in Experiment 1 can be seen, with a decrease in the perceived peak depth with increasing smoothness coefficient. It can be seen from visual inspection that the data for the condition with half occlusions (HO) is frequently within error margins of the no half occlusion condition (NHO). This is backed up with a repeated measures ANOVA, which shows that there is no significant difference between the two conditions (f= (1,3) = 0.452, p=0.459)

Participant C



Participant E



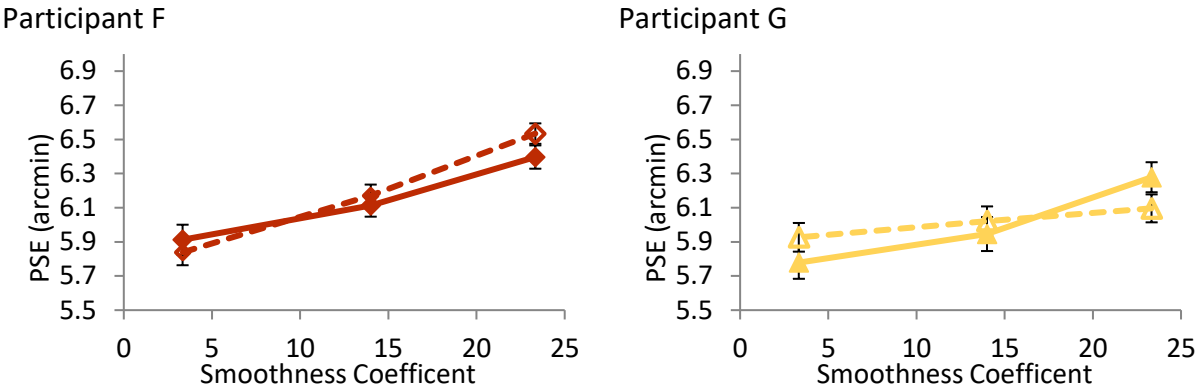


Figure 4.10: The PSEs (perceived depth) with SC for all participants in experiment 2. Dashed lines and hollow symbols = Half occlusions, Solid lines and solid symbols = No half occlusions. Error bars are one SEM (included in figure but sometimes too small to resolve). Note the scale change from Figure 4.6.

Thresholds showed some variation between participants, as shown in Figure 4.11. However, within subjects the thresholds were not significantly between the two conditions (repeated measures ANOVA, $F(2,3) = 0.001$, $p=0.975$).

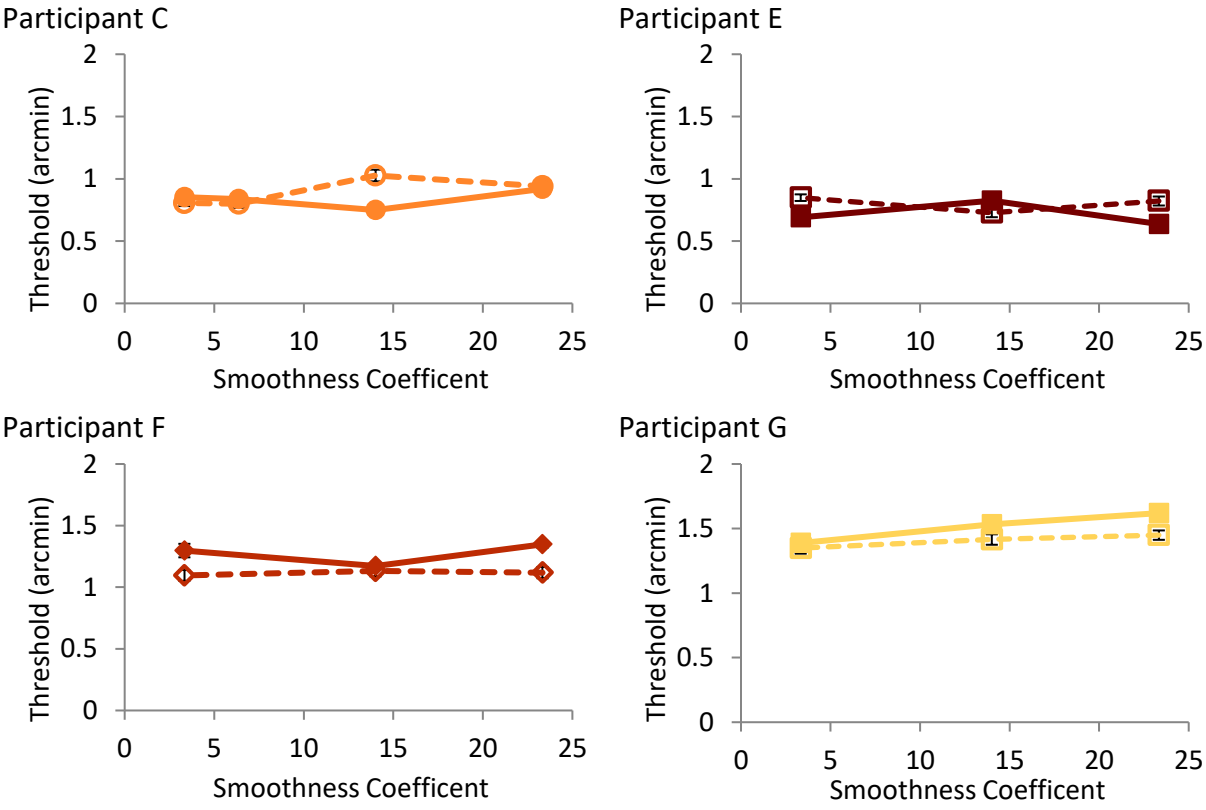


Figure 4.11: The thresholds with SC for all participants in experiment 2. Dashed lines and hollow symbols = Half occlusions, Solid lines and solid symbols = No half occlusions. Error bars are one SEM and in some cases too small to see.

4.4.3 Discussion

In Experiment 1 we had compared a smooth object with no half occlusions to a sharp object with half occlusions. As half occlusions have been known to create a percept of depth (Harris, 2010; Harris & Wilcox, 2009; Tsirlin et al., 2010), we hypothesized that the absence of half occlusions in the smooth stimulus may contribute to the observed decrease in perceived peak depth. To test this, we designed a stimulus that has half occlusions or no half occlusions depending on the orientation of the object, but which was identical in all other respects. If half occlusions contributed to the percept of depth in our objects, we expected a difference in the perceived peak depth between these two conditions.

This experiment found no significant difference between the perceived peak depth in the condition with and without the half occlusions. We therefore have no evidence that the half occlusions are used by the participants to aid in this depth judgement, indicating that for our objects the decrease in peak depth with smoothness coefficient is due to a different cause. It is already known that half occlusions only create a percept of depth in certain circumstances even when they are presented in isolation from any other cues (Harris & Wilcox, 2009). In this experiment, we have a strong depth cue from disparity with only a few elements of the random dot stereogram forming the half occluded surface. Perhaps it is not surprising that we cannot detect any effect of half occlusions on the perception of our object.

4.5 Overall discussion of experiments 1, 2, and future directions

In this Section, we:

1. Summarise of the implications of this Chapter on the first strand of the thesis.
2. Compare the two experiments, considering what the difference in participants' performance can tell us about the mechanism causing the decrease in perceived peak depth of the object.
3. Consider how these results relate to the current models of disparity extraction,
4. Discuss object based mechanisms that might cause the observed bias in perceived peak depth.
5. Consider a cue conflict between disparity and the texture gradient on the smooth edge of the object.
6. Compare the object's shape to the edge shape used in the Craik-O'Brien-Cornsweet illusion.
7. Look at the relevance of the results here for the second strand of the thesis, where we look at the hypothesis that binocular vision is used to break camouflage.
8. Round up with a discussion of the future directions we will take in the first strand of the thesis.

In the first strand of this thesis we consider how binocular vision contributes to object perception. In this Chapter, we have experimented using an object that has a poorly defined edge, to see if object processing influences the perceived depth of the object. In Experiment

1, we discovered that if there was a poorly defined edge between the object and the background then the object would be perceived as having a decreased peak depth. In Experiment 2, we investigated if this could be due to the presence of half occlusions in the stimulus, but found no evidence of any effect. We conclude that the decrease in peak depth is due to disparity processing, potentially due to the disparities within a segregated object being combined to give an overall estimate of the object's depth.

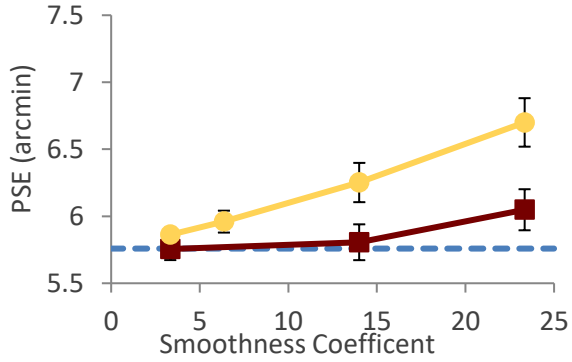


Figure 4.12: Comparison of average participant results in Experiment 1 (yellow circles) and experiment 2 (red squares). Error bars are the SEM based on observer variation.

The second experiment also replicated the results of Experiment 1 – that a greater smoothness coefficient decreases the perceived peak depth of the object. However, there is less overall bias in the PSE at higher smoothness coefficients (see comparison of averaged participant data between the experiments in Figure 4.12). While this could be due to inter-participant variation, it is of interest as in this experiment only one edge was smooth, meaning there were fewer elements between peak and zero-disparity (this can be seen mathematically by the Eq.4.3 for the stimulus in only being dependent on σ rather than σ^2). As our segregation and averaging theory above involves averaging over the intermediate disparities present in the smooth edge, then in this stimulus with fewer intermediate disparities we would expect a smaller decrease in the perceived peak depth. This is exactly what we observe, further indicating that the segregation and averaging hypothesis may hold merit.

The results here are not clearly explained by the current models of disparity extraction. This is because the current understanding of disparity extraction relies on comparing small regions of the image in two eyes, and using this comparison to estimate the disparity of the small region (see Chapter 2.2.5). Typically, this process causes averaging over small scales due to the finite minimum size of the regions estimating disparity e.g. (Norcia & Tyler, 1984; Tyler & Julesz, 1980). However, the averaging caused by disparity extraction is typically on scales of 3-6arcmin or lower (Filippini & Banks, 2009; Harris et al., 1997) – much smaller than the 100+arcmin region over which we have disparities near the peak depth. The small 5arcmin region should therefore extract the correct peak disparity for a large area of the stimulus, thus correctly estimating the peak depth of the object. However, this is not what we observe: in order for the edges to have an effect on the perceived peak depth we need a mechanism that averages disparities over a region larger than 100arcmin – far larger than the 5arcmin regions used for fine scale disparity extraction.

We hypothesize that the long range interaction of disparity is due to an object based mechanism. This is inspired by Gestalt grouping (see Section 2.2.6) experiments such as those

conducted by Deas & Wilcox (2014, 2015), and Pizlo et al. (2005a). These experiments also used very large stimuli (Deas and Wilcox used an object of 2.2^{o} by 3.3^{0}), and found that the perception of depth was impaired when elements were grouped into a single object. This indicates that there could be a large scale mechanism that is causing a bias in the perception of depth within objects. It is inspired by this work that we explore the potential of averaging taking place across the extended region of the object itself in Chapter 5.

4.5.1.1 Potential cue conflict

Unfortunately, the stimulus itself may introduce an additional complication. By maintaining a constant dot density across the RDS, we have created a potential cue conflict on the smooth edges of the object. If the random dots were a texture on the surface of the object, then the density of dots should increase when the surface slants more steeply away from the observer (Hillis et al., 2004). As we wish to avoid introducing any monocular cues to the experiment, we could not replicate this effect, thus causing a cue conflict that may decrease participants' sensitivity or introduce bias to the task (Backus et al., 1999; Gillam, 1968; Gillam, Blackburn, & Brooks, 2007; Gillam & Ryan, 1992; Hillis et al., 2004; Knill & Saunders, 2003; Ryan & Gillam, 1994). However, as the top of the object is approximately flat, then the region over which participants are making a judgement will not have this cue conflict, so the effects on the perceived peak depth of the object should be limited and do not explain the effects observed here.

It is interesting to compare the decrease in peak depth of the object to the Craik-O'Brien-Cornsweet illusion (Anstis et al., 1977) discussed in Section 2.2.6. When two planes of identical depth are joined together using a Cornsweet profile (see bottom Figure 4.13 reproduced from Figure 2.11) then there is a bias in the depth difference between the two planes. This is of particular interest, as the Craik-O'Brien-Cornsweet illusion is causing a change in the perceived depth of a surface due to the shape of the edge between the two surfaces. This is very similar to our results, where a smooth edge is altering the perceived depth between the surface of the background and the surface at the top of the object. While the Craik-O'Brien-Cornsweet illusion causes a difference in perceived depth between the two planes, and ours causes a decrease. The difference between these effects is possibly because of the presence of a discontinuity in the smooth depth change in the Craik-O'Brien-Cornsweet profile. It would be interesting to see a study that investigated if the perceived depth of a smooth edged object could be manipulated by introducing a discontinuity in the profile similar to that in the Craik-O'Brien-Cornsweet illusion.

The shapes explored in this Chapter are very relevant to the second strand of the thesis which will explore the hypothesis that binocular vision is used to break camouflage – here we have effectively found a shape that may allow an animal to camouflage itself to a binocular predator. If an animal resting on the background was approximately cylinder shaped, then on its edges it would have depth discontinuities similar to our sharp edged object. The animal, however well camouflaged would stand out to a binocular predator. However, if the animal changed its shape to be similar to our smooth edged object by smoothing itself into the background, then it would decrease its apparent peak depth. This manipulation might allow the animal to blend better into the contours of the background and, with its decrease in peak depth, pop out less from the background making it harder to spot. In Chapter 8 we use visual search experiments to explore if the smooth object is harder to spot.

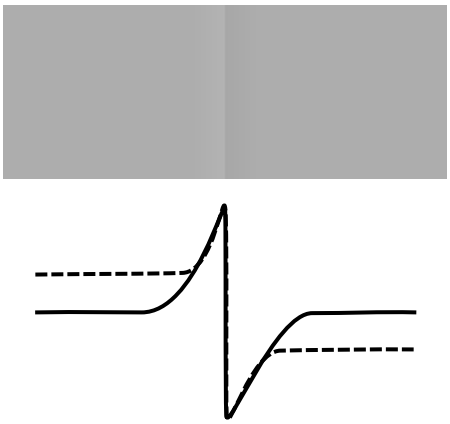


Figure 4.13: Top: Cornsweet illusion in the luminance domain. Both sides have identical luminance. Image reproduced with permission, (Fibonacci, 2005).
Bottom: solid line: cross-section of the Cornsweet profile. Dashed line: representation of the viewer's perception.

To shed light on why the smooth edged object is perceived with less depth, we now turn to the alternative explanation discussed after Experiment 1 (Section 4.3.4.1): that poor object segregation causes some of the non-peak disparities from the smooth edge to be segregated from the background as part of the object. Subsequent averaging to reduce disparity noise over the segregated object would cause disparities from the smooth edge to reduce the average disparity calculated for the entire object, pulling down the perceived depth of the object. This effect is discussed and explored in detail in the next Chapter.

5 How do we segregate an object from the background?

- **Strand 1, How does depth from binocular vision contribute to object perception?**
- **Investigating: If object based disparity processing involving segregating the object from the background influences the perceived depth of the object.**
- **Task: Create a computational model which can replicate participants' performance.**
- **Manipulation: Size and shape of the object segregated by the model.**
- **Results 1: Model is a good fit to existing results from Experiments 1 and 2.**
- **Results 2: Model can predict participants' performance on a new stimulus.**
- **Conclusions: Visual system uses properties of the object to segregate it from the background, then estimates disparity via averaging across the segregated region.**

Figure 5.1: A camouflaged soft coral crab (*Hoplophrys oatesi*, centre of the image). Note how the crab has grown protrusions to imitate its background – outside of the coral this camouflage would be ineffective. Image reproduced with permission, (Hobgood, 2005)

5.1 Introduction

This Chapter continues the first strand of the thesis, exploring how disparity defined depth interacts with objects perception. There are many current models of disparity extraction in the early levels of the visual system which can describe many of the visual effects observed in humans e.g. (Allenmark & Read, 2011; Banks et al., 2004; Chen & Qian, 2004; Filippini & Banks, 2009)(see Sections 2.2.4 and 2.2.5 for an in depth discussion). Although they are very powerful, these models typically only describe the early stages of disparity extraction and do not address the behaviour involving disparity combination for object perception. Indeed, Section 2.2.4 detailed an in depth discussion of the importance of the later stages of the visual system, where the first evidence of disparity processing on a similar scale to our objects occurred in V4 (Umeda et al., 2007). In order to better understand how a depth defined object is perceived and how this interacts with camouflage, we are not interested in the exact anatomy or mechanisms used in the visual system, but rather a general understanding of the which aspects of our stimuli are governing the perception of an object's depth. We therefore take a step back from these complex models of the visual system and create a simpler quantitative model to explore what interactions of disparity can cause the altered perception of the peak depth of the object found in the previous Chapter.

Quantitative models of the visual system are simple to calculate, understand and implement, meaning we can characterise the interaction of the perceived depth of the object with the shape of our stimuli in a way that would be very hard using more complex models. In turn, these insights could inform the current literature on disparity processing, enabling a more directed approach to the construction of complex models: it is easier to construct a complex model when we already have a basic understanding of the interactions involved in the perception of disparity in our stimuli. We therefore expect this model to be complementary to the techniques currently used for studying the visual system without directly competing with them.

This Chapter consists of three main Sections:

1. First, in Section 5.2, we discuss the overall approach and basis behind the model we will be using, and why it might account for the observed bias in perceived peak depth.
2. The next two Sections (5.3 and 5.4) concern themselves with modelling Experiments 1 and 2 from the previous Chapter. Interestingly, we find that the best fit for the model is to segregate and average over a square region the same size as at the distance between the inflection points (plateau size) of the smooth object.
3. In Section 5.5 we take the developed model and use it to predict the performance of a new experiment. We have two different predictions, depending on exactly how the visual system is segregating the smooth object. We run the experiment, and find a good fit between the predicted and experimental results when segregation is based properties of the shape of the object.

5.2 Model basis & overall approach

As discussed above, we consider a quantitative model of the interaction of our stimulus and perceived depth, and consider what processing would have to happen to create our observed results in Chapter 4 – that an object with smooth edges is perceived with less peak depth than one with sharp edges. We start from the underlying equation that defines the shapes of our objects and investigate if a simple lateral disparity interaction (see previous Section for an explanation) would be required to create the reduction of perceived depth observed in the smooth object.

Here, we aim to discover how the edges of the smooth object could interact with the perception of the peak depth of the object. Perhaps the simplest form interaction would be if the perceived depth of the object was calculated as the average disparity over an area that has been segregated out as the object. Averaging of this form would be advantageous overall as we know the first stages of disparity extraction are error prone (Tsai & Victor, 2003), and therefore perceiving the disparity of a single point as being exactly the estimated disparity at that point could cause misinterpretations of object shape and texture. An average over an area would increase the precision and accuracy of the depth estimate for opaque continuous objects, thus conferring an advantage for object perception when implemented after disparity extraction.

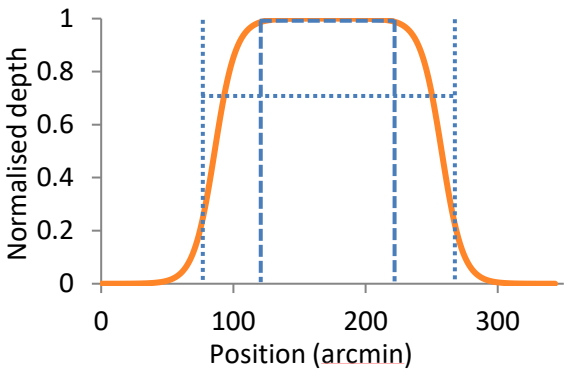


Figure 5.2: Averaging over a small region (vertical dashed lines) gives an estimate similar to the peak depth (horizontal dashed line), but averaging over too large a region (vertical dotted lines) gives a much lower value (horizontal dotted line).

If we consider averaging over a small but extended area, as between the dashed lines in Figure 5.2, then the average depth will be similar to the peak depth of the object, regardless of the shape of the edges. However, with an object like our smooth object, if the averaging area is too large, such as between the dotted line in Figure 5.2, it will include below-peak disparities contained within the edges of the object. Averaging over these below-peak disparities causes a decrease in the average peak depth and therefore the perceived object depth. The modelling and experiments presented in this Chapter primarily concern themselves with testing this hypothesis that the decrease in peak depth is caused by averaging over an extended region, most probably to improve signal-to-noise ratio in the object.

5.3 Exploring the results of Experiment 1 using modelling

5.3.1 Introduction

In Section 5.1 we speculated that the peak depth of the smooth object was perceived as being lower than the sharp object due to some form of lateral disparity interaction at a stage of the visual system beyond disparity extraction – probably via some form of disparity averaging after object segregation. In this Section, we present a simple quantitative model that explored this idea to see if averaging over a segregated area of the object could replicate the participants' performance in Experiment 1.

We implemented the quantitative model via the equations that generate the shape of the object. As discussed in Section 5.1 we are interested a characterising and understanding how the perception of depth is influenced by the shape of our object. We therefore create a model based on the equations that govern the stimulus; this model that should complement the complex approach of the models of the visual system such as those by Allenmark and Read (2010, 2011) and others (see Section 2.2.5 for a discussion).

In Section 5.1 we speculated that the visual system was segregating an area of the background out as the object, then averaging over the object in order estimate the overall disparity of the object. In order to implement a model that averaged over the segregated object, we had to choose the shape of the segregated region that was averaged. As we speculated that averaging occurs during the depth estimation of objects, we thought that that the shape of the averaged area might be dependent on the shape of the object – in our case square. We contrast this square based averaging to a circle and a Gaussian to explore the effect of segregated shape on the model's performance.

5.3.2 Square window model

In this sub-section we consider a computational model where the shape of the area being averaged (called the averaging window), is square. This was because all the objects used in Experiment 1 are approximately square-based. The predicted perceived peak depth is the average disparity within this window.

5.3.2.1 Implementation – integrals

The fastest and simplest way to implement an average over an area of our object is divide the integral of the function that defines the underlying shape of the object by the area that is averaged over. The function used to define the smooth object is: (Eq. 4.1, repeated here for clarity).

$$f(x,p) = \frac{1}{2}\left[\tanh\left(\frac{1}{\sigma}\left(x - \frac{w-p}{2}\right)\right) - \tanh\left(\frac{1}{\sigma}\left(x - \frac{w+p}{2}\right)\right)\right] \qquad \text{Eq. 5.1}$$

Where f is a function of the position x and the plateau size p. Plateau size determines the distance between the points at half depth. σ is the smoothenss coefficient, and determines the edge profile. Finally, w is the overall width of the object. The average of this function

across a square area can be found by dividing the two dimensional integral by the area covered by the integral:

$$\delta_{average} = \frac{\delta_p \int_{x_1}^{x_2} \int_{y_1}^{y_2} f(x,p)\, f(y,p) dy\, dx}{(x_2 - x_1)(y_2 - y_1)} \quad \text{Eq. 5.2}$$

Where x_1 and x_2 are the minimum and maximum x values of the window – likewise with y_1 and y_2. So the bottom left of the averaging window is located at (x_1, y_1) and the upper right at (x_2, y_2).

For the moment, we consider only the numerator of this equation, which calculates to the total disparity within the square area:

$$\delta_{total} = \delta_p \int_{x_1}^{x_2} \int_{y_1}^{y_2} f(x,p)\, f(y,p) dy\, dx \quad \text{Eq. 5.3}$$

As x and y are orthogonal, the integrals are separable:

$$\delta_{total} = \delta_p \int_{x_1}^{x_2} f(x,p)\, dx \int_{y_1}^{y_2} f(y,p) dy \quad \text{Eq. 5.4}$$

Solving these integrals separately, and substituting in $p = \frac{w}{2}$ as in the first experiment we obtain:

$$\int_{x_1}^{x_2} f\left(x, \frac{w}{4}\right) dx = \int_{x_1}^{x_2} \frac{1}{2}\left[\tanh\left(\frac{1}{\sigma}\left(x - \frac{w}{4}\right)\right) - \tanh\left(\frac{1}{\sigma}\left(x - \frac{3w}{4}\right)\right)\right] dx \quad \text{Eq. 5.5}$$

$$\int_{x_1}^{x_2} f\left(x, \frac{w}{4}\right) dx = \frac{\sigma}{2} \ln\left(\frac{\cosh\left(\sigma^{-1}\left(\frac{w}{4} - x_2\right)\right) \operatorname{sech}\left(\sigma^{-1}\left(\frac{w}{4} - x_1\right)\right)}{\cosh\left(\sigma^{-1}\left(\frac{3w}{4} - x_2\right)\right) \operatorname{sech}\left(\sigma^{-1}\left(\frac{3w}{4} - x_1\right)\right)}\right) \quad \text{Eq. 5.6}$$

For brevity, we define a function:

$$g\left(x_1, x_2, \frac{w}{4}\right) = \ln\left(\frac{\cosh\left(\sigma^{-1}\left(\frac{w}{4} - x_2\right)\right) \operatorname{sech}\left(\sigma^{-1}\left(\frac{w}{4} - x_1\right)\right)}{\cosh\left(\sigma^{-1}\left(\frac{3w}{4} - x_2\right)\right) \operatorname{sech}\left(\sigma^{-1}\left(\frac{3w}{4} - x_1\right)\right)}\right) \quad \text{Eq. 5.7}$$

So, substituting this back into equations 5.4 and 5.2 respectively, we find that:

$$\delta_{total} = \delta_p \sigma^2 g\left(x_1, x_2, \frac{w}{4}\right) g\left(y_1, y_2, \frac{w}{4}\right) \quad \text{Eq. 5.8}$$

$$\delta_{average} = \frac{\delta_p \sigma^2 g\left(x_1, x_2, \frac{w}{4}\right) g\left(y_1, y_2, \frac{w}{4}\right)}{(x_2 - x_1)(y_2 - y_1)} \quad \text{Eq. 5.9}$$

As we are considering a square area to average over, we know that $x_2 = y_2$ and $x_1 = y_1$, and $x_2 - x_1 = l$, where l is the edge length of the square. Additionally, as the square is centred on the peak depth of the stimulus, we set $x_2 = w - x_1$, so that the edges of the averaging window are the same distance from the edges of the stimulus on both sides. This gives us the final equation for calculating the average disparity within an arbitrary square of edge length l on our smooth stimulus:

$$\delta_{average} = \frac{\delta_p \sigma^2 g\left(x_1, w - x_1, \frac{w}{4}\right)^2}{l^2} \quad \text{Eq. 5.10}$$

We are calculating the peak depth of the object, which is located centrally, so we therefore consider the square window to be centred on the object, so that $x_1 = \frac{w-l}{2}$. Substituting this into Eq. 5.10:

$$\delta_{average} = \frac{\delta_p \sigma^2 g\left(\frac{w-l}{2}, \frac{w+l}{2}, \frac{w}{4}\right)^2}{l^2} \quad \text{Eq. 5.11}$$

The only unknown variable in this equation is therefore the size of the averaging window to be used, l. This is our only free variable, which we fit to each participant's data by generating the predicted perceived peak disparity for each smoothness coefficient at one window size. The goodness-of-fit of the window is calculated using the chi-squared test statistic:

$$\chi^2 = \sum_{\sigma} \frac{(O_\sigma - E_\sigma)^2}{(\delta O_\sigma)^2} \quad \text{Eq. 5.12}$$

Where O_σ is the observed value (i.e. the PSE from Experiment 1, Section 4.3.3) of the peak depth at smoothness coefficient σ, E_σ is the expected value from the model. δO_σ is the error in the observed PSE, obtained from bootstrapping when fitting the psychometric function to the participant's raw data, see Section 3.2.1 for a discussion and Section 4.3.3 for an example fit. The advantage of this test statistic is that it takes into account the size of the error bars in the experimental data – an important consideration as the participants vary inconsistently in their performance at different smoothness coefficients (see Figure 4.6). This procedure is repeated for many different window sizes, and the window with the lowest chi squared test statistic is considered the best fit and retained (Andrae, Schulze-Hartung, & Melchior, 2010). The model is applied to each participant separately. We only accept the model as having successfully fitted participant's data if the model produced a value of χ^2 under 25.0 in fewer than 1000 iterations.

To get an idea of how well the test statistic fits the observed data, we turn to the simplest of regression theory – the R^2 test statistic, defined by:

$$R^2 = 1 - \frac{\sum_\sigma (O_\sigma - E_\sigma)^2}{\sum_\sigma (O_\sigma - \bar{O})^2} \qquad \text{Eq. 5.13}$$

Where $\bar{O}$ is the average observed value for all smoothness coefficients. Effectively, a R^2 value of 1 indicates that the model can account for all the variation in the observed data, and a value of 0 indicates that it can account for none of the variation (subject, of course, to other statistical tests). Finally, a result of -1 or less indicates that the data would be better fitted by a linear fit than the proposed model.

5.3.2.2 Results

The model performs well, producing fits to all participants with an average R^2 of 0.92, indicating that the model can account for 92% of the variance in the data (see data table in Figure 5.4 for a breakdown for all participants). The observed PSEs for all participants is plotted with the best model prediction for fits to the data in Figure 5.3.

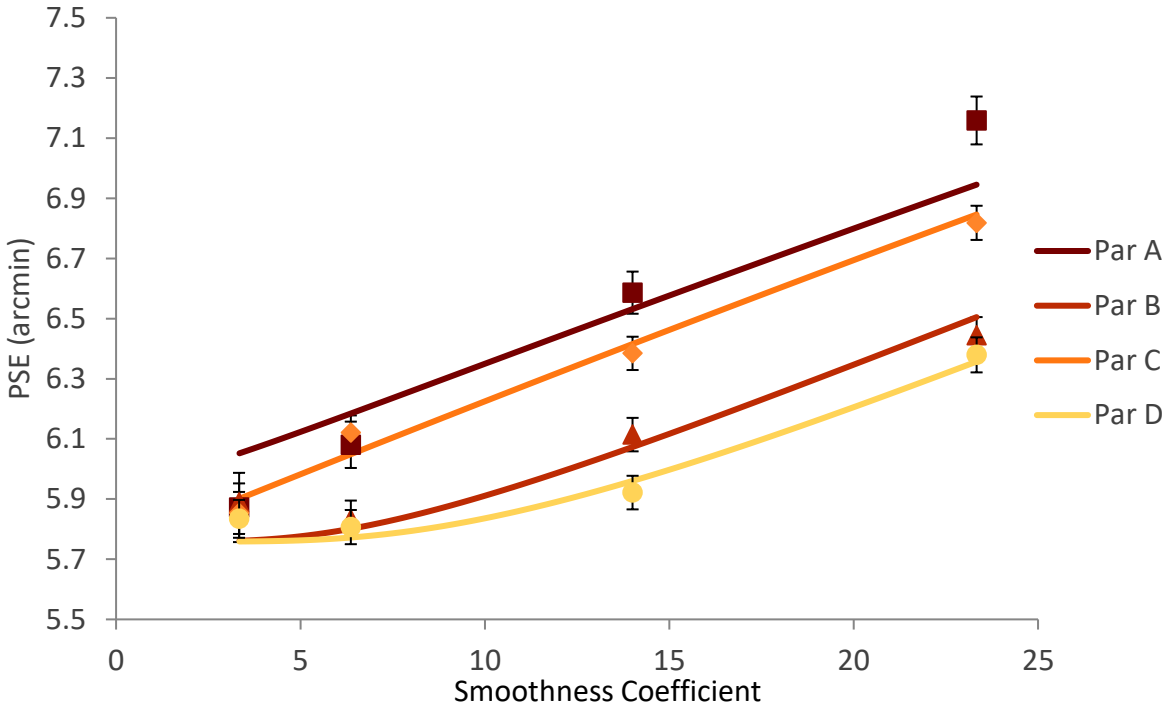
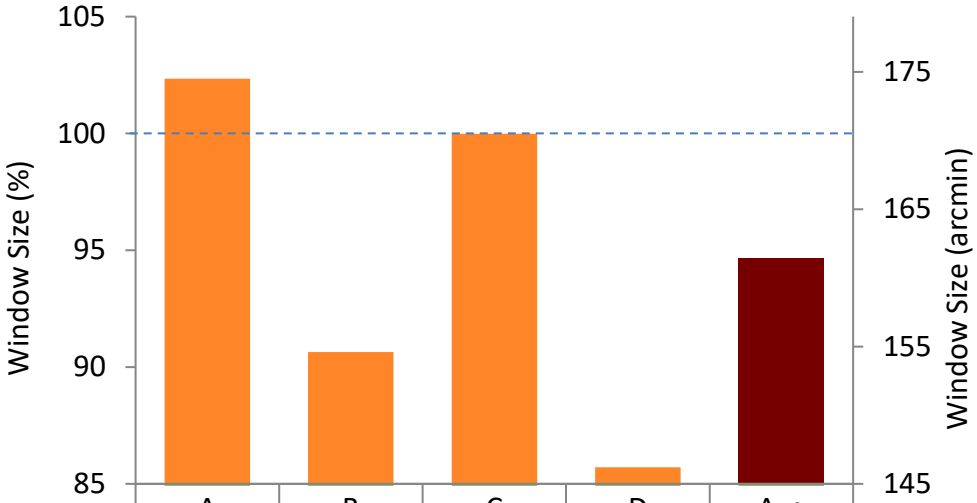


Figure 5.3: Square model fit (solid lines) to the participants' data (symbols). Error bars are one standard error.



	A	B	C	D	Avg
Window Size (%)	102	90	100	85	94
Window Size (arcmin)	174.5	154.6	170.5	146.2	161.4
R^2	0.81	0.94	0.98	0.96	0.92

Figure 5.4: Bar graph: window size used for each participant and the group mean. Dotted line is plateau size of the smooth object. Data table: Window size (arcmin) and the R squared test statistic for each participant. Window size (%) is the size of the fitted window divided by the plateau size (171 arcmin) as a percentage.

5.3.2.3 Discussion

In this Section, we set out to model the participants' performance in Experiment 1 to gain an understanding of why the participants have a bias in their perceived peak depth of the smooth object. We speculated that after object segregation, the segregated object area would be averaged to decrease noise in the estimate of the depth of the object. If the object was square, and the visual system segregated based on the shape of the object then we would expect the visual system to average over a square area (called an averaging window). This model could account for the decrease in perceived peak depth because if the smooth object was segregated part way down the smooth edges, then the averaging would include disparities below the peak value, thus decreasing the estimate of the depth of the object. We created a model to average over different square windows to test if these could fit the experimental data.

The model describes the data well, with the fits produced accounting for over 80% of the variance in the data for each participant (separately). It is important to note we only had one free parameter, and the shape of the window over which we average was fixed. This suggests that the square shape used by the model is likely to be a fundamental part of the peak disparity extraction process. Also of interest is the size of the windows fitted to the model: in data table in Figure 5.4 the range of fitted window sizes is between 146 and

175arcmin. These values are very close the size of the plateau used in creating the sharp and smooth objects – 171 arcmin, further indicating that the shape and size of the averaging window is dependent on the properties of the object.

Here we offer some speculation to why it might be an advantage to average over a flexible sized area of the object. As we discussed in Section 2.2.4, disparity extraction is an inherently noisy process - if, when judging the disparity of a point in space, then the extracted disparity of that point was the estimate would have a high degree of error. As objects in the scene are typically extended across significant regions in space, and objects often vary smoothly in depth, it is not unreasonable for the visual system to make the assumption that similar nearby disparities are part of the same object, and therefore are of similar magnitude to the requested point judgement. Therefore, by segregating an area of the object with approximately constant disparity from its surroundings, and then combining the disparities over the segregated region it should be possible to improve the precision of the depth estimate of the object. One way of combining disparities over a region in order to reduce the error in the estimate of the peak depth would be to take the average depth.

We speculate that this is the mechanism that we are replicating in this model: that the visual system is segregating out a central area of the object, and then averaging over it in order to reduce the error in the estimate of peak depth. If this is the case, then the shape of the averaging window should be dependent on the object's shape. In order to confirm or refute this suspicion we investigate if other window shapes can reproduce a good fit in Section 5.3.3. By performing modelling on the results from the second experiment (see Chapter 5.4), we can confirm if any findings are coincidences from this specific shape used in Experiment 1, or if there is an underlying behaviour governing the disparity averaging area.

5.3.3 Circular window model

In the previous Section, we successfully fitted participants' data for Experiment 1 to a model that segregated over a square region (called a window) and then averaged over all disparities within this area. Given the model has only one free parameter, it created a good fit to participants' data. We speculated that the square shape of the averaging region was due to object segregation, and therefore based off the shape of the object. Here, we test this hypothesis by considering a model where the region averaged over is circular: the object used in experiment 1 is square-based so we would expect this model to fail if the averaging window is dependent on the shape of the object. We expect the circular model to be unable to emulate it for human performance, but include it for completeness.

5.3.3.1 Implementation of the circular and Gaussian models

As in the previous Chapter, we initially calculated the integral of the shape of the object for a circular window. However, the integral is much more complex, and involved considerable calculation time using Wolfram Mathematica®'s Integrate function. An attempt to run this in the same fitting program used in the previous Section resulted in unreasonable runtimes. We therefore turned to a simpler technique for calculating the best fit for a circular window

– we calculated the average disparity by first generating a disparity matrix across the object, then averaged all the disparities within the desired region. This was originally the approach used in fitting the square window model, however the technique has a poor trade-off for calculation speed vs accuracy. To ensure that the accuracy for this different technique was correct and comparable to the previous square model, we took the resulting window size fit and ran it numerically though the integration process in Wolfram Mathematica ®.

We now describe the method for calculating the average disparity within a circular window. A disparity matrix D was created, such that each element i, j corresponded to the disparity calculated from Equation 5.1 at a point on the object $(x = i, y = j)$. To calculate the average disparity for a window of a given radius, r we take a second matrix A such that $A_{ij} = 1$ if the element was within the window, and A_{ij} = 0 if the element is outside of the window:

$$A_{ij} = \begin{cases} 1 \; if \; \sqrt{\left(x - \frac{w}{2}\right)^2 + \left(y - \frac{h}{2}\right)^2} \leq r^2 \\ 0 \; if \; \sqrt{\left(x - \frac{w}{2}\right)^2 + \left(y - \frac{h}{2}\right)^2} > r^2 \end{cases} \qquad \text{Eq. 5.14}$$

We then do elementwise multiplication (∘) of the disparity matrix D with matrix A, and sum all resulting elements – this provides the total disparity contained within the circle. To obtain the average disparity, we then divided by the total number of non-zero elements (i.e. the sum of A):

$$\delta_{circle} = \frac{\sum(D \circ A)}{\sum(A)} \qquad \text{Eq. 5.15}$$

As before, we performed this calculation for a range of window sizes and for all smoothness coefficients used in Experiment 1, and compared the model's predictions to the participant's performance. The fit with the smallest chi-squared test statistic was considered the best fit, and we then calculated the R^2 test statistic to give us an idea of how well the prediction fitted the observed data. This is repeated separately for each individual participant.

5.3.3.2 Results

Results for each individual participant of their best fit for the circular model and the observed data are shown in Figure 5.5, with the results and test statistics summarised in Table 5.1. It can be seen from Figure 5.5 that the predicted model fits are far from the experimental data – this is backed up from the test statistics seen in Table 5.1,

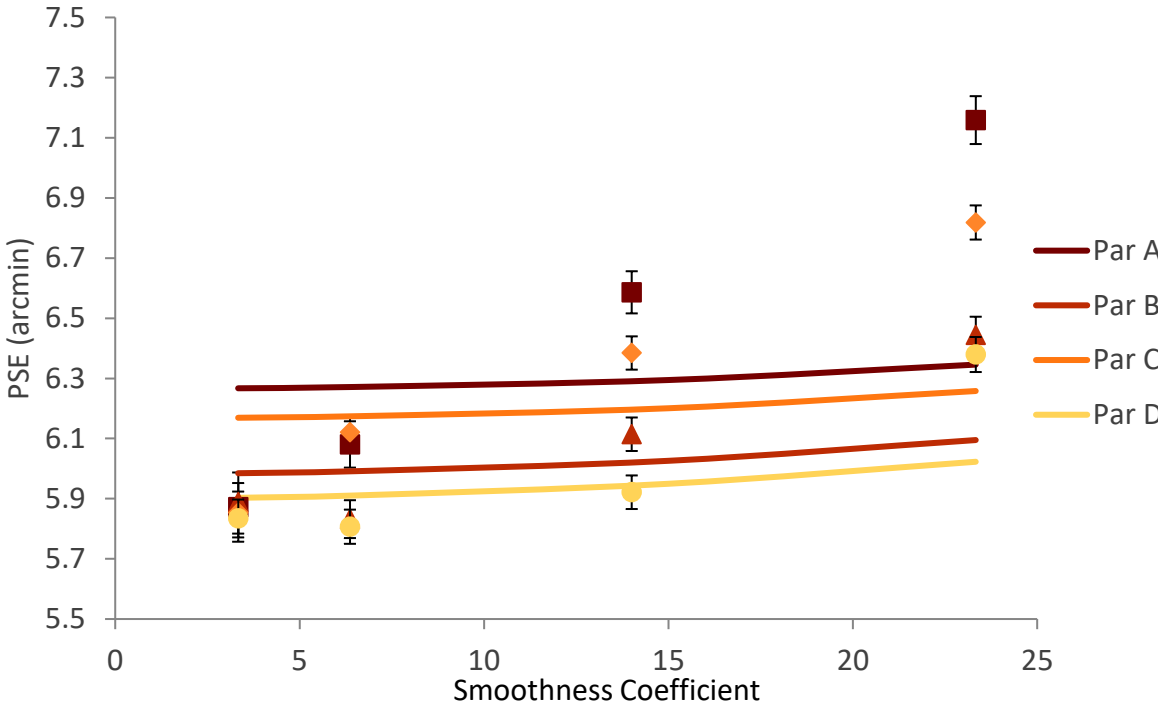


Figure 5.5: Circular model fit (solid lines) to the participants' data (symbols). Error bars are one SEM.

Participant	Window Size (arcmin)	R^2
(a)	165.3	-239
(b)	146.0	-20
(c)	158.3	-89
(d)	139.5	-15
Average	152.4	-91

Table 5.1: Window size and R^2 test statistics for the circular model fit.

5.3.3.3 Discussion

We developed a circular model to help confirm if the averaging region was dependent on having a square shaped averaging window. Having run a simple circular window model, it can clearly be seen that the predicted perceived peak disparities are far outside of the

observed perceived peak disparities, with the test statistics indicating that the data is far better fitted with a straight line rather than with the circular model. Compared to the square window model in Chapter 5.3.2 this model is extremely poor, indicating that shape of the window which is averaged over is important to predicting the perceived peak depth, not just the act of averaging over an extended area.

5.3.4 General discussion – Experiment 1 modelling performance

Here, we developed a model to characterize the quantitative approach that the visual system was using to estimate the peak depth of the smooth objects. We found that a model that averages over a square central region of the object performed very well, being able to account for an average of 92% of the variance in the data. This model was compared favourably to similar versions that performed the same calculation using circular windows: the circular model delivered an extremely poor fit.

In Section 5.3.2.3, we speculated that the success of the square window indicated that the brain segregated the square object from the background and then averaged over the segregated square region. This may confer an advantage by reducing the errors present in the disparity signal from disparity extraction. This segregation and averaging increases the precision of the estimate of the depth of the object, however when applied over objects such as ours with smooth depth edges it risks including disparities that are different due to the physical depth differences rather than depth extraction errors, thus introducing the observed bias in the perceived depth of the object. If the brain is truly segregating an area of the object to average over, then the good performance of the square window should be unique as it is based on the shape of the object. However, this mechanism and explanation is very speculative, especially given our simplified quantitative model of the visual system. It would be interesting to see further work extending this model into a computational theory of the visual system that explored the origins and reasons for a mechanism that causes an average of the surface of an object.

In the next section we test how well our quantitative model enables us to understand the interaction of the object shape with the perception of depth in the object by applying the model to Experiment 2.

5.4 Exploring the results of Experiment 2 using modelling

5.4.1 Introduction

In the previous Section, we found that a model based on averaging disparities over a square window performed very well at fitting participants' performance on the smooth edged stimuli in Experiment 1. However, the model has only been tested against one stimulus – in order to understand how well this model applies to the generic processing in the visual system we must test its performance against other objects, with different shapes and characteristics. In the Experiment 2, we used a more complex stimulus with one smooth edge and one sharp edge. This stimulus is still square based, so we would expect that the

square averaging model developed in Section 5.3.2 should be able to fit the performance of the participants'. The stimulus and results from Experiment 2 will therefore make a good test of how general the model is, as the stimulus is similar in shape but results in different observed PSEs. The only alteration we need to make to the model is to the change the equation defining the shape of the object to that used for the second experiment.

5.4.2 Model Implementation – integrals

The same process as used in Section 5.3.2.1 can be used to calculate the average disparity over an area using integrals. However, due to the presence of conditionals in the stimulus, we end up with two different equations that govern the model, depending on whether the square window is large enough to incorporate the sharp edges of the object:

$$x_2 - x_1 \leq w, \qquad \delta_{prediction} = \frac{f(x_2, x_1)}{2(y_2 - y_1)} \qquad \text{Eq. 5.16}$$

$$x_2 - x_1 > w, \qquad \delta_{prediction} = \frac{f(x_2, x_1)w}{2(y_2 - y_1)(x_2 - x_1)} \qquad \text{Eq. 5.17}$$

There are two stimuli used in Experiment 2, however these stimuli are identical in their equations but with x and y switched. Therefore, the predicted flattening for both conditions is identical.

5.4.3 Results

The integral equations predict that the participants should not perceive a difference between the two conditions – this is what we see in the results of Experiment 2 (Section 4.4.2).

Model fits are plotted against each participant's performance individually in Figure 5.6. Test scores and fitted window sizes are summarised in Table 5.2. The fits are good in general, although the fit for participant C only accounts for 53% of the variance – from inspection of their performance however it can be seen that they have unusually large error bars for this condition compared to other participants, suggesting that they found this condition of the experiment particularly hard. The R^2 test statistic cannot take into account the error in the participant's performance, so the poor score on this statistic is likely due to their uncertainty in response.

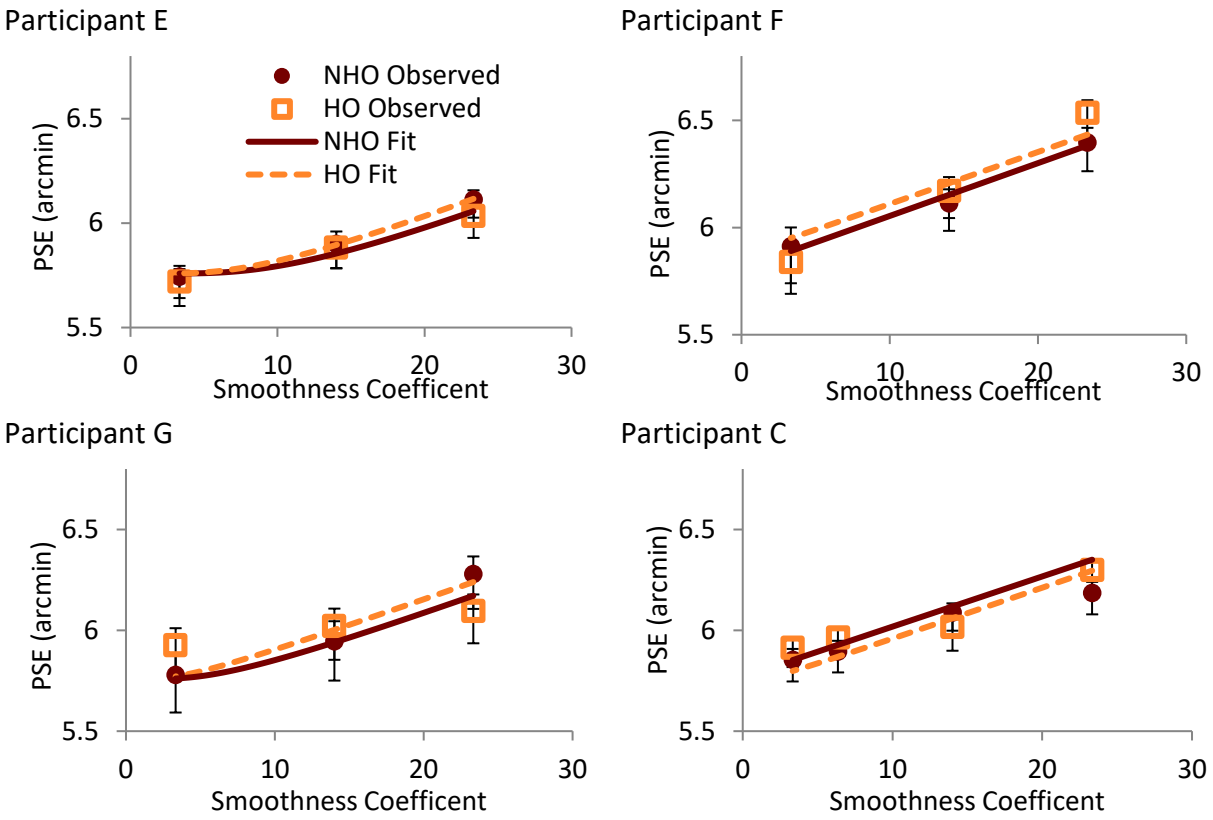


Figure 5.6: Participants' performance and square model fits for all participants in Experiment 2. Participant C participated in Experiment 1. NHO = No Half Occlusions, HO = Half Occlusions.

Participant	Condition	Window Size (arcmin)	Window Size (% of sharp-edged patch)	R^2
E	NHO	151.7	89	1
E	HO	145.1	85	0.94
F	NHO	172.2	101	0.98
F	HO	173.3	101	0.79
G	NHO	163.2	95	0.96
G	HO	157.2	92	0.53
C	NHO	167.8	98	0.98
C	HO	171.4	100	0.88
Average	NHO	164.2	96	0.98
Average	HO	162.3	94	0.87*
Average	Total	163.2	95	0.93*

Table 5.2: Window size and R^2 test statistics for the square model for Experiment 2. NHO has no half occlusions, HO has half occlusions. *Average R^2 excludes participant C condition 2.

5.4.4 Discussion

When we modelled participants' performance for Experiment 1, we found that the model that best fit participants' performance was an average over a square region. We speculated this was due to the square based nature of the object . To backup this characterization of the interaction of the stimulus and the perceived peak depth of the object, we created an analogous square based model on the stimuli used in Experiment 2, where it should perform as well due to the square based nature of the stimuli in Experiment 2.

The model fits very well to the data shown; with the average window size used being similar to the modelling of Experiment 1 (Section 5.3) and having a high R^2 value. As predicted, this indicates that the model may be generalizable to many different square-based stimuli.

Interestingly, the results from the window sizes used in the experiment are similar to those found in the previous modelling of Experiment 1 (which has an average window size of 152.4 arcmin, see Section 5.3.2.2). These window sizes are very similar to the plateau size (distance between the half-depth points on the stimulus) of the smooth object, and the width of the sharp object. In the modelling of the first experiment we tested several different object shapes, and found that the square shaped window gave the best fit to the participant data. The square shaped window was most similar to the shape of the objects displayed, and the size of the window was similar to the size of several key characters of the objects displayed. We speculated that this indicated a link between the size and shape of the averaging window and the size and shape of the object.

This speculation that there is a link between window size and object size is backed up by the performance of participant C, who completed both experiments. Modelling for this participant fitted their performance using a window size of 170.5 arcmin in Experiment 1, and 167.8 and 171.4arcmin for NHO and HO conditions respectively in Experiment 2 with a R^2 of 0.98 for all fits. These window sizes are very close to each other with a consistently good fit, indicating that the model and the participant are performing similar calculations in both experiments.

To consider this possible link between the averaging window and the object, we look at what aspects of the stimulus display may cause the averaging window to be similar in size and shape to the objects displayed. There are two main possibilities, discussed in depth in Section 5.3.2.3 and again in the introduction of the next Section, 5.5.1:

1. Shape based averaging: That the properties of the smooth object are determining the shape and size of the area to be segregated then averaged.
2. Template based averaging: The shape and size of the sharp object is providing a template that determines the shape of the smooth object.

We can distinguish between these two theories by changing the plateau size of only the smooth object and seeing if the size of the averaging window changes. If changing the plateau size alters the size of the averaging window, then the averaging window must be

determined by the shape based averaging – if not then the template is determining the shape and size of the averaging window.

5.5 Experiment 3: Testing model predictions using a new experiment

5.5.1 Introduction

We have created a simple descriptive model in Section 5.3 that is able to fit our observations of previous experimental results to a very high degree. However, this does not show that the understanding we have gained from this model is applicable in novel circumstances. Here, we test the model by predicting the results of an experiment that has not been run, and then run the experiment to test how well the model stands up.

For this experiment, we predict what the window size is based on the stimulus, then calculate the predicted perceived peak depth, then run the experiment. The model will have no free parameters, as the only parameter in Sections 5.3 and 5.4 was window size. Holding all the parameters fixed means that we cannot account for individual participant variation and our prediction is for the group behaviour. Zero free parameters also means we expect the prediction to be much poorer than the previous fits, but this prediction will indicate if the behaviour of the model is generalizable to other situations.

Additionally, the current model fits have found that the window size used in averaging is coincident with two major properties of the experimental stimuli:

1. The edges of the window lie near the inflection points (and the full width half depth) of the smooth stimulus. This indicates that the visual system may be using these properties to determine the edge of the object, then using them to segregate the object from the background and average over the area that has been segregated out as the 'object'. We call this hypothesized mechanism shape based averaging.
2. The size of the window is close to the size of the sharp stimulus, which is always present and the participant is requested to compare the smooth stimulus to the sharp stimulus. The sharp stimulus may be used as an indication of the size of the smooth object, and therefore the visual system uses the sharp object as a template to segregate the smooth object from its background. This effectively causes segregation and averaging may be taking place over a template dictated by the sharp edged object. We call this template based averaging.

These two predictions have very different implications – one where the shape of the object is governing the behaviour, implicating our hypothesized object segregation mechanisms, and the other due to the setup of the experiment. In this experiment we vary the size of the plateau of the smooth object (see Figure 5.7 for a 3D diagram of this manipulation), to test which of these behaviours is driving the averaging we observe. If this manipulation makes a difference to the size of the averaging window, then this indicates that the visual system is using shape based averaging as opposed to template based averaging.

It is worth noting that, the fitted window size is approximately 90% of the size of the plateau size used for both the smooth and sharp stimuli. We are not predicting that the width of the plateau on either the smooth or sharp objects is exactly where the edge of the averaging occurs - it is highly unlikely that the averaging occurs over a distinct, hard edged region, however as the models of the previous two experiments show (Sections 5.3 and 5.4), it is a good approximation to the true area averaged.

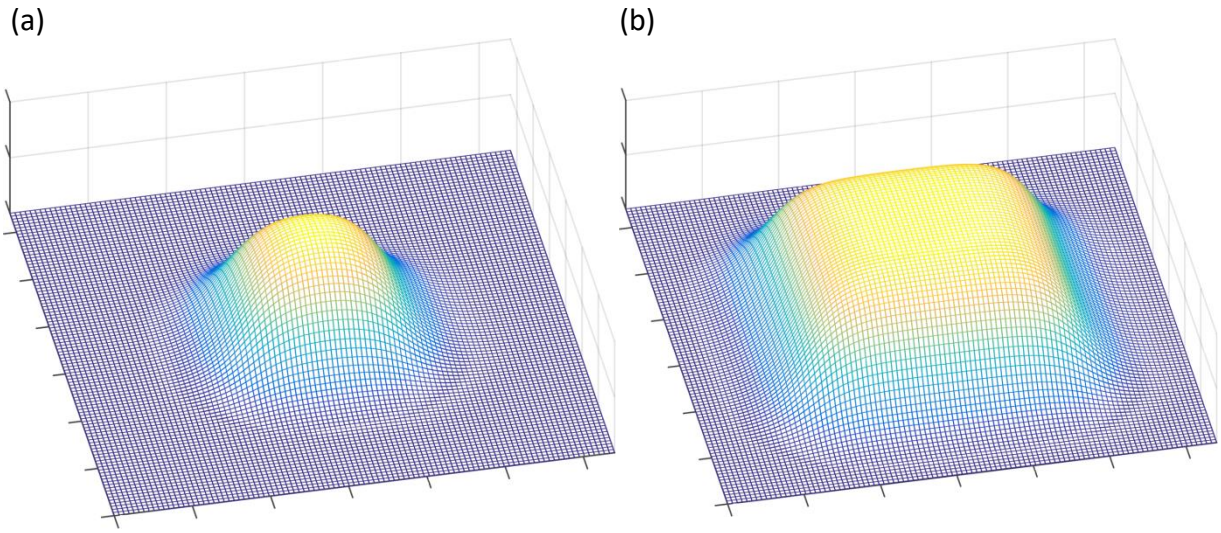


Figure 5.7: A 3D representation of the stimuli used in experiment 3. (a) Small plateau size, (b) Large plateau size.

5.5.2 Modelling, and model predictions

We took the equation governing the predicted peak depth from the model and altered it to the shape of an object with variable plateau size. However, this time we had two predictions: that the shape of the object defines the size of the averaging window, called shape based averaging; or that it is set by the size of the sharp edged stimulus, called template based averaging. Here, we address the predictions of each in turn. For both predictions, we consider our equation of the peak depth (average disparity) of the object, repeated here in a generalised form for simplicity:

$$\delta_{average} = \frac{\delta_p \sigma^2 g(x_1, x_2, p) g(y_1, y_2, p)}{l^2} \qquad \text{Eq. 5.18}$$

Where δ_p is the peak disparity of the object, σ is the smoothness coefficient, $g(x_1, x_2, p)$ is the underlying integrated function of the smooth edges (see Eq.5.7), p is the plateau size and l the width of the averaging window – defined as $x_2 - x_1$. (x_1, y_1) is the bottom left hand corner of the averaging window, (x_2, y_2) is the top right hand corner.

5.5.2.1 Shape based averaging

In the case of shape based averaging, the shape of the object defines the size of the averaging window. Therefore, we fix the size of the averaging window, $l = p$, the plateau

size. The positions of x_2 and x_1 are dependent on the plateau size, as shown in Eq. 5.19 and Eq. 5.20

$$x_1 = y_1 = \frac{w - p}{2} \quad \text{Eq. 5.19}$$

$$x_2 = y_2 = \frac{w + p}{2} \quad \text{Eq. 5.20}$$

These were subsisted into Eq.5.18 to give Eq.5.21, predicting the perceived peak depth at an arbitrary plateau size p and smoothness coefficient σ.

$$\delta_{shape} = \frac{\delta_p \sigma^2 g\left(\frac{w-p}{2}, \frac{w+p}{2}, p\right)^2}{p^2} \quad \text{Eq. 5.21}$$

Figure 5.8 shows a 3D plot of the effect of changing smoothness coefficient on predicted PSE at different plateau sizes. Note how the predicted PSE increases dramatically at low smoothness coefficients – this is the effect of the objective peak depth being decreased by the smooth edges encroaching on the peak of the object. The effect of changing plateau size on the predicted PSE alone is shown in Figure 5.11a (page 121).

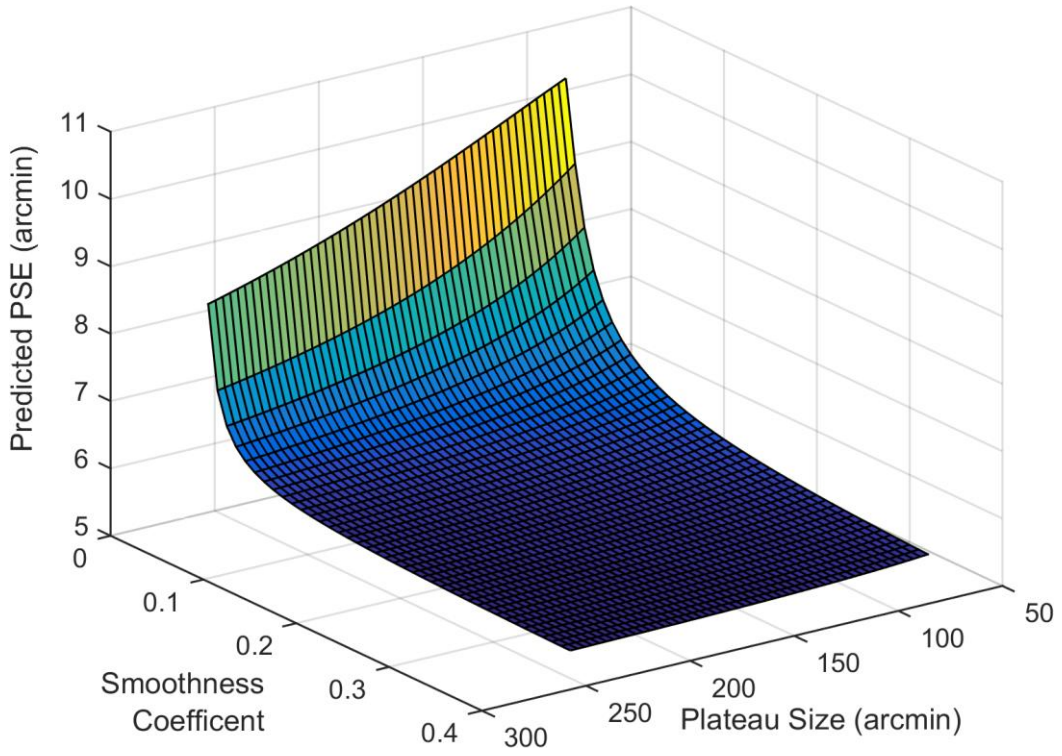


Figure 5.8: The effect of smoothness coefficient and plateau size on the predicted PSE for shape based averaging.

5.5.2.2 Template based averaging

In template based averaging, we expect that the averaged area of the window to be dependent on the size of the sharp object. This means we are considering x_1 and x_2 to be constants, fixed to the edge positions of the sharp object at $w/4$ and $3w/4$. We can substitute this straight into Eq.5.18 giving us Eq. 5.22.

$$\delta_{template} = \frac{\delta_p \sigma^2 g\left(\frac{w}{4}, \frac{3w}{4}, p\right)^2}{p^2} \qquad \text{Eq. 5.22}$$

The effect of changing plateau size and smoothness coefficient on the predicted PSE is shown in Figure 5.9. When comparing to the shape based averaging in Figure 5.8, we can see that at large plateau sizes (150 to 250 arcmin) the predictions of both models is very similar. However, at smaller plateau sizes (below 150 arcmin) the predictions differ, with the PSE greatly increasing with decreasing plateau size for the template based model – this is unlike the shape based model, where PSEs do not change much with plateau size.

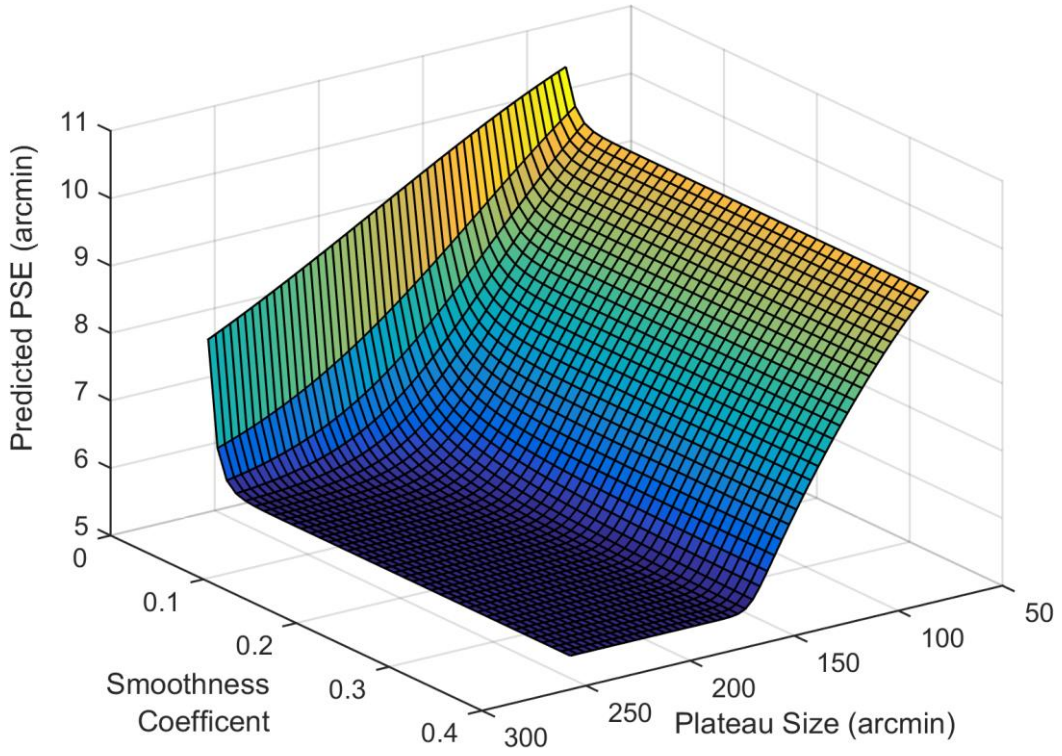


Figure 5.9: The effect of smoothness coefficient and plateau size on the predicted PSE for shape based averaging.

5.5.3 Stimuli

We need to consider which smoothness coefficient to use for this experiment. As we have two predictions that we wish to discriminate between, we will therefore select the smoothness and plateau size where the difference between the predictions is greatest. We plot the difference $\delta_{shape} - \delta_{template}$ against plateau size and smoothness coefficient in Figure 5.10.

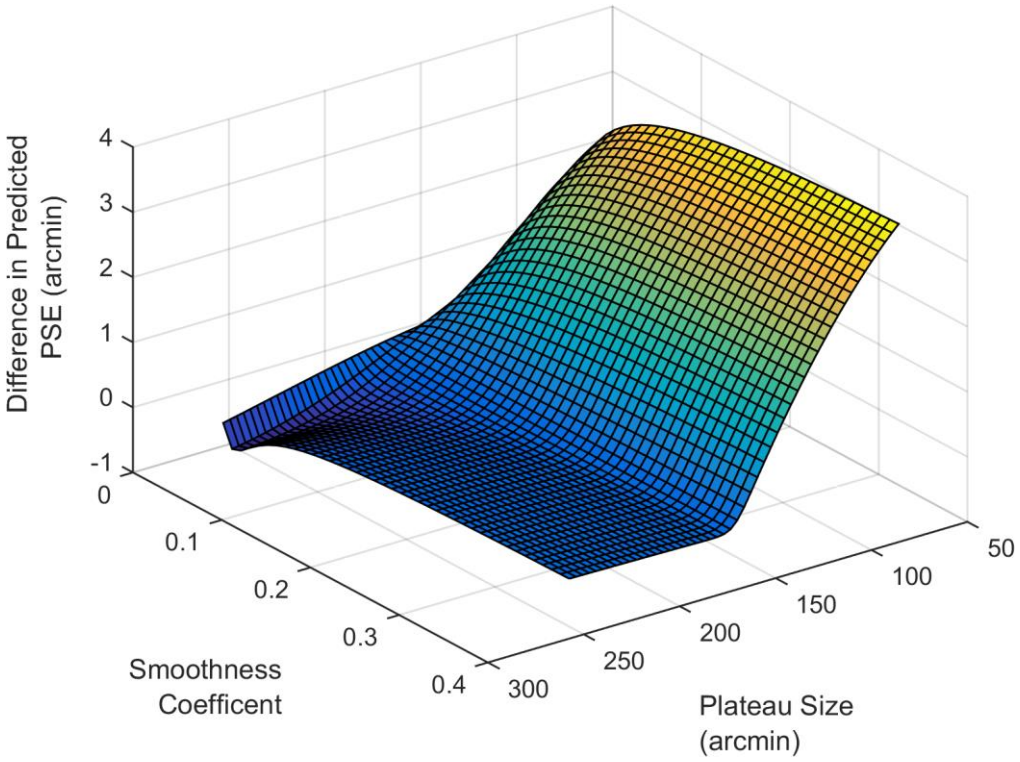


Figure 5.10: The difference between the two model predictions, note the rotation of the graph from the previous two figures to give a better view of the data.

The difference between these model predictions is maximal at maximal smoothness coefficient, however as we decrease plateau size the edges of the smooth object are brought closer together. This creates a problem: having too large a smoothness coefficient and too small a plateau size will cause the smooth edges of the object to overlap, decreasing the objective peak depth of the object. To avoid this, we must trade-off between maximising the difference between the predictions of the models, and ensuring that the object is displayed with a constant peak depth. After exploring models, we decided the best compromise was a range of plateaus of 107.3 to 193.2 arcmin across, with a smoothness coefficient of 14. The predictions for the perceived peak depths are displayed in as the dotted red (template) and solid blue lines (shape based) in Figure 5.11(a).

In summary, the sharp edged object was identical to those used in previous Chapters. However, the smooth edged stimulus had an additional manipulation – this time the plateau size was displayed as 107.3, 139.5 or 193.2 arcmin, whilst the smoothness coefficient was kept constant at 14. As shown Figure 5.11(b), this manipulation changes the distance between the inflection points (see Figure 5.7 for a 3D representation). As before, the sharp stimulus was of constant disparity and plateau size of 171.1arcmin. Due to an increased range of expected values of the PSE, seven disparities were shown between 4 and 10arcmin. Figure 5.12 shows an example screenshot of the stimulus. We will compare participants' performance to a pair of model predictions, which are generated with zero free parameters.

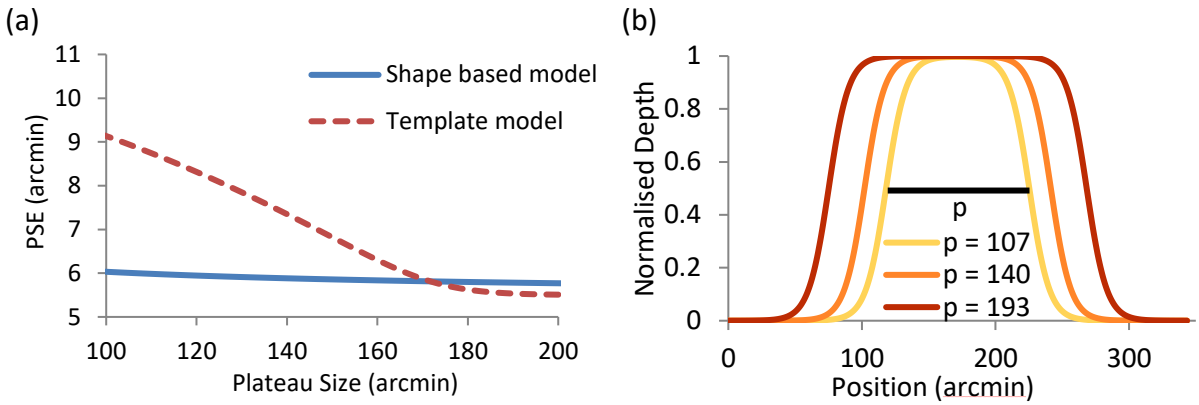


Figure 5.11: (a) The predictions for the two models at a smoothness coefficent of 14. (b) The effect in cross-section of manipulating plateau size on the shape of the object – the three plateau sizes illustrated are those used in the experiment.

All other experimental details are the same as in Experiment 1 (Section 4.2).

5.5.4 Results

Out of 10 participants, 9 successfully completed the experiments, with one excluded due to delivering a flat psychometric function. Participant F has a PSE for the plateau of size 103arcmin outside of the measured range, but is included for completeness. The results for all participants are shown in Figure 5.13, alongside the two predictions, while χ^2 analysis is presented in Table 5.3. The model predictions have no free parameters, and the window size was fixed before performing the Experiment as discussed in Sections 5.5.2.1 and 5.5.2.2 for all participants, so cannot account for participant variation. Seven of nine participants appear to follow the shape based model (blue solid line), with χ^2 values between 1 and 5.5 (excluding Par A at 49 and F at 18), indicating that the model gave an excellent to acceptable fit but did not account for all sources of error. The shape based model performs considerably better than the template based model (red dotted line), which has χ^2 values between 140 and 276 for all but Participants A (8) and F (3). None of these participants follow the prediction exactly, but this is to be expected as we have created a predictive model with no free parameters.

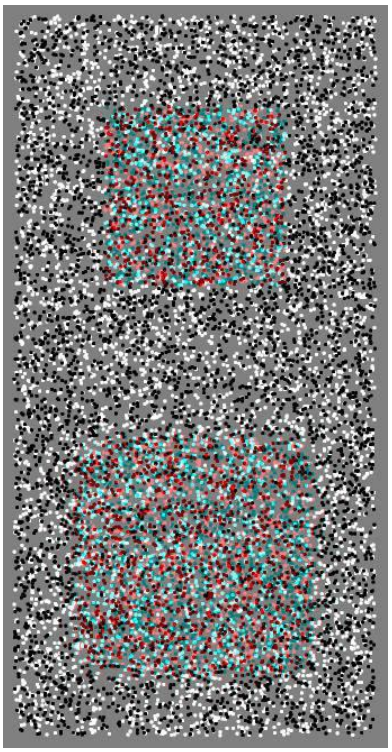


Figure 5.12: Anaglyph of stimulus for experiment 3, with the lower smooth stimulus having a plateau size of 193.2 arcmin

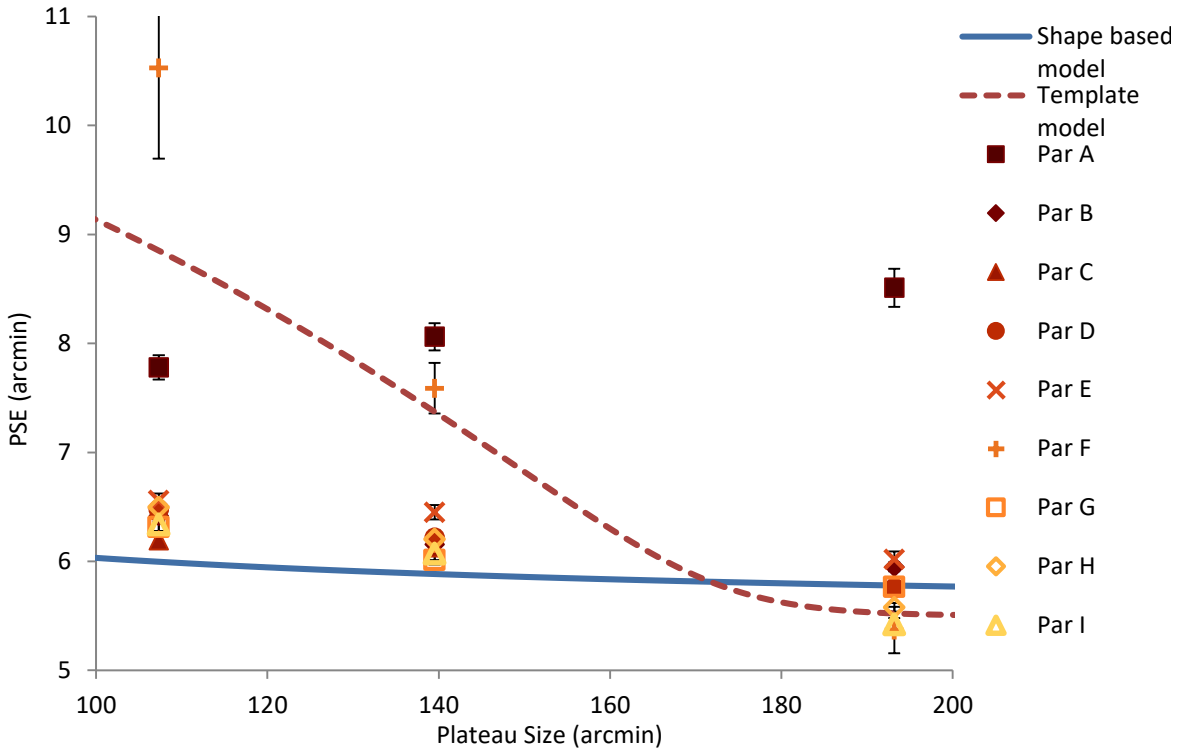


Figure 5.13: Participant data for Experiment 3 alongside the template and shape based predictions. Error bars are 1SEM.

	Shape based	Template
Par A	49.2	7.9
Par B	1.9	236.3
Par C	1.1	229.2
Par D	2.8	179.1
Par E	4.2	141.4
Par F	18.7	33.5
Par G	1.7	275.2
Par H	5.4	276.6
Par I	3.8	218.2
Average	1.0	215.0

Table 5.3: χ^2 values for the nine participants that completed Experiment 3. A χ^2 of 1 indicates an excellent fit, greater χ^2 indicate progressively poorer fits. The model has no free parameters, so the model parameters are identical for all participants – all variation in χ^2 reflects variation in participant performance. Average is the X^2 measured to mean for each plateau size of all participants' data.

There are two outlier participants, however both of these participants had thresholds over three times those of other participants (Figure 5.13, Par A: 2.3 to 2.9 arcmin, Par F: 7 to 4 arcmin, the remaining participants had thresholds of 0.5 to 1.3arcmin). This indicates that they found this experiment much harder, and were likely resorting to different techniques to complete the experiment (possibly because of poor stereothresholds). Both these

participants have a very poor χ^2 value (see Table 5.3). Participant F (Orange plus) appears to be following a trend similar to the template model but has considerably larger error margins than the other participants. Despite the apparent goodness of fit to the template model however, analysis shows their performance to be a poor fit for either model $\chi^2 = 34$ for the template model, $\chi^2 = 18$ for shape based modelling). Participant A (Solid red square) seems to be doing something completely different, with their perception of the smooth stimulus being considerably flattened and almost independent on the plateau size.

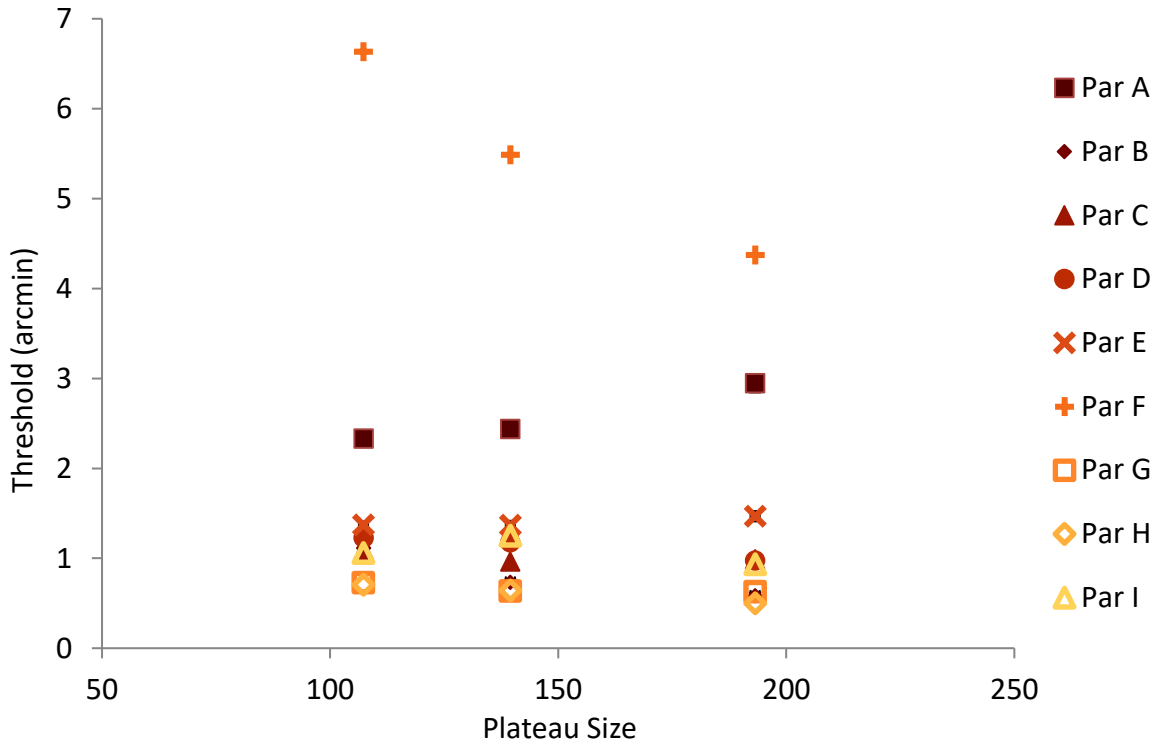


Figure 5.14: Threshold data for Experiment 3, error bars are too small to be visible on this scale.

5.5.5 Discussion

This experiment was designed to test if the size of the averaging window was determined from the smooth object (shape based averaging hypothesis) or from the sharp edged object (template based averaging hypothesis). We modelled the effect of each of these hypothesises when changing the plateau size of the smooth object to obtain predictions, and then compared these predictions to participants' performance to find the model prediction that best matched human performance. Overall, the majority of participants have a good fit to the prediction of the shape based modelling and very poor fits from the predictions of the template based modelling. We therefore conclude that it is most likely that the visual system is basing the size of the averaging window on the shape of the observed object.

The observed PSEs are slightly greater than predicted for the shape based modelling – some deviation from the prediction is to be expected as the model had no free parameters as

window size was fixed prior to taking experimental results. The model therefore cannot compensate for variation in participants' performance, as discussed in the introduction to this Section.

5.6 Overall discussion of model performance

For the first strand of this thesis, we are considering what a disparity defined object is. This Chapter explored an important aspect of object recognition – how the visual system behaves when an object has a smooth edge that is hard to segregate from its background. In the previous Chapter, we found that smooth edge caused a bias in the perception of depth of the object. Here we used those results to create a quantitative model of the visual system in order to better understand how the shape of our stimulus was interacting with the perceived peak depth of the object. Having developed a model that could fit participants' data, we then tested it by attempting to predict the results of Experiment 3 before we ran it on participants. We found that the model performed well, given that it had zero free parameters.

Here, we briefly review the model's performance by considering how it has performed at each of the three tasks: modelling Experiment 1; modelling Experiment 2; and predicting participant performance in Experiment 3. We discuss what this tells us about how the visual system is processing the peak depth and averaging across large regions and consider how this relates to the current literature on disparity averaging. Finally, we relate this back to animal camouflage, and consider how the discoveries of the model could be used to help an animal attempting to camouflage itself from a stereoscopic observer.

Given the simplicity and quantitative approach of the model, the model has performed well, and is able to explain over 80% of the variance for most individual participants when fitted to their experimental data from Experiments 1 and 2 (Figure 5.4 and Table 5.2). We were then able to alter the model to predict group behaviour prior to experimentation, with zero free parameters, the majority of participants' performance at a novel stimulus (Section 5.5).

Perhaps the most interesting aspect of the model for the first experiment (Section 5.3) was the shape of the window model used – the only window shape that could fit the data well was square (Section 5.3.4). As both the smooth and sharp objects were square-based, this strongly suggested that the shape of either the smooth or sharp object was prompting a form of segregation across the top of the object, followed by averaging to obtain an estimate of peak depth. This speculation was further strengthened by the best fit window sizes for both Experiments 1 and 2, which showed that the size of the averaging window used was close to the plateau size of both the objects (Figure 5.4 and Table 5.1) – a parameter that determined the distance between the inflection points in the smooth object and the distance between the edges in the sharp object.

We therefore had two alternative sources of the size of the averaging window: either we had a shape-based averaging, where the visual system segregated the edge of the smooth

object from its background and then averaged over the area to obtain a depth estimate; or template based averaging, where the coincident display of the well-defined sharp object was used as a template to determine the area of the smooth object over which to average. In Section 5.5.2 we created two models with no free parameters, one to predict the performance of a template based approach and one to predict a shape based approach. By altering the plateau size of the smooth object, we could create distinct predictions for these two cases, and compare them to participants' performance. We found that participants were closest to the shape based averaging, where the area averaged over is based on the object's own size and shape (Section 5.5.4).

The averaging aspect of our shape based averaging is not that remarkable in itself: it is well known (as discussed in the background theory, Section 2.2.4) that the early stages of disparity extraction result in the averaging across small scales e.g. (Allenmark & Read, 2010, 2011; Filippini & Banks, 2009; Goutcher & Hibbard, 2014; Tyler & Julesz, 1980). These effects are typically thought to occur at the scale of the finest scale disparity detectors at around 5arcmin across (Filippini & Banks, 2009; Harris et al., 1997). However, what we are hypothesising here that the averaging is occurring over much larger areas – at scales of over 150arcmin. Averaging effects have been observed over this length scale, for example in overlapping transparent planes (Kaufman et al., 1973; Parker & Yang, 1989; Stevenson et al., 1991; Tsirlin et al., 2008), however in this situation the elements of different disparities in close proximity, potentially allowing small scale averaging effects between adjacent elements to account for the large scale percept (Harris, 2014). Here, we have averaging occurring over a large area which is dependent on the shape of an object, with a large lateral separation between elements of different disparities. While averaging specifically over an object has not been proposed before, the long range effects share similarities to some studies on Gestalt grouping (Section 2.2.6). In particular, Deas and Wilcox found that perceptually grouping joining lines into objects causes a reduction in the perceived depth (Deas & Wilcox, 2014) – an effect that could be caused by the averaging mechanism hypothesised here.

The results of this model have potential ramifications on the second strand of this thesis, which explores if depth perception from binocular vision enables camouflaged objects to jump out from their background, making them trivial to detect. Here, we have found an interaction of the stimulus with the visual system that may enable an animal to confound one of the mechanisms of depth perception in a predator, thus removing this advantage. If an animal were to colour itself such that it would be segregated as many objects (called disruptive colouration, see Section 2.1.1 (Cuthill et al., 2005; Osorio & Srinivasan, 1991; Ruxton & Sherratt, 2004)), then each part of the animal that appeared as a separate object may be individually averaged over, giving each a separate depth estimate. This could potentially create the impression of many different objects, each divorced in depth from the other. A stereoscopic observer would therefore no longer see the animal as a distinctly

continuous shape standing above the background, thus ridding the stereoscopic observer of its advantage.

While this is an interesting speculation on the interaction of luminance and disparity, we do not know how the processes responsible for segregating an object based on disparity cues interact with other cues to object segregation. While work has been done on the perception of depth when disparity is placed in conflict with other cues to depth such as shape from shading e.g. (Chen & Tyler, 2015; Lovell et al., 2012), they do not look at interaction of depth perception with cues to object segregation. In the next Chapter we explore this gap: how do disparity based object segregation mechanisms interact with a luminance cue to object segregation?

6 How does depth segregation interact with luminance segregation?

- **Strand 1, How does depth from binocular vision contribute to object perception?**
- **Investigating: If luminance cues influence segregation of depth defined objects.**
- **Task: Which of two objects has a greater peak depth?**
- **Manipulation: One object has a luminance cue to segregation.**
- **Results: Luminance only has an effect when the disparity cue is poor.**
- **Conclusion: Luminance and disparity do interact, but in an RDS it is very hard to make the disparity cue weak enough for the luminance cue to have an effect.**

Figure 6.1: Two moths resting on the bark of a Jackfruit tree (one central, one bottom right). Note how brown luminance edges break up the moth into several smaller sections that are not moth-shaped. Image reproduced with permission, (Baliga, 2014)

6.1 Introduction

In the first strand of this thesis, we are exploring how disparity-defined depth influences object perception. So far, we have been exploring the perception of purely disparity defined objects – in Chapter 4 we have investigated the effect of a smooth edged object on the perception of disparity defined objects. We found that the objects were perceived with a decreased peak depth relative to the objective peak disparity. In Chapter 5 we used computational modelling to explore the mechanisms that could be causing this effect, and concluded that after object segregation the disparities within the object were averaged, possibly to remove errors from the disparity extraction process. However, poor segregation over the smooth object resulted in the inclusion of disparities that were lower than the peak, resulting in a perceptual decrease in their peak depth.

In this Chapter, the third and final experimental Chapter of the first research strand, we are interested to see if object segregation and subsequent averaging is purely driven by the disparity cues to depth, or if other cues to segregation are important. Luminance has been found to improve the judgement of the shape of rectangles (Regan & Hamstra, 1994), and to have an effect on human depth perception (Didyk, Ritschel, Eisemann, Myszkowski, Seidel, et al., 2012; Richards, 1977), but we do not know if this is a combination of luminance and disparity before or after object segregation and averaging as proposed in Chapter 5. To explore this, we introduce a luminance cue to object segregation, and investigate if there is a change in perceived peak depth with the size of the luminance cue. If the perception of depth in the object is then altered by the luminance cue to segregation, this will indicate that the luminance and disparity cues were combined to segregate the object prior to disparity averaging.

Adding luminance segregation to the disparity defined object created a surprising number of problems with the perception of the object. In this Chapter, we present several experiments:

1. In Experiment 4, we tested if luminance segregation of the object is a sufficiently strong cue to alter the perceived peak depth of a sharp object. We used the model from Chapter 5 to predict the results, but found that the addition of different luminance dots causes the experiment to be too hard for naïve participants to compete.
2. In Experiment 5, we explore a series of alterations to attempt to improve participants' performance while still testing the interaction of disparity and luminance cues to object segregation. This Experiment is the presentation of a series of interesting stimulus manipulations which lead to the creation of Experiment 6, rather than a single standalone experiment.
3. Experiment 6 combines the improvements of Experiment 5 with the luminance segregation methods of Experiment 4. Here, we find that when the disparity signal is

well defined and the luminance signal poorly defined (because we are only altering the luminance of the dots making up the RDS), then the luminance cue to segregation has no effect on the perceived depth of the object.

4. In Experiment 7, we investigate if decreasing the reliability of the disparity cue we could result in an interaction of luminance and disparity segregation. We add random noise to the disparity signal, and find that the perceived peak depth of the object is altered by the size of the luminance window, indicating that segregation based on luminance influences the area segregated and averaged over by the object based disparity processing.

6.2 Experiment 4: Does a luminance cue to segregation change the perception of a sharp object?

6.2.1 Introduction

Our previous experiments have always compared an object with smooth depth edges to an object with discontinuous (sharp) edges, as we were exploring the effects of object shape on perceived depth. However, if luminance is a sufficiently powerful cue to object segregation, it is possible that the luminance cue could cause the segregation of the object over the region defined by luminance independent a disparity cue to object segregation. If this is the case, then a luminance cue to segregation on the sharp object would cause the disparity averaging in the sharp object to occur over a region defined by the luminance cue.

If the luminance defined object was larger than the disparity defined object, then there would be elements of zero disparity contained within the luminance defined object. Segregation across the luminance object would then cause these zero disparity elements to be segregated into the same object as the non-zero elements of the disparity defined object. If luminance segregation of the object occurs before the disparity averaging modelled in Chapter 5, then the depth of the object would be calculated by averaging over both the zero and non-zero disparity elements. This would result in a decrease in the perceived peak depth of the sharp object. We therefore start by comparing two sharp objects to explore if a luminance edge larger than the sharp object can provoke a decrease in peak depth indicating that disparity averaging is influenced by luminance segregation.

6.2.2 Experimental stimuli

The two sharp objects used are identical in form to those in previous experiments, using white and black dots on a grey background. However, we wish to vary the luminance of the object in order to create a luminance cue to segregation. The sharp comparison object has a square luminance window with stimulus dots displayed across the full range of luminance, from white to black (12.24 cd/m^2 and <0.01cd/m^2). To create a luminance edge in the sharp test object, the luminance of the foreground dots are manipulated, either the white dots are made darker (light grey, 8.64 cd/m^2) or the black dots lighter (dark grey, 3.6 cd/m^2)

luminance. By altering the luminance of either the black or white dots, the overall luminance of the region will increase or decrease creating a luminance window. This change also alters the contrast which can effect perceived depth (Legge & Yuanchao, 1989; Rohaly & Wilson, 1999), so this window provides a cue to segregation of the object via luminance and contrast (Micheson contrast: target, 0.7; background, 1).

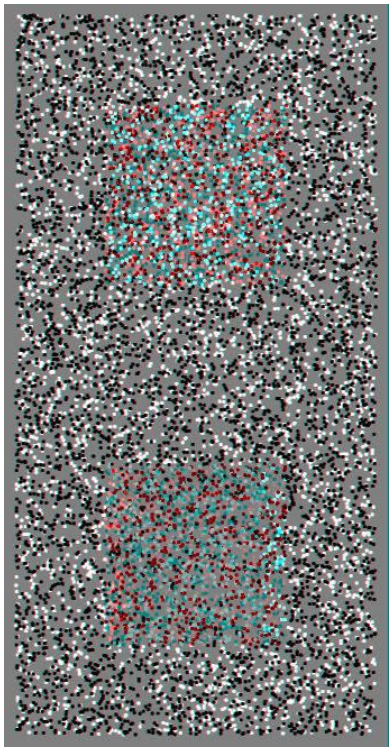


Figure 6.2: Stimulus for Experiment 4, with the sharp object top and bottom, but the bottom object has a low luminance window of 171.3 arcmin.

In order to create a luminance window, we must either increase or decrease the luminance of the dots. As there is no clear indication if we should either increase or decrease the overall luminance, each trial was randomized to either increase the luminance (high luminance condition) of the black dots within the window (thus increasing the overall luminance) or to decrease the luminance (low luminance condition) of the white dots (thus decreasing the overall luminance). We recorded on each trial if the luminance was high or low.

When determining the size of the luminance window we had to balance two conflicting problems: make the luminance window small and too few background elements would be included to measurably decrease perceived peak depth; make the window size too big and the luminance window would not appear to be connected to the disparity defined sharp object, and therefore could be disregarded when segregating the disparity defined object from its background.

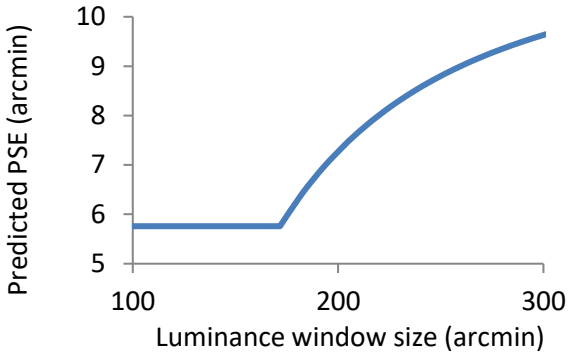


Figure 6.3: Predicted increase in PSE for a sharp edged object with different sized segregation windows.

To determine if the luminance window was large enough to create a significant difference in predicted results, we took the model from Chapter 5. In Chapter 5 the model averaged the disparities over a variable sized region of the smooth object to calculate a predicted perceived peak depth. Here, we altered it to average the disparities in the sharp object over a set range of window sizes. For simplicity, we assume that the segregation for averaging will be completely dependent on the position of the luminance edge, producing the prediction shown in Figure 6.3.

In order for this experiment to work, the visual system has to link the disparity defined sharp object and the luminance defined object as being the same object. During pilot experiments we found that if the region of altered luminance was too large then it appeared to be unrelated to the disparity defined object and was ignored. This creates a trade-off: If the luminance windows are too similar to the width of the object, then the model does not predict a measurable change in PSE from the background elements present in the luminance window. If the luminance window is too large, then the participant will discount it. We balanced this trade-off using pilot experiments to see if the participant perceived the luminance window to be part of the object, and chose the largest window size that appeared to be consistently perceived as part of the object.

We chose a luminance widow of 1.2 times the size of the disparity defined sharp object (206.1 arcmin). This was the largest window (and therefore the window size with the greatest increase in predicted perceived peak depth) which consistently appeared to be connected with the sharp disparity defined object to a couple of lab member pilots. The model predicted a PSE of 7.5arcmin, a detectable decrease of approximately the same magnitude the smoothest object in Experiment 1 (Section 4.3.3, Figure 4.6).

We use two sharp edged objects of identical objective peak depth that are displayed simultaneously. One object (the test) has altered dot luminance of the background, creating a luminance window 0.8 or 1.2 times the width of the disparity defined object. If the luminance window has an effect we expect the 0.8 window size to be the within uncertainties of the objective peak depth, and the 1.2 to be perceived with less peak depth. As in Experiments 1 and 2, the comparison object had a constant disparity of 5.8arcmin.

The experimental setup was the same as in Section 3.1 and experimental procedure was the same as in Section 4.2.2. As a reminder, the participants were asked to indicate using the up and down arrow keys “which object has a greater peak depth”. Participants’ stereoacuity was tested with a TNO test (“TNO Stereotest, Richmond Products,” 2014), and were rejected if they could not see depth in the sixth plate. Participants then completed a demonstration version of the experiment, which started with trials taking 10s, reducing to the 2s presentation that was used during the experiment. The data from the demonstration was checked to ensure that the participants could complete the task. The experiment itself was presented in blocks of approximately 300 trials, taking 10-15min to complete each block with a forced 60s break between blocks. Each trial was a 2s presentation of the stimulus, followed by a response prompt screen. After response, a fixation cross was displayed with a prompt to press the space bar to continue. There was only one experimental session which comprised of the demonstration and three blocks.

6.2.3 Results

Psychometric functions were fitted and the PSEs of the participants were extracted. Out of eight participants, six completed the demonstration with good enough results to continue, and only three had results good enough to fit psychometric functions (the other three did

not converge) and one of these three has a very flat psychometric function (participant C), such that the uncertainty on the measurement at a fractional window size of 1.2 is over 1,000 arcmin.

Figure 6.4 shows the mean PSEs for all three participants with extractable PSEs. Note the large amount of individual variation, the omitted error bars for Participant C for clarity (errors are in excess of 550arcmin in all conditions but the high luminance 0.8 window size). Results were not as expected, with all participants showing a decrease in perceived peak depth when the stimulus is displayed with a small window (0.8 times the size of the disparity defined sharp object). Here, the model predicted there should be no effect of luminance as all disparities within this luminance window are at peak disparity. Results for the large luminance window (1.2 times the size of the disparity defined sharp object) were very variable, with Participant A perceiving the same depth as the standard stimulus, Participant B and Participant C (whose results we should regard with scepticism due to the large uncertainty on this measurement) showing a lower PSE and thus an increase in perceived peak depth.

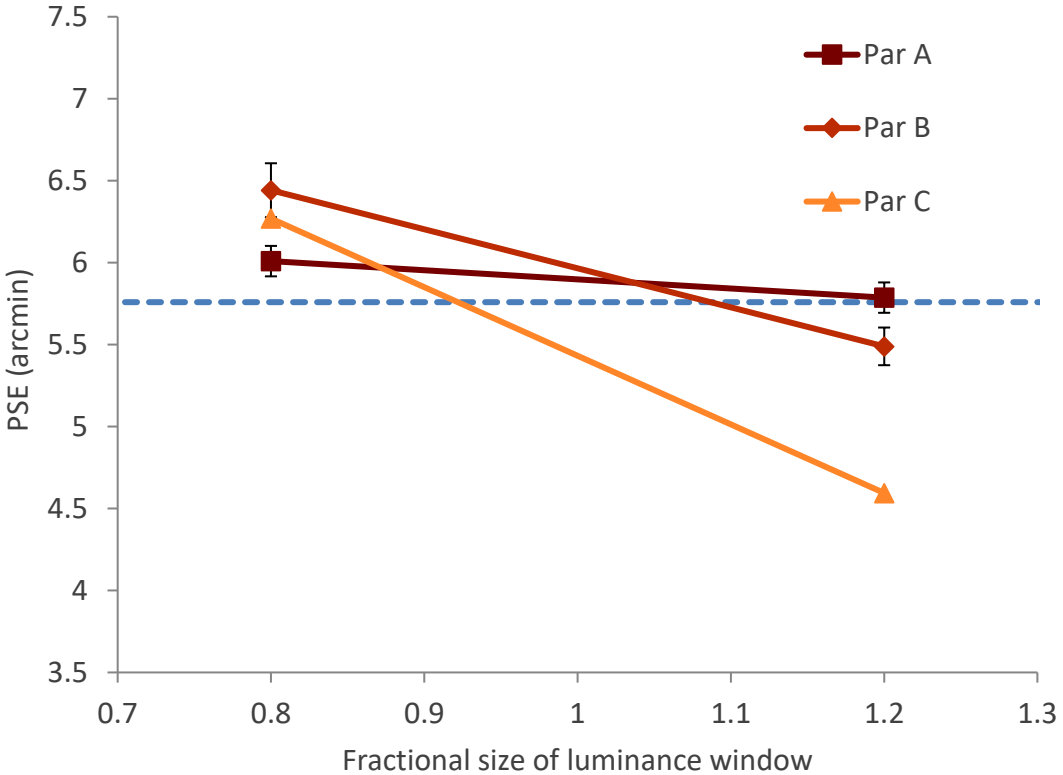


Figure 6.4: The PSEs for all three participants, the dotted line is the veridical depth estimate. Error bars are omitted for participant C, as they were all in excess of 200 arcmin.

As the window luminance was counterbalanced between high (light) and low (dark) luminance, we extracted PSEs separately for the light (high luminance) and dark (low luminance) conditions to ensure this there was no difference between these two conditions.

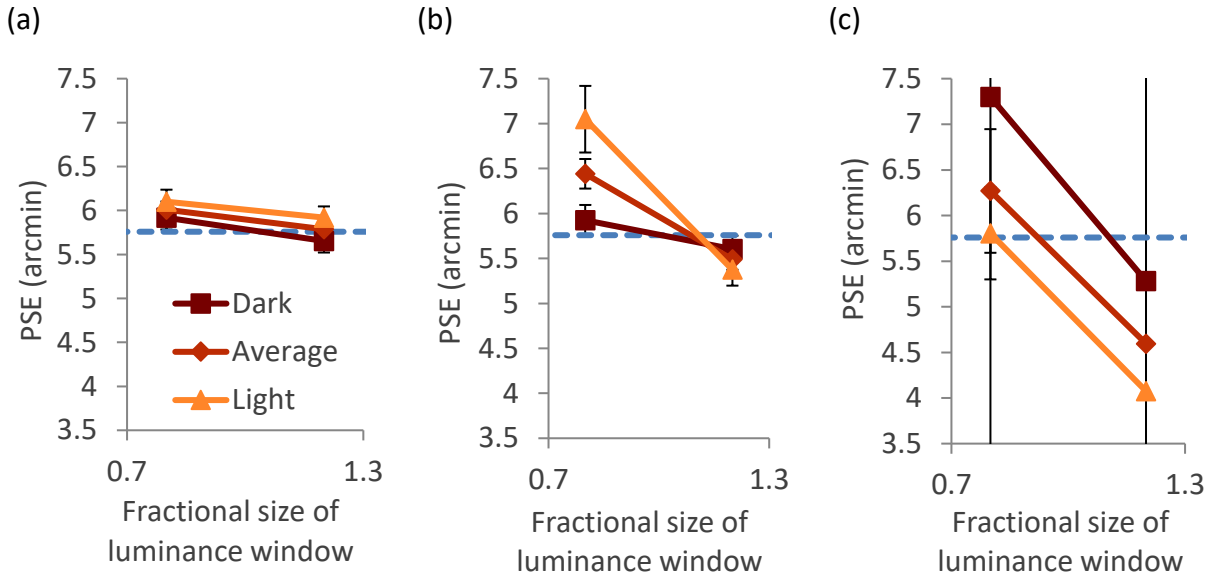


Figure 6.5: PSE's individually presented for each participant. (a) participant A, (b) participant B and (c) participant C.

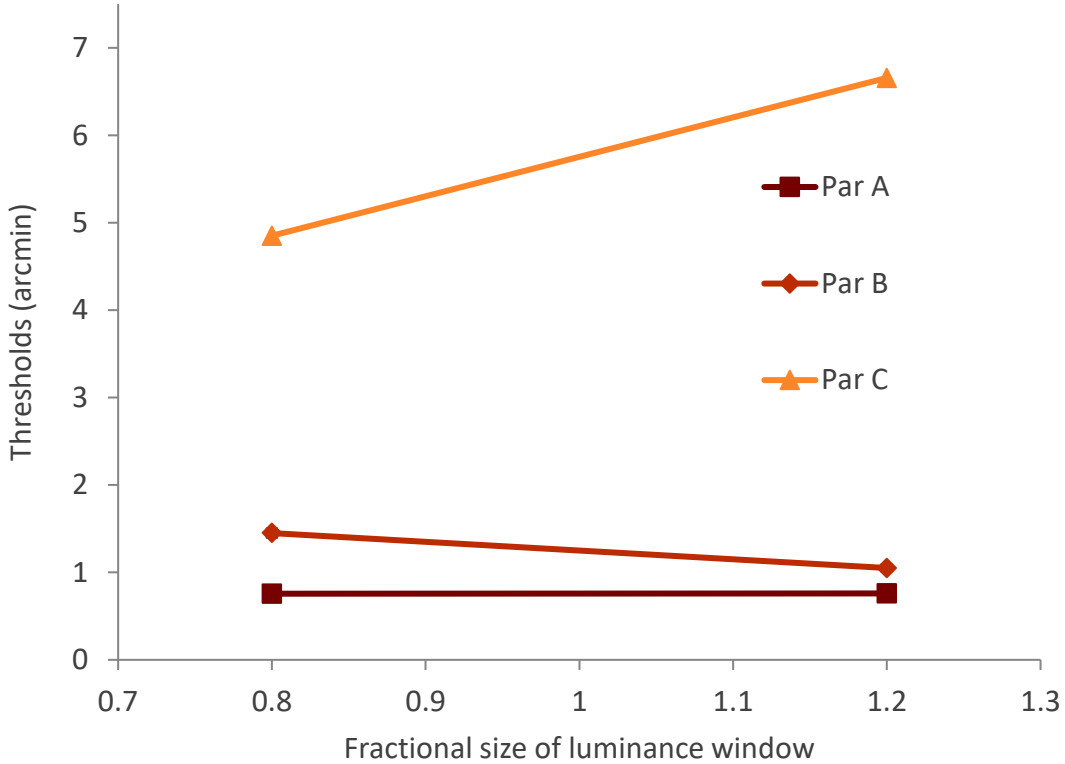


Figure 6.6: Thresholds for all three participants. Error bars are omitted for participant C, as they were infinite. Error bars of one SEM are present for participants A and B, but are too small to be seen clearly.

As the luminance of the dots should be above threshold, we know of no reason that there should be any performance difference between the light and dark luminance windows. Figure 6.5 shows the results compared to the combined (average) condition presented above.

Results show a large amount of participant variation, with Participant B surprisingly showing a very marked difference between the light and dark conditions.

Due to the shallow response curves, the threshold data is very variable, particularly for Participant C whose curves were so flat the error bars on the thresholds were infinite. For completeness, we include threshold data in Figure 6.6, with error bars stripped off for clarity.

6.2.4 Discussion

Overall, this experiment has shown some very odd and unexpected results, with extreme variation in the participants' performance, both in their results and their ability to complete the experiment. Given that only 3 of 8 participants had extractable PSEs, and of those only three only two have acceptable error margins, this suggests that the introduction of dot modulated luminance makes the experiment extremely hard to complete. We therefore suggest that while the data can be explored for trends, a more reliable experiment is needed to draw any conclusions.

Given the very consistent results of Participant A, it suggests that this participant ignored, or potentially could not see the luminance window for both conditions. Participant B's performance varied depending on whether the window was darkly or light coloured. The lighter luminance window appears to cause this participant to use the luminance window to segregate the object (as we hypothesized), decreasing the apparent depth of the object. Participant C clearly found the task very difficult, and with error margins in the thousands it is hard to evaluate their data.

It is unexpected to find a difference between the participants' performance in the high (black dots are displayed as dark grey, 3.60 cd/m^2) and low (white dots are displayed as light grey, 8.64cd/m^2) luminance conditions as the contrast of the windows was almost identical for both conditions (0.70 for high luminance, 0.71 for low luminance). We know that the luminance and contrast of the stimuli can alter the perception of disparity defined depth (Legge & Yuanchao, 1989; Richards & Foley, 1974; Schor & Howarth, 1986; Schor & Wood, 1983). When looking specifically at RDS, the luminance of the screen used is sometimes as low as 10cd/m^2 (Schneider et al., 1999; Tsirlin et al., 2008), but in these cases the dots are white on a black background, and it is not uncommon to use a luminance in excess of 60cd/m^2 (Harris et al., 2012; Stevenson et al., 1991; Yang & Blake, 1995). Read and Cumming (2003) used a random dot stereogram with white and black dots on a grey background, with a maximum luminance of 41cd/m^2. In order to ensure that the luminance of our screen is

not causing issues with the perception of depth, in the next experiment we increase the screen's brightness to its maximum, with white dots then being displayed at 37.09 cd/m^2.

The interaction of luminance and disparity raises other considerations about the experimental stimulus that we used. By defining the luminance edge via the dots, we have created a luminance edge which is signalled by either the white dots being made darker (low luminance condition) or the black dots being made lighter (high luminance condition). This means that the luminance edge is only signalled by half of the dots, and the other half are unchanged. As the dot density is 20%, this means that only 10% of the pixels on the screen have an altered luminance across the luminance defined edge, perhaps making the edge luminance poorly defined to a naïve observer. In grouping terms, there is not much more similarity between the elements of the luminance window than between elements of the background and the luminance window (see Section 2.2.6). In the next Experiment we make the luminance edge a property of the background rather than the dots, enabling us to form a distinct and continuous edge.

Although it cannot explain why there is so much participant variation, it is possible that the disparity edge being so sharp makes the object too reliably defined by disparity for a poorly defined luminance edge to contribute. This is a common observation in cue combination literature, where a reliable cue will often override the perception of an unreliable cue e.g. (Knill, 2007; Knill & Saunders, 2003; Lovell et al., 2012; Ryan & Gillam, 1994). This speculation is supported by the Participant A's data, they do not perceive any difference between the different sizes of luminance window. In order to make the disparity edge less well defined, we will return to comparing a sharp object to the smooth object, which has a poorly defined disparity edge between the object and the background – potentially increasing the amount that the visual system will rely on the luminance window as a cue to segregation.

6.3 Experiment 5: Exploration of segregation on other objects via background luminance

6.3.1 Introduction

In the previous experiment, we found that varying the luminance of a central area to create a luminance edge had odd and variable effects on participants' performance. It was thought that there could be three main contributing factors:

1. The luminance edge was too weakly defined by only modulating the luminance of the dots, and not the background.
2. The dot luminance was too similar to the background, making the dots hard to perceive.
3. The distinct edge of the sharp object is too reliable a cue to object edge for the luminance information to be considered.

In this Section, we attempt to rectify these problems by:

1. Creating a well-defined luminance edge by modulating the luminance of the background instead of the dots.
2. Applying two solutions to increasing dot luminance and visibility:
    a. Increasing the overall luminance of the screen, enabling us to have a much larger difference between the luminance values used.
    b. Changing the background to being black and the dots to all white, making the dots more distinct from the background (each dot now has a maximum contrast of 1, rather than 0.5 as in the previous experiment).
3. We use the smooth edged object, which has a continuous depth edge. This should reduce the reliability of the disparity edge as a cue to segregation, allowing for luminance information to be taken into consideration in segregating the object.

As most of the experiments presented in this Section are exploratory with few to no pilots and no full experiments, this Section is more of an overview of the attempts made to explore the issues and possible solutions to improve participants' performance.

6.3.2 Exploration of stimuli with a square background

This Section contains a number of pilot stimulus manipulations, each of which gradually changed the overall stimulus in an attempt to improve people's ability to perform the task. However, as none of the changes made a dramatic difference we will discuss them as a single alteration. All manipulations performed here used a constant sharp stimulus as a comparison to the smooth stimulus, with an increased maximum screen luminance of 37.09 cd/m^2.

In this Section, we explore pilots with lab members and colleagues as participants, rather than naïve participants. This enables easy recruitment, and these people were willing to repeatedly come back and pilot small changes to stimuli, enabling us to monitor the effects of multiple small changes to the stimuli.

The stimuli still comprised of random dot stereograms of the same density. We still asked participants to judge which of two objects has a greater peak depth. Both the objects had the same disparities as presented previously in this Chapter and in Chapter 4, with the comparison sharp object having constant disparity and the test smooth object having variable disparities.

6.3.2.1 Moving luminance edge from dots to the background

The first change was to remove the luminance edge from the dots and instead modulate the luminance of the mid-grey background to being either light or dark grey within a central square area (similar to Figure 6.7). While this increased the visibility of the luminance cue to segregation, it introduced an additional problem that we dub the 'curtain effect'. This is where the disparity of the edges of the square luminance region dictate the overall perceived depth of the entire luminance window, but not the dots – effectively the square luminance edges appeared to be part of a depth plane located behind the dot-disparity defined object. This was a serious issue, as it divorced the luminance-defined background from the smooth edged object, meaning they appeared as two independent objects. This potentially introduces effects explored in stereo transparency literature as discussed in Section 2.2.6, such as poorer depth thresholds (Wallace & Mamassian, 2004).

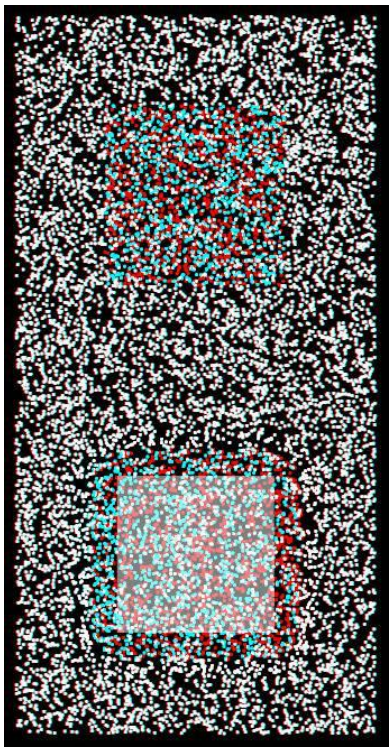


Figure 6.7: Anaglyph of square stimuli in Experiment 5. Sharp object top, smooth with a square luminance window of zero disparity bottom. Note how the constant luminance square appears behind the dots which we call the 'curtain effect'.

6.3.2.2 Alteration of dot luminance

The second change was to alter the object such that the background was black (<0.01 cd/m^2) except for a mid-grey luminance square (18.67cm/m^2) with all dots being white (37.09cd/m^2) – see Figure 6.7 for a screenshot. The luminance window now has a Micheson contrast of 0.5, with a background contrast of 1. While this helped with the visibility of the dots and luminance window (compare Figure 6.7 to Figure 6.2) the curtain effect became more pronounced, with the square now clearly being distinctly perceived as having a separate depth, and therefore being a separate object to the dot defined surface.

6.3.2.3 Reducing the curtain effect

We attempted to reduce the curtain effect by setting the disparity of the luminance window to being equal to that of the edge of the object. For example, for a luminance window of 171 arcmin (the plateau size of the smooth objects used), then disparity is set to half the peak depth. Unfortunately, for the smooth object, lines of constant disparity are not square in shape –therefore giving a square a constant disparity did not help, as the corners of the square appear to stick out above the dot defined surface as shown in Figure 6.8 (despite the apparent small scale in the diagram this was visible when undertaking the experiment). This is potentially also part of the problem with the previous experiment in Section 6.2, where a square luminance window was applied to the dot luminance. The creation of the correct

shape to match the shape of the smooth object as a uniform background patch is non-trivial in MATLAB®, so we turn to a different shape object where this is not a problem – circular stimuli.

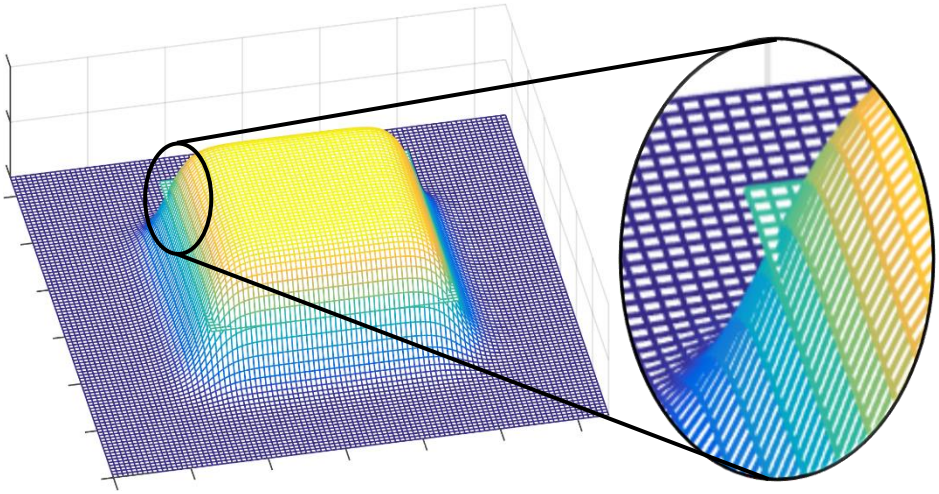


Figure 6.8: Representation of the flat square with disparity equal to the half depth of the smooth object. Note how the constant disparity of the square means disparity is noticeably larger than the smooth objects at the corners.

6.3.3 Circular stimuli & monochromatic dots

6.3.3.1 Stimuli

Circles are very easy to display using the Psychophysics toolbox in MATLAB® so we altered the formula for the smooth edged object (see Equations 6.1 and 6.2) to generate a smooth curricular object using plain polar coordinates. Unfortunately, when integrated the circular smooth stimulus is only dependent on σ^{-1} not σ^{-2} (due to only one $f(r,p)$ term in Eq. 6.2), so we expect a smaller decrease in perceived peak depth than for the square-based stimulus. Therefore, in order to trial the circular objects, we initially piloted the experiment with two sharp edged objects to test if naïve participants could complete the task given circular objects and circular luminance windows. If pilot participants could do the task with circular stimuli, then we would explore the effect of smooth circular objects to see if there was a measurable decrease in perceived peak depth in these stimuli.

$$f(r,p) = \frac{1}{2}\left[\tanh\left(\frac{1}{\sigma}\left(r - \frac{w-p}{2}\right)\right) - \tanh\left(\frac{1}{\sigma}\left(r - \frac{w+p}{2}\right)\right)\right] \quad \text{Eq. 6.1}$$

$$\delta(r,\theta) = \delta_p f(r,p) \quad \text{Eq. 6.2}$$

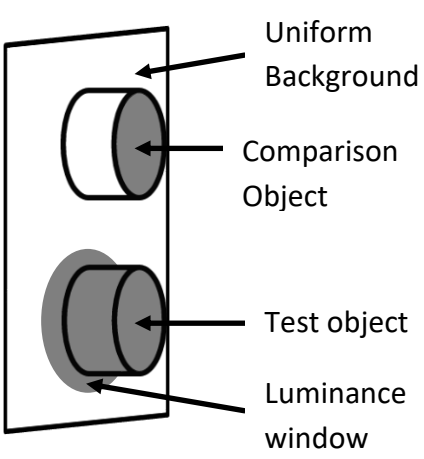


Figure 6.9: Diagram of the displayed depth of the luminance windows for the circular object.

Both sharp edged disparity defined objects were of identical size (radius 96arcmin, black outlined cylinders in Figure 6.9). The test object has a luminance window with a radius of either 74arcmin (0.8 times the size of the disparity defined object, as shown in in the anaglyph in Figure 6.10), or 111.6arcmin (1.2 times the radius of the disparity defined object), as shown diagrammatically as the dark grey region on the lower cylinder in Figure 6.9. The comparison object has a luminance window of radius identical to the disparity defined object (96arcmin), as represented by the dark grey circle on the top cylinder in Figure 6.9.

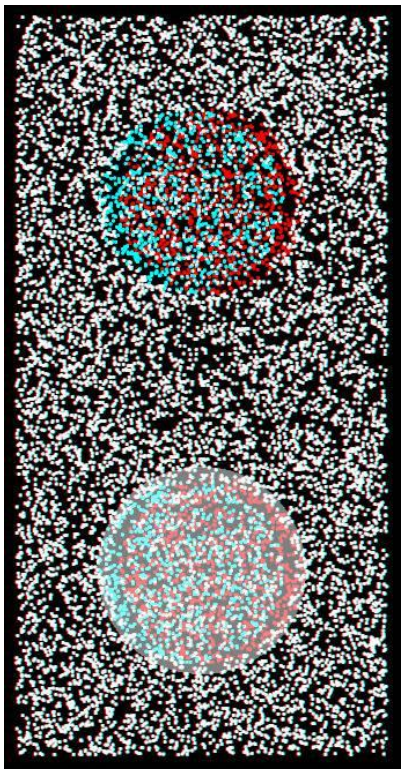


Figure 6.10: Anaglyph of circular stimulus in Experiment 5. Comparison object (top) and test object with luminance window of 1.2 times object size (bottom). Note that the region with no dots is only due to the exaggerated disparity used in this screenshot.

In order to reduce the curtain effect, we added disparity to the luminance window that was consistent with the disparity of the dots at the edge of the window. The window with a fractional size of 1.2 has zero disparity, while 0.8 has the disparity of the peak of the test object. The disparity of the window with fractional size of 0.8 and the peak depth of the test object were stepped between seven values equally spaced between 5.8 and 8.4 arcmin. The luminance window on the comparison object and the peak depth of the comparison object had a disparity of 5.8arcmin.

Pilots were able to complete this experiment with extractable PSEs without reporting a perception of the curtain effect. Therefore, we ran this experiment on naive participants. Our main objective is for the experiment to be able to be completed by the majority of participants as we expect the disparity edge of the sharp object to be too well defined for the luminance edge to strongly influence segregation. However, it would be interesting to confirm this hypothesis, if the experiment was doable by naïve participants. If the experiment was successful, we would have then moved onto smooth circular stimuli.

6.3.3.2 Results – circular stimuli

Figure 6.11 shows the results for circular stimuli, where the background was black (<0.01 cd/m^2) with a mid-grey (18.67cd/m^2) luminance window and white (37.06cd/m^2) dots. Out of eight participants only seven were able to complete the demo stage. Out of these, two had PSEs for both luminance window sizes outside of the measured range of disparities, and therefore was rejected. From the remaining five participants, only one, Participant D had extractable PSEs for both the 0.8 and 1.2 luminance windows. The remaining three were capable of completing the experiment with the smaller 0.8 luminance window (Participants E, F, G and H see Figure 6.11). Out of these, all but Participant F had PSEs far outside of the displayed disparity range for a window 1.2 times the diameter of the stimulus. Note that normally participant F's data point at 1.2 should be excluded, but as it is only just outside of the range it is included for interest. Additionally, several participants self-reported the circular luminance window as appearing as a separate object to the disparity defined object.

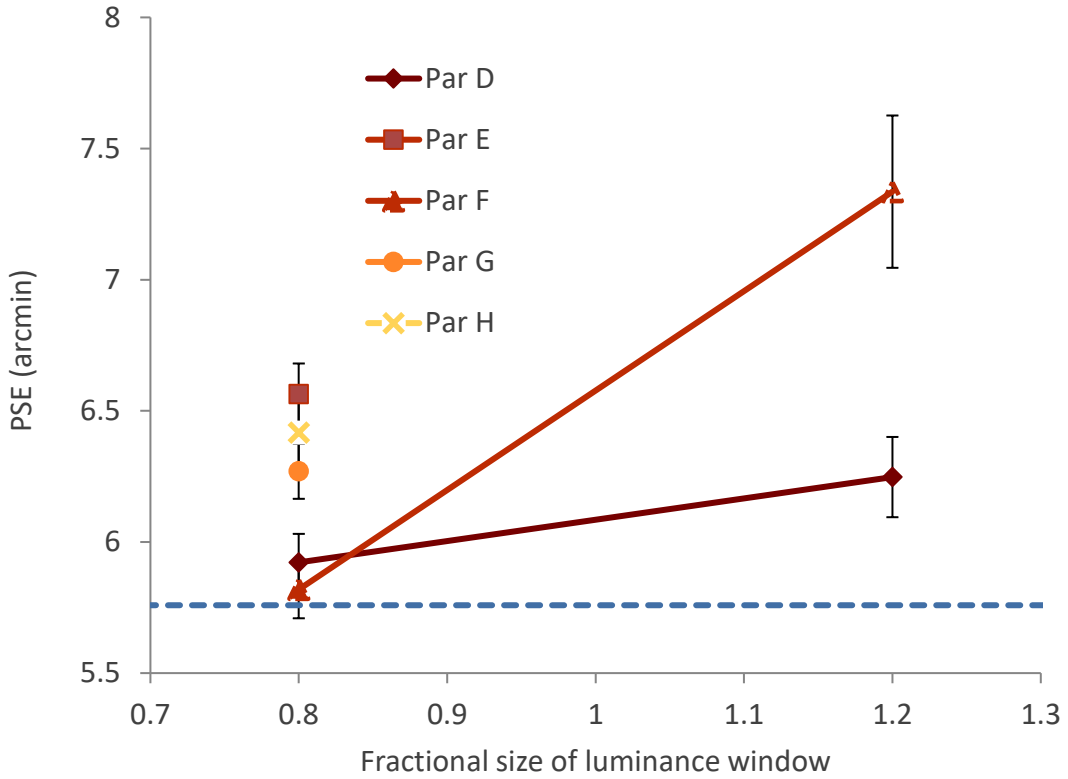


Figure 6.11: The PSEs for five participants. No PSEs were extractable for participants E, G and H for the 1.2 luminance window, and the 1.2 for F is extrapolated, not interpolated and therefore likely has additional errors over the calculated standard error bars.

As the data is in the majority of cases completely flat, we cannot extract meaningful thresholds (typically these were on the order of infinity).

6.3.4 Discussion

By using a circular window with white dots and a mid-grey luminance window on a black background, we appear to have made the task easier than in Experiment 4 as some naïve

participants can complete the experiment with meaningful results. However, we still have a low success rate for a psychophysical task – only one of eight participants were able to complete the entire experiment with usable results. Four had good performance on only the smaller luminance window.

Interestingly three participants consistently chose the comparison stimulus almost all the time for the larger luminance window. One interpretation is that they are viewing the stimulus with the larger window as very much flatter; however, given the poor past performance at the luminance window experiments it is probable that these participants were ignoring the dots, comparing the disparities of the luminance windows. The disparity of the 1.2 luminance window was 0 arcmin, and the disparity of the 0.8 luminance window was identical to the peak depth of the test object. If participants were comparing the disparity of the luminance window to the comparison object, then we would expect participants to always select the comparison object as having a greater peak depth when the test object was displayed with a luminance window of 1.2. When the test object was displayed with a luminance window of 0.8, then we would expect the participants to act as unbiased observers. This is approximately what these three participants did, although it does not explain the minor flattening effect observed in the 0.8 radius stimulus. Given the number of participants that could not complete the experiments at all, the only solid conclusion we can take is that these participants were unable to correctly complete the task as requested, and were each following a different strategy.

The circular objects provide multiple more problems: the continued reporting of the curtain effect when viewing a smooth edged circular object; the expectation of a decreased flattening effect when compared to square objects; and being much more complex to model. We will therefore return to the square based stimuli in future experiments. However, this experiment and its pilots seem to indicate from the increased performance (and self-reporting from pilots) that the use of white dots on a black background is easier to perceive than the mid-grey background with white and black dots. We will therefore continue using this regime in future to increase clarity of the disparity defined object.

Potentially the most likely avenue to produce a usable result without the curtain effect is to revert to displaying the luminance edges using the dots themselves. As the dots make up the disparity defined objects, then it should be clear that the luminance boundary is a property of disparity defined object, and not a second object which is individually defined in depth. A similar technique to this failed in Section 6.2, but the change in dots to all being white and the increased overall luminance may alleviate the problems found there – we discuss this in detail in the next Section.

6.4 Experiment 6: Segregation on smooth objects via dot luminance

6.4.1 Introduction

We have previously tried several different iterations of the experiment, attempting to find a way to successfully cause a luminance segregation cue to be associated with the disparity cue. In the previous experiment, we have tried changing the luminance of the background to create a luminance edge near where disparity segregation occurs. Unfortunately, this has consistently created a percept of a shaded shape that is unrelated to the disparity defined object which we call the curtain effect. However, other changes such as the use of a black background and white dots helped stabilise the percept of the object. Here, we attempt to fuse the luminance edge with the disparity defined object by changing the luminance of the dots that make up the disparity defined object. This should clearly associate the luminance as a property of the disparity defined object, potentially allowing the luminance edge to interfere with the disparity segregation processes.

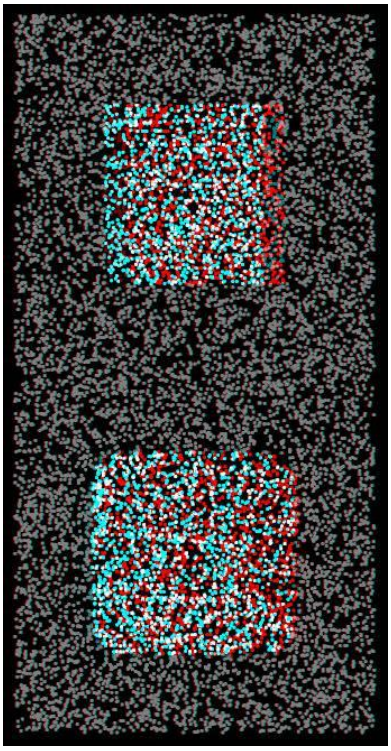


Figure 6.12: Anaglyph of stimuli in Experiment 6. Sharp object (top) and smooth (bottom), with a 1.125 fractional luminance window on the lower object

Although in Section 6.2 we found that adding the luminance cue to the dots caused the experiment to be very hard to complete, we have many improvements that should help:

1. We are using a black background (<0.01cd/m^2) with mid-grey (18.67cd/m^2, Michelson contrast 0.5) and white dots (37.09cd/m^2, Michelson contrast 1), which should ensure both object and background dots visible.
2. Originally in Experiment 4 (Section 6.2), only half of the window dots (either the white or black dots) had a different luminance to the background dots. In this experiment, all the dots within the window will be white and all the dots outside will be mid-grey, making the luminance window better defined, with no dots in the window having the same luminance as dots present in the background.
3. We are going to reduce the sizes of the luminance windows to the minimum possible that still creates a sufficient predicted difference to be experimentally detectable – this could further increase the relevance of the luminance cue to segregation.
4. Rather than applying the luminance increase over a set square area, we are going to display dots as white if they are above a certain disparity and black if below. This

means that the shape of the luminance window will correspond fully to the shape of the smooth object. We discuss this below.

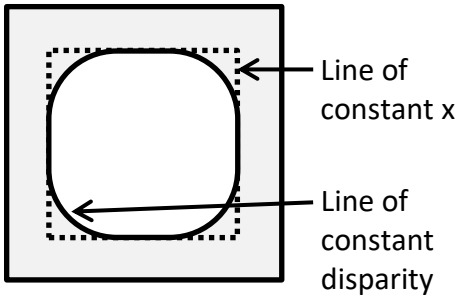


Figure 6.13: Difference between the square luminance window in Experiment 4 (dotted line, Section 6.2) and the constant disparity defined luminance window used here (solid line)

We choose a horizontal distance from the centre of the window, and calculate the disparity of the smooth object at this point. We then draw a line of constant disparity, which follows the curved shape shown in Figure 6.13. We use this shape as our luminance window, which means that the shape of the luminance window now matches the shape of the smooth object.

6.4.2 Stimuli and methods

We return to the previous square shaped sharp and smooth objects. However, we keep the black background and white dots from the previous circular experiment as this appeared to increase the visibility of the object and the luminance window. The central region of dots inside the luminance window is white (37.09cd/m^2, Michelson contrast 1), while the dots outside of the window are mid-grey (18.67cd/m^2, Michelson contrast 0.5) on a black background (<0.01cd/m^2) as shown in Figure 6.12. Dots outside the window are still easily visible, but the difference of dot luminance of 18.42 cd/m^2 across the should make it clear that the luminance window is part of the disparity defined object, not a second object that is hanging behind a transparent disparity defined curtain. Additionally, the sharp object had the same increase in luminance that coincided with the disparity edge (square, side length of 171.1arcmin) to reinforce the luminance window as a cue to object segregation.

Smooth objects are displayed at smoothness coefficients of 3, 14 and 26 with seven different peak depths as in Experiments 1 and 2, and compared directly to a sharp object of constant disparity. The luminance cue on the sharp object always corresponded to the disparity edge (171.2 arcmin), while the luminance cue on the smooth object was either 150.2, 171.2 (the same as the plateau size) or 193.2 arcmin across (fractional plateau size 0.875, 1 and 1.125).

We ran the different luminance window sizes through the model developed in Chapter 5 to obtain the predicted perceived peak depth for the objects with every combination of luminance windows size and smoothness coefficient. The model was run as if the luminance window dictated the area of the disparity defied object that was segregated and then averaged – this is the maximum effect that the luminance window could have on the perceived peak depth. These predictions are plotted in Figure 6.14.

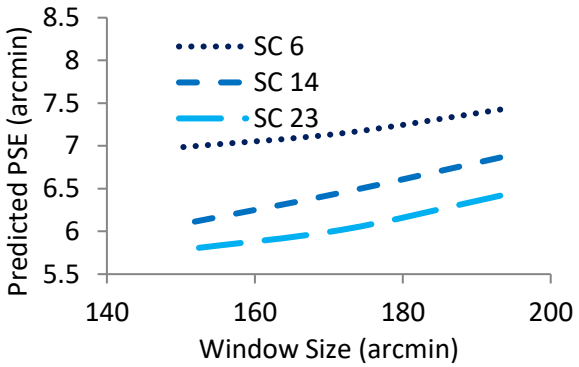


Figure 6.14: Predicted perceived peak depth for objects with each smoothness coefficient with the size of the luminance window

The experimental setup was the same as in Section 3.1 and experimental procedure was the same as in Section 4.2.2. Participants were tasked to press either the up or down arrow buttons to indicate which object had a greater peak depth. Participants completed seven blocks of approximately 300 trials each, each of which took 10-15min to complete with a forced 60s break between the blocks. These blocks were completed in two sessions: In the first hour long session, participants completed the TNO test, then a demo, participants that could not complete these successfully (see Section 4.2.2) were rejected from further study. Participants then completed three blocks. In the second session Participants completed the remaining four blocks.

6.4.3 Results

Psychometric functions were fitted and the PSEs were extracted for each combination of smoothness coefficient and window size for each participant. We appear to have made the experiment much simpler to complete, with only one participant unable to complete the demo and two having PSEs outside of the extractable range, leaving five out of eight participants completing with usable data.

Each results graph in Figure 6.15 plots PSE against luminance window size for an object of a single smoothness coefficient. By comparing between the three graphs, we can see that the PSEs are greater in Figure 6.15c (SC14) than Figure 6.15a (SC3) and b (SC0), and greater in Figure 6.15b (SC3) than Figure 6.15a (SC0). Effectively, objects of greater smoothness coefficient have a greater PSE. This replicates the data in Chapters 4 and 5, where an increase in smoothness coefficient results in a decrease in the perceived peak depth of the object.

Looking at the change of window size with PSE within each of the three graphs in Figure 6.15 there appears to be no relation between the model-predicted PSEs (dotted line) and the observed data. Statistical analysis shows that there is no significant effect of window size on observed PSE at any smoothness coefficient (two way repeated measures ANOVA for

window size, F(1.094,4.374) = 0.501, p=0.531 for window size and smoothness, F(1.139,4.555) = 0.501, p=0.534 with a Greenhouse-Geisser correction for invalidation of the assumption of sphericity p<0.05).

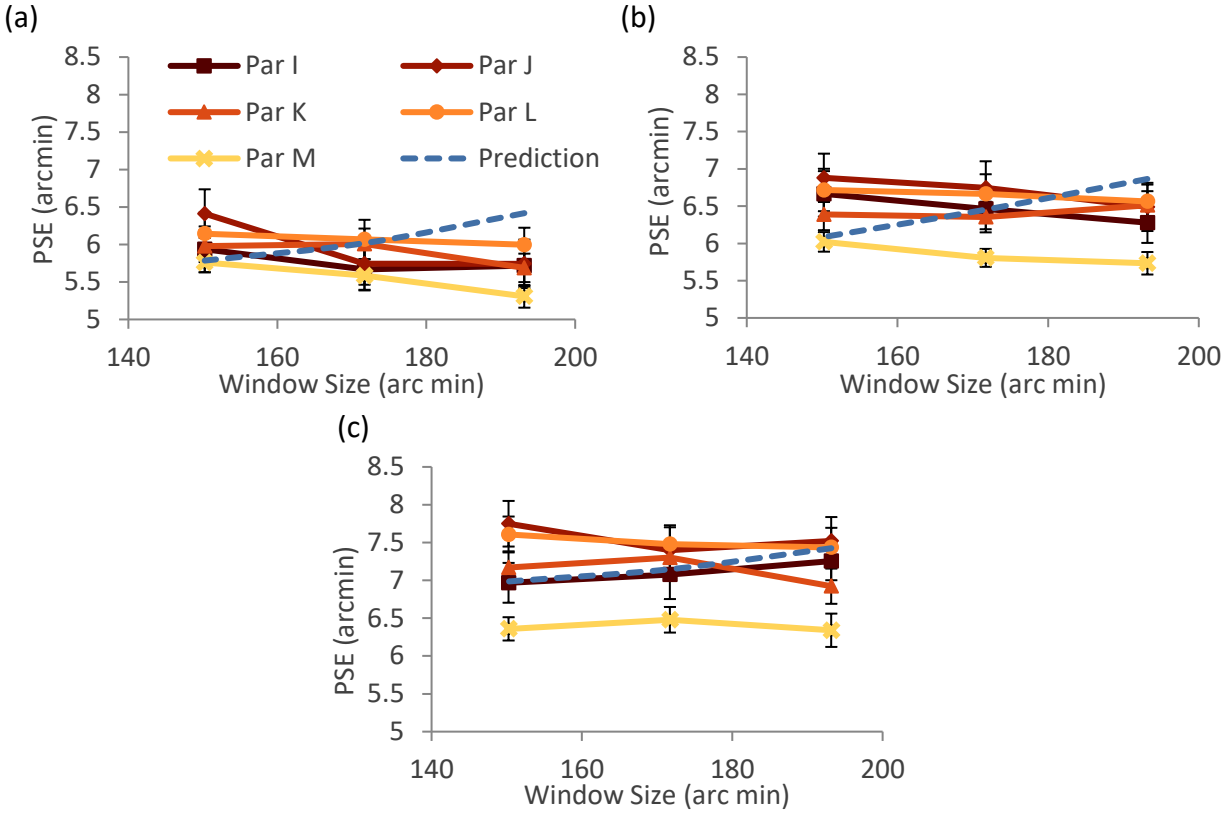


Figure 6.15: Predictions alongside participant performance with smoothness coefficients (a) 26 (b) 14 (c) 6.

6.4.4 Discussion

We successfully made the experiment simple enough that participants were able to complete it without too much difficulty, although the error bars and dropout rates (2 out of 7 have extractable PSEs) are still a little large in comparison to the experiments in Chapters 4 and 5. However, the surprising result is that there is no effect at all of the luminance window size on the perceived disparity. This indicates that the luminance window is not affecting the area over which disparity segregation is taking place.

As the disparity is very well defined, even the smooth edge may provide a strong cue to object segregation. The luminance cues we have presented may be comparatively poor cue to the object's edge as the luminance was delivered only via the positions of the random dots. As these dots are randomly placed, the edge of the luminance window appears uneven rather than a distinctive straight line, potentially weakening the strength of this cue to segregation. This may mean that the visual system is using the disparity cue to segregate the object and ignoring the more unreliable information provided by the luminance window.

A similar problem was encountered by Lovell et al (2012) when investigating the combination of shape from shading and shape from disparity – they found that the disparity cue could be strong enough that the shape from shading cue was ignored. We will therefore try the next step that they took in alleviating this problem – adding random noise to the disparity cue to make the disparity unreliable and harder to judge (Harris & Parker, 1992), thus forcing the visual system to rely more on the presence of the luminance cue for segregation.

6.5 Experiment 7: Does random disparity noise cause a luminance edge to change the perception of peak depth?

6.5.1 Introduction

In the previous experiment, we managed to create a stimulus such that the majority of participants could complete the experiments. However, despite this success the experiment delivered a null result. We hypothesise that the luminance edge did not have an effect on the disparity segregation due to the respective noise in the signals – the disparity is very well defined on a continuous surface for all dots. The luminance edge is comparatively poorly defined, with the fact that the dots convey the cue causing a very ragged edged square shape, as can be seen from the screenshot in Figure 6.12. It is commonly known in cue conflict and cue combination literature that if one signal is very reliable, then other unreliable cues to depth will be reduced or even discounted e.g. (Chen & Tyler, 2015; Hillis et al., 2004; Knill, 2007). In our case this potential asymmetry in cue strength may have caused the visual system to ignore the segregation from the luminance in favour of the disparity cue.

In order to test this hypothesis, we present an experiment where the disparity has random noise added in order to degrade the disparity signal (Norman, Lappin, & Zucker, 1991). The luminance cue remains identical – therefore the addition of disparity noise could push the visual system towards reliance on the luminance cue for segregation. As before, we compare the perceived peak depth of the objects to the predicted perceived peak depths if the luminance window dictated segregation (generated using the model from Chapter 5).

6.5.2 Stimuli and methods

Stimuli are presented identically to the previous experiment (Experiment 6, Section 6.4.2), except this time we use disparity noise as in Lovell et al (2012). In order to weaken the disparity signal, after generation of the left and right halves of the RDS, each dot in each half of the RDS is moved to the right by an angle randomly generated from a normal distribution with mean 0 arcmin and standard deviation of 0.58arcmin. Random disparity noise was added to each dot in each eye individually, which could decrease the reliance on disparity defined segregation and increase reliance on the luminance window for object segregation.

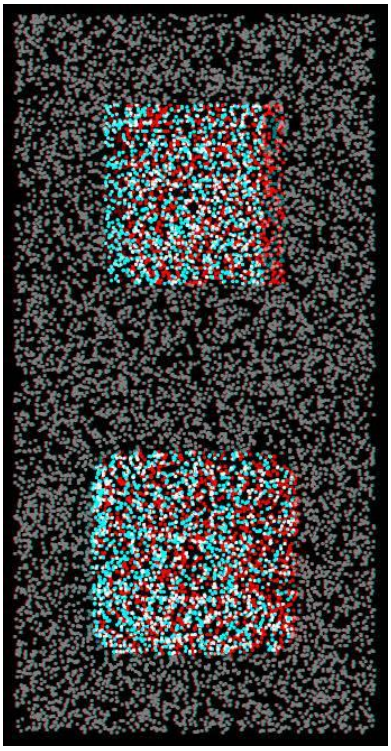


Figure 6.16: Anaglyph of Experiment 7. Sharp object (top) and smooth 10% disparity noise (bottom) with a 1.125 fractional luminance window on the lower object.

As in the previous experiment, we use the model in Chapter 5 to predict the expected PSEs for the different window sizes if the area segregated then averaged was completely defined by the luminance window. As can be seen from Figure 6.14, the objects with smoothness coefficients 3 and 26 have minimal difference between their predicted PSEs at the extremes of window size. Therefore, we only run this experiment with a single object of SC14. This brings the total experiment duration down to one hour, consisting of the TNO test and demo followed by three blocks of approximately 300 trials each, with a forced 60s break between each block.

6.5.3 Results

Eight out of nine participants completed the experiment, with one being excluded by the TNO test. Participant H could not correctly see depth at a window size of 193arcmin (with a PSE of around 35arcmin and a bootstrapped SEM of 3000arcmin) – the data is included for interest however. The remaining participants are clustered into two groups, as can be seen in Figure 6.17, firstly Participants A D E and F, and a second group comprising of participants B C and G (indicated with dashed lines). Due to this clustering, we separate the participants into two groups – those with an increase in PSE with window size (A,D,E and F) and those with a decrease in PSE with window size (B,C and G) and analyse them individually.

Participants A D E and F showed the expected trend similar to the blue dashed model prediction, with a significant change in PSE with window size (repeated measures ANOVA on Participants A, D, E and F only, $F(2,32)=0.01$, $p=0.001$). It can be seen from Figure 6.17 that, as expected, the PSE increases with increasing window size – in line with the prediction that

segregation and averaging are affected by the luminance window. However, the magnitude of this effect is extremely variable between these participants.

Participants B, C and G (the three with the smallest PSE at a window size of 190 arcmin, marked by dashed lines in Figure 6.17) were less effected by window size than the previous group. For this group, the window size has no significant difference on their perception of the object (repeated measures ANOVA on Participants B, C and G only, $F(2,4.3) = 0.10$, $p=0.101$).

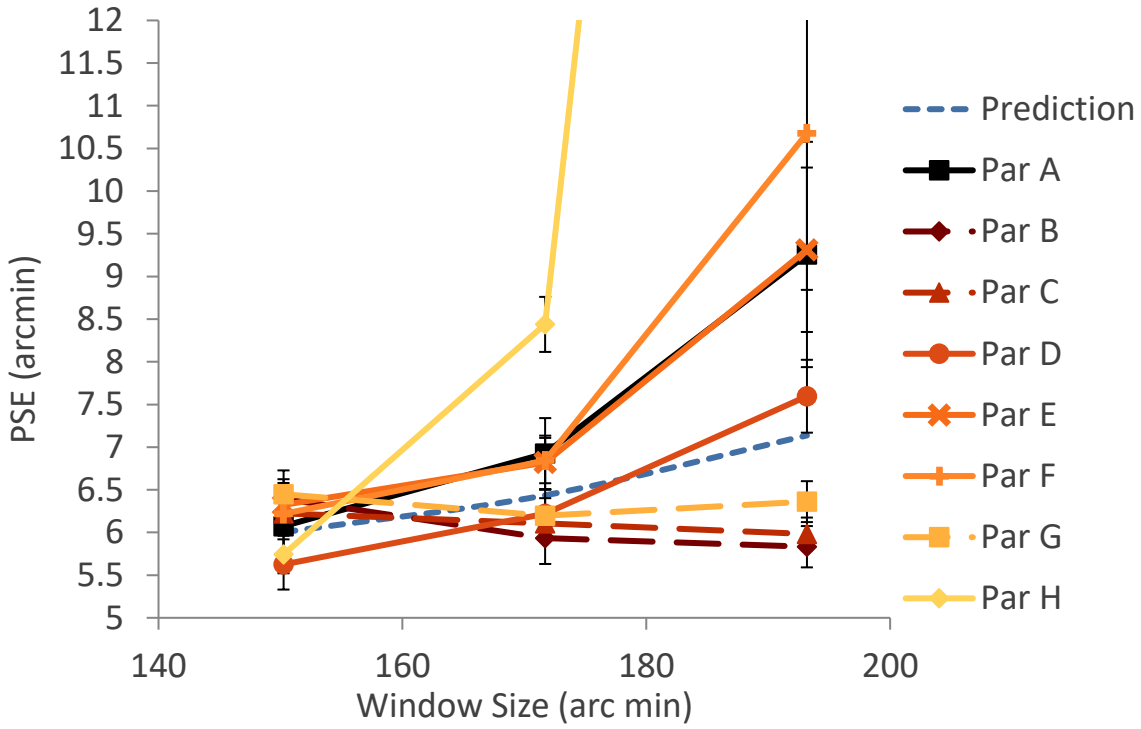


Figure 6.17: The PSEs with window size for all participants in Experiment 7, with prediction. Error bars are one SEM.

Looking into threshold data with the same participant split (Figure 6.18), there is no significant effect of window size on thresholds for participants B C and G (repeated measures ANOVA, $F(2,0.12) = 0.988$, $p=0.99$). Participants A D E and F however show a significant effect (repeated measures ANOVA, using the Greenhouse-Geisser correction as the assumption of sphericity is violated ($p=0.02$) we find $F(1,13) = 0.03$, $p=0.033$). Unlike manipulation of smoothness coefficient, this implies that the increase in luminance window cue is making the task harder for those affected by it as well as creating a bias.

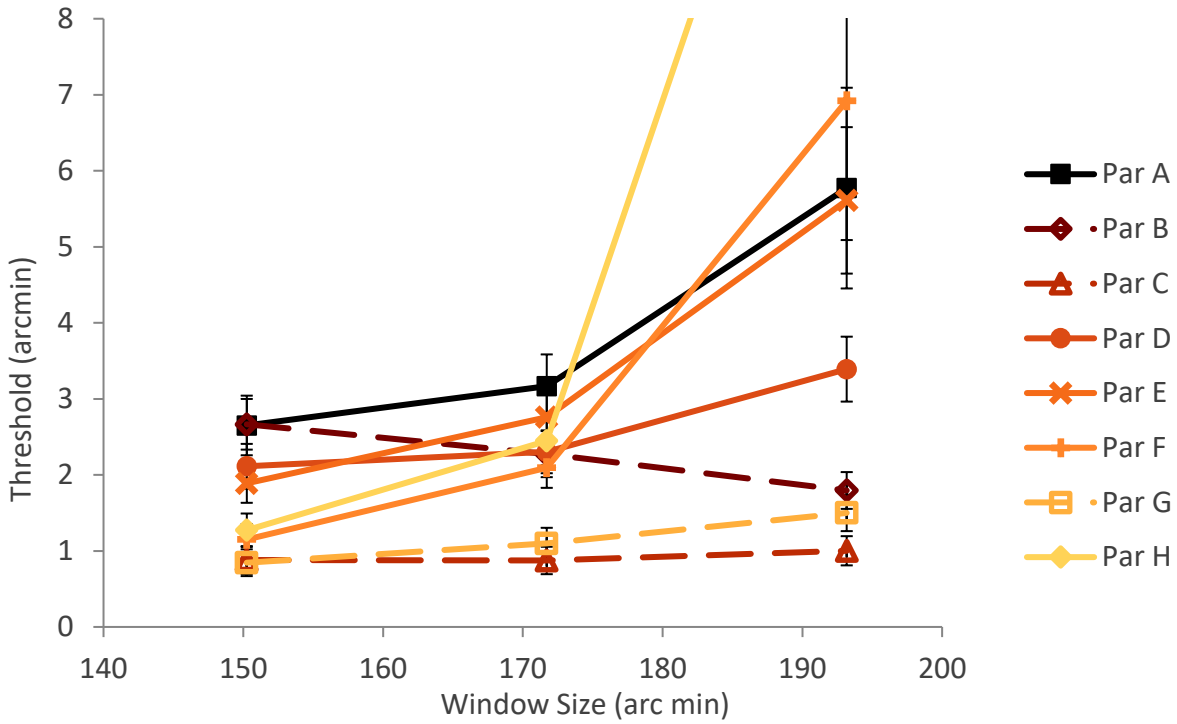


Figure 6.18: The thresholds as a function of window size for all participants in Experiment 7. Error bars are one SEM.

6.5.4 Discussion

In this experiment we have demonstrated the effect we hypothesised at the beginning of the Chapter – that an additional cue to object segregation can cause a bias in the perceived peak depth, with a larger segregation cue causing a decrease in the perceived depth of the object. This effect appears to be subject to a very large degree of individual variation, with approximately a third of participants not experiencing the effect, and two thirds having an effect.

We altered the model from Chapter 5 to predicted the perceived peak depth of the object if the object was segregated using the luminance window alone, then the disparities within the luminance window were averaged. We found that the participants that were affected by the luminance window had a greater decrease in perceived peak depth than that predicted by this simple averaging model. This degree of individual variation between participants and the presence of changing thresholds with window size indicates that other mechanisms are having an influence that were not observed in the previous experiments. We speculate that this is due to the individual differences in weighting the contribution of the luminance and disparity cues in the cue combination stage of visual processing. For example, Hills et al. (2004) paper on cue differences in observers that they speculated that some individual differences were due to individual participants weighting cues differently – see Section 2.2.6 for a brief discussion of cue combination.

Future investigation into this area would be valuable, as it would be very interesting to know exactly how the cue combination is happening and why there is so much variation in the effect of luminance on disparity estimation. However, the difficulty of getting any result at all using these stimuli indicates that they are far from the optimal way to investigate these effects.

6.6 General Discussion

Measuring the interaction of disparity segregation and luminance segregation turned out to be remarkably tricky when using RDS to create a disparity defined objects. The primary problems were that the majority of methods of adding a luminance cue were perceived as forming a separate object from the disparity defining dots. This confusion made the experiment impossible to perform with any degree of accuracy or interparticipant reliability.

However, we found that by increasing the contrast and overall luminance of the screen, then altering the dots to be either grey or white with a black background (see Figure 6.12) enabled participants to complete the experiments (Section 6.4). When the luminance edge was being perceived as part of the object in Experiment 6, the luminance edge did not have any effect on the perceived depth (Figure 6.15). In order to make the luminance cue have an effect we introduced eye-dependent disparity noise to force more reliance on the luminance cue (Harris & Parker, 1992; Lovell et al., 2012). Even then, in Experiment 7 we only found that two thirds of the participants showed a decrease in perceived peak depth with a larger luminance window. Additionally, for these participants, the larger luminance cue greatly decreased their sensitivity to peak depth.

6.6.1 Speculation on the interaction of disparity and luminance

As the luminance cue to segregation of the object influences the perceived peak depth of the object, this offers some interesting but speculative insights into where in the visual hierarchy segregation and averaging takes place. If the luminance cue could not influence the perceived peak depth of the object, then this would suggest that object segregation and averaging in the binocular domain was conducted separately to object segregation from luminance. However, as luminance segregation has influenced the perceived disparity, this indicates that these cues are combined further up the visual hierarchy, perhaps in a mid-level shape processing area (Loffler, 2008). As such, we agree with Richards 1977, that monocular cues are essential in considering depth perception in objects (Richards, 1977).

Interestingly, as the disparity is averaged over the object this suggests that further processing of disparity and perhaps other cues such as luminance take place after being combined in the visual system. This is perhaps not surprising, given the number of similar effects that have been found that are analogous between luminance and disparity (Anstis et al., 1977; Didyk, Ritschel, Eisemann, Myszkowski, & Seidel, 2012; Lunn & Morgan, 1995). For example, Lunn and Morgan (1995) found that the Hermann grid illusion depicted in Figure 6.19 creates a similar illusion when using disparity-defined depth instead of luminance (see Section 2.2.6 for a discussion of similarities between luminance and disparity). This has ramifications on the work of many modellers of the visual system as their models may need to take into account processing much later than the current models of the early visual system.

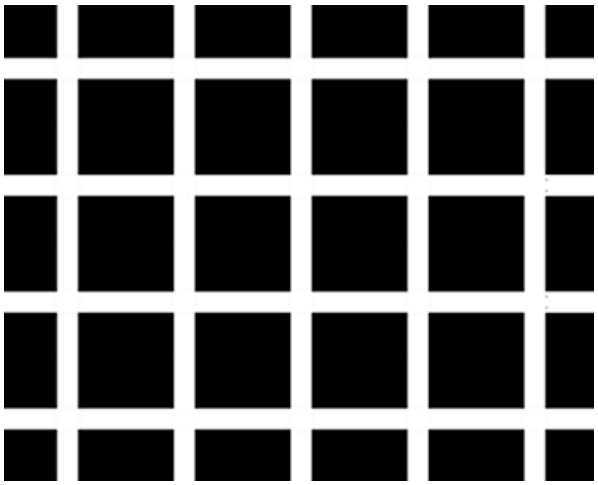


Figure 6.19: Hermann grid illusion – there appear to be dark patches in the intersections of the horizontal and vertical white lines. Reproduced with permission, (Paaliaq, 2007).

While the exploration of the interaction of luminance segregation and disparity averaging over the surface of the object has been very interesting, the sheer difficulty and interparticipant variation in the experiments presented here indicates that this is the wrong paradigm to study this effect in detail. We therefore propose that future exploration of this area would best be left to another project which can dedicate time and resources to designing an experiment with a stable interparticipant percept to study the effect. An interesting point to start may be by using stimuli similar to Lovell et al. (2012), where a random dot stereogram is combined with self-contained shape from shading cues. By using shape from shading rather than sharp luminance edges, it would be possible to create a strong percept of an object of similar shape to the disparity defined shape, but created from luminance. With both luminance and disparity providing a good cue to the object it may then be easier to specifically investigate the interaction of disparity averaging with object cues from shape from shading.

6.6.2 Effect of luminance and disparity interactions on camouflage

As an aside, this Chapter overlaps with the second strand of the thesis considering Julesz' assertion that objects "jump out from their background when stereoscopically fused" (Julesz, 1971) . If luminance cues to object edges alter the disparity based object segregation mechanism, then an animal using disruptive camouflage (see Section 2.1) could exploit use this monocular camouflage technique to also fool a binocular predator, and still appear as

multiple different objects. This would be compounded if depth from disparity was then averaged over the miss-identified objects, as they would appear to have different depths, making it hard to identify as a prey item. Interestingly, this observation acts against the hypothesis that binocular vision is used to break camouflage – an optimal system for detecting disruptively coloured targets would be for disparity processing to operate independently from luminance and colour, thus providing an additional method of detecting animal shape. It is probable that other uses of binocular vision (as discussed in Section 2.1.3) and the risk of misidentifying object based solely off one cue means that combing cues is more valuable than the single use of detecting camouflage prey.

The objective of camouflage is not to specifically decrease the perceived depth, but to make the object harder to spot. The first few experiments may give an indication of a potential camouflage technique that is slightly different to the proposed one in the introduction to this Chapter: by having luminance edges that are not coincident with disparity edges, it appears make it very hard to extract depth from the object. While this effect may be due to the highly abstract nature of our RDSs, it is possible that adding luminance edges that are irrelevant to the depth boundaries on the object, may disrupt the perception of the prey as object, making it harder to spot.

7 Does disparity defined depth assist in breaking camouflage?

- **Strand 2: Does depth perception break camouflage?**
- **Investigating: Does disparity defined depth trivialise camouflaged target detection.**
- **Task: Search for a camouflaged target.**
- **Manipulation: Target sometimes has a disparity cue**
- **Results: Disparity assists in target detection, but does not trivialise it.**
- **Conclusions: Depth helps break camouflage, but detection of disparity defined targets is not trivial:**
- **Further work: Do smooth objects hinder disparity defined object detection?**

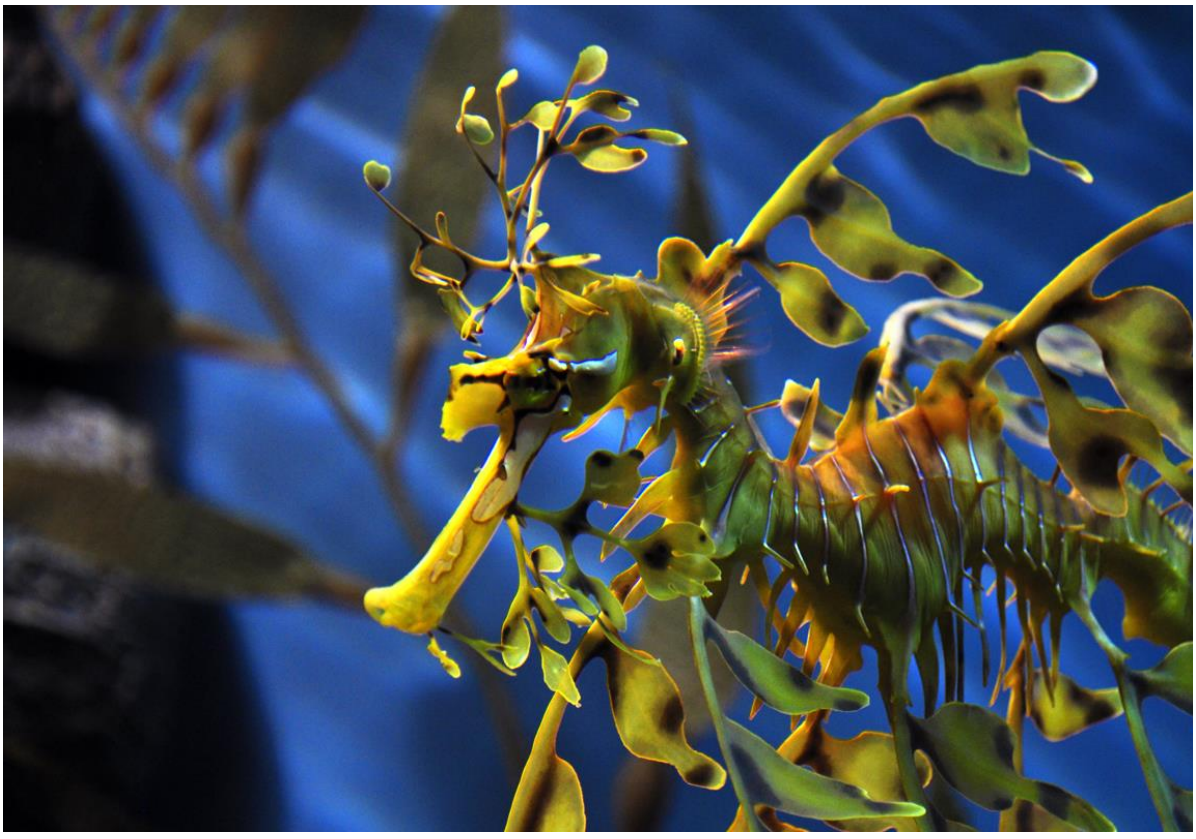


Figure 7.1: The head and body of the leafy sea dragon, *Phycodurus eques* showing excellent masquerade. Image reproduced with permission, (©Ta-graphy, 2010)

7.1 Introduction

In this thesis, we are exploring the interaction of disparity defined depth and camouflage via Julesz's original assertion in 1971 that "even under ideal monocular camouflage, the hidden objects jump out in depth when stereoscopically fused" (Julesz, 1971). We broke this statement into two strands to explore: first, investigating what a disparity defined object is, and second, the interaction of stereoscopic vision with camouflage. In the previous three Chapters, we have explored the perception of an object with a smooth depth edge (Chapter 4), gained an understanding of how the edge of the object is detected (Chapter 5) and explored the interaction of luminance and disparity when segregating the object (Chapter 6).

In this and the following Chapters, we explore the second strand of this assertion: that the objects jump out in depth when stereoscopically fused, thus rendering monocular camouflage useless. In order to do this, we turn away from psychophysics, which compares two stimuli, and towards visual search experiments, where the participant must find a target object in a visual display. By making the target very similar to the background of the display, we can make visual search tasks very similar to searching a section of the environment for a target camouflaged via background matching. As discussed in Section 2.1.2 and 2.3, we can use two main metrics in a visual search experiment to measure camouflage:

1. Reaction time – the length of time it takes to find the target object. The longer it takes to detect a target, the more likely the predator will give up searching and move on, leaving the target undetected. Therefore, we take a longer reaction time as an indicator that the target is better camouflaged e.g. (Cuthill et al., 2005; Lovell et al., 2015; Penacchio et al., 2016; Penacchio, Lovell, Sanghera, et al., 2015).
2. Accuracy – how many of the targets are correctly identified. A target that goes undetected more often is thought to be better camouflaged e.g. (Bond & Kamil, 1998; Cuthill et al., 2005; Stevens et al., 2007).

To measure these two metrics, we need to design an experiment where the participant searches a visual field for a camouflaged target object. We then need to measure the reaction time – the length of time it takes the participant to detect the target. Alongside this, we need to measure accuracy, to ensure that the participant is genuinely detecting the target, and as a measure of camouflage. As discussed in Section 3.2.2, we use only a single target object that is always present as multiple target objects or present/absent searches are prone to creating biased results (Cain et al., 2013; Duncan & Humphreys, 1989; Lovell et al., 2008).

In Experiment 8 presented in this Chapter, we present an experiment to explore if the presence of depth from binocular disparity does cause a camouflaged luminance defined object to be detected more quickly than if depth is not present. We continue using random dot stereograms, as these present the ideal monocular camouflage that Julesz is proposing,

and ask participants to search for a square target (defined by luminance) among rectangular distractors. On two thirds of trials, the target has an added square-shaped depth from disparity – if binocular depth does break camouflage then the participants should detect the object faster than when depth is not present. As a preview, we find that binocular disparity does assist in target detection.

In this Chapter we:

1. Present the new methods and stimuli used for all visual search experiments.
2. Present Experiment 8, in which we tested if a target object with a disparity cue was easier to detect than the target object without disparity.

In the following two Chapters, we continue to use the general methods described in the next Section (7.2), but test two new ideas:

> In Chapter 8 (Experiment 9) we investigate if the decrease in peak depth and decreased perceptual grouping of the smooth object (see Section 4.5 and 5.6) could impair the camouflage breaking properties of binocular disparity discovered in Experiment 8. In preview, we find that the smoother object is slower to detect.
>
> In Chapter 9 we investigate how the results of Experiment 9 hold up in naturalistic conditions. In Experiment 10, we use the methods described in Section 7.2, but add naturalistic disparity noise to the background, to investigate if the smooth object is still harder to detect. In Experiment 11, we create a real-world analogue of the visual search experiments, using a unique experimental setup and techniques.

7.2 General methods for visual search experiments

In the previous three Chapters, we have been using an identical setup for all experiments. However, this setup is not suited for the visual search experiments in this and the next two Chapters. In this Section, we discuss the new experimental equipment, setup and methodology used for the visual search experiments. Exact details of the stimuli used are discussed in the individual Experimental Sections.

Here, we present a brief overview of the visual search experiments. In these visual search experiments, we measure the time it takes for a participant to find a target placed at a random location in a RDS that covers the majority of each half of the screen. The longer it takes the participant to find the target, the better camouflaged the target. In order to test the effect of different manipulations of target shape and size on the target's camouflage, we use several different target conditions – either adjustments of the target's disparity or it's smoothness (profile in depth). Each experiment is blocked in to sections, in each experimental block, we choose to display a certain number of repeats of each condition, and interleave these between the trials, displaying only one target of one condition in each trial.

The visual search experiments we use follow a simple procedure: first the observer is presented with a black screen with a central fixation cross. After one second, this proceeds to the RDS. Within this RDS is the target object, defined by either luminance or disparity or a combination of these two cues. The participants' task is to press a button as soon as they spot the target, then indicate on a response screen (discussed in Section 7.2.2) where they saw the target.

7.2.1 Experimental equipment

In order to use as large a display as possible, the visual search experiments used a mac monitor which was placed at 1.1m away from the participant. The monitor has a resolution of 2560 by 1600, and a size of 32 by 19 degrees in visual angle. Pixels are at 0.747 arcmin across (making the individual 2x2 pixel square dots of the RDS have a side length of 1.49 arcmin). A random dot stereogram and Wheatstone stereoscope (Section 3.1) are still used. Maximum luminance (white) for this screen is 268.8 cd m^{-2}, and minimum luminance (black) is 0.4 cd m^{-2}.

We replaced the keyboard input method with the mouse, for reasons explained in the next Section.

7.2.2 Experimental methodology

The visual search experiments require a different procedure to the previous experiments as they are asking different questions. Participants search in a RDS for an object, with no limits to the length of time they can spend searching. A participant who is fast at the task will therefore progress through trials faster than a slow participant (unlike in the previous experiments, where each trial was set to take 2s). We allow participants to proceed through blocks of the experiment at their own rate, completing as many blocks as they can during the one-hour session. In order to ensure a reasonable degree of precision, we require that the participant completes a minimum number of blocks (three in Experiment 8).

We reduce the number of trials in each block (to fifty in Experiment 8), aiming at an average of 10 minutes per block for each participant. As different conditions are interleaved within each block, we wish to avoid participants stopping partway through a block at the end of the session, and then having completed a different number of trials for each experimental condition. These 10-minute blocks should allow us to fit as many blocks into the one hour as possible without breaking blocks partway though at the end of the one-hour session.

Participants are required to correctly complete at least the sixth panel in the TNO test. Additionally, participants complete a visual search demo to familiarise themselves with the task before undergoing the experiment. Participants were informed that their task was to find a square-shaped object in the field of random dots. The demo consisted of 12 trials (each condition was displayed to the participant 3 times), and were identical to the experimental blocks – the demo was for practice and to ensure the participants had

correctly understood the task. Participants unable to complete the demo with at least 80% success rate at identifying the target objects were rejected from further study.

7.2.3 Experimental procedure

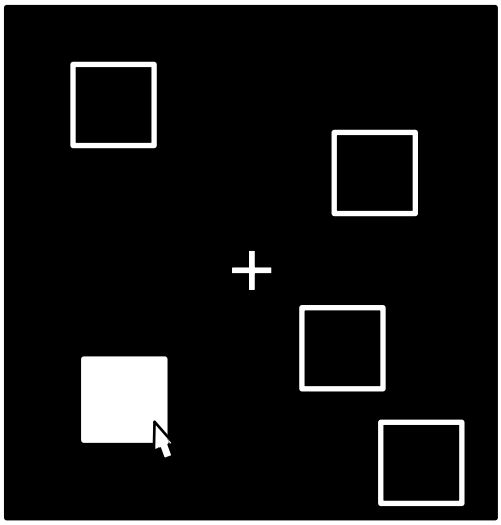


Figure 7.2: Mock up response screen. The mouse is initially at the fixation cross, with all possible response positions as hollow white squares. When the participant hovers the mouse over a square, it is highlighted white to indicate the participants' selection. The participant clicks the square they think was at the target location.

The procedure of the experiment is significantly changed from the previous three Chapters. At the beginning of the experiment, the participant is shown an alignment screen consisting of a box around the area that will contain the visual search stimuli. Participants are asked to ensure that they can see the entire box in each eye individually, and only one box with comfortable fusion when both eyes are open. If needed, they are shown how to adjust the angle of the stereoscope mirrors to achieve this.

The experiment then starts with a mouse click, initiating the display of the visual search stimulus. During the presentation, the mouse pointer is hidden, and the participant is given as long as they wish to spot the target – as soon as the participant identifies the target they click the mouse button. The time taken between the mouse clicks is measured by the program as their reaction time.

In order to ensure that the participant correctly identified the location of the target, the participant is then presented with a response screen containing a set of white square outlines as in Figure 7.2, one square corresponding to the location of the target object. The mouse pointer is revealed at the fixation cross, and the participant requested to click which of the squares corresponded to the location of the target. No two squares were allowed to be within 74.7 arcmin of each other, called the padding, to ensure that participants could not get confused about the location of the target with respect to the response patches. The participant moves the mouse to the square where they thought the target was. This square will turn white to indicate the participant's selection – the participant presses the button to confirm and move onto the next trial. The participant can mouse over as many selections as they wish before making a selection. Once a square is selected, the fixation cross was displayed for one second before the next stimulus onset. No feedback was given.

We switched from a keyboard input in the previous three Chapters to a mouse in these visual search Chapters because there is no clear way of selecting between the five response boxes using a keyboard. Using a mouse, selecting the correct box is intuitive for the participant.

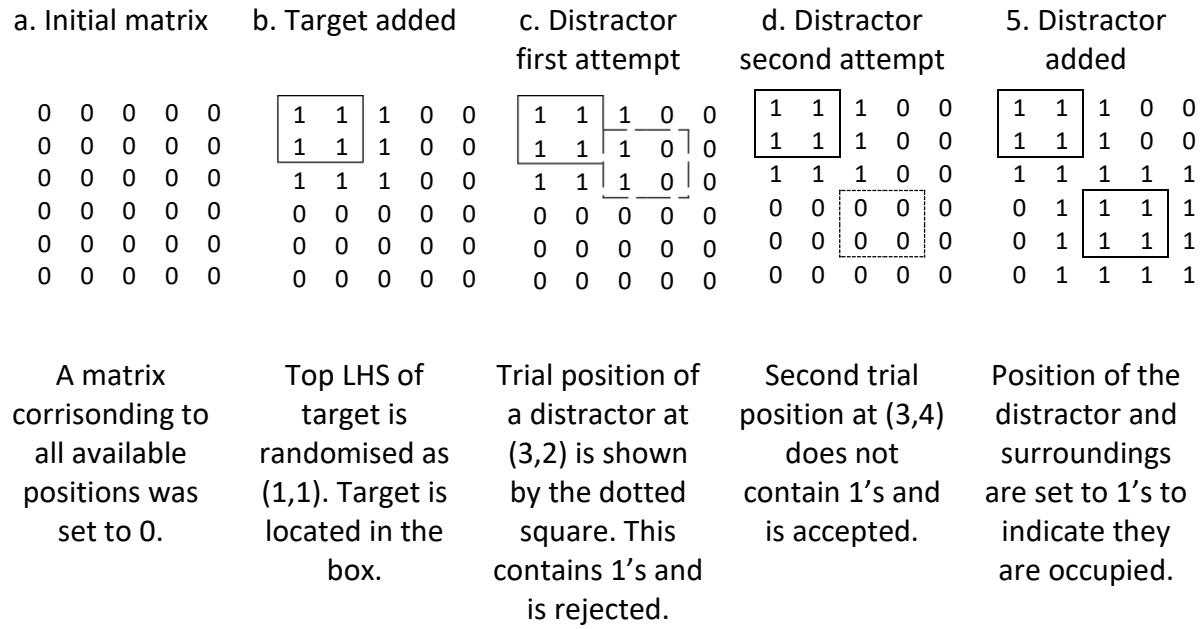


Figure 7.3: An example small matrix of how the screen was populated with objects. The objects in this demonstration are 2 by 2, with a padding of 1. Solid lines indicate the position of stored objects, dotted lines the attempted positioning of objects.

To ensure that search times were not dictated by the initial fixation position the position of the target was pseudo-random. The target was not allowed to appear within 74.7 arcmin of the fixation cross to ensure that participants had to search the stimulus to locate the target. Additionally, the target could not appear within 74.7 arcmin of the edge of the random dot stereogram.

Randomised position within these constraints was obtained by using a matrix, each element of the matrix corresponded to one pixel on the screen where the target object could be placed. Every element of the matrix was set to zero, to indicate that it was an allowed position, as in Figure 7.3a. The top left corner of the target was calculated by a randomised (x, y) coordinate, and all elements that would be covered by the target and the banned region around it (which stops objects being too close to each other) were assigned a '1' as in Figure 7.3b. A second random (x, y) coordinate was then generated as the top left of the first distractor. If any of the elements within the area of the second object are ones (as in Figure 7.3c), then this position is rejected and a new one generated until the object contains only zeros as in Figure 7.3d (this is only attempted 100 times before the entire matrix is reset). All elements covered by the new object and the area around it are assigned to one, as in Figure 7.3e. As can be seen in Figure 7.3e, there is only one randomised top left position left that could result in the placement of an object at (4,1) – once an object is placed here we cannot fit any more in. This typically takes place at around 55% coverage (Brosilow, Ziff, & Vigil, 1991). This not an issue with our stimuli however, as they only have approximately 1% coverage, so all objects can be reliably positioned randomly in the scene.

7.3 Experiment 8: Does disparity information decrease reaction time?

7.3.1 Introduction

First, we must test the hypothesis that to a stereoscopic observer, it is easy to spot a monocularly camouflaged target when disparity information is present, as proposed by Julesz in 1971 (Julesz, 1971). This is a highly persuasive argument and has been often accepted but not rigorously tested (Heesy, 2009; Isbell, 2006). In this Chapter, we wish to test that the presence of disparity defining a camouflaged shape makes the object easier to spot than when displayed without a disparity cue.

Here, we use a visual search task where the target is defined by a luminance cue, and with an additional disparity cue for two thirds of the trials. Participants were not informed of the presence of the additional disparity cue. In theory, if disparity defined depth assists with breaking monocular camouflage, then the targets with the disparity cue should be faster to detect than those without. A luminance defined object is trivial to detect in a uniform luminance RDS, so in order to make the target camouflaged, we add luminance defined distractors. The target is then camouflaged because it matches distinct features of the background (see Section 2.1.1).

7.3.2 Stimuli

The general procedure was detailed in Section 7.2.2 and we used the same experimental setup as in Section 7.2.1. In this Section, we detail the specifics of stimuli we display using these techniques.

The target was a 112 arcmin square, with six rectangular distractors of 146 by 78 arcmin with a randomly defined orientation of either 0 or 90 degrees (see Figure 7.4). Dot density was 217 dots per square degree. The objects were created by selecting dots within these areas, and increasing their luminance to white (268.8 cd m^{-2}), whilst background dots were dark grey (24.6 cd m^{-2}) on a black background (0.4cd m^{-2}). When disparity was present in the target it was either crossed or uncrossed (added to each dot, in front or behind the screen, see Section 2.2.2) of 11 arcmin.

The stimulus consists of one square target with six rectangular distractors, as shown in Figure 7.4. The participants are informed that the task is to find the square object within the rectangular distractors. The participants are not informed that on two thirds of trials the target has non-zero disparity: one third of trials 11 arcmin crossed, one third 11 arcmin uncrossed, and one third zero disparity. As half occlusions (regions of the background viewable by only one eye, see Sections 2.2.3 and 4.4) can contribute create differences between crossed and uncrossed stimuli (e.g. (Becker et al., 1999; Kim, 2013)) we use the smooth object at smoothness coefficient 3 (see Figure 4.3, page 63), which has no perceived change in peak depth relative to a sharp-edged object and no half occlusions to remove any potential bias in perception due to half occlusions. If a disparity cue helps in the detection of

an object, the participants should spot the targets that contain disparity faster than those without.

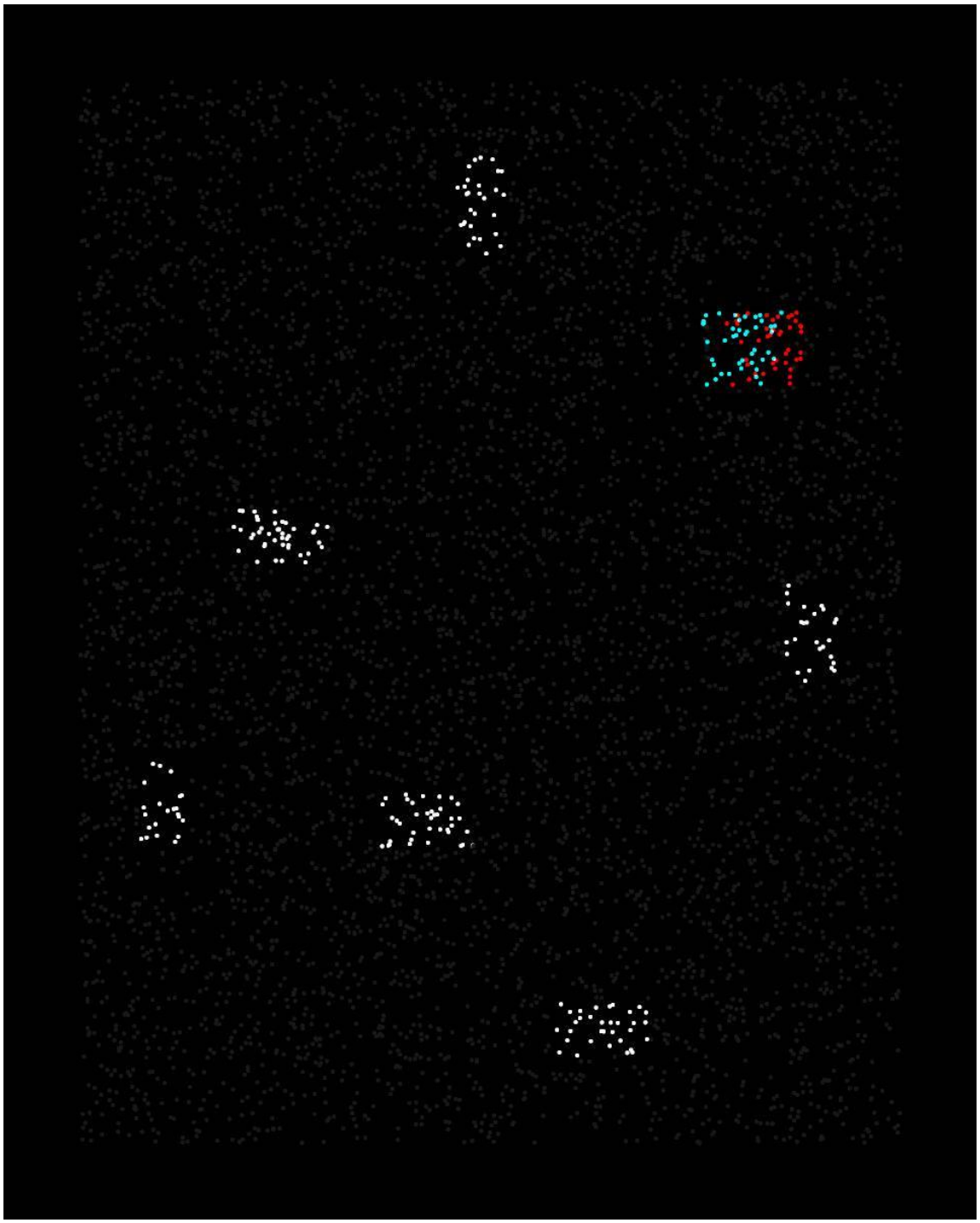


Figure 7.4: Anaglyph showing the visual search stimuli. Target square is located at the upper right of stimulus, with a disparity cue. Note that the dots in the background are dark-grey and may not show up clearly on non-calibrated devices.

7.3.3 Results

Out of twelve participants, one did not complete the demo and was excluded. All remaining participants completed the minimum of 3 blocks (150 trials for each condition), with a

maximum of 7 blocks completed (350 trials per condition). Mean reaction times, taken across correct trials only (percentage correct is discussed later, alongside Figure 7.8b) are shown in Figure 7.5 and demonstrate a clear trend, with the zero disparity target taking the longest to detect. There are two exceptions: Participants A and B (indicated by dotted lines in Figure 7.5).

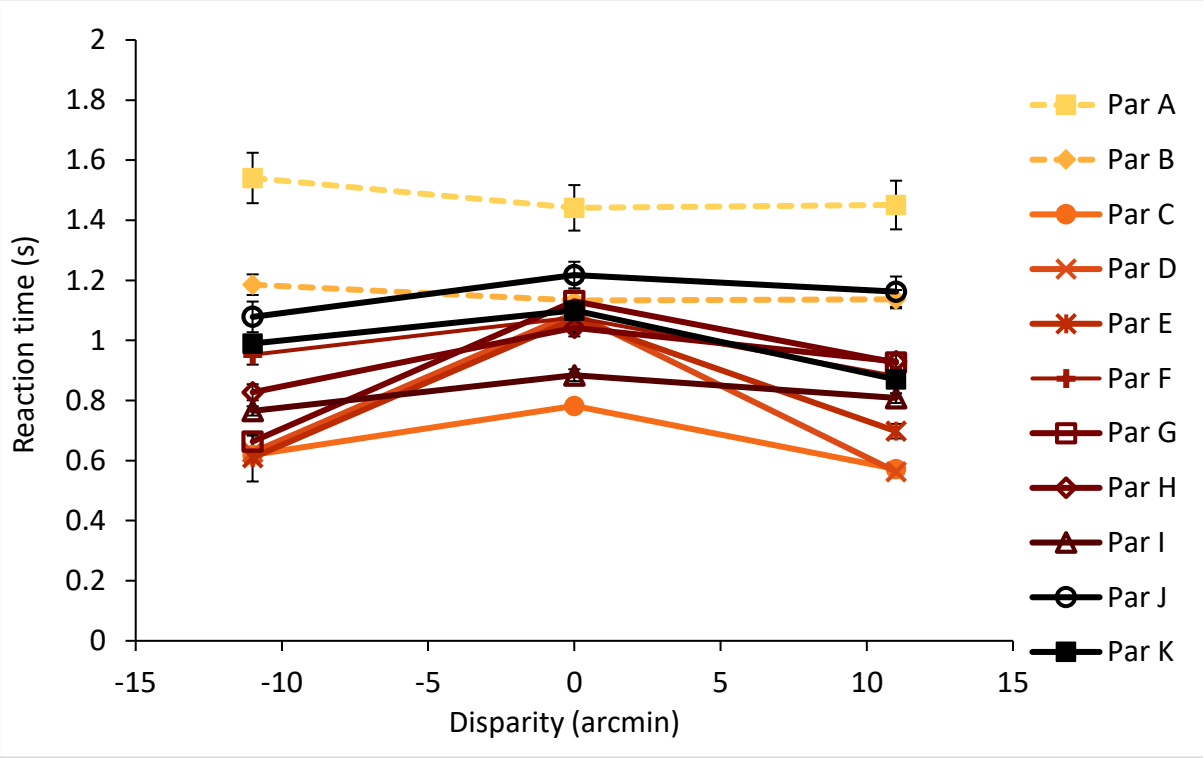


Figure 7.5: Mean reaction times for correct trials to detect the square target object. Objects with crossed disparity (which stand out from the background) are negative.

The mean reaction times are non-normal (Shapiro-Wilk, $p<0.05$), and therefore statistical analysis was done using the Wilcoxon signed rank test, as a non-parametric alternative to the t-test that does not assume normality. We compare three conditions, meaning there are three pairwise comparisons, so we used a Bonferroni correction to the significance value to avoid type 1 errors from multiple comparisons. Statistical significance is therefore taken at the $p=0.017$ level.

Statistical analysis using the Wilcoxin signed rank shows differences between disparity absent and disparity present trials to be highly significant on a group level for both positive and negative disparities (0 to 11 arcmin, $Z=-16.9$, $p<0.0005$; 0 to -11 arcmin, $Z = 18.4$, $p<0.0005$). Group mean data is plotted in Figure 7.6 Testing on an individual level replicated these results except for Participants A and B, who showed no significance for any pair of conditions ($p>0.30$) and Participant F who showed no significance between zero and 11 arcmin disparity ($p=0.11$).

The difference between positive and negative disparity present trials was only significant for some participants (C,D,E,G,H,K), and was not significant on a group level (11 to -11 arcmin, Z = -2.2, p=0.031 which is insignificant due to the Bonferroni correction). For the participants with a significant difference, three found the -11 arcmin faster to detect and three found 11 arcmin faster to detect. This discrepancy between disparity present trials is probably due to either chance or to individual differences.

Despite the highly significant results, it is hard to see the advantage afforded by disparity from the mean reaction times in Figure 7.5, although we can see the effect when looking at the mean reaction time across all participants in Figure 7.6. This difficulty in seeing the trend in individual data is due to the individual variation in reaction times between participants. To reduce participant variation in reaction time, we normalized the results for each participant by dividing the mean reaction for each condition by that participants' mean reaction time in the zero disparity condition. This gives us a measure we call the gain – a number greater than 1 indicates a decrease (speed up) in detection time compared to the zero disparity conditions, a number less than 1 indicates an increase (slow down) in detection time. In Figure 7.7 we plot the gain against the displayed target disparity. The gain shows the trend in the mean data more clearly, with the majority of participants displaying a noticeable gain when disparity information was present. Analysed across significant participants (discussed above) the mean gain was 22%, with a standard deviation of 14%. Across all participants, mean gain was 17%, with a standard deviation of 17%.

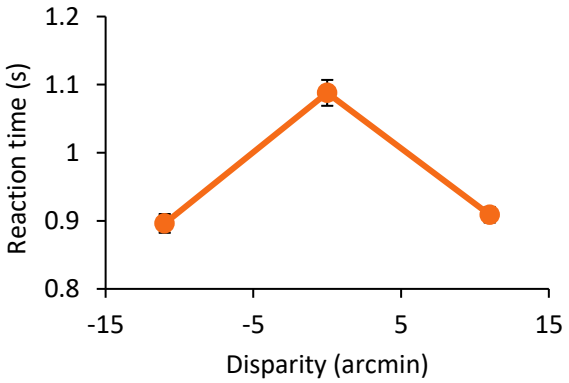


Figure 7.6: Mean reaction times for correct trials averaged (mean) across all participants. Error bars are one standard error.

Mean reaction times, while informative to test how well camouflaged the target is can be skewed by a few outlying results. In Figure 7.8(a) we can see that the median reaction times replicate the trend shown by the mean, indicating that the results were not caused by outliers.

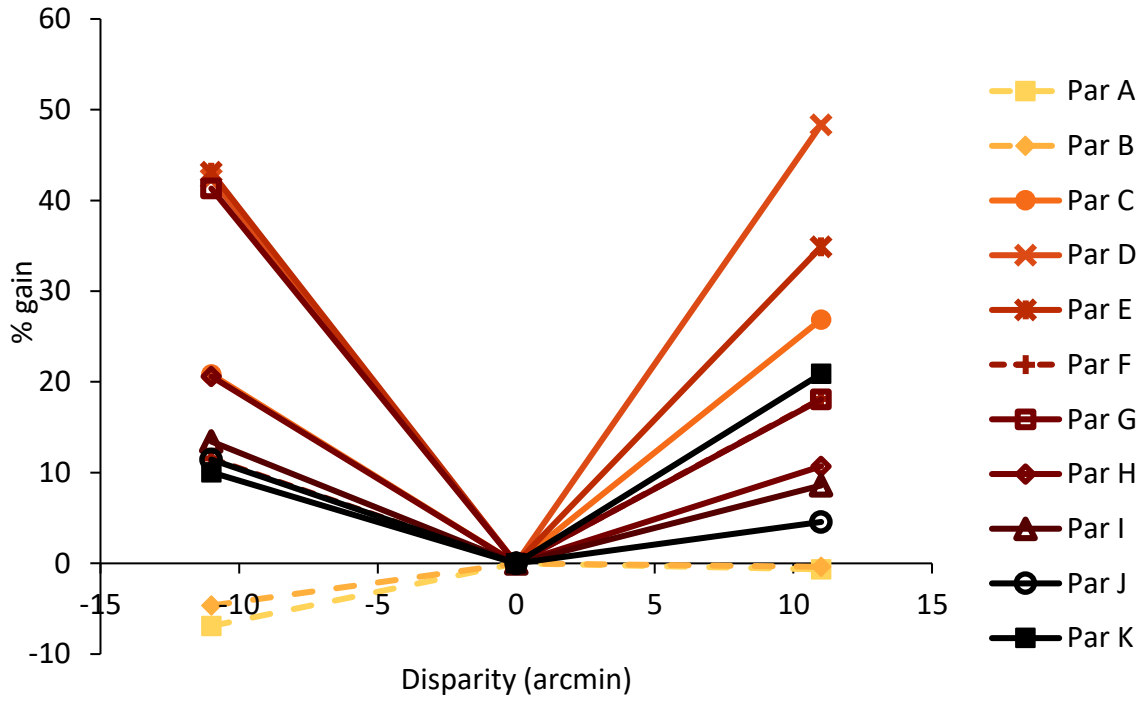


Figure 7.7: Percentage gain of reaction time with the presence of a disparity cue – a 10% gain means the participant was 10% faster than the disparity absent trial.

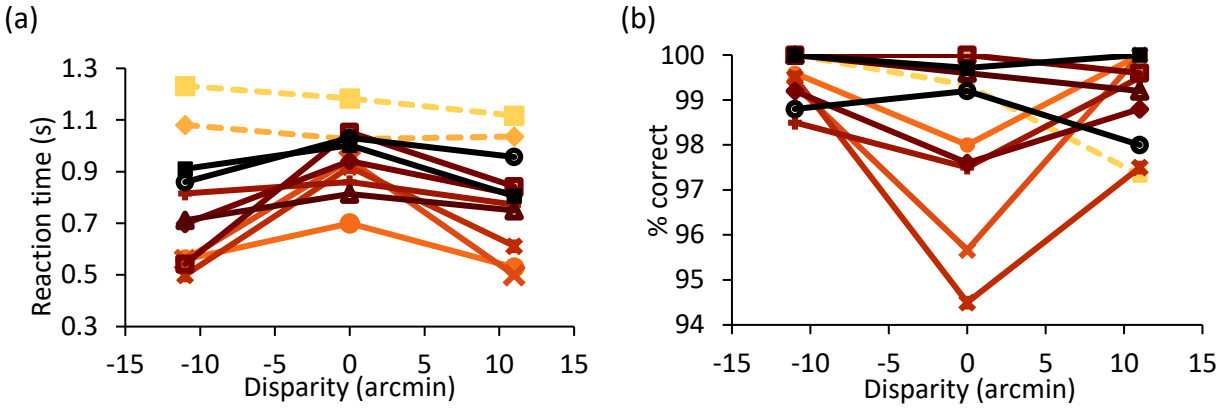


Figure 7.8: Participants' lines coloured identically to Figure 7.7(above). (a) Median reaction times for all participants. (b) Percentage correct. Participants with dotted lines had non-significant effects.

Finally, for this experiment, we investigate the percentage correct of participants, shown in Figure 7.8(b) (note error bars are omitted, due to being around 5-8 percentage points and thus covering the entire graph). All participants performed above 94% correct, with there being no clear link between disparity and percentage correct. Some participants seem to have poorer performance on the zero disparity trials but this is within the calculated

standard errors (and typically corresponds to getting 5 trials incorrect out of a minimum of 150).

7.4 Discussion

Here, we consider the implications of this experiment on the theory that disparity defined depth is used to break camouflage. We then discuss the lack of a difference between the crossed and uncrossed cues in reference to current literature. Finally, we consider exactly what it means to have no disparity, and wrap up with a discussion of the next steps to test the interaction of binocular depth perception and camouflage.

In order to test the theory that the presence of disparity defined depth would assist in the detection of camouflaged prey, we created an experiment where the target sometimes contained disparity information. As predicted, we found that the presence of disparity in the target object significantly decreases reaction times for eight out of ten participants, however the presence of disparity did not make them trivial to spot, with the average decrease in reaction time being 22%. This shows that the disparity in the target assists with detection, even when participants are not informed about the presence of disparity.

Given that participants were not informed that disparity was present in the stimulus it is interesting that the presence of the cue could helped nine out of eleven participants detect the target. As we did not inform participants about the presence of disparity this makes the effect of decreasing reaction times robust – the participants are unlikely to be exhibiting any top-down behaviour looking for the presence of disparity first, then searching for the square luminance cue.

There was little evidence of a difference between crossed and uncrossed disparities (depth in front and behind the background plane) unlike some of the literature that has investigated uncrossed vs crossed disparities specifically e.g. (O'Toole & Walker, 1997; Patterson et al., 1995; Patterson, Moe, & Hewitt, 1992). A study by Becker et al. (1999) studied the asymmetry between the perception of crossed and uncrossed stimuli by manipulating fixation relative to the background plane. They found that there was no difference between uncrossed and crossed conditions when both conditions were displayed in front or behind of the background plane an experiment that was replicated by Kim (2013). However, they observed a difference between the conditions when crossed stimuli were presented in front of the background plane, and uncrossed behind. Becker et al. (1999) suggested that the asymmetry was due to the difference in occlusions when the stimuli were presented in front or behind the background plane, rather than the crossed or uncrossed nature of the stimulus. This suggestion is consistent with our findings – although we display the objects in front or behind the background plane, we use a smooth object to eliminate any occlusions, thus removing the asymmetry observed in earlier studies.

On a pedantic note, all of these presentations have a disparity cue – having zero disparity is not the same as an absence of disparity. However, as both the distractors and the target

have zero disparity, then the disparity cue here contains no information and so cannot be used to detect the target. While we could consider comparing monocular and binocular presentation, this is not as simple as it sounds: monocular presentation can alter the perception of flat objects when compared to binocular presentation (Prinzmetal & Gettleman, 1993), and it is possible that two eyes are used to improve signal to noise in the display (Heesy, 2009; Jones & Lee, 1981). To add an additional level of complexity, some predators with overlapping regions in both eyes will may use these regions to enhance signal, but not to calculate depth e.g. (Martin, 2015; Ott et al., 1998). It would be interesting to see a set of experiments based around comparing monocular and binocular presentations. In the meantime, we think that making the disparity cue give no information is sufficient to give us an idea of the difference in detection rates that binocular vision provides.

In 1971 Julesz proposed the idea that depth from binocular disparity would cause even well camouflaged objects to be detected easily. Whilst widely accepted, the presence or magnitude of this effect had never been formally tested. Here, we found that a target that is defined by both disparity and luminance is, on average, detected 17% faster than a target defined by luminance alone. This was a significant decrease in reaction time for the majority of participants, indicating that disparity defined depth can assist in detecting camouflaged targets. In the next Chapter, we investigate if the advantage of disparity defined depth could be counteracted by manipulating the perception of depth in the target object.

8 Can certain shapes create stereoscopic camouflage?

- **Strand 2: Does depth perception break camouflage?**
- **Investigating: Do smoothed reduce the advantage afforded by binocular vision.**
- **Task: Search for a target object.**
- **Manipulation: Target has different edges smoothness, rendering it easier or harder to segregate.**
- **Results: Smoother objects are slower to detect.**
- **Conclusions: Smoother objects are harder to detect than sharper objects. This could be used as camouflage against a binocular observer.**
- **Further work: Do these effects carry over to real world tasks?**

Figure 8.1: The feathery edge around this common baron caterpillar (*Euthalia aconthea*) may make it hard to segregate the edge of the caterpillar from the mango leaf. Image reproduced with permission, (Auswandern;, 2007)

8.1 Introduction

In the second strand of this thesis, we are considering the interaction of camouflage and depth perception, particularly with respect to Julesz's claim that objects with disparity defined depth would be trivial to detect (Julesz, 1971). In this Chapter, we investigate if certain shapes, such as the smooth objects in the first strand of this thesis could counteract the advantage of binocular vision and become hard to detect, forming a kind of stereoscopic or binocular camouflage.

Although depth perception enables the extraction of depth across the scene, there are additional levels of processing to segregate an object from the background that cause misperceptions of depth– as discussed in the first strand of the thesis. We hypothesize that these additional levels of processing and the resultant biases in the perception of depth (modelled in Chapter 5) may slow down detection of disparity defined targets. This hypothesis is supported by the findings of Deas and Wilcox (2015). They found that participants detected lines of disparity defined dots faster when they were more strongly grouped via good disparity continuation than when disparity continuation was weakened by adding disparity noise to the dots (see Section 2.2.6). This indicates that the better the grouping of the elements making up an object, the faster it can be segregated from the background and identified as the target. In this Chapter we extend the work of Deas and Wilcox (2015). We investigate if 3D objects that are well grouped (in our case, sharp objects) are faster to detect than those whose shape causes problems with grouping and cause a misperception of depth (in our case, smooth objects, the relative grouping of sharp and smooth objects is discussed in Section 5.6).

If smoother objects are harder to detect than sharper ones, then this indicates limitations in disparity processing that animals may be able to exploit these limitations of disparity processing by altering their shape, therefore taking longer to detect and being better camouflaged (Merilaita, 2003). This has interesting implications for the statement that binocular vision can break camouflage – if it is possible to adopt certain strategies to make the object is harder to detect, then it is effectively displaying a form of camouflage. This camouflage technique is similar to crypsis, where the target is avoiding being detected as an object separate from the background. We dub this effect binocular or stereoscopic camouflage (Cammack & Harris, 2016).

In this Chapter, we present an experiment that uses the smooth and sharp objects from the first strand of the thesis to investigate if certain objects shapes take longer to detect using depth from binocular disparity. As we found in Chapter 5, the smooth object is perceived with less peak depth than the sharp object, which we interpreted as a problem with object segregation. We speculated that the elements that make up the smooth object are poorly grouped together, and are grouped with the background via good continuation (see Sections 2.2.6 and 4.2.1), making it hard to segregate the object from the background. Deas and Wilcox (2015) performed an experiment that found that objects that were strongly

grouped were detected faster than objects with weak grouping. We therefore expect that the smooth object will take longer to detect than a sharp object that is strongly grouped with itself via good closure.

We compare the reaction times of three different objects: the sharp object with half occlusions, a smooth object with no half occlusions and no decrease in perceived peak depth (SC3) and a smooth object with a no half occlusions and a decrease in perceived peak depth (SC14, see Section 4.3.3). This should enable us to separate any effects of reaction time due to the shape of the smooth edge and the processing and perception of the object:

- If the lack of half occlusions in the smooth objects has an effect on reaction times, then the both smooth objects will have altered reaction times compared to the sharp object.
- If the impaired perceptual grouping or the decreased perception of depth in the smooth object has an effect on reaction times, then we expect to see altered reaction times between the smooth object with no decrease in perceived peak depth and the smooth object with a decrease in perceived peak depth.

8.2 Experiment 9: Do smooth objects take longer to detect?

8.2.1 Methods

We investigated if it takes longer to detect an object with a poorly defined edge between the object and the background. To do this, we measured the detection time of a disparity-defined smooth object (see Figure 8.2b) embedded in RDSs and looked at the effect of smoothness coefficient (Section 4.2.1) on reaction time. The general setup and experimental procedure were the same as those detailed in the general description of visual search experiments in Section 7.2. In this Section, we detail the specifics of the stimulus used in this experiment.

Unlike the previous experiment (Section 7.3), we did not change the luminance of the dots, so providing no monocular cues to segregation of the object and thus forming the perfect monocular camouflage discussed by Julesz (Julesz, 1971). The experimental stimulus is an RDS that covers 14.1 by 18 degrees containing white dots (268.8 cd m^{-2}) at a density of 217 dots per square degree on a black background (0.4 cd m^{-2}). The background dots have zero disparity, with a single target object of peak depth of 5 arcmin.

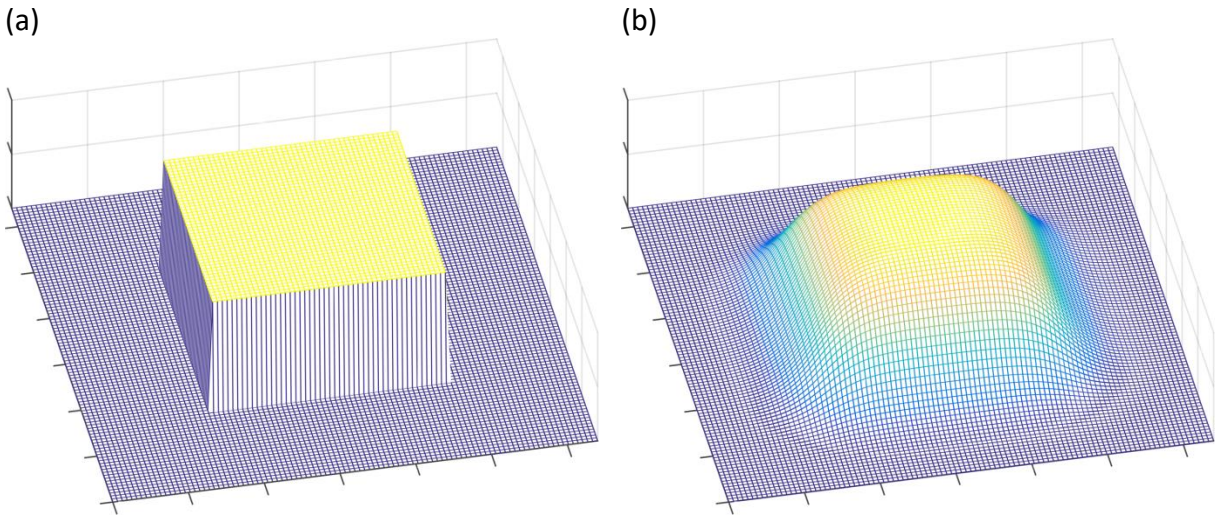


Figure 8.2: A 3D representation of the stimuli used in Experiments 1 and 9. (a) Sharp object with half occlusions, (b) smooth object with no half occlusions.

We compared the detection times of the objects used in the first strand of the thesis - the sharp edged object with a discontinuous edge and half occlusions; and the smooth object whose edge is continuously joined to the background and has no half occlusions (see Section 4.2.1 for a discussion of these objects, and Figure 8.2 for a 3D representation). For the smooth object, we use two smoothness coefficients - 3 (SC3) and 14 (SC14, see Section 4.2.1, page 63, Figure 4.3 for cross-sections). The object with SC3 is included as it has a smoothed edge but no perceptual decrease in perceived depth (see Section 4.3.3, Figure 4.6), while the SC14 object has a much smoother edge which causes a perceived decrease in peak depth (see Section 4.3.3, Figure 4.6). Both of these objects are not as strongly grouped than the sharp object, but only the SC14 object has significant decrease in perceived peak depth, which we speculated was due to poor segregation (Chapter 5) and weak grouping between object elements.

Due to the range of object shapes, there is no simple choice that will make a good distractor, so do not use distractors in this experiment (the advantages of no-distractor experiments are discussed in Sections 2.1.2, 2.3 and 3.2.2). The angular size of the object is made small enough (74.7 by 74.7 arcmin) that participants take a more than a second searching the screen to detect the single disparity defined object in the zero disparity background. To ensure that the participants have detected the object, the response screen is displayed with five possible locations of the target (as in Figure 7.2). The participant must click on the correct location of the target, if they fail to, the data for that trial is rejected (response screen is discussed in more detail in Section 7.2.2).

8.2.2 Results

A total of twelve participants were recruited, one was rejected for performing at chance level (20%) and a second for not completing sufficient trials (completing only 2 blocks, not the required 3 blocks, or 75 trials for each object). Mean reactions times excluding all

incorrect responses (percentage correct is shown in Figure 8.4b and discussed later) for the remaining ten participants are shown in Figure 8.3. The majority of participants show an increase in reaction times at the highest smoothness coefficient.

Inspection of the mean reaction times in Figure 8.3 seems to show two main groups of participants – the main cluster of seven who have a mean reaction time for smoothness coefficient 14 that is approximately double that of their reaction times for smoothness coefficient 3 and the sharp object (smoothness coefficient 0). Although it may appear that the smoothness coefficient 3 is slightly faster to detect than smoothness coefficient 0, this is not statistically significant (on a group level, Wilcoxin signed rank test, Z=-1.45, p=0.147) whereas the increase from smoothness coefficient 0 to smoothness coefficient 14 and smoothness coefficient 3 to smoothness coefficient 14 is highly significant (on a group level, Wilcoxin signed rank test, SC0 to 3: Z=-13.03, p<0.0005, for SC0 to 14: Z=-13.99, p<0.0005).

Three participants (B, H and I) do not show such a dramatic increase in reaction times, however when analysing these Participant I still shows a significant increase between smoothness coefficient 14 and the other two conditions (see Table 8.1). Note that this participant performs the task particularly fast, making the difference between smoothness coefficient 14 and the other conditions hard to see on the scale of the graph.

Only two participants (B and H) did not follow the group statistics, and are indicated by dashed lines in the graphs. We consider these participants' performance in the discussions.

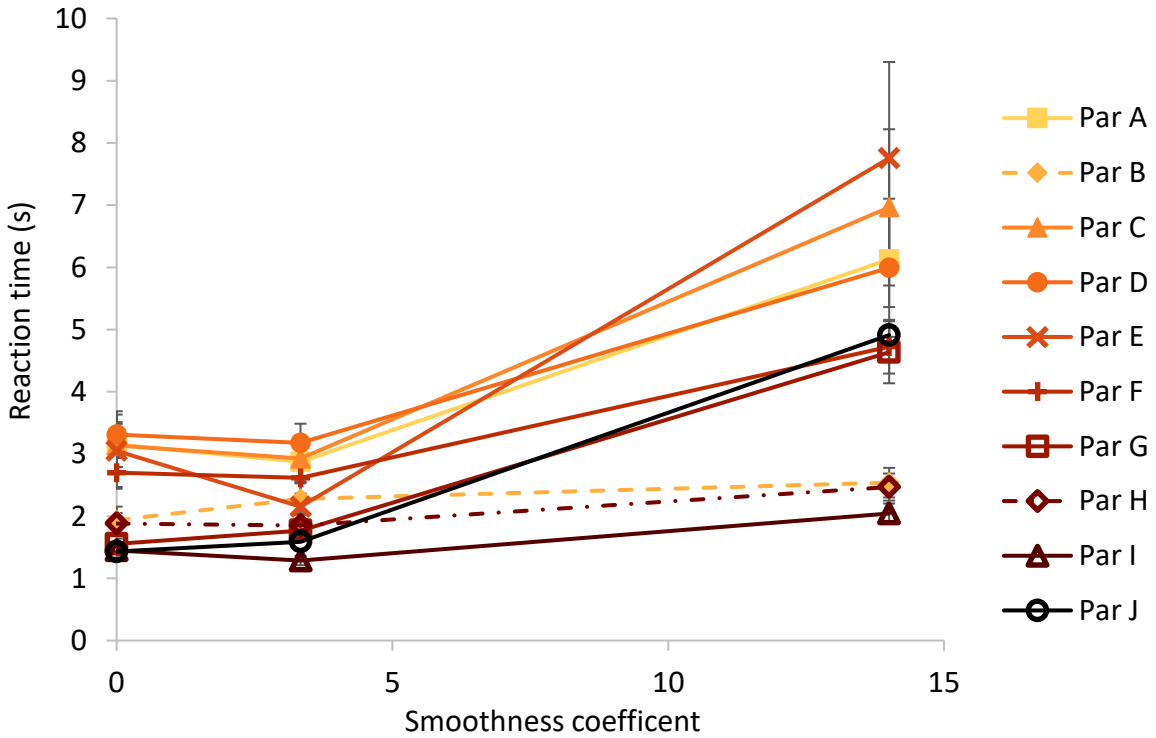


Figure 8.3: Mean reaction times for correct trials to detect the object in experiment 9. Dotted lines are participants' B and H, whose performance is considered individually in the discussion.

Participant	0 - 3	0 - 14	3 - 14
A	0.479	<0.005	<0.005
B	0.432	0.075	0.175
C	0.144	0.001	<0.005
D	0.841	0.015	0.006
E	0.917	<0.005	<0.005
F	0.923	<0.005	<0.005
G	0.132	<0.005	<0.005
H	0.465	0.029	0.005
I	0.035	0.007	<0.005
J	0.923	<0.005	<0.005

Table 8.1: Per participant Wilcoxin signed rank test. Green cells indicate results below 0.017, the Bonferroni corrected significant level for three comparisons. Note that here we are effectively making thirty comparisons (as there are ten participants), so the individual data is included for interest and not to draw conclusions from.

We also considered the median reaction times to ensure that the results were not being driven solely by outliers, shown in Figure 8.4(a). Median reaction times are much faster (due to the non-normality of the data), however show a similar trend, with the reaction time at the largest smoothness coefficient being typically greater that at the sharpest.

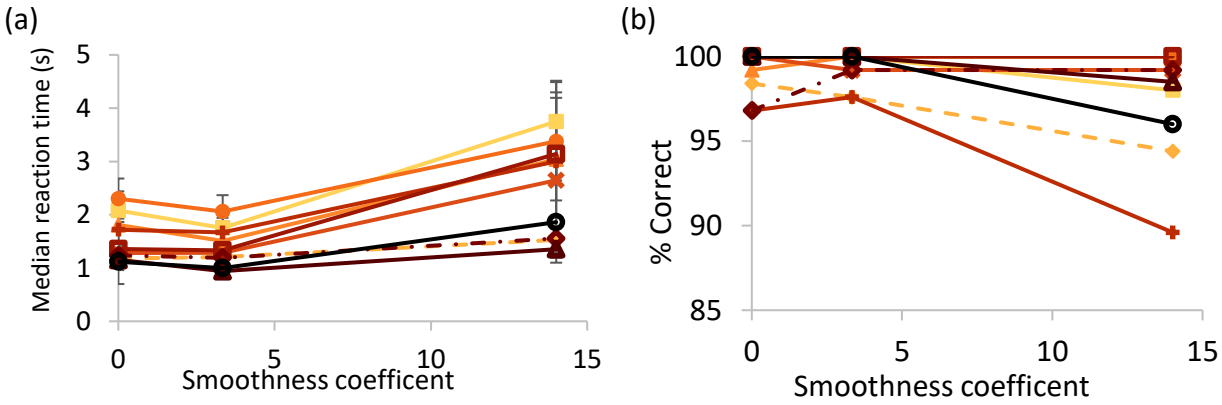


Figure 8.4: See Figure 8.3 for legend (a) Median reaction times for correct trials for all participants. (b) Percentage correct.

Percentage correct is analysed and shown in Figure 8.4(b). The lowest percentage correct among the remaining four participants is 88% at smoothness coefficient 14, with the majority of participants above 95% for all conditions. This indicates that the participants were able to do the task on all conditions with few mistakes. Note that all incorrect trials were excluded from the reaction time measures.

Additionally, we recorded the position of the target object on the screen. This enables an analysis of the change of reaction time with the radial distance from fixation cross (called

eccentricity) for each smoothness coefficient. To do this, we take all the results between 0 and 100arcmin eccentricity for one smoothness coefficient, then calculate the mean reaction time. We then take the mean for all results for one smoothness coefficient between 100 and 200arcmin, then between 200 and 300arcmin and so on, until we have a value for mean reaction time for all ranges of eccentricities present in the display. This is repeated for all smoothness coefficients, enabling us to plot the mean reaction time against smoothness coefficient and eccentricity in Figure 8.5.

Eccentricity data shows a similar trend to the reaction time data, with an increase reaction time with smoothness coefficient at all eccentricities. As expected, the further from fixation the longer it takes to detect the target, with a fairly consistent increase in reaction time with eccentricity.

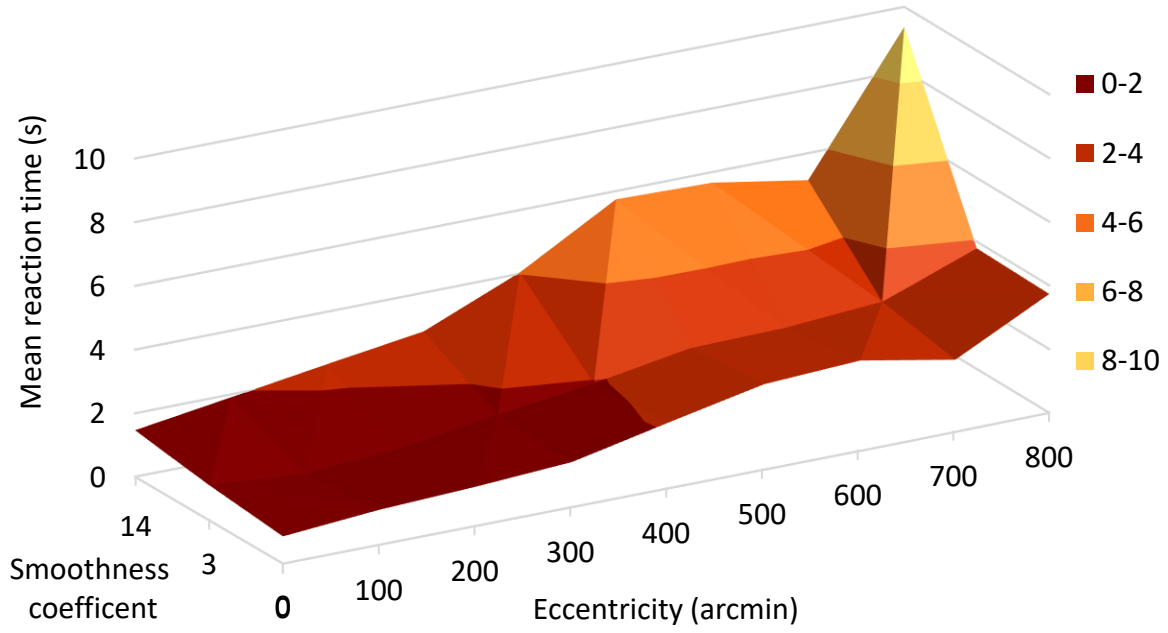


Figure 8.5: Effect of eccentricity on the reaction time for different smoothness coefficients in Experiment 9.

8.3 Discussion

In this discussion, we start with a general overview of this experiment and its findings. In the following sub-sections, we:

1. Discuss why we can consider the smooth object as being better camouflaged than the sharp object.
2. Consider if the results in this experiment are applicable to the real world.

3. Ask the question: why do smoother objects take longer to detect? We present some speculative experiments that may help understand in more detail the non-camouflage implications of this experiment.
4. Briefly conclude the experiment.

In Chapter 7 we started investigating the second strand of this thesis: exploring the interaction of binocular vision with camouflaged objects. We investigated the extent to which disparity defined depth assisted with detection of a camouflaged luminance defined target. We found that disparity defined depth decreased reaction times by an average of 17%. In this Chapter we investigated if certain shapes, such as the smooth object, could counteract the advantage of binocular vision and become harder to detect, thus forming a kind of camouflage designed to counter binocular vision.

In this experiment, we found evidence that the smoother the edge of the object, the longer it takes to detect. An object with smoothness coefficient 14 took approximately 2.3 times longer to detect than objects of smoothness coefficient 0 and smoothness coefficient 3, despite a decrease in peak perceived depth of approximately 15% (Chapter 5, Figure 4.6). This overall trend was replicated in the median data, indicating that it is not due to a few outliers. We analysed the effect of reaction time on eccentricity and found that the more eccentric the presentation, the longer the object took to find (Figure 8.5). This is as expected, as the participants commenced search from a fixation cross in the centre of the display. The smoother the object, the longer it took to find at greater eccentricities, but this is probably due to the mean reaction time for a smoother object being longer.

On an individual level, two participants did not follow the overall trend. Participant H shows a significant difference between SC0 and SC14 but not SC3 and SC14. As an aside, this participant self-reported systematically searching from top left to bottom. We discuss Participant B's performance later in this Section.

8.3.1 Are smoother objects better camouflaged?

We conclude that smooth objects are better camouflaged, as the smoothest object takes longer to detect than the sharpest object. We can conclude this despite all participants being equally accurate at detecting the object at all smoothness coefficients. This is because in this experiment, participants were incentivised to keep on searching the display for the target as they knew it was always present. Without this knowledge, participants would have eventually given up and decided to move on to the next trial – the longer the target takes to detect, then the more likely participants would have been to give up before detecting the target. This effect is discussed in depth in Section 2.1.2.

An alternative experiment to explore smoothness as a form of binocular camouflage would be to use genetic algorithms which mimic natural selection, such as those used by Burg and Alais (2015) and others (Geisler & Diehl, 2002, 2003). These algorithms display the observer with a set of targets with random properties (in our speculated experiment, we would use

random smoothness coefficients). Observers given a search task to detect the targets, and all those targets that are detected are considered to be 'dead'. They combine the properties of pairs of the surviving stimuli with a random mutation (a random number added to these properties) to create a set of 'offspring' of equal number to the original set. These offspring are then displayed to the observer, and the dead ones removed. This is repeated many times, meaning that the final set of offspring will have a set of properties that has been optimised by selection to be hardest to detect. It would be interesting to apply this technique to our stimuli, and allow properties such as the smoothness and total volume to change with different generations of stimuli. (Volume should change rather than just height or width as animals are typically the size they are for reasons other than camouflage (Clarke & Sardesai, 1959; Sand, Cederlund, & Danell, 1995; Warren & Lawton, 1987)). It would be interesting to see if the 'optimal' camouflage technique was to be objectively flat and smooth, or if there was a trade-off in shape that we have not detected by only being able to investigate three different smoothness coefficients.

With both of the above experiments, our task is performed in an artificial environment, leading to questions about the applicability of these visual search tasks to the real world.

8.3.2 Are these findings applicable to the real world?

The background of the stimuli used in this experiment is very artificial with a flat zero disparity background and shape and form defined only by small dots, therefore we may observe effects that are not applicable to the real world. For example, the flat background may have enabled observers experienced with RDSs to pick out the target very easily (for some observers, reaction time was under a second for all trials) because the target was the only item in the search task with non-zero disparity. This effect is potentially behind Participant B's non-significant results – they were an experienced stereoscopic observer, so they may have found it easier to detect anything with non-zero disparity, making the shape of the object less relevant. Experienced stereoscopic observers may not display this advantage in a naturalistic setting, as the background would be rough with varying disparity – meaning that the target would not be the sole feature with non-zero disparity.

The ease with which stereoscopic observers can detect non-zero disparities in RDS displays exemplifies a problem in linking our findings to camouflage: an RDS is a very unnatural environment, and findings in simple environments do not always extend to more complex scenes (Goutcher & Mamassian, 2005; Harris, 2014). To properly test how well camouflaged an object is we must relate our findings more closely to the real world.

In order to link these abstract experiments back to the real world, in the next Chapter (9) we take two steps towards reality. In Experiment 10, we introduce a naturalistic disparity noise in the background to emulate an object sitting on a real, bumpy, surface. In Experiment 11, we print a 3D model of the objects and place them on a real surface to see if the objects are camouflaged in a real task. In the next sub-section, we consider the non-camouflage implications of the experiment presented in this Chapter.

8.3.3 Why do smoother objects take longer to detect?

It is hard to distinguish exactly which aspect of the object with smoothness coefficient 14 causes an increase in reaction time. Here, we present two potential interpretations:

1. The increase in detection time is due to the decrease in perceived peak depth. Objects with smoothness coefficient 14 have a lower perceived peak depth than objects with smoothness 0 or 3, which are perceived with similar peak depths (Section 4.3.3). In the experiment presented in this Chapter, we found that objects with lower smoothness coefficients were detected faster by participants than objects with higher smoothness coefficients.

   While there have been no other studies on the effect of perceived depth on reaction times, similar effects have been found when looking at objective differences in disparity defined depth. Wardle et al. (2010) found that disparity was most effective at helping detect a visual overlay when displayed with a disparity of over 6arcmin. Finlayson et al. (2013) found search for a luminance defined T shape amongst distractors was faster when the T shape had a disparity of 1.5arcmin or greater. This indicates that decreasing objective depth makes an object slower to detect, so the same may be true for perceived depth. It may be desirable for a camouflaged animal to decrease perceived rather than objective depth because there are other requirements that dictate an animal's size (Clarke & Sardesai, 1959; Sand et al., 1995; Warren & Lawton, 1987).

2. The increased difficulty of segregating smoother objects (Chapter 5) directly leads to an increased detection time. This is supported by the work of Elder and Zucker, and Deas and Wilcox who found that well grouped objects (which are easier to segregate) were faster to detect than poorly grouped objects (Deas & Wilcox, 2015; Elder & Zucker, 1993). Neider et al. found that participants only improved at detecting a camouflaged object when trained on detecting camouflaged rather than non-camouflaged, targets. This led them to conclude that improvements were due to processes such as object segregation (Neider, Ang, Voss, Carbonari, & Kramer, 2013). These studies indicate that if the object is harder to segregate from the background, then it will take longer to detect the object.

With the evidence at hand, we conclude that we cannot distinguish between these two factors. We now discuss a speculative starting point for an experiment to distinguish between these two factors.

There are several potential experiments that may help distinguish between these two effects, but here we consider the idea behind an eye tracking experiment. Eye tracking is a common way to understand the visual behaviour of a participant in search experiments e.g. (Cain et al., 2013; Foulsham et al., 2014; Neider & Zelinsky, 2006; Nodine & Kundel, 1987).

This potential experiment uses the same setup and stimuli as the current one (with the participant searching for sharp or smooth objects), except with the inclusion of a calibrated eye-tracker to monitor the observer's fixation position relative to the screen. The observer's fixation position would be tracked, and the time between first fixation on the target area and their mouse click to indicate detection of the target would be measured. We do not know how long this time should be, but the two interpretations provide different predictions on the length of time between first fixation and detection for the smooth and sharp objects:

For the first interpretation, the sharp object takes time between commencing the trial and being fixated for the first time, but is detected as soon as it is fixated. Smooth objects with a lower peak depth take longer before first fixation, but are also detected when fixated. Therefore, in the first interpretation we would expect the time between first fixation and detection to be similar for both the smooth and sharp objects. However, there will be a longer period between commencing the trial and first fixation for smoother objects than for sharp objects.

In the second interpretation, the smooth object is harder than the sharp object to segregate out from the background. This means that once fixated upon, it may take time for the smooth object to be segregated and detected as the target. The sharp object, being easier to segregate, would take less time to be segregated and detected. Therefore, in this interpretation we would expect the time between first fixation and detection to be longer for the smooth object than for the sharp object. However, there will be a similar time between commencing the trial and first fixation for both objects.

8.3.4 Conclusions

We conclude that the smooth object is better camouflaged because it takes longer to detect than the sharp object (detection time as a measure of camouflage was discussed in Section 2.1 and 2.3). However, the experiment presented here uses a very abstract flat background, which may cause visual behaviour that is not applicable to the real world. In the next Chapter, we investigate if smoothness as a form of camouflage is applicable in the real world by comparing the detection time of smooth objects to sharp objects in realistic environments.

9 Does stereoscopic camouflage work in natural conditions?

- **Strand 2: Does depth perception break camouflage?**
- **Investigating: Do our findings work in real life?**
- **Task: Search for a camouflaged target on a screen and in a sandpit.**
- **Manipulation: Target has different edge profiles, as in previous Chapters.**
- **Results: Smooth edged objects are still harder to detect.**
- **Conclusions: By making an object poorly defined it is possible to camouflage it in a real world search task.**

Figure 9.1: A crab spider (*Misumena vatia*) with wasp. An excellent example where the real world must be considered: to our eyes the spider is camouflaged, but to an insect the spider is highly visible in the UV spectrum (Heiling et al., 2003). Image reproduced with permission, (Leillinger, 1998)

9.1 Introduction

In the second strand of this thesis, we are exploring the claim that the addition of disparity defined depth to a camouflaged object makes it easy to detect. In the previous two Chapters, we first (in Chapter 7) established that the addition of disparity information to a luminance defined target object assisted with the speed of detection of the target. However, the additional of disparity information did not make the detection of the target trivially fast. In Chapter 8 we explored why the disparity information did not make detection of the target trivial using the smooth objects from the first strand of the thesis. We found that an object with a more poorly defined edge (the smoothest object) took significantly longer to detect than objects with sharply defined edges. This indicates that the need to process disparity information to identify the object as separate from the background may be making it harder to detect – effectively rendering it camouflaged to a stereoscopic observer, compared to an object with a sharply defined depth edge.

In this Chapter we continue investigating the interaction of binocular vision and camouflage, however, our investigations so far use extremely abstract objects and backgrounds. We wish to address this issue and create experiments how well camouflaged the smooth object appears under levels of increased realism.

1. In the Experiment 10, we address a major problem with our artificial background – it is flat. The majority of surfaces in the real world are not flat, so it would be interesting to see if binocular vision still confers an advantage when searching for an object in a depth-textured background, or if it is merely of use when detecting a disparity defined patch on a flat background.
2. Experiment 11 did not use computer stimuli, but rather the objects and backgrounds existed physically in the world. The experiment was performed with real 3D objects placed in a large sandpit and tested if smooth objects are still better camouflaged than sharp objects in a real world experiment. We believe that this experiment will provide much needed reference back to the real world, as this experiment will consider other cues that are not present in our RDS (Richards, 1977).

9.2 Experiment 10: Adding a natural background

9.2.1 Introduction

In Experiment 9 (Section 8.1), we present evidence that a smooth object is slower to detect. This indicates that smoothly segregated shapes could be used as a form of binocular camouflage, requiring only minor adjustments of shape to result in a greater survival rate. However, as discussed at the beginning of this Chapter all the previous experiments are still very far from real life, with our object being displayed on a perfectly flat front-to parallel plane. Here, we look at how object shape interacts with binocular vision in a more realistic environment by introducing a background that has a naturalistic depth distribution.

We used a visual search experiment to explore the effect of the background on the perception of the three objects used in the previous two experiments. We anticipated two effects on reaction time of participants searching for the object:

1. Variation in background depth will hinder search times for all objects, as the objects will all be camouflaged within the background variation, making them harder to spot. This effect may compound with the objects being displayed on sloped background surfaces, hindering sensitivity to depth (Glennerster & McKee, 1999) and potentially slowing down search times for all objects.
2. The variation in background depth is more similar to the smooth object than the sharp object – therefore we would expect additional slowing for the smoother object over the previous experiments with flat background. This effect will cause a greater increase in reaction times for objects with higher smoothness coefficient than those with lower smoothness coefficient.

9.2.2 Methods

The experimental procedure and setup are the same as the visual search experiments described in Section 7.2. The target stimuli are identical to the previous experiment, described in detail in Sections 8.2.1 and 4.2.1. The target placement algorithms are also identical – the target is a square disparity defined object of side length 72arcmin, displayed with a smoothness coefficient of either 0, 3 or 14 (see Figure 4.3, page 63 for cross-sections) and a peak disparity of 6 arcmin. The background and target are formed of a RDS with white dots and a black background (dot size and luminance are identical to the previous experiment, detailed in Section 8.2.1). The major change in this experiment is the introduction of naturalistic disparity noise to all dots in the RDS – we detail how this is delivered below. We also discuss a minor alteration to the procedure of the experiment.

9.2.2.1 Naturalistic background

In order to generate a naturalistic noise background, we used a toolbox (McClure, 2014) performing an algorithm called fractal terrain generation. Fractal terrain generation generates a set of randomly located dots, each with a disparity that is dependent on the disparities of the nearby dots.

In short, the algorithm works by selecting three randomly selected points, then assigning each of them a random number – here used as the disparity of each point. Then the algorithm goes on a second pass, selecting three times the number of random points than in the previous iteration (so for the second iteration 9, for the third 27 and so on). The current depth of these new points is calculated by interpolation between the three closest previously generated points. The depth of these new points is then further randomised by a set fraction of the random depth added in the previous iteration (called the roughness, for example, 1/3rd of original noise on the second iteration, 1/9th on the second and so on). This process is repeated with an increasing number of points, each of which is shifted in depth by a decreasing amount of randomly generated depth – hence the fractal nature of the

generation. The process is repeated until the process forms enough dots to create the needed detail. We found that 8 iterations provided the density of dots required for our RDS.

The spatial frequency profile of fractal noise is approximately 1/f (Olsen, 2004), which is similar to the noise profile present in contrast and luminance in natural scenes e.g. (Simoncelli & Olshausen, 2001; Torralba & Oliva, 2003). In the absence of large data sets of natural disparity distributions (some small ones of limited scenes exist e.g. (Liu, Bovik, & Cormack, 2008; Scharstein & Szeliski, 2003), but use too limited a range of environments), we speculate that 1/f noise may be similar to natural disparity noise distribution. This speculation is backed up as fractal terrain generation is frequently used to generate realistic textures such as 3D landscapes (McClure, 2014) as shown in Figure 9.2.

Fractal terrain generation is a convenient way of creating random disparity noise as it generates a set of randomly positioned dots in three dimensions. We can use the (x, y) location of each dot as the position in the RDS, and the z dimension as the disparity of the dot. Fractal terrain generation results in z dimension is generated as a number between 0 and 1, which we rescale to provide dots with disparities between -8 and 8 arcmin (discussed later in this Section). Once calculated, disparities are added to the generated dots of the RDS as in previous experiments (see Section 3.1).

Figure 9.2: Example of the use of fractal terrain generators to create a realistic landscape. Image reproduced with permission, (Huber, 2004)

Fractal terrain generation creates a variable number of dots within the selected area, therefore we set a limit of between 217 and 434 dots per square degree. The code randomly removes any dots above this density, or regenerates with a new RDS if there are insufficient dots present. To generate this number of dots, we needed to use 8 iterations of the fractal terrain generator, with a roughness and iterative roughness of 0.2.

As with the previous Experiment (9), we compare participants' reaction times for objects with smoothness coefficient 0, 3 and 14, objects were 74.7 by 74.7 arcmin with a peak depth of 6 arcmin. In each trial, only one object is displayed. The placement of this object objects is randomly decided as in the previous two visual search experiments (see Section 7.2).

To combine the disparity defined shape of the object and the disparity defined fractal noise background, we take the (x, y) location of the dots where the object is to be placed. We then calculate the disparity contribution from the object for the location of these dots (using the same equations for the smooth and sharp objects as in the rest of this thesis, as described in Section 4.2.1). The disparity contributions from the shape of the object and the fractal noise background are combined additively, so it appears as if the object is sitting on the background. This has the side effect that if the object is placed on a depth-textured region, then the top of the object will have the same non-smooth fractal depth as the background. A 3D representation of the finished stimulus is shown in Figure 9.7.

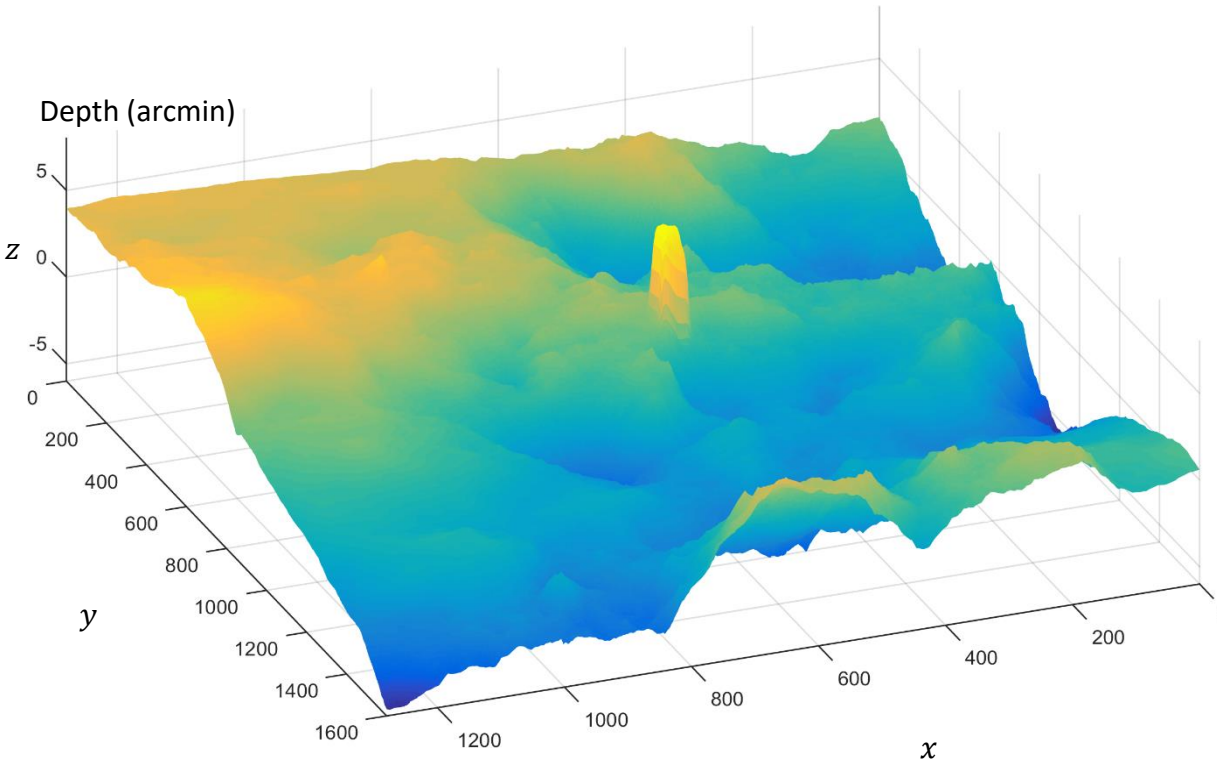


Figure 9.3: 3D representation of the target (located 600,600) embedded within the generated fractal disparity noise (z-axis). Depth is greatly exaggerated for clarity. Yellow indicates greater positive disparity, blue greater negative disparity.

This experiment is aiming to create a realistic background, to test if our objects are camouflaged in a more realistic environment than a flat background. We speculate that with a flat background, observers experienced with RDSs may be able to pick out a few dots that have non-zero disparity and identify these as the target without having to perceive the overall object. In this experiment, we wish to remove this advantage

If the target object is displayed with greater disparity than the variation in disparity in the background, then the point of greatest disparity would always be where the target is located. This presents a problem similar to that discussed in Section 8.3.2 of the previous experiment – the observer could just pick out the point with maximum disparity in the stimulus, and always be correct without identifying the object. This effect would be a feature of the experimental setup, not of camouflage – a camouflaged animal would not necessarily have the greatest disparity in the area that was being searched. Therefore, to be as realistic as possible we must ensure that the highest disparity in the scene is not always the target. The terrain was therefore generated with a range of depth between crossed and uncrossed disparities of 6arcmin, with approximately 90% of points appearing in the central 8 arcmin of disparity (between 4 arcmin crossed and 4 arcmin uncrossed).

Additionally, the noise was checked to ensure that it introduced as few false targets as possible. While some shapes in the scene could appear to resemble the smooth edged object, the magnitude of the disparity across an area the size of the targets was always much less (under a third) than the peak disparity of the target object. This should ensure no features in the random noise resemble the targets, as they will either have less depth, or be three times as large as the targets.

9.2.2.2 A change in procedure

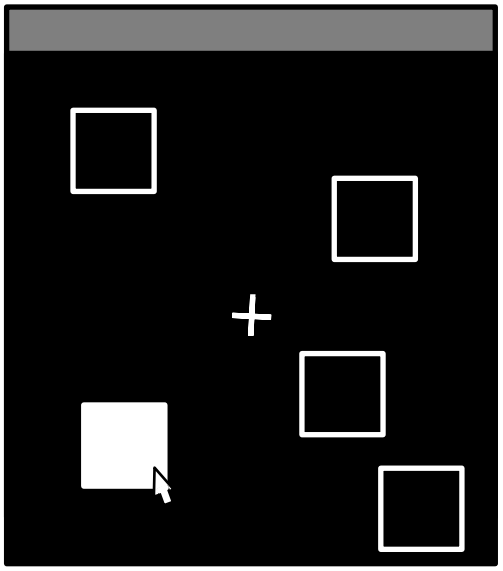


Figure 9.4: Mock up response screen. The option mouse over turns white to indicate where the viewer is responding. There is a grey bar at the top of the screen where participants can click to indicate that none of the available options are where they thought the target was located.

We want to be able to distinguish between cases when the participant gives up (or mistakes a feature for a potential target) and knows when they are wrong from the cases when they get confused and believes the target to be located somewhere else. To do this, we added a grey bar at the top of the response screen shown in Figure 9.4. Participants were instructed to click this bar if, and only if there was no option to select the target to be located where they thought they had seen it. This enables us to distinguish between when the participant knows they were incorrect (clicks the grey bar) and doesn't (they indicate that the is target located somewhere it is not present).

Participants are most likely to click the grey bar when they become bored of the task and give up, so they click the mouse button to move on to the next trial. They will then know they are wrong, and hit the grey bar.

9.2.3 Results

Nine participants out of ten completed the experiment - one was rejected as they could not complete the demonstration correctly. Out of these nine participants, four were recruited from participants that had completed the previous Experiment (9) within the last month – participants D, I, J and H. For clarity, we have continued lettering participants from Experiment 9 (Section 8.2).

The overall mean reaction time results for correct trials (percentage correct is displayed in Figure 9.8) as a function of smoothness coefficient can be seen in Figure 9.5. It can clearly be seen that an increase in smoothness coefficient increases reaction time, as in the previous experiment (Section 8.2.2, Figure 8.3). This change in reaction time is highly significant on a group level ($p=0.0005$, Wilcoxon Signed Rank Test) for all participants for smoothness coefficient 0 to 14 and smoothness coefficient 3 to 14, and on an individual level (Table 9.1). Additionally, there is a significant decrease in reaction time between smoothness coefficient 0 and smoothness coefficient 3 for Participants N and G (see Table 9.2), but no significant change on a group level ($p>0.017$). The decrease in reaction time is larger than in the previous experiments (see Section 9.2.4.2 for a comparison), and, overall, participants were slower than in Experiment 9. Several of the participants who participated in this and the

previous experiment self-reported that the task was harder. As before, the median follows the same trend, as shown in Figure 9.6.

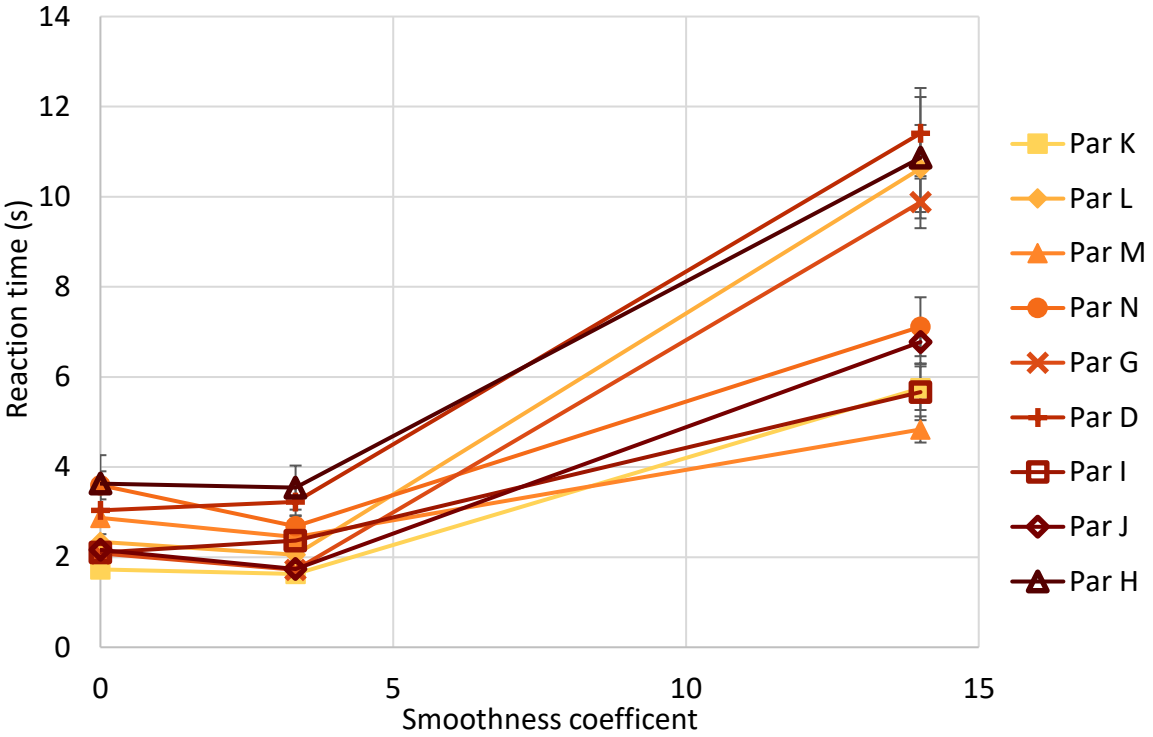


Figure 9.5: Mean reaction times for correct trials for participants in Experiment 10.

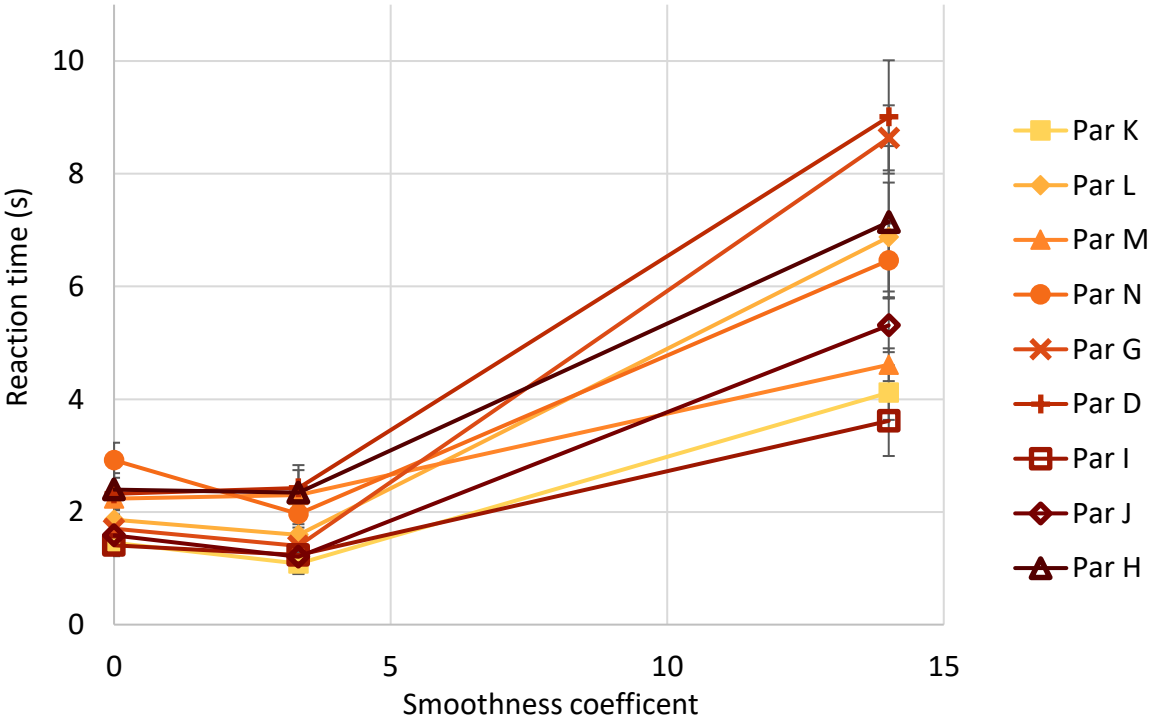


Figure 9.6: Median reaction times for correct trials for participants in Experiment 10.

Participant	0 - 3	0 - 14	3 - 14
K	0.071	<0.005	<0.005
L	0.83	<0.005	<0.005
M	0.117	<0.005	<0.005
N	0.003	<0.005	<0.005
G	0.011	<0.005	<0.005
D	0.405	<0.005	<0.005
I	0.842	<0.005	<0.005
J	0.04	<0.005	<0.005
H	0.786	<0.005	<0.005

Table 9.1: Per participant Wilcoxin signed rank test. Green cells indicate results below 0.017. Note that we are making 27 comparisons, so the individual data is included for interest but should be interpreted with caution.

Looking at the mean reaction times (Figure 9.5) there appears to be a split in participant performance between faster and slower participants at smoothness coefficient 14. However, this split is not reflected between participants at smoothness coefficient 0 or 3. It is possible that this split indicates a difference in search strategy when the search task is made difficult, although this split is not reflected in the median data in Figure 9.6. To investigate this further, we normalise the data for each participant by dividing the mean reaction time for each smoothness coefficient by the reaction time at smoothness coefficient 0. This gives us a measure of how many times slower the participant was to detect the smoother objects, plotted in Figure 9.7. From this data we can see that the apparent grouping of participants is no longer visible, suggesting that the grouping is due to chance.

The majority of participants were above the 80% correct mark (see Figure 9.8), although there is a noticeable decrease in the percentage correct for some participants at the highest smoothness coefficient. Additionally, the miss rates are much higher than the previous experiment, with the worst participant at smoothness coefficient 14 in Experiment 9 being 90% correct (Section 8.2.2, Figure 8.4b), and the best at this Experiment being 96% correct at the same smoothness coefficient. Participants M and N have much lower percentages correct than their peers at the highest smoothness coefficient, with participant M at 49% correct and N at 65% correct (chance level is 20% as in the response screen there is one correct option and four incorrect, see Figure 9.4).

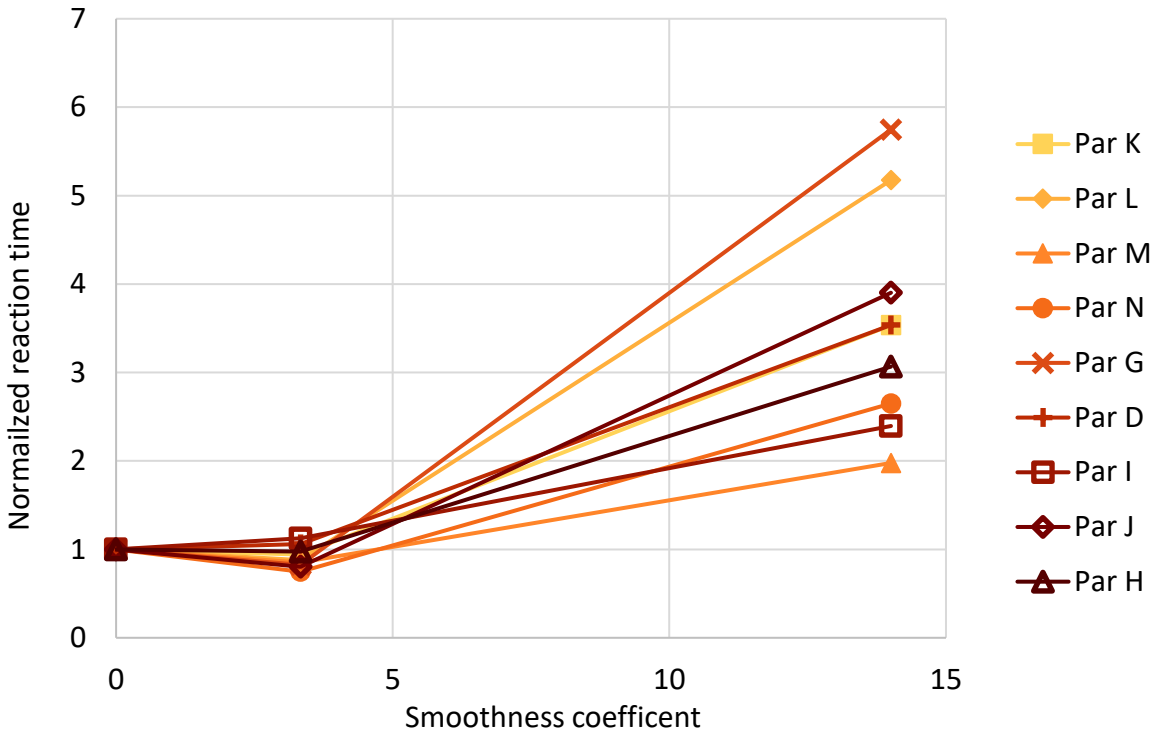


Figure 9.7: Normalized reaction time for correct trials by dividing the reaction time at each SC by the reaction time at SC0, giving a measure of how many times slower the participant was at each SC relative to SC0.

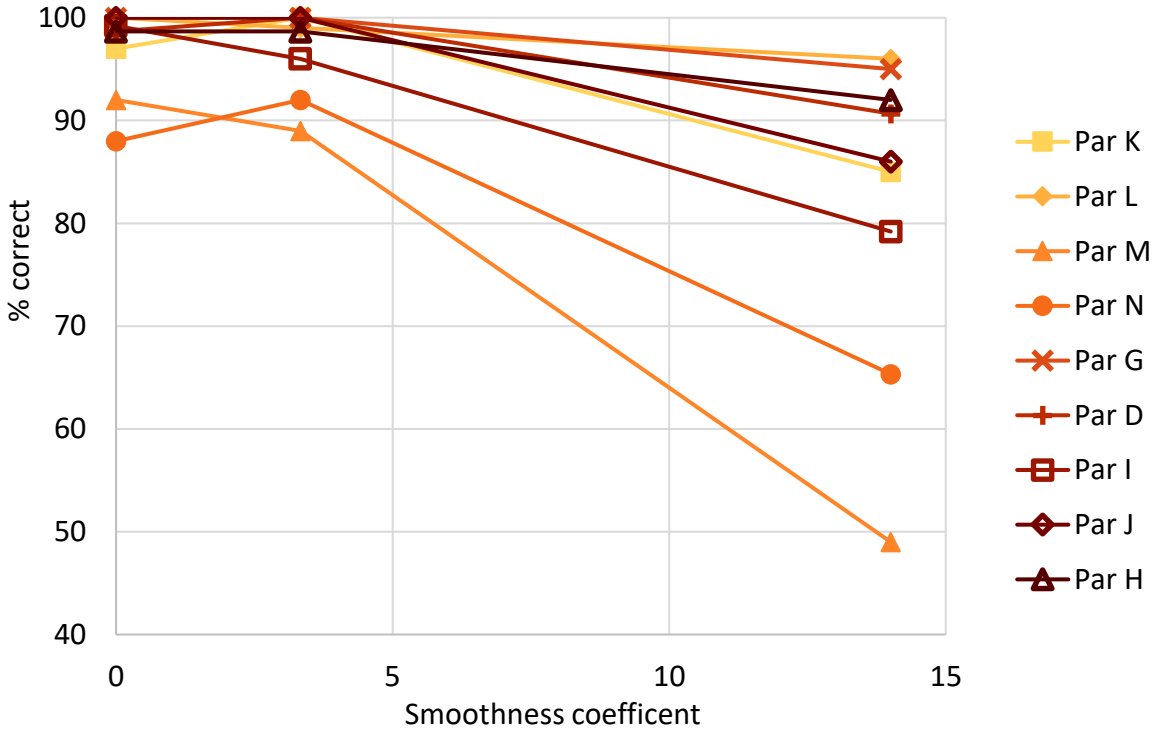


Figure 9.8: Percentage correct for participants in at Experiment 10

We have further detail available from when analysing percentage correct: the percentage known wrong (i.e. the percentage of trials on which participants knew they had not identified the target and chose the grey bar, see Section 9.2.2.2) shown in Figure 9.9a; and percentage unknown wrong (i.e. percentage of trials participants did not know they had not identified the target and chose the wrong location) shown in Figure 9.9b. Note we are considering the percentage incorrect in this breakdown. For the majority of participants, when they were incorrect they knew they were incorrect and have indicated so. However, there is a clear pair of outliers, one in each graph, who we discuss in detail below.

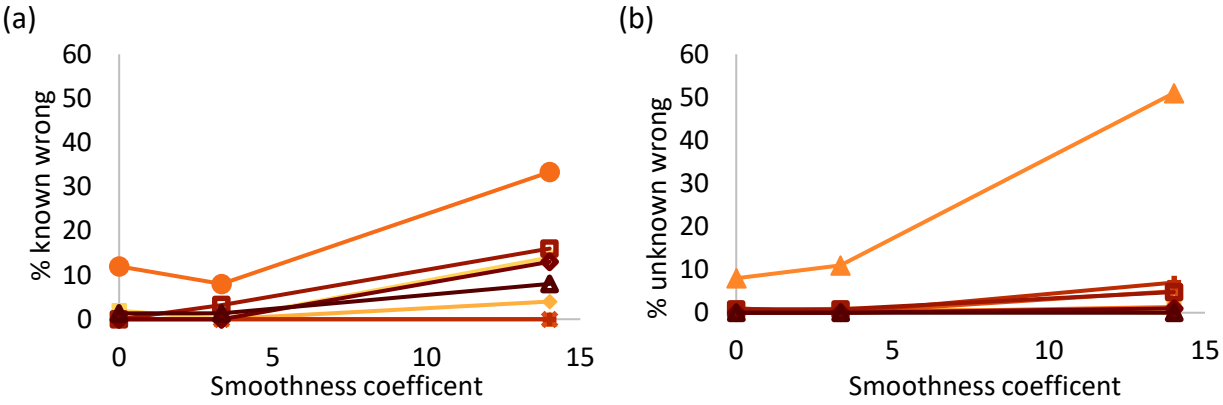


Figure 9.9: See Figure 9.8 for legend (a) Percentage known wrong (b) Percentage unknown wrong.

Participant M (the outlier in Figure 9.9b, orange circles) appears to have had a different approach than most participants, as when analysing their data individually they appear to have a quitting threshold (the length of time participants will search for before giving up and guessing the target's location) of around 10s, and only searched for more than 10s on 5 trials out of 300. It appears this may be a unique approach on this participants' part, where they searched the across the entire stimulus evenly, then guessed if they did not detect the stimulus. This participant never used the grey bar to indicate that they knew they had got the trial wrong.

On the flip side, participant N (the outlier in Figure 9.9a, orange triangles) did not show the same behaviour on search times. Instead, they frequently indicated they had incorrectly identified the target responses for the highest smoothness coefficient, opting to indicate they knew they were wrong on 25% of trials. This potentially indicates a different trade off, where they did not spend so much time verifying if they had correctly identified the target but rather accepted a lower percentage correct in the well camouflaged cases (there was no incentive for the participant to get a high percentage correct).

The remaining participants typically realised if they had incorrectly identified the object, with known incorrect being chosen considerably more frequently than the participants clicking an incorrect response box. Mean reaction times on known incorrect trials was much longer than on correct trials, with an overall mean for all participants of 14.6s and standard

deviation of 17.7s and median of 8.9s, with individual means varying from 10s to 22s. The data here is sparse, with between 0 and 40 data points per participant and a standard deviation between 5s (for the participant with a 10s mean) and 33s (for the participant with a 22s mean). This suggest this is indicative of the quitting threshold (discussed in Section 2.1.2 and 8.3), where the participants will only search the RDS for the target until giving up and moving onto the next trial.

9.2.4 Discussion

This experiment was designed to start to bridge the gap between the very abstract visual search task in the previous experiment and a real world detection of camouflage. We introduced a naturalistic disparity noise background to emulate placing the object on a real surface such as a beach or tree bark. The background succeeded in being naturalistic, with several participants commenting how natural and aesthetically pleasing the it appeared. Overall, we observed the same effect as in the previous experiment, with the mean reaction times increasing significantly for the highest smoothness coefficient stimulus. This strongly indicates that the technique of making the object smoother can cause an effect of stereoscopic camouflage in natural scenes, although apart from half occlusions making no difference we cannot conclude exactly which aspect of the smooth object makes it harder to detect (see Section 8.3.3). The differences between this experiment and the previous one are interesting and address our hypothesis at the beginning of the Chapter, so we discuss them in depth in Section 9.2.4.2. First, we briefly consider some of the standalone results we have found from this experiment.

9.2.4.1 Specific discussion of Experiment 10

We find on a group level that there is no significant difference between the smoothness coefficient 3 target with no half occlusions and smoothness coefficient 0 target with half occlusions, despite the creation of a background that varies smoothly in depth. On an individual level, two participants (N and G) had a significant increase in the reaction time of the smoothness coefficient 3 target compared to the smoothness coefficient 0 target. We cannot read too much into this as we are comparing nine participants and there is no significant effect on a group level. However, the absence of a difference on a group level is very interesting, as the stimulus is smooth-edged, indicating that the sharp sides and half occlusions in the smoothness coefficient 0 stimuli are not conferring a disadvantage despite the creation of a background that is smoothly varying in depth. This strongly indicates that half occlusions are not a particularly salient stimulus when the observer is searching in the scene for the target object. It would be interesting to explore this in more detail; however, as discussed in more detail in Section 8.3, this is another branch that will be left for future research.

Error rates increase with increasing smoothness coefficient (see Figure 9.8), with the majority of Participants knowing when they had got the location of the target incorrect. We speculate that the increase in error rates at higher smoothness coefficients is due to

adoption of a quitting threshold: typically, the smoothest objects take the longest to find, therefore these are most likely to be the objects where the participant quits their search with a guess of object location. This hypothesis is supported by two aspects of the results: Firstly, the behaviour of Participant M who has a clear quitting threshold of 10s (they only searched for greater than 10s on 5 trials out of 300, see Section 2.1.2) and a decrease of percentage correct with increasing smoothness coefficient. Secondly, the incorrect trials showed considerably longer reaction times than the correct trials, with the mean incorrect reaction time being 14s and the mean correct reaction time at 4s. This is indicative of a quitting threshold, with the majority of incorrect trials being quitted without knowing the location of the target, rather than getting the target location wrong.

This experiment provides an addition to the investigation of realism on 3D vision. Currently, there is very little research looking at the factors affecting binocular depth perception in the context of naturalistic scenes. Most current work with naturalistic stimuli looks at other aspects of 3D vision such as shape from shading cues. One such study by Lovell et al. uses photographs of stones, and shows either self-shadows or cast shadows (Lovell et al., 2009). This study is primarily interested in how shadows are processed, and if a stone with an incongruent shadow is faster to detect. This paradigm could potentially be adapted to see if the addition of shape from shading cues speeds up visual search for realistic objects. Some other studies look at established visual search effects in more realistic scenes, such as those of Neider et al (2010). It would be interesting to further increase the reality of these studies by adding binocular disparity, and see what influence this has on participants' search performance.

In the next Section, we discuss the original predictions and motivations of this experiment by comparing it to the previous experiment with a flat background. We then go on to present another experiment that is designed to be a real world analogue of the experiment presented here.

9.2.4.2 Comparison of Experiments 9 and 10

At the beginning of Experiment 10, we hypothesized that the introduction of a naturalistic background would have two effects: Firstly, an overall slowdown in reaction times; Secondly, a greater slowdown for the smoothest object. Fortunately, five participants completed both Experiments 9 and 10, allowing us to have a look at the difference between their performances in the two experiments. In Figure 9.10, we compare the mean reaction times between the two experiments, both individually and on a group level. We can see that the reaction times for the naturalistic background (Experiment 10, dotted lines) are typically greater than that for the flat background (Experiment 9, solid lines).

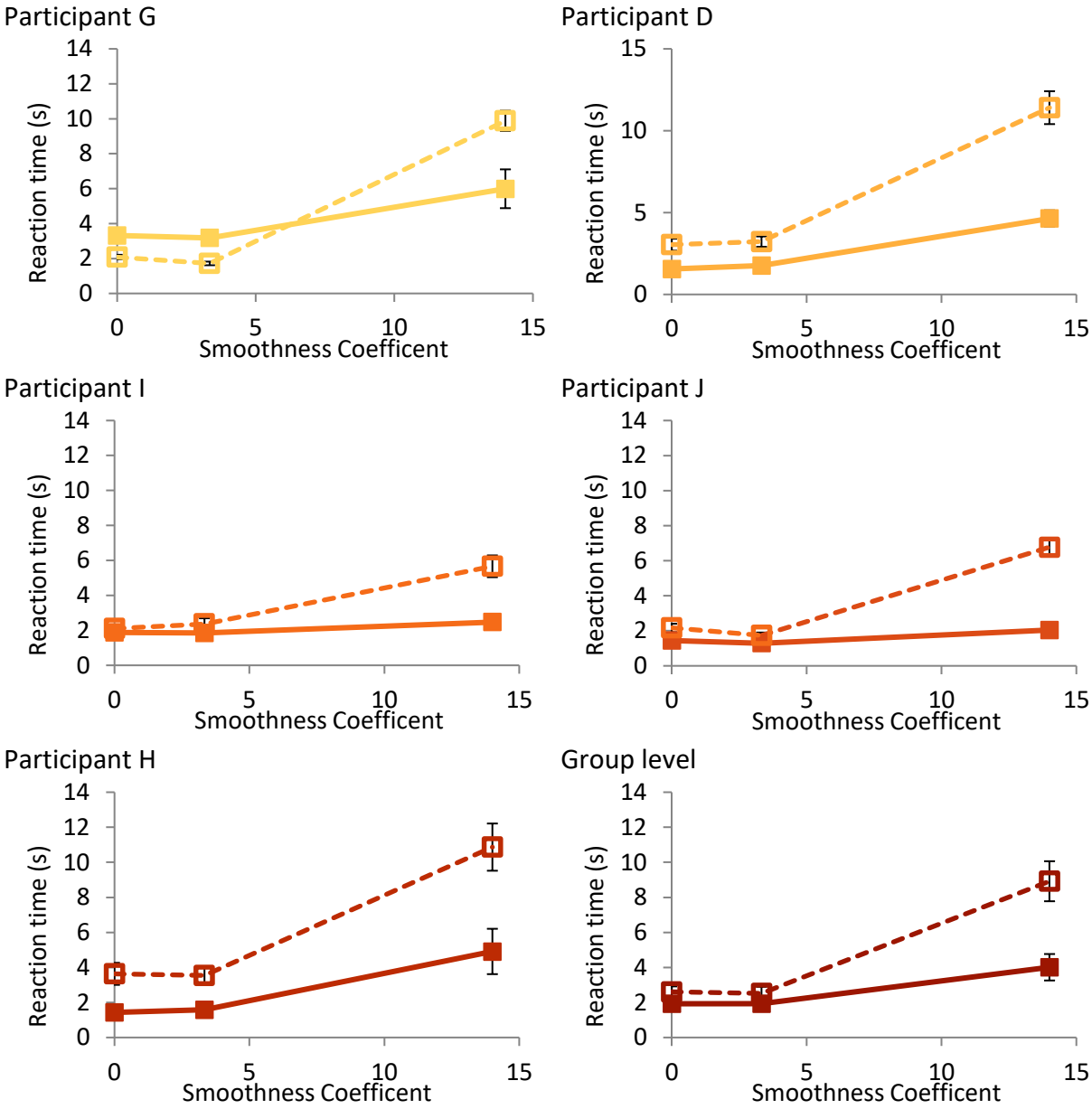


Figure 9.10: Differences in performance between Experiments 9 and 10, for all 5 observers and a comparison of the group mean. Error bars are one standard error.

Participant	SC0	SC3	SC14
D	0.585	0.684	0
G	0.262	0.299	0
H	0.005	0	0
I	0.017	0	0
J	0	0.121	0

Table 9.2: Individual statistics from the Wilcoxon Signed rank test comparing between the smoothness coefficients for the flat and fractal Experiments. As before, take care drawing conclusions from individual statistics due to the number of comparisons made.

On a group level, there is a significant difference between the two experiments with smoothness coefficient 0 (Wilcoxin signed-rank test, $p=0.005$, $Z=-2.782$) and 14 ($p<0.0005$, $Z=-11.6$) but no significance between the smoothness coefficient 3 conditions ($p=0.717$, $Z = -0.363$). Individually, only participant J followed this trend, with Participants H and I having significant differences between the experiments for all smoothness coefficients, and Participants D and G only between the SC14 and the other conditions (see Table 9.2). This indicates quite a large degree of individual variation due to the introduction of the naturalistic background, suggesting that some participants found it easier to isolate the target than others. All participants showed a decrease in percentage correct in Experiment 10 over Experiment 9.

At the beginning of Experiment 9, we made two hypothesises; the first was that naturalistic depth background would hinder search times for all smoothness coefficients. The difference is not consistently observed, with only two participants showing a significant slowing in reaction time across all conditions. This inconsistency is probably due to individual differences such as search techniques and differences in depth perception. For some observers, the difference in disparity between the base and the peak of the object will still be great enough to be easily identifiable from the background, causing the introduction of the background alone to be insignificant to search times. For observers that are poor at detecting disparity, the background may interfere more with the perception of stereoscopic depth, making it harder to distinguish the objects from the background and causing a slowdown in reaction time. It would have been interesting to measure stereo-thresholds in a similar RDS to these stimuli and see if there was any link between stereo-thresholds and search times.

An additional effect that may cause individual variation in detection of the object is the way in which the object was added to the background. The object was inserted by summing the object's disparities with the disparity of the naturalistic background, meaning that the top surface of the object was modulated with a depth profile identical to the background. Observers relying on the surface being flat from their experience of the previous experiment could find it harder to identify the target object, whereas more flexible observers may be more able to discount this difference.

The second hypothesis, that the smoothest object would show a greater increase in reaction times has been supported by these Experiments 9 and 10, with the difference being highly significant on both a group and individual level ($p<0.0005$ for all participants). This is most likely caused by the increase in difficulty of segregating out the object from an undulating than flat background. Here, we discuss this effect and a couple of other possible explanations:

1. The perceptual grouping of the smooth object with the smoothly changing background is enhanced compared to a flat background, as we have increased the similarity of the object to the background thus perceptually grouping them (see

Figure 2.12). This would increase the difficulty of segregation of the object from the background, but because of the background characteristics as opposed to the difficulty of processing the object in isolation. This could be considered a form of background matching, as it is the similarity between the object and its specific background that makes the object harder to detect. This explanation highlights the importance of considering the environment when drawing conclusions about visual search tasks.

2. It is possible that the perceptual decrease in peak depth of the smoothest object causes it to be perceived as being similar to the background objects. This seems unlikely, as a smoothness coefficient 14 object only has a perceived peak depth decreased by approximately 8.5% (Section 4.3.3), whilst the background variation was less than a third of the depth of the object over the length scale of the object (Section 9.2.2.1). Thus, we consider it to be unlikely for the target to be misidentified as a background feature. However, the decrease in perceived peak depth may interact with explanation 1, and help form slightly better background matching than the sharper objects, making the smooth object have better camouflage for this display.
3. Additionally, the projection of the background undulations to the front of the object will have the greatest effect on the smoothest object as it is joined continuously to the background. The projection of the undulations onto the smooth edges will make the smoother object more similar than the sharp object to the background.

Explanation 1 is supported by the literature showing that perceptual grouping effects search times e.g. (Deas & Wilcox, 2015) (see Section 2.2.6 for a discussion), however we have insufficient evidence to draw any conclusions. In the next Section we propose a speculative future experiment that may begin to tease apart these explanations.

Interestingly, there is an increase in error rates for Experiment 10 over Experiment 9, particularly for the SC14 stimulus. It is likely to be caused by a mix of two different effects depending on the response times:

1. The increased search time led participants to reach their quitting threshold, and they thus clicked to continue onto the next trial. Participants may have felt incentivised to quit more readily in Experiment 10 than Experiment 9 due to the presence of an “I don’t know” option. This explanation is supported by the fact that there were longer response times in incorrect (mean 14s) than correct trials (mean 4s, see end of the results in Section 9.2.3).
2. Participants mistook undulations in the natural texture background to be the target, and realised they were wrong when there was no response option at their selected location. This explanation would explain the short response times where participants

were incorrect (causing a standard deviation in the mean incorrect response time of 17.7s).

9.2.4.3 Future directions

The grey bar used in Experiment 10 gave us an idea of why the participant was incorrect on different trials. However, further experiments could untangle the reasons for participants' performance. For example, it is hard to know if participants used the grey bar because they had misidentified the target or quitted the search. To distinguish between these two, the number of alternative targets in the response screen could be increased until it was saturated, so that all locations on the RDS had a potential response target nearby (see Section 7.2.2). In the case that Participants had misidentified a background undulation as the target, then this would ensure a response target would be nearby for them to click, which would then give an 'unknown incorrect' trial. In the case of quitting threshold, the participants would still be using the 'known incorrect' option (maybe now better labelled the 'I quit' option), enabling the determination of whether participants were misidentifying the target or quitting.

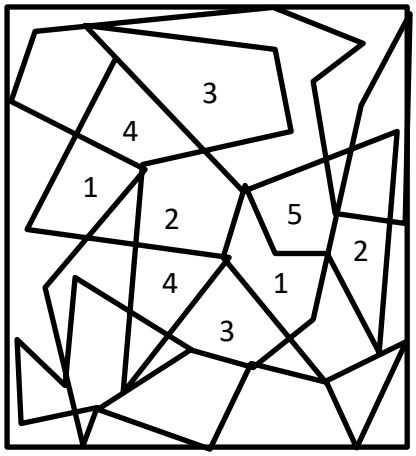


Figure 9.11: The display is broken up into polygonal areas, each of which is assigned a constant disparity (numbers are example disparities). The change in depth between the areas is sharp, meaning the edges indicated by black lines have a similar depth profile to the edges of the sharp object.

We have three speculative reasons (see end of Section 9.2.4.2) to why the smoothest object was found to take longer to detect in the naturalistic background than in the flat background (compare experiments 9 and 10). We offer a possible experiment that may start to distinguish between these three explanations. If we generated a flat background that was broken up into polygonal areas with sharp changes in depth between them (see Figure 9.11), then the changes in depth in the background would be most similar to the sharp object. The sharp object would then have the best background matching, and the smooth object the poorest. If the sharp object then takes longer to detect compared to Experiment 9 than the smooth object, this will indicate that the change in reaction times is due to the background matching rather than the alternative explanations.

Overall, we conclude that the environment in which the visual system is operating is an important consideration – here we have found that the camouflage of the smoothest object is significantly enhanced in a naturalistic background. This illustrates how we need to anchor our abstract experiments back to natural or naturalistic environments when drawing conclusions about the function of different aspects of the visual system – even with a minor increase in reality we have altered our results considerably.

It would be possible to continue adding realistic elements to the experiment using computer generated imagery. Using tools such as Radiance, it would be possible to test the performance of these shapes of target in a naturalistic environment whilst maintaining control of the cues to depth present in the scene. This would enable the comparison of binocular vision to other cues to 3D shape when spotting our three dimensional shapes, and to see how effective stereoscopic camouflage from shape was at reducing search times.

With a wealth of directions to go in, we decided the most informative was to test the combined effects of the additional cues on the stereoscopic camouflage displayed by these objects by going straight to full reality. In Experiment 11 we introduce a real world analogue of Experiment 10 to see how well our objects perform in a realistic camouflage search task with both monocular and binocular cues to depth present.

9.3 Experiment 11: Naturalistic analogue using a sandpit

We have found in Chapter 7 that while disparity defined depth speeds up target detection, it does not make the task of detecting a camouflaged object trivial. In Chapter 8 we investigated this further, and found that the more poorly defined the edge of the object, the slower it was to spot. Both these Chapters have used very abstract stimuli from which we are making conclusions about how camouflage in the real world works. In this Section, we draw inspiration from a paper by Foulsham et al. (2014) who compared between lab and real world visual search experiments (paper is discussed in Section 2.3). We investigate our findings so far by using real objects in a real world experiment to see if the effects found in the computer based experiments still hold.

So far, we have introduced a random noise background similar to continuous background textures found in real life. We found (in Section 9.2) that the smooth edged object was much harder to detect than sharp edged object. This replicated the results we have found previously, although the smoothest object was significantly harder to find in the natural noise background than the flat background (see Section 9.2.4.2 for a discussion).

In this experiment, we create a real world analogue of the screen-based visual search tasks used in the Experiments 8-10. We replace the computer screen with a sandpit, photographed in Figure 9.12. We have created a 3D printed version of each of the three targets used in the previous experiments (with smoothness coefficients 0, 3 or 14, see Figure 4.3) page 63. Each of these targets has been camouflaged by coating it in sand. Each trial consists of one of these objects being placed in the sandpit, and the participant attempting to find the target as fast as possible. As before, we look at the effect of smoothness coefficient on reaction times. Due to the abundance of visual cues other than binocular disparity introduced by presenting the experiment in the real world, it is not possible isolate if any effects are due to binocular disparity specifically, instead we are primarily interested in whether the reaction times are slower for smoother objects – this will allow us to test if the smooth objects are better camouflaged in a real environment.

Figure 9.12: Photograph of the sandpit used to hide real world objects in. Note the white markers around the edge, and the T shaped tape on the floor that the participants stood with one foot each side of the stem.

9.3.1 Methods

9.3.1.1 Experimental setup

In order to create a real world analogue to the previous experiments, we have created a controlled but naturalistic environment in a laboratory setting. As would be expected, the equipment has changed dramatically from the previous experiments.

The participant, instead of searching for a target object on a screen now searches for a target object in a 1.8m square sandpit, as photographed in Figure 9.12. The sandpit has been lined with black pond liner, then filled with approximately 75kg of play pit sand, leading to a sand depth of 1-2cm. Due the pond liner not lying completely flat at the edges of the sandpit, the stimulus is never placed within 18cm of the edge of the pit. This is similar to previous visual search experiments, where the stimulus could not appear in the 74arcmin closest to the edge of the screen.

The stimuli used in the previous experiments were translated into a real world analogue via 3D printing. The stimuli were designed to be the same angular size (74.7 arcmin) and disparity (6 arcmin) as the on-screen stimuli when viewed (at an angle of 32.5^{o}) by an average height participant (1.57m) from a (ground) distance of 75cm. When placed closer in the sandpit to the observer, the object will have a greater angular size, and when further away it will be smaller. To ensure that observers were central and 75cm from the centre of

the sand pit, we stuck tape in a T shape to the floor – as can be seen in Figure 9.12. Observers were asked to place one foot each side of the stem of the T (which was located midway along the side of the sandpit) with their toes on the top of the T, was located at 15cm from the sandpit.

We used Matlab to translate the three objects (smoothness coefficient 0, 3 and 14, see Figure 4.3, page 63 for a cross-section) into 3D printable STL files (Sven, 2008). The printed objects had a total side length of 37 by 37mm, and a height of 3mm. The 3D printer used white nylon to an accuracy of 0.4mm (3DPrintUK, 2016). In order to establish camouflage of the object in the sand, we coated the 3D printed shapes with sand. PVA glue was diluted to approximately 1 part in 2.5 water, and brushed thinly over the 3D shape. The object was then covered in sand and left several hours to dry. This process was repeated 3 times, resulting in a smooth covering of sand with no white patches visible.

The PVA glue caused the sand to be a slightly darker colour than the loose dry sand. Under strong lighting, the difference was visible, but under dim lighting the difference was much harder to spot. We therefore used diffuse lighting provided by a 60W halogen lamp which was pointed at the ceiling above the sandpit (the lamp was located approximately 1m from the ceiling and 3m from the sandpit) – this technique also minimised shadows on the object. Additionally, the difference in colour is the same for all three objects, so any decrease in detection time from the colour will affect all objects identically.

As the stimuli were real 3D objects, we had two copies of each of the shapes – one used in the experiment and one for backup. Unfortunately, PVA glue adheres relatively weakly to nylon plastic, so after two weeks of use the original objects had to be rotated out and repaired, while the backup versions were used in their place. These backups were identical except for small differences in surface texture from when they were coated in sand. This was not noticeable prior to wear and tear – we had to label the back of the objects to identify which objects were backups.

To 3D print the smooth objects, all objects were printed with a 1mm thick backing 37cm across. This edge was made thicker by the addition of the sand coating, and made the objects easy to detect when placed on the surface of the sand. To avoid the objects being detected by the edge of their backing rather than the shape of the surface the experimenter followed the following procedure to conceal the edge between the backing of the 3D printed object and the sandpit sand, with the number of each stage referring to a numbered photograph in Figure 9.13:

1. Object was placed on the surface of the sandpit.
2. The object was buried in sand.
3. A small foam brush was used to remove all loose sand on and surrounding the object.

4. A large foam brush was used to brush off the sand from the backing, leaving sand slightly covering the edge of the backing.

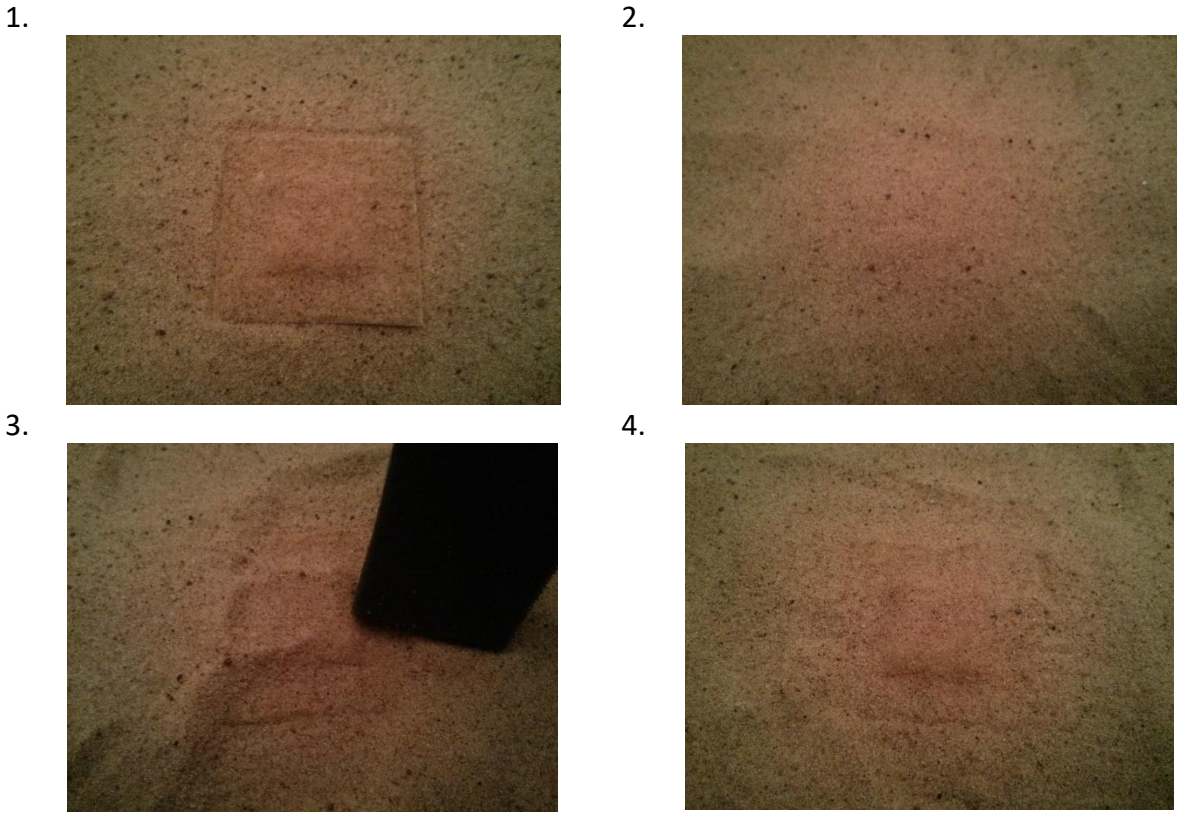


Figure 9.13: Step by step illustration of hiding the object in the sandpit. Numbers refer to the step-by-step guide in the text.

This technique effectively removed the edge of the backing as a cue, although it necessitates the interval between trials to be around 20-30s between trials. For this reason, participants were allowed to use their phone to occupy themselves between trials in order to alleviate the boredom of waiting. However, despite concerns of boredom, many participants commented that the experiment was more interesting than most they participated in.

The use of the foam brush, and placing and picking up the object left small ridges and textures in the sand. These were very tricky to remove and create a smooth sandy background, so the entire sandpit was deliberately textured using the large foam brush. This made any artefacts from placing and interacting the object match the overall texture of the sandpit, and therefore not indicate the position of current or past objects to the participant.

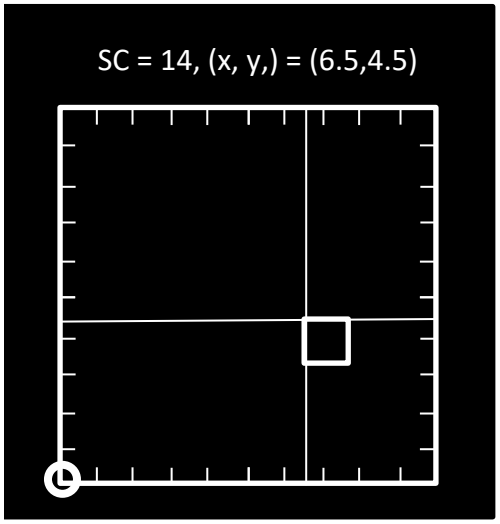


Figure 9.14: Mock-up of the object placement screen.

We used an altered version of the computer program that ran visual search experiments to determine where the object should be placed, and to keep track of data trials. Figure 9.14 shows a mock-up of the object placement screen shown to the experimenter. At the top of the screen, the smoothness coefficient of the object to be used was displayed, next to the randomised position of the object in (x, y) coordinates. To assist in placing the object, this location was displayed on a diagram of the sandpit – each tick on the sandpit diagram corresponded to a white marker on the edge of the sand pit. This enabled the experimenter to accurately place the object without marking the sandpit in a grid pattern which might have assisted a participant in searching the sandpit (Gilchrist & Harvey, 2006; Scinto & Pillalamarri, 1986). In order to ensure that the experimenter always related the coordinates of the sandpit correctly to the screen, a large O was placed in one corner of both the diagram of the sandpit and the sandpit itself.

9.3.1.2 Experimental procedure

The methodology used during experiments was intended to be similar to the procedure in the previous computer based experiments (described in Section 7.2) to keep the results as comparable as possible. The start of an experimental session was therefore identical – after signing the participant consent form, participants completed the TNO stereotest ("TNO Stereotest, Richmond Products," 2014). Participants were required to complete until at least the 6th plate or they were rejected from the experiment.

Participants able to complete the TNO stereo test were then shown a demonstration target object. This target object was not one of the objects used in the experiment – instead it was a cuboid of the same dimensions and sand coating as the target objects, but without the backing. By asking participants to search for this object then it ensured they did not search for the combination of object and the backing, but the object alone.

On the floor in front of the sandpit was a white marker indicating where the participants should stand (discussed in Section 9.3.1.1). Participants were asked to stand one foot each side of this marker, then turn around and face away from the sandpit. The experimenter then placed the object using the computer screen and technique described in Section 9.3.1.1.

Once the object was placed and the experimenter had returned to the computer, then the participant was asked to turn around and to search for the object in the sandpit. As they turned around the experimenter clicked the mouse button to start the timer. Once spotted,

the participant verbally reported seeing the object and the experimenter pressed a mouse button to stop the timer. The participant then indicated with a laser pointer the location of the object, and the experimenter recorded if they were correct using the (Y)es and (N)o buttons on the keyboard. No feedback was given. The participant was then asked to turn around, and the experimenter picked up the old target. This procedure was repeated for every trial.

Trials were then continued for a full hour. Unknown to the participants, the hour session was blocked in groups of 12 trials, each of which had four repetitions of object of smoothness coefficients 0, 3 and 14 randomly interleaved. By running an integer number of blocks we ensure that the participants complete an equal number of the different conditions, but managed to complete as many trials in one hour as possible. Participants completed two 1 hour sessions on different days. Participants that did not come to both days were rejected, as they had insufficient results to give us useful data.

For this experiment, trials took much longer than previous visual search experiments, so despite running the experiment for two hours only have a mean of 26 trials per participant per condition – much lower than the previous 75 trials. To alleviate problems caused by slow data collection for each participant, we ran a total of 17 participants, out of which 14 completed both hours of experimentation. All participants passed the TNO test, but one participant did not come back for the second hour of participation, and two were ill during testing and had to stop prematurely.

9.3.2 Results

As with the previous visual search experiments, we are running the experiment to measure how well smoothing the edge of the object makes the object camouflaged. We do this by measuring the detection time of the object – an increase in detection time indicates an object that is better camouflaged.

Mean reaction times for correct trials by each individual participant are shown in Figure 9.15. On a group level, all conditions are significantly different from each other (Wilcoxin signed rank test, $p<0.0005$, SC0 to SC3: $Z=-7.611$, SC0 to SC14: $Z=-12.774$, SC3 to SC14: $Z=-8.075$). On an individual level, shown in Table 9.3, all but one participant (U) has significance between the SC0 and 14 conditions. However, Participant U's data should be regarded with caution, as they were only right 37% of the time for the SC14 condition, resulting in a total of 12 correct trials.

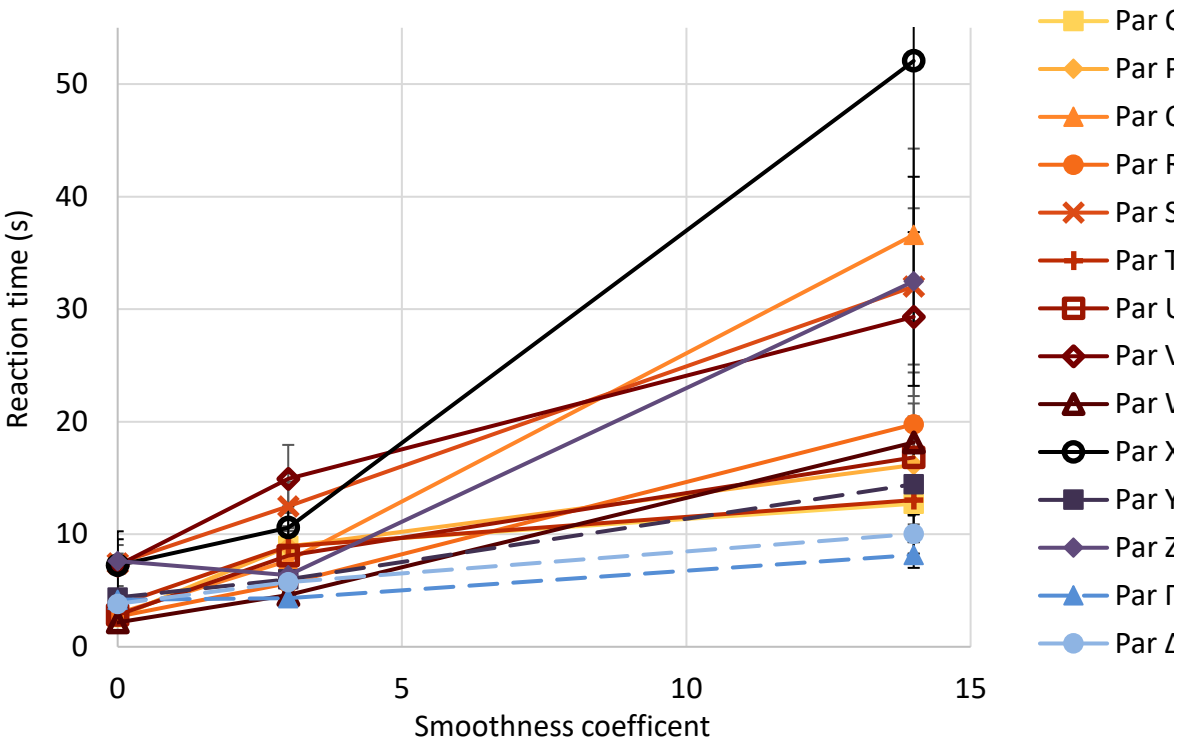


Figure 9.15: Mean reaction times for correct trials for participants in Experiment 11.

Participants' were more variable on an individual level, although we should be very cautious drawing conclusions from individual statistics, as we are doing 45 difference comparisons, greatly increasing the chances of false negatives and false positives. To limit the crowding on the figures, we split the participants into individual series on a bar chart in Figure 9.16 – this enables us to easily compare the performance between smoothness coefficients for each individual participant. Statistically, there were two major groups: Participants Q, R, S, X, Y, Z, Δ and Γ did not have significance between smoothness coefficients 0 and 3. Participants O, P, T, U, V, Δ and Γ had no significance between smoothness coefficients 3 and 14. For the former participants the results approximately replicate the previous experiments, where only a very smooth stimulus made a difference to detection rate. For the latter participants, these participants could be using strategies and extra cues not available in the previous experiment. We discuss what may be causing these differences in the discussion.

The data in this experiment was more skewed than previously. This effect can be clearly seen by comparing the median reaction times in Figure 9.17 to the median reaction times in Figure 9.6 (page 163). Whilst this shows a difference on an individual level between the mean and median for some participants, the overall effect is the same – reaction times increase in increasing smoothness coefficient.

As with previous visual search tasks, the percentage correct is informative. We plot percentage correct in Figure 9.18, and can see that participants are typically correct all the time for Smoothness coefficient 0 and most of the time for smoothness coefficient 3.

smoothness coefficient 14 is more variable, with the percentage correct being between 60 and 100% correct. We looked into this in more detail, and found that participants typically had a quitting threshold in the region of 30s. After this point, almost all trials were incorrect, likely caused by participants getting bored with the task and guessing the object's location.

Participant	0 - 3	0 - 14	3 - 14
Par O	0	0	0.241
Par P	0	0	0.028
Par Q	0.028	0	0
Par R	0.024	0.002	0.017
Par S	0.072	0	0
Par T	0.001	0	0.61
Par U	0.001	0	0.024
Par V	0.014	0.004	0.03
Par W	0.009	0	0
Par X	0.034	0	0.004
Par Y	0.313	0	0.003
Par Z	0.201	0.005	0.003
Par Γ	0.775	0.028	0.034
Par Δ	0.142	0	0.086

Table 9.3: Per participant statistical analysis of the real world experiment. Green cells indicate a significance p<0.017, the Bonferroni corrected significance level for three comparisons.

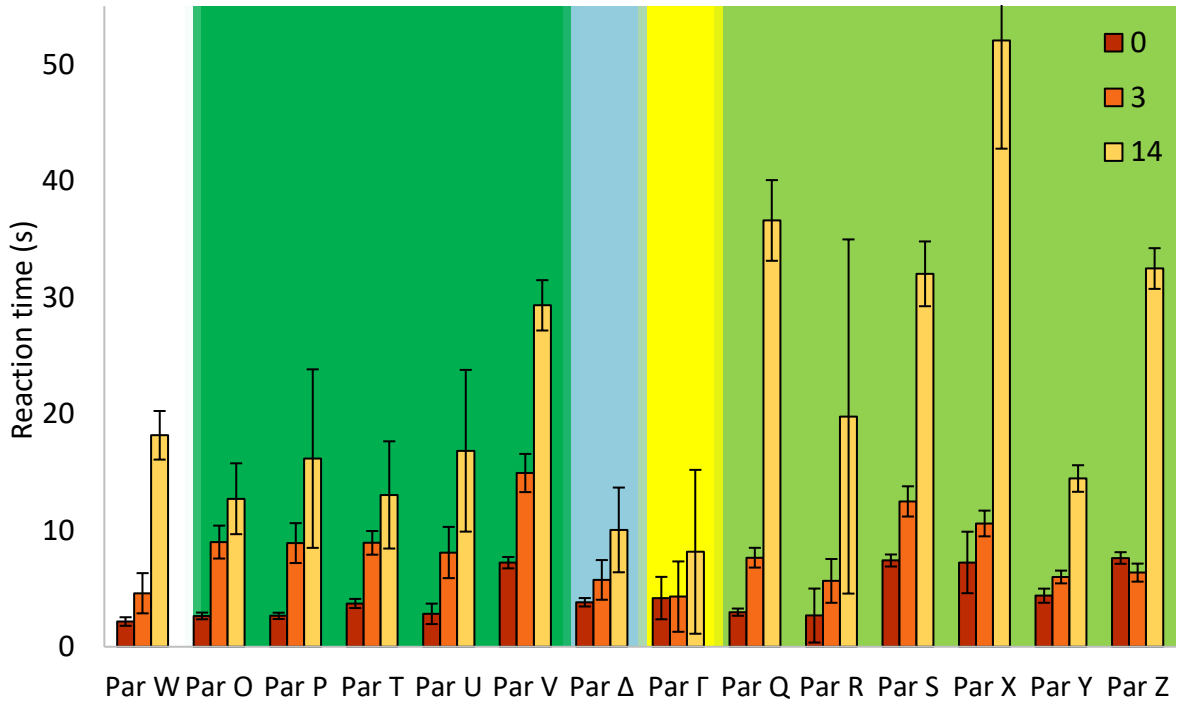


Figure 9.16: Mean reaction times for correct trials for each participant in Experiment 11. White (Par W) is significant between all conditions, dark green (Pars O-V) are not significant between smoothness coefficients 3 and 14, light blue (Par Δ) is only significant between smoothness coefficients 0 and 14, yellow (Par Γ) is not significant for any condition and light green (Pars Q-Z) are not significant between smoothness coefficients 0 and 3.

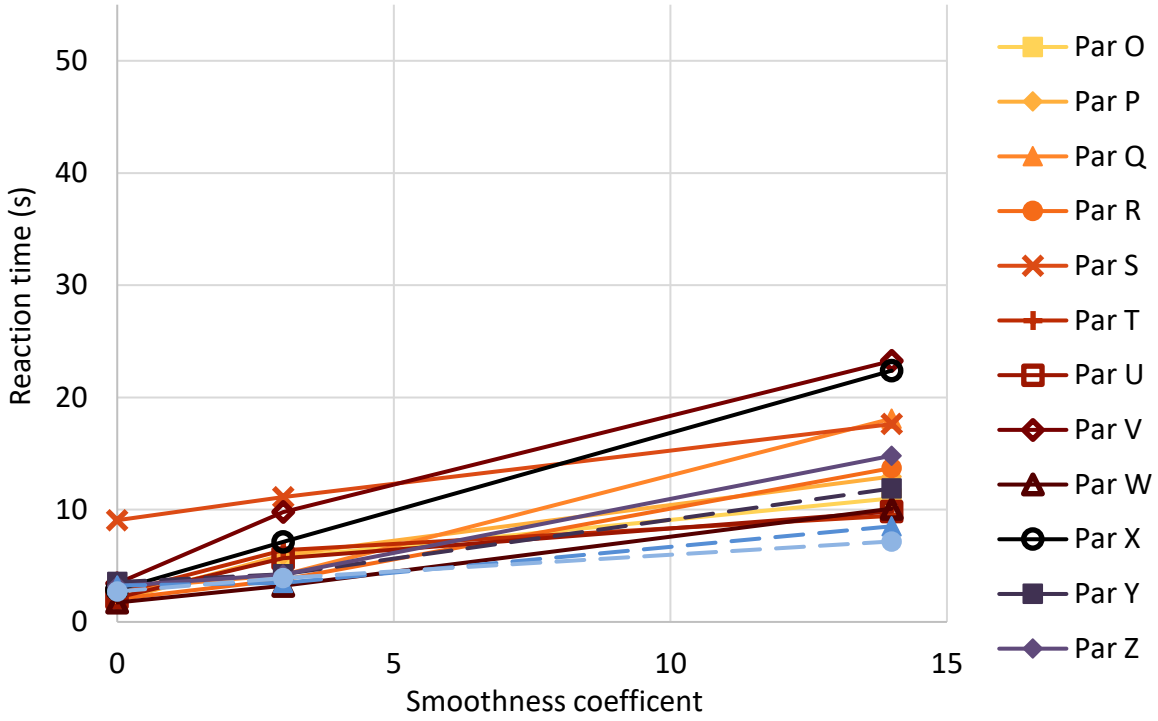


Figure 9.17: Median reaction times for participants in Experiment 11.

.

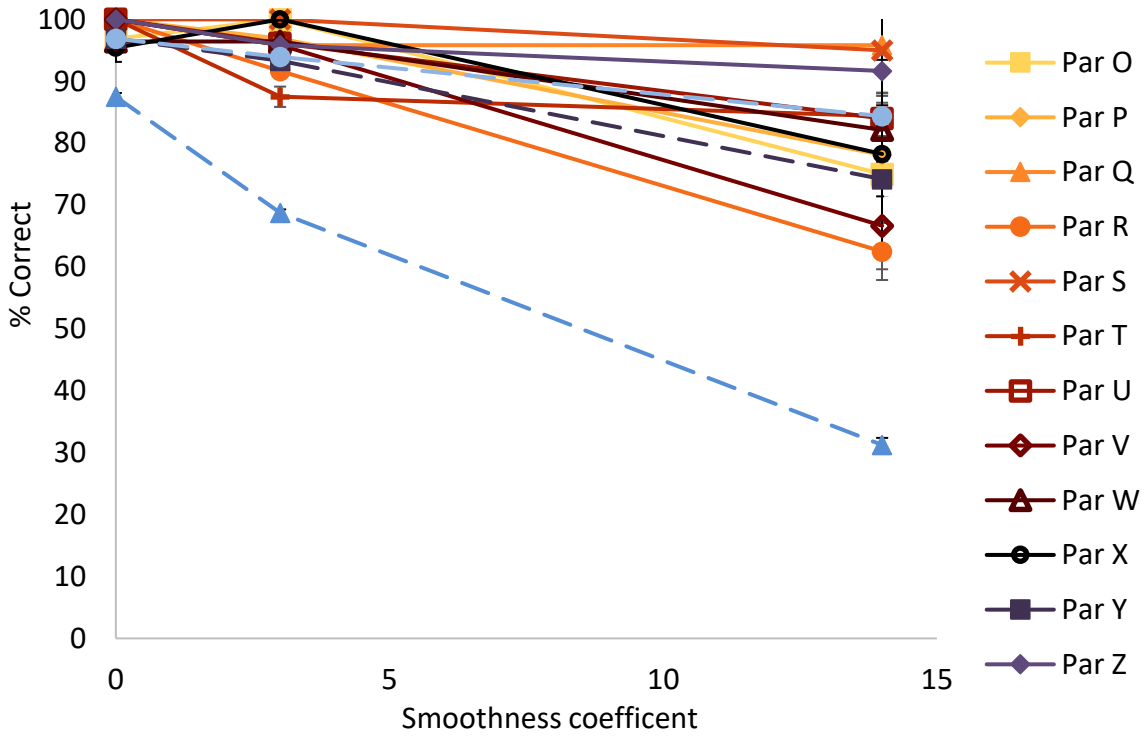


Figure 9.18: Percentage correct for participants in Experiment 11.

We also considered if the position of the object in the sandpit or the trial number had any effect on the reaction time. With so few trials, it was hard to analyse the effect of trial number, and we concluded that we did not have enough samples to obtain a coherent analysis. We also divided the sandpit up into a grid and looked at the mean reaction time for different positions in the sandpit. The position in the sandpit was dominated by an asymmetric distribution of different smoothness coefficients across the sandpit, but again no meaningful analysis could be done because of the small number of trials.

We analysed a few additional aspects of the raw data in this task to see if we could extract any additional trends, but found nothing of note.

9.3.3 Discussion

In this thesis, we have been experimenting on highly abstract RDS displays and drawing conclusions about camouflage in a real world setting, where there are many more cues and complexities to extracting depth. We agree with Richards in 1977 that there is a tenuous link between RDS and behaviour in the real world (Richards, 1977). Despite the lack of attention to this conclusion (Regan, n.d.), we decided to conduct a real world experiment to test the grounding of our results in RDS in the real world.

Overall, we have created a real world analogue of a visual search task, that, although slow has given us some interesting insights into the behaviour of these objects in the real world and given us some interesting comparisons between real world and virtual search tasks. We restrict the discussion of real world vs virtual experiments to the next Section (9.4). Here, we discuss the effects that may explain the results observed in this experiment: We consider the difference between smoothness coefficient 0 and smoothness coefficient 3 in relation to self-shadow concealment (the use of colouration or shape to conceal shape from shading cues, see Section 2.1.1). Finally, we round up with a discussion of some of the problems faced when creating a real world experiment.

9.3.3.1 Self-shadow concealment

One of the monocular cues to shape and form that we have added to the experiment by conducting it in the real world is shadows and shading, caused by a non-homogeneous (although diffuse) light source. These shadows can be used to extract depth information, and thus forming a monocular cue to 3D shape and form (Norman et al., 2006; Todd, 2004). Self-shadow concealment is a camouflage technique (discussed in Section 2.1.1) where the animal changes its shape and/or colouration in order to reduce or remove the presence of shadows so they are not available as cues to object shape (Penacchio, Lovell, Sanghera, et al., 2015; Rowland, 2009; Ruxton et al., 2004; Stevens & Merilaita, 2009). Here, we discuss if the smooth object may be displaying self-shadow concealment, thus making it harder to spot.

The significant difference on a group level between SC0 and SC3 clearly shows an effect that has not been present in the previous experiments. We speculate that this significant

difference is due to the presence of shadows in the stimuli that were not present before. Despite our attempts to make the lighting in the room as diffuse as possible, there were still faint shadows on the object, which were more distinct with the object with sharp edges. The two smoother objects had little shadowing in comparison – thus their shape may have been acting as a form of self-shadow concealment, another form of camouflage (Penacchio et al., 2016; Penacchio, Lovell, Sanghera, et al., 2015; Penacchio, Lovell, Cuthill, et al., 2015; Rowland, 2009; Ruxton et al., 2004).

We see a significant difference between smoothness coefficients 3 and 14. For brevity, we will refer to 'smoothness coefficient' as 'SC' for the rest of this discussion. While there is a weaker shadow on the SC14 than the SC3 object, it is hard to quantify. Under the diffuse lighting in the room when standing from the participants' viewpoint we thought there was perceptually less difference in shadowing between the SC14 and SC3 objects than between the SC0 and SC3 objects. Additionally, the slowdown between SC3 and SC14 is much greater, with taking on average 4.7 times slower to detect SC14 over SC0, and only 1.7 times slower to detect SC3 over SC0. The SC14 object does not appear to have much less shadowing than the SC3 object, but had a much greater reaction time than SC3, compared to the difference between the SC0 and SC3 objects. We therefore speculate that differences other than shadowing between the SC3 and SC14 object are also making it much harder to spot, such as an increase in difficulty of segregating the object (for detail or more possibilities, see discussions in Section 9.2.4).

On and individual level, we had two major groups of participants – those who had no significance between SC0 and SC3 and those with no significance between SC3 and SC14. Due to the large number of comparisons being made, a proper Bonferroni reduction in the level of statistical significance would render almost none of these comparisons significant. However, we tentatively speculate that those who have no significance between SC3 and SC14 may be segregating the object using shadows (as there appears visually to be little difference between the shadows on the SC3 and SC14 objects) – hence there is little difference in the time it takes to detect SC3 and SC14 objects. When SC0 and SC3 are not significant this may indicate participants who are using other cues such as disparity to segregate – for these cues there is little difference between the SC0 and SC3 objects. However, this is a very tenuous speculation, and would require in depth study with a specifically designed experiment and a method of measuring the strength of the shadowing cue to object shape to find any evidence.

An interesting extension to this experiment would be to have the participants perform it with one eye shut to test if binocular disparity was assisting participants in detecting the target object. This would remove binocular disparity cues, although it would also remove binocular information that was not due to depth from disparity, as discussed in Section 7.4. An example is that the overlapping views from two eyes helps boost signal and reduce noise, thus having only one eye would reduce the strength of cues we think of as

monocular, such as luminance (Formankiewicz & Mollon, 2009; Heesy, 2009; Jones & Lee, 1981; Martin, 2015; Ott et al., 1998; Prinzmetal & Gettleman, 1993). However, removing binocular information could enable us to start to determine how much difference depth from disparity and half occlusions make to detecting camouflaged objects in a real world environment. We think it is improbable that shutting one eye would stop the smoothest object being the hardest to spot, due to the effects of self-shadow concealment and the increased difficulty of segregation of the smooth object for multiple cues (see Section 9.4 for an in depth discussion).

9.3.3.2 Future directions of real world experiments

Needless to say, conversion of a virtual search task to a real world search task was an exacting process, but one which we can consider a success. However, it is not without its problems, and there are more experiments to be done to control more variables in a real world experiment. The most noteworthy of these are:

1. Finer control of lighting: Particularly how diffuse the lighting was. This would enable us to test if the difference between SC0 and SC3 is indeed driven by shadows, and if the effect between SC3 and SC14 is still robust in the absence of shadowing.
2. Better colour matching: Weak diffuse lighting rendered the difference in colour between the objects and the sandpit to be subtle, we would have preferred the colours and textures to be identical between the target and the object to make the monocular camouflage as good as possible.
3. Faster experiment: While the error bars are remarkably small given the small number of trials it would have been valuable to have more repeats and less downtime between the trials.
4. Control of viewing angle: Participants are of variable height, and stand to one side of the sandpit, leading to the participants' viewing the object from an angle. With a taller room, we could have placed participants on a pedestal or even balcony above the experiment, meaning that participants would look straight down on the objects.
5. Multiple viewing distances: Allison et al. (2009) found the presence of environmental cues altered selectively enhanced stereoscopic depth at certain distances (4.5 and 9m). Combined with control of viewing angle it would be interesting to make several 3D printed objects, so we could vary viewing distance whilst keeping disparity constant to see if the different objects were faster to spot at certain viewing distances.
6. Participant behaviour: We do not know how much time we measure between the mouse clicks that the participant is actively searching the sandpit, if they are using their eyes or head to search the visual field (in the previous experiments, head movements were controlled), or how much time they are spending verifying the target.

The real world visual search experiment that we used partly as inspiration for this study by Foulsham et al. (2014). Foulsham et al. fitted participants with a mobile eye tracker, then had the participant search a mail room for a pigeon hole with a certain name on it. The pigeon hole was made more visible (salient) by making it coloured. The eye tracking data enabled Foulsham et al. to investigate the change in participants' head and eye movements when they were informed that the colour was important versus when they were naïve to the presence of the colour cue. The nature of Foulsham's experiment meant that several of the issues we faced were not a problem. However, other problems, such as the participants' visual behaviour at searching the sandpit could be informed by using their eye tracking software. This enabled Foulsham et al. to analyse the participants' behaviour while searching, and identify trends such as how much time the participant spent searching by moving the head as opposed to moving the eyes.

Other explorations of the perception of the object in the real world could be made by moving away from visual search experiments. To better compare the real world and virtual objects we could create a setup similar to the experiments of McKee and Taylor's (2010) and Frisby et al. (1996). In McKee and Taylor's experiment, participants made judgements in depth between a pair of vertical rods both with and without adjacent objects and textures. A similar experiment, using our objects instead of rods could inform us on the effect of the textured sandpit on the perception of depth in the object. We could also vary the texture of the objects to measure the effect of the sandy coating on the perception of the object.

This regime could be taken further – we could create the object out of LED lights, similar to the work of Welchman et al. (2004). Welchman et al. tested how participants judged if an approaching object would hit them. They originally used binocularly defined dots moving in depth on a computer screen, but compared this to real life performance by using a movable LED light in a dark room. This regime allowed them to measure the effects of cues present in the real world, such as changing accommodation, whilst maintaining the moving object as a point of light. We could emulate this idea, and create a 3D display of LEDs to create real world objects defined by points of light in a dark room – a display that would be very similar to our computer based tasks. When compared to the solid 3D objects this would enable us to distinguish between effects from bringing the object into the real world and effects from adding additional cues such as shape from shading.

9.4 Comparison of virtual and real world experiments

In this Chapter of the thesis, we aimed to investigate how well camouflaged the smooth object would be in a more realistic setting. In the first Experiment (Section 9.2) we added a naturalistic disparity noise background to investigate if the increased difficulty of finding the targets due to the background noise would drown out any effect of the objects' shape. In the second Experiment (Section 9.3) we realised this experiment into a real world equivalent, using 3D printed objects and a sandpit. Here, we consider the implications of the differences between these two experiments.

Perhaps the most striking difference between our real world search task and our virtual search tasks, shown in Figure 9.10, is the difference in reaction times between SC0 and SC3 objects. In the virtual task, we have no significant difference in this comparison. Previous experiments had shown that the smooth object was not harder to find purely because of the presence of smooth edges and lack of half occlusions, but rather due to the interaction of the smooth edge with the extracting of depth for the disparity defined object (Chapters 4 and 5). Despite this, in our real world search task we observe a significant difference between SC0 and SC3, although the effect is smaller than the difference in reaction times between SC3 and SC14 (see Figure 9.10). We speculate that the significant difference between SC0 and SC3 is due to the effects of self-shadow concealment (see Section 2.1.1), where the smooth edges of SC3 have a decreased shadowing compared to the SC0 objects, as discussed in Section 9.3.3.

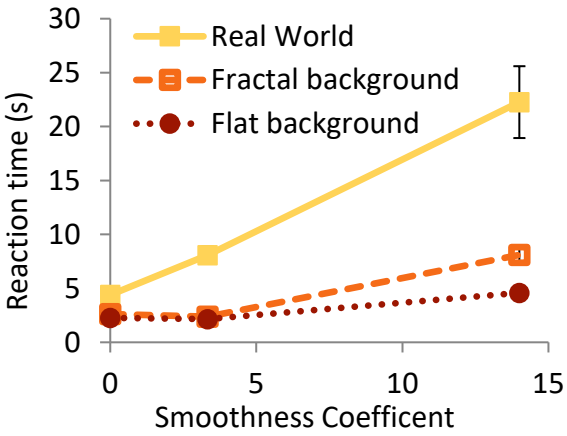


Figure 9.19: Group mean of the three visual search Experiments testing the effect of smoothness coefficient.

There is a very large increase in the reaction times between the virtual and the real world experiment as shown in Figure 9.19. This effect increased with increasing smoothness coefficient. An increase in reaction time itself is not surprising, as during pilots the real world experiment was deliberately made harder by the decrease of luminance in the room for three reasons:

1. With a long setup time and few total results, it is valuable to have longer trials, as this magnifies the absolute reaction time difference between conditions, making any differences easier to detect above noise.
2. The dim lighting obscured colour differences between the object and the sandpit, but made the object harder to detect.
3. Participants in the real world experiment made head movements often being made to properly search the full area, which may have slowed search down. Head movements were restricted in the virtual experiments.

However, these effects do not explain why the SC14 object has a much greater difference in detection times than the SC0 or SC3 objects – for the virtual experiment SC14 is only 3.4 times slower to detect than SC0, but it is 4.8 times harder in the real world.

One hypothesis is as to why smoothest object was so much harder to spot in the real world experiment relates to the shape of the background. It is possible that in addition to

increasing the difficulty segregating the object, the depth profile of the smooth object matched the background undulations of the real world experiment much better than the fractal background of Experiment 10. This would mean the object could exhibit a depth based background matching (Section 2.1.1), where the object is harder to detect because it matches elements of the background. Additionally, we know that camouflage is more effective in a visually complex environment (Merilaita, 2003) – it is possible that the increase in difficulty of detecting all objects is representative of the increased difficulty detecting a camouflaged object in its environment.

We also hypothesise that the smoothest object is so much harder to spot than the sharper objects due to the presence of Gestalt grouping cues. Until now, we have only been considering the effect of the smooth object as causing a difficulty in grouping due to binocular disparities – primarily due to good continuity between the smooth object and the background. However, we have not considered the effect of the reduced closure (see Figure 9.20, or Section 2.2.6 for a discussion) of the smooth object due to the smooth edge not having clear closure when compared to the sharp object, which has a distinct, closed edge. It may be that the reduction of good closure also applies to other cues that are now available in the real world experiment – such as differences in luminance, colour and shadowing. The object is likely being segregated by a combination of these cues (Johnston et al., 1994; Regan & Hamstra, 1994; Richards, 1977) – thus increasing the detection time over purely using binocular vision. If the smooth object is making all these cues less reliable at grouping the object, then we might expect a greater increase in detection times than when the object was being detected via one cue – as observed here.

4. Closure

Dots are perceived as a closed ring even when occluded.

Figure 9.20: Excerpt from Gestalt grouping in Figure 2.12

Despite the differences between the two experiments it is surprising how similar the results of the two experiments were, given the change in the number of cues to object detection, both to the shape of the object and monocular cues to its location. This is particularly remarkable given:

1. In the original experiments, the objects were viewed top-down; in the real world experiment they were viewed from an angle (32.5^{o} from vertical for an average observer). This gave the participants a different perspective, and would make object segregation different as the edge of the smooth object facing away from the observer would appear steeper than the edge facing the observer. Despite this, we see an effect of smoothness on the detection times.
2. There are monocular cues present in the real world experiment that are not present in the virtual experiment. As discussed earlier (Section 9.3.3.1), further investigation could be made by comparing reaction times with both eyes open to with one eye

shut to see how much binocular cues to perception affected the detection of the target. Some of the non-visual search experiments discussed in Section 9.3.3.2 could also inform us on the effect of monocular cues.

Despite all the complexities of creating a real world analogue of our virtual experiment, we think it is very interesting exercise, especially given our motivations are the study of camouflage – a real world phenomenon which has to interact with the many complex cues available in the environment. When making claims about the real world from RDS we agree that it is valuable to relate findings back to more realistic displays (Regan, n.d.; Richards, 1977). We would encourage future experimental designs to consider extending their research into a real world task, as it is an illuminating and rewarding experience which adds an extra degree of relevance to the experimental findings.

10 Summary

In this thesis, we examined Julesz's statement that "even under ideal monocular camouflage, the hidden objects jump out in depth when stereoscopically fused" (Julesz, 1971). While widely accepted as a use of stereoscopic or binocular vision, this hypothesis had not been rigorously tested. We investigated this observation, splitting it up into two sub-questions for detailed study:

1. What is a disparity defined object? What happens when the boundary between the background and the object is poorly defined?
2. Does binocular vision enable us to break camouflage? Are there ways in which prey animals can break binocular depth perception, thus reforming their camouflage?

In order to discuss these two individual strands, we use the format shown in the flowchart at the beginning of the thesis (Figure 1.3), but tabulated for clarity. For detailed discussions of the results for each Experiment and Chapter, see their individual discussion Sections. Numbers in brackets indicate the Sections relevant to each statement.

In the first three experimental Chapters we explored the first strand of this thesis: How does depth from binocular vision contribute to object perception? The results were illuminating, summarised here Chapter by Chapter:

Chapter	Expt	Focus	Results	Conclusions
4: What is a depth defined object?	1	Smoothed object boundary	An object with a smooth edge was perceived with less depth than the sharp object with a distinct edge (4.3.3).	The perception of depth at the centre of the object is influenced over a long range by the object's edges (discussed in 4.3.4).
	2	Half occlusions and boundaries	An object with half occlusions was perceived to have the same depth as an object with the same depth shape, but without half occlusions (4.4.2).	The difference in perception between the smooth and sharp edged objects could be due to the presence or absence of half occlusions (4.4.3).

5: How do we segregate a depth defined object from its background?	-	Modelling object segregation	A computational model that segregated out a square area of the stimulus as the 'object' and averaged over this area was able to fit participants' perceived peak depth in Experiments 1 and 2 (5.3 and 5.4).	The decrease in depth in the objects was probably due to a mechanism that averages over a large region of the segregated object, probably to give a consistent depth estimate for the object (5.3.45.4.4).
	3	Testing model predictions	A zero free-parameter model that created a prediction based on averaging over a region dependent on the size of the object (5.5.2.1) predicted the perceived peak depth of a variable plateau size object better (5.5.4) than a model with a fixed size averaging region (5.5.2.2).	The size of the segregation area was dependent on the size of the object, supporting the hypothesis that the visual system segregates the object from the background then averages over the object to improve the precision of the predicted depth (5.5.5). We speculate that cues to object segregation may alter the perceived peak depth of the object.
6: How does depth segregation interact with luminance segregation?	4 – 6	Luminance and well defined depth objects	These three experiments found that manipulating a luminance cue to segregation on the disparity defined objects made the task extremely difficult for participants (6.2,6.3 and 6.4).	With a good disparity cue and a poor luminance cue, we found no evidence that the luminance cue made affected the depth averaging mechanisms (6.4.4).
	7	Luminance and poorly defined depth	We added disparity noise to the smooth object, and found that the luminance cue to segregation altered the perceived depth of the object (6.5.3).	This strongly indicates that binocular and monocular cues are combined to segregate the object from its background, then the depth of the object is estimated via depth averaging (6.6).

In summary of the first strand, we found that that objects with a poorly defined depth edge were perceived with a decrease in peak depth. We found we could predict perceived peak depth using a model that segregated a region of the object based on the shape of the object itself, then averaged over this region. We investigated the interaction of binocular disparity with luminance, and found that a luminance cue to segregation altered the perceived peak depth of the object. This implies that object segregation occurs by combining multiple cues, before further processing of an object's properties such as its depth.

These Chapters present several new findings and techniques over those used in the current literature – indeed the results and methods of Chapters 4 and 5 have been published (Cammack & Harris, 2016). Typically, when investigating the perception of depth in an object, experiments look at highly abstract objects on featureless backgrounds e.g. (Deas & Wilcox, 2014, 2015; Kim, 2013). In Chapter 4, we investigate the effect of viewing objects with extended edges that join into the background, and find that this has an effect on the peak depth of the object. This effect seems to have links to the literature on Gestalt grouping (Elder & Zucker, 1993; Pizlo, Li, & Francis, 2005b; Wertheimer, 1923), with several similarities between our results and those of Deas and Wilcox (2014).

The results of Chapter 4 were not clearly explained by current models of disparity averaging, which could only explain disparity interaction on a short range, about 3-6 arcmin (Filippini & Banks, 2009; Harris et al., 1997). This implied that there was a disparity averaging mechanism taking place over longer scales than currently modelled. Given the body of literature that had previously found effects from Gestalt grouping that could not be explained by these models (Elder & Zucker, 1993; Pizlo et al., 2005b; Wertheimer, 1923), we thought that the long range effect may be due to further object based disparity processing.

Inspired by the words of Marr (1982), we created a novel computational model in Chapter 5 to study the overall computation used in the visual system when processing our object. This model found the results from Experiments 1,2 could be fitted by long range disparity averaging across a segregated region of the object – an effect not proposed before. We believe this model could inform development of current models of disparity extraction, e.g. (Allenmark & Read, 2010, 2011; Filippini & Banks, 2009; Goutcher & Hibbard, 2014). Using our results, these models could be extended to perform segregation after disparity extraction based on local changes in disparity such as inflection points. Subsequent averaging over the segregated region would create a model computationally similar to our model, but on algorithmic or hardware level. Comparing these new models to human performance could then inform us on how the brain may be processing the disparity signal for object perception.

There appeared to be a link between object shape and the area that was segregated and averaged – in Chapter 6 we investigate if other cues to object segregation, specifically luminance, could alter the area segregated and averaged. While other studies have investigated the effects of luminance and disparity e.g. (Anstis et al., 1977; Didyk, Ritschel,

Eisemann, Myszkowski, & Seidel, 2012; Lunn & Morgan, 1995), our work indicates that luminance and disparity cues to object processing may be implemented before disparity processing is complete.

The next three experimental Chapters explored the second strand of the thesis, considering how depth perception interacts with camouflage. We used different experimental techniques from the first strand but the experiments were informed by results and conclusions of object perception in the first strand.

Chapter	Expt	Focus	Results	Conclusions
7: Does depth perception break camouflage?	8	Does disparity decrease detection time?	A luminance defined target object was camouflaged amongst similar distractors (7.3.2). We found that the addition of disparity defined depth assisted with identifying the target, with up to a 20% decrease in reaction times (7.3.3).	Disparity defined depth gives a moderate improvement in detecting monocularly camouflaged targets (7.4).
8: Does poor segregation create stereoscopic camouflage?	9	Is a poorly defined object harder to spot?	Smooth objects were monocularly camouflaged in a RDS (8.2.1). We found that the smoother the edge of the object, the longer it took to detect (8.2.2).	An object with smooth edges takes longer to spot than one with sharp edges, potentially forming stereoscopic camouflage (8.3).
9: Does stereoscopic camouflage work in natural conditions?	10	Adding a natural background	We investigated if the addition of a naturalistic disparity noise background would make all objects equally hard to spot, as they would all be much harder to segregate from their background (9.2.1). The previous findings were upheld: that the more poorly defined the edge, the harder an object is to spot (9.2.3). Overall detection times were increased relative to Experiment 9 (9.2.4.2).	The smooth object is still harder to spot in naturalistic disparity backgrounds, indicating that smooth edges could be used to in the real world to create stereoscopic camouflage (9.2.4).

	11	Real world experiment	The previous experiment was used as the basis of a real world experiment (9.3.1). Here, we found that the smoother the object, the longer it took to detect (9.3.2). Unlike in previous experiments, low smoothness coefficient objects (relatively sharp edges) took longer to spot than objects with sharp edges (9.4).	Smooth objects are better camouflaged than sharper objects even in real world experiments (9.3.3). However, the addition of other cues not present in the computer experiment, such as shape from shading (9.3.3.1) seem to make a significant difference to detection time of objects with intermediate levels of smoothness (9.4).

In summary of the second strand, we found that objects that were defined by both luminance and disparity were faster to detect than objects defined by luminance alone, suggesting that depth from disparity could be used to find monocularly camouflaged objects. Objects with smoother edges took longer to detect than objects with sharper edges. This is most likely because segregation of a depth defined object from its background is not trivial and causes a decrease in the perception of depth in the smooth object, making it harder to detect. As a smoother object is harder to detect, this indicates that animals wishing to hide from stereoscopic observers could adapt their shape to make themselves harder to spot, thus forming camouflage. This camouflaging potential was robust, even under a unique real world analogue of our abstract computer experiments.

Since Julesz (1971) proposed the idea that binocular depth perception was used to break camouflage, it has often been assumed to be the case without formal testing e.g. (Heesy, 2009; Isbell, 2006). In Chapter 7, we establish that depth from disparity can and does assist in the detection of monocularly camouflaged objects, even when the observer is not aware of the disparity cue. However, the speedup in detection of camouflaged objects was not as dramatic as expected, and in Chapter 8 we investigated if the objects from the first strand may influence the detection time via the added difficulty of segregating poorly grouped objects (Deas & Wilcox, 2015; Elder & Zucker, 1993; Troscianko, Benton, Lovell, Tolhurst, & Pizlo, 2009b). We found the smooth objects took longer to detect, concurring with current literature and potentially forming a new kind of camouflage which we dub ‘stereoscopic’ or binocular camouflage (Cammack & Harris, 2016; Stevens & Merilaita, 2009).

Finally, in Chapter 9 we test our observations in more realistic environments to test if the results hold outside of abstract vision experiments (Richards, 1977). This is a very rare step, with only a few other studies relating results directly back to the real world e.g. (Foulsham et al., 2014; Frisby et al., 1996; McKee & Taylor, 2010; Welchman et al., 2004) and even

rarer in visual search e.g. (Foulsham et al., 2014; Smith et al., 2008). We successfully create a real world Experiment that replicates some of our results, but shows some interesting new effects, such as the potential of an interaction between binocular camouflage and self-shadow concealment (Penacchio et al., 2016; Penacchio, Lovell, Sanghera, et al., 2015; Penacchio, Lovell, Cuthill, et al., 2015; Rowland, 2009; Ruxton et al., 2004). We hope that the methods and success of this real world experiment will inspire others to relate their virtual experiments back to the real world.

In conclusion, we find that Julesz's claim that "even under ideal monocular camouflage, the hidden objects jump out in depth when stereoscopically fused" (Julesz (Julesz, 1971), p145) should be viewed with caution. We find that depth assists with the detection of camouflaged objects, but does not make detection trivial. Additionally, the profile of the object can make a binocularly defined object hard to segregate from the background, making it difficult to detect. This potentially forms a kind of stereoscopic camouflage that, in the real world could be complemented by other camouflage techniques such as self-shadow concealment and disruptive colouration.

Acknowledgements

I would like to thank these people for their invaluable contributions:

- This thesis would not have been possible without Julie Harris, my supervisor, for her invaluable support, advice, knowledge and guidance through the three years of this PhD. Also for her endless patience with my dyslexic writing.
- Tomas Otto, my second supervisor, for his guidance, and his support and advice giving me the confidence to make my presentation and graphical skills what they are today. I now enjoy giving presentations!
- Helen Cammack, my wife. Without her moral support I would never have made it here.
- All the lab members for their advice, ideas and insight, including:
    - Ana Porskalaite, a final year project student who took Experiment 8 as her final year project, piloted, tweaked and collected all the data presented here.
    - Andrew Chua, a summer student for spending weeks sorting out the details of Experiment 11, running and collecting the data presented here.
    - Andrew Mackenzie, for all the wisdom passed down from the heights of a PhD student a year above me. Also for all the long discussions of experiments at conferences, many of which are presented in this thesis.
    - Benjamin Portelli, Abigail Lee and Andrews Mackenzie, for all the patient hours of constructive criticism and piloting.
    - Olivier Penacchio, for his support and insights into modelling and camouflage.
    - The suggestions and advice of Martin Giesel.
    - Xavier Otzau, for his interest and attempts at creating a detailed neuronal model of my stimulus.
- All the vision lab members, your insight and contributions around the department and especially at Journal Club let me avoid so many pitfalls. A special callout to Giedre Zlatkute for the idea of using sand. Also to Christopher Gillespie for his advice and predictions – he was more right about the third year of the PhD than I ever thought possible.
- Everyone at the School office and Mary Latimer and Helen Sunderland, who were always there to help in any way they could.
- Sheila Baillie, the NHS and Eric Bowman, for their help and advice dealing with Dyslexia and APD. You made it so much easier to write this.
- The discussions and ideas everyone who talked to me at conferences, namely VSS annual general meeting (2015), SVG (2014,2015 and 2016) and AVA annual general meeting (2014).
- Everyone in my PhD office, and all the invaluable hours of discussions, piloting and sympathy.

- My friends and family, for their support, advice and a sympathetic ear and a hot chocolate whenever things went pear-shaped.

Without the support of these bodies, this research could never have been conducted:

- EPSRC for funding my 600th Anniversary scholarship, equipment and other costs.
- University of St Andrews and the School of Psychology and Neuroscience, for funding, office and laboratory space.
- The BBSRC for funding.
- Grants from EPS, CAPOD, the School of Psychology and Neuroscience at the University of St Andrews and the HPC consortium.

These programs made large contributions to this thesis:

- SPSS statistics, for most statistical analysis ("IBM SPSS Statistics for Windows," 2013).
- MATLAB, for most data analysis and experimental programming ("MATLAB and Statistics Toolbox," n.d.).
    - Psychophysics toolbox, for display of stimuli (Brainard, 1997; Kleiner et al., 2007).
    - Palamedes toolbox for psychometric function fitting (Prins & Kingdom, 2009).
- Wolfram Mathematica, for integration and modelling ("Mathematica," n.d.).
- Stereophotomaker, for the creation of the anaglyphs presented in this thesis (Suto & Sykes, 2014)

Glossary

Term	Definition
2AFC	Two alternative force choice – an experimental paradigm where the participant has to choose between two alternative options.
Background matching	A camouflage technique where the patterning, colours and other features of the animal resemble the background so well that the animal cannot be distinguished from the background, see Section 2.1.1.
Binocular	Visible or viewed with both eyes simultaneously. See Section 2.2.1.
Binocular vision	Using two eyes simultaneously to see the same area of visual field. See Section 2.2.1.
Detection time	The length of time it takes an animal to find a camouflaged object, be it predator or prey. This is only loosely defined in literature, but we mean the time between the start of a trial and the participant indicating they have found the object.
Disparity	The angular difference in the position between the images of an object in each of the two eyes. We always mean relative crossed disparity in this thesis unless otherwise stated, see Section 2.2.2.
Disruptive Colouration	A camouflage technique where an animal uses bold contours or colours to appear to several different unrelated segments or objects, and thus not be identified as an animal, see Section 2.1.1.
Cryptic	An animal that is sufficiently well camouflaged via background matching so as to appear not to be present, see Section 2.1.1.
Half-occlusions	Regions of a scene that are visible by only one eye (monocularly) due to an opaque surface obscuring the view of the other eye. See Section 2.2.3.
Luminance	The intensity of light emitted from the surface per unit area.
Masquerade	A camouflage technique where the animal attempts to appear to be something in the environment that is not of interest to the observer, e.g. Figure 7.1. See Section 2.1.1.
Mimicry	A camouflage technique where an animal attempts to appear to be a different animal, typically one that is well defended or not worth predating. See Section 2.1.12.1.
Model	A computer program or equation designed to emulate or predict the behaviour of a real-life system.
Monocular	Visible or viewed only with one eye.
PSE	Point of subjective equality – the numerical value when the participant cannot perceive the difference between the displayed stimuli, even though they may be displayed as being different. See Section 3.2.1.
Plateau size	A value determining the distance between the inflection points of the smooth object. This is also equal to the full width half maximum of the function. See Section 4.2.1.
Quitting threshold	The length of time a participant will search for the target before giving up. See Section 2.1.2.
Real world experiment	An experiment where the visual stimulus is physically present, rather than displayed on a computer screen.

RDS	Random dot stereogram – a display consisting of a randomly distributed set dots, such that when viewed stereoscopically (usually using apparatus such as a stereoscope) depth can be perceived within the stereogram. See Section 3.1.
Reaction time	The length of time it takes for the participant to indicate they think they have detected the target. Unless otherwise stated, this is only analysed for correct trials. See Section 3.2.2.
SC	See smoothness coefficient
Self-Shadow concealment	A camouflage technique where an animal obscures cues to its shape and location by decreasing the visibility of shadowing caused by the animal. See Section 2.1.1.
Segregation	Isolating a region of space as being different, in this thesis it typically refers to where an object is perceived as being different to the background.
Smoothness coefficient	A value determining the first (gradient) and second differential of the object shape – perceptually this makes the object seem smoother or sharper. A higher number is considered to be smoother. See Section 4.2.1
Stereoscopic camouflage	A shape or patterning designed to impair binocular vision in the observer, rendering it harder to spot the animal or object using depth cues.
Stereopsis	A strong impression of depth from the scene, but not necessarily using two eyes. Julesz used stereopsis to refer to what we call here binocular vision.
(Wheatstone) Stereoscope	A device consisting of four mirrors used with a computer screen to display depth. See Section 3.1.

References

©Ta-graphy. (2010). The leafy sea dragon, Phycodurus eques, is a marine fish in the family Syngnathidae. *CC BY 3.0*. Retrieved from https://commons.wikimedia.org/wiki/File:Leafy_sea_dragon_by_Ta-graphy.jpg

3DPrintUK. (2016). 3D printer specifications. Retrieved January 1, 2016, from https://www.3dprint-uk.co.uk/machines-maximums-and-minimums/

Abrams, P. A. (2000). The Evolution of Predator-Prey Interactions : Theory and Evidence. *Annual Review of Ecology and Systematics*, *31*, 79–105.

Akerstrom, R. A., & Todd, J. T. (1988). The perception of stereoscopic transparency. *Perception & Psychophysics*, *44*(5), 421–432.

Allenmark, F., & Read, J. C. A. (2010). Detectability of sine- versus square-wave disparity gratings : A challenge for current models of depth perception. *Journal of Vision*, *10*, 1–16. http://doi.org/10.1167/10.8.17.Introduction

Allenmark, F., & Read, J. C. A. (2011). Spatial stereoresolution for depth corrugations may be set in primary visual cortex. *PLoS Computational Biology*, *7*(8), 1–23. http://doi.org/10.1371/journal.pcbi.1002142

Allison, R. S., Gillam, B. J., & Vecellio, E. (2009). Binocular depth discrimination and estimation beyond interaction space. *Journal of Vision*, *9*(1), 10.1-14. http://doi.org/10.1167/9.1.10

Anderson, B. L., & Nakayama, K. (1994). Toward a general theory of stereopsis: Binocular matching, occluding contours, and fusion. *Psychological Review*, *101*(3), 414–445. http://doi.org/10.1037/0033-295X.101.3.414

Andrae, R., Schulze-Hartung, T., & Melchior, P. (2010). Dos and don'ts of reduced chi-squared, 1–12.

Anstis, S. M., Howard, I. P., & Rodgers, B. (1977). A Craik-O'Brien-Cornsweet illusion for visual depth. *Vision Research*, *18*(2), 213–217.

Anzai, A., Chowdhury, S. A., & DeAngelis, G. C. (2011). Coding of stereoscopic depth information in visual areas V3 and V3A. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *31*(28), 10270–10282. http://doi.org/10.1523/JNEUROSCI.5956-10.2011

Anzai, A., & DeAngelis, G. C. (2010). Neural computations underlying depth

perception. *Current Opinion in Neurobiology*, *20*(3), 367–375. http://doi.org/10.1016/j.conb.2010.04.006

Auswandern;, W. (2007). Euthalia aconthea on Mango Leaf. *CC BY 2.0*. Retrieved from https://www.flickr.com/photos/wohinauswandern/4146066451

Backus, B. T., Banks, M. S., Van Ee, R., & Crowell, J. A. (1999). Horizontal and vertical disparity, eye position, and stereoscopic slant perception. *Vision Research*, *39*(6), 1143–1170. http://doi.org/10.1016/S0042-6989(98)00139-4

Baliga, V. (2014). Big and small - Moth Camouflage. *CC BY 2.0*. Retrieved from https://www.flickr.com/photos/vipinbaliga/13918606197

Banks, M. S., Gepshtein, S., & Landy, M. S. (2004). Why is spatial stereoresolution so low? *The Journal of Neuroscience*, *24*(9), 2077–2089. http://doi.org/10.1523/JNEUROSCI.3852-02.2004

Barlow, H. B., Blakemore, C., & Pettigrew, J. D. (1967). The neural mechanism of binocular depth discrimination. *The Journal of Physiology*, *193*(2), 327–342. http://doi.org/10.1113/jphysiol.1967.sp008360

Beck, D. M., & Palmer, S. E. (2002). Top-down influences on perceptual grouping. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(5), 1071.

Becker, S., Bowd, C., Shorter, S., King, K., & Patterson, R. (1999). Occlusion contributes to temporal processing differences between crossed and uncrossed stereopsis in random-dot displays. *Vision Research*, *39*(2), 331–9.

Bertamini, M., & Lawson, R. (2008). Rapid figure - Ground responses to stereograms reveal an advantage for a convex foreground. *Perception*, *37*(4), 483–494. http://doi.org/10.1068/p5728

Bian, Z., & Andersen, G. J. (2010). The advantage of a ground surface in the representation of visual scenes. *Jounal of Vision*, *10*(8), 1–19. http://doi.org/10.1167/10.8.16.Introduction

Blake, R., & Wilson, H. (2011). Binocular vision. *Vision Research*, *51*(7), 754–70. http://doi.org/10.1016/j.visres.2010.10.009

Blakemore, B. C. (1970). The range and scope of binocular depth perception in man. *Journal of Physiology*, *211*(3), 599–622.

Bond, A. B., & Kamil, A. C. (1998). Apostatic selection by blue jays produces balanced polymorphism in virtualprey. *Letters to Nature*, *395*(October),

594–596.

Bond, A. B., & Kamil, A. C. (2002). Visual predators select for crypticity and polymorphism in virtual prey. *Nature*, *415*(6872), 609–613. http://doi.org/10.1038/415609a

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436.

Bredfeldt, C. E., & Cumming, B. G. (2006). A Simple Account of Cyclopean Edge Responses in Macaque V2. *Journal of Neuroscience*, *26*(29), 7581–7596. http://doi.org/10.1523/JNEUROSCI.5308-05.2006

Bredfeldt, C. E., Read, J. C. A., & Cumming, B. G. (2008). A Quantitative Explanation of Responses to Disparity-Defined Edges in Macaque V2. *Journal of Neurophysiology*, *101*(2), 701–713. http://doi.org/10.1152/jn.00729.2007

Brenner, E., & Landy, M. S. (1999). Interaction between the percieved shape of two objects. *Vision Res*, *39*, 3843–3848.

Brenner, E., Smeets, J. B. J., & Landy, M. S. (2001). How vertical disparities assist judgements of distance. *Vision Research*, *41*(25–26), 3455–3465. http://doi.org/10.1016/S0042-6989(01)00206-1

Briscoe, A. D., & Chittka, L. (2001). The evolution of color vision in insects. *Annual Review of Entomology*, *46*, 471–510.

Brosilow, B. J., Ziff, R. M., & Vigil, R. D. (1991). Random sequential adsorption of parallel squares. *Physical Review A*, *43*(2), 631–638.

Bülthoff, H. H., & Mallot, H. A. (1988). Integration of depth modules: stereo and shading. *Journal of the Optical Society of America A*, *5*(10), 1749–1758. http://doi.org/10.1364/JOSAA.5.001749

Burg, E. Van Der, & Alais, D. (2015). Evolving the stimulus to fit the brain : A genetic algorithm reveals the brain ' s feature priorities in visual search, *15*, 1–15. http://doi.org/10.1167/15.2.8.doi

Burge, J., Peterson, M. a, & Palmer, S. E. (2005). Ordinal configural cues combine with metric disparity in depth perception. *Journal of Vision*, *5*(6), 534–542. http://doi.org/10.1167/5.6.5

Burt, P., & Julesz, B. (1980). A disparity gradient limit for binocular fusion. *Science*, *208*(4444), 615–617.

Cain, M. S., Adamo, S. H., & Mitroff, S. R. (2013). A taxonomy of errors in multiple-target visual search. *Visual Cognition*, *21*(7), 899–921. http://doi.org/10.1080/13506285.2013.843627

Cammack, P., & Harris, J. M. (2016). Depth perception in disparity-defined objects : finding the balance between averaging and segregation. *Philosophical Transactions of the Royal Society B*, *8*(371). http://doi.org/10.1098/rstb.2015.0258

Caro, T., Izzo, A., Reiner, R. C., Walker, H., & Stankowich, T. (2014). The function of zebra stripes. *Nature Communications*, *5*, 3535. http://doi.org/10.1038/ncomms4535

Catania, K. C. (2009). Tentacled snakes turn C-starts to their advantage and predict future prey behavior. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(27), 11183–11187. http://doi.org/10.1073/pnas.0905183106

Cave, K. R., & Wolfe, J. M. (1990). Modeling the role of parallel processing in visual search. *Cognitive Psychology*, *22*(2), 225–271. http://doi.org/http://dx.doi.org/10.1016/0010-0285(90)90017-X

Chandrasekaran, C., Canon, V., Dahmen, J. C., Kourtzi, Z., & Welchman, A. E. (2007). Neural correlates of disparity-defined shape discrimination in the human brain. *Journal of Neurophysiology*, *97*(2), 1553–1565. http://doi.org/10.1152/jn.01074.2006

Chen, C., & Tyler, C. W. (2015). Shading Beats Binocular Disparity in Depth from Luminance Gradients: Evidence against a Maximum Likelihood Principle for Cue Combination. *Plos One*, *10*(8). http://doi.org/10.1371/journal.pone.0132658

Chen, Y., & Qian, N. (2004). A Coarse-to-Fine Disparity Energy Model with Both Phase-Shift and Position-Shift Receptive Field Mechanisms. *Neural Computation*, *1577*, 1545–1577.

Churchland, P. S., & Sejnowski, J. (1988). Perspectives on Cognitive Neuroscience. *Science*, *242*, 741–745.

Clarke, K. U., & Sardesai, J. B. (1959). An Analysis of the Effects of Temperature upon the Growth and Reproduction of Dysdercus fasciatus Sign. (Hemiptera, Pyrrhocoridae). I.—The Intrinsic Rate of Increase. *Bulletin of Entomological Research*, *50*(2), 387–405.

Collett, T. S. (1985). Extrapolating and Interpolating Surfaces in Depth. *Proceedings of the Royal Society B*, *244*(1234), 43–56. http://doi.org/10.1098/rspb.1985.0020

Cooper, J. M., & Allen, J. A. (1994). Selection by wild birds on artificial dimorphic prey on varied backgrounds. *Biological Journal of the Linnean Society*, *51*(4), 433–446. http://doi.org/10.1111/j.1095-8312.1994.tb00972.x

Cornsweet, T. (1970). *Visual Perception*. Academic Press, Inc.

Cott, H. B. (1940). *Adaptive coloration in animals*. Methuen and Co; London.

Craik, K. J. W. (1966). *The nature of psychology: a selection of papers, essays, and other writings*. (S. L. Sherwood, Ed.). Cambridge University Press.

Cumming, B. G., & DeAngelis, G. C. (2001). The Psyiology of Steriopsis. *Annual Review of Neuroscience*, 203–38.

Cumming, B. G., Johnston, E. B., & Parker, A. J. (1991). Vertical disparities and perception of three-dimensional shape. *Nature*, *354*, 411–413.

Cumming, B. G., & Parker, a J. (1997). Responses of primary visual cortical neurons to binocular disparity without depth perception. *Nature*, *389*(6648), 280–3. http://doi.org/10.1038/38487

Cumming, B. G., Shapiro, S. E., & Parker, a J. (1998). Disparity detection in anticorrelated stereograms. *Perception*, *27*(11), 1367–77.

Cuthill, I. C., Stevens, M., Sheppard, J., Maddocks, T., Parraga, C. A., & Troscianko, T. S. (2005). Disruptive coloration and background pattern matching. *Letters to Nature*, *434*(March), 72–74. http://doi.org/10.1038/nature03325.1.

Darwin, C. (1859). *On the Origin of Species*.

Dawkins, M. (1971). Perceptual changes in chicks: Another look at the "search image" concept. *Animal Behaviour*, *19*(3), 566–574. http://doi.org/http://dx.doi.org/10.1016/S0003-3472(71)80113-6

Dawkins, R., & Krebs, J. R. (1979). Arms Races between and within Species. *Proceedings of the Royal Society B*, *205*(1161), 489–511.

de la Rosa, S., Moraglia, G., & Schneider, B. A. (2008). The magnitude of binocular disparity modulates search time for targets defined by a conjunction of depth and colour. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*. Educational Publishing Foundation. http://doi.org/10.1037/1196-1961.62.3.150

DeAngelis, G. C. (2000). Seeing in three dimensions : the neurophysiology of stereopsis. *Trends in Cognitive Sciences*, *4*(3), 80–90.

DeAngelis, G. C., Ohzawa, I., & Freeman, R. D. (1991). Depth is encoded in the visual cortex by a specialized receptive field structure. *Nature*. http://doi.org/10.1038/352156a0

Deas, L. M., & Wilcox, L. M. (2014). Gestalt grouping via closure degrades suprathreshold depth percepts. *Journal of Vision*, *14*(9), 1–13.

http://doi.org/10.1167/14.9.14

Deas, L. M., & Wilcox, L. M. (2015). Perceptual grouping via binocular disparity: The impact of stereoscopic good continuation. *J Vis*, *15*(11), 11. http://doi.org/10.1167/15.11.11

Didyk, P., Ritschel, T., Eisemann, E., Myszkowski, K., & Seidel, H.-P. (2012). Apparent stereo: the Cornsweet illusion can enhance perceived depth. *Proc. SPIE*.

Didyk, P., Ritschel, T., Eisemann, E., Myszkowski, K., Seidel, H.-P., & Matusik, W. (2012). A luminance-contrast-aware disparity model and applications. *ACM Transactions on Graphics*, *31*(6), 1. http://doi.org/10.1145/2366145.2366203

Dobias, J. J., & Papathomas, T. V. (2013). Recovering 3-D shape: Roles of absolute and relative disparity, retinal size, and viewing distance as studied with reverse-perspective stimuli. *Perception*, *42*(4), 430–446. http://doi.org/10.1068/p7409

Dukas, R., & Kamil, A. C. (2001). Limited attention: The constraint underlying search image. *Behavioral Ecology*, *12*(2), 192–199. http://doi.org/10.1093/beheco/12.2.192

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

Edmunds, J., & Edmunds, M. (1974). Polymorphic Mimicry and Natural Selection : A Reappraisal. *Evolution*, *28*(3), 402–407. Retrieved from http://www.jstor.org/stable/2407161

Egri, A., Blahó, M., Kriska, G., Farkas, R., Gyurkovszky, M., Akesson, S., & Horváth, G. (2012). Polarotactic tabanids find striped patterns with brightness and/or polarization modulation least attractive: an advantage of zebra stripes. *The Journal of Experimental Biology*, *215*(Pt 5), 736–45. http://doi.org/10.1242/jeb.065540

Elder, J., & Zucker, S. (1993). The effect of contour closure on the rapid discrimination of two-dimensional shapes. *Vision Research*, *33*(7), 981–991. http://doi.org/10.1016/0042-6989(93)90080-G

Endler, J. A. (1981). An overview of the relationships between mimicry and crypsis. *Biological Journal of the Linnean Society*, *16*(1), 25–31. http://doi.org/10.1111/j.1095-8312.1981.tb01840.x

Endler, J. A. (1984). Progressive background in moths, and a quantitative measure of crypsis. *Biological Journal of the Linnean Society*, *22*(3),

187–231. http://doi.org/10.1111/j.1095-8312.1984.tb01677.x

Endler, J. A. (2006). Disruptive and cryptic coloration. *Proceedings of the Royal Society B*, *273*(1600), 2425–2426. http://doi.org/10.1038/nature03312

Fahle, M., & Westheimer, G. (1988). Local and global factors in disparity detection of rows of points. *Vision Research*, *28*(1), 171–178. http://doi.org/http://dx.doi.org/10.1016/S0042-6989(88)80016-6

Feltmate, B. W., & Williams, D. D. (1989). A test of crypsis and predator avoidance in the stonefly Paragnetina media (Plecoptera: Perlidae). *Animal Behaviour*, *37*, *Part 6*, 992–999. http://doi.org/http://dx.doi.org/10.1016/0003-3472(89)90143-7

Fibonacci. (2005). Modified from Cornsweet Illusion. *CC BY-SA 3.0*. Retrieved from https://commons.wikimedia.org/w/index.php?curid=1788932

Filippini, H. R., & Banks, M. S. (2009). Limits of stereopsis explained by local cross-correlation. *Journal of Vision*, *9*(8), 1–18. http://doi.org/10.1167/9.1.8

Finlayson, N. J., Remington, R. W., Retell, J. D., & Grove, P. M. (2013). Segmentation by depth does not always facilitate visual search. *Journal of Vision*, *13*, 1–14. http://doi.org/10.1167/13.8.11.doi

Fleet, D. J., Wagner, H., & Heeger, D. J. (1996). Neural encoding of binocular disparity: energy models, position shifts and phase shifts. *Vision Research*, *36*(12), 1839–57. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8759452

Formankiewicz, M. A., & Mollon, J. D. (2009). The psychophysics of detecting binocular discrepancies of luminance. *Vision Research*, *49*(15), 1929–1938. http://doi.org/10.1016/j.visres.2009.05.001

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. http://doi.org/10.1037/cep0000004

Frisby, J. P., Buckley, D., & Duke, P. A. (1996). Evidence for good recovery of lengths of real objects seen with natural stereo viewing. *Perception*, *25*(2), 129–154.

Gantz, L., & Bedell, H. E. (2012). Variation of Stereothreshold with Random-Dot Stereogram Density. *Optom Vis Sci*, *88*(9), 1066–1071.

http://doi.org/10.1097/OPX.0b013e3182217487.Variation

Garding, J., Porrill, J., Mayhew, J. E. W., & Frisby, J. P. (1995). Stereopsis, vertical disparity and relief transformations. *Vision Research*, *35*(5), 703–722. http://doi.org/10.1016/0042-6989(94)00162-F

Geisler, W. S., & Diehl, R. L. (2002). Bayesian natural selection and the evolution of perceptual systems. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *357*, 419–448. http://doi.org/10.1098/rstb.2001.1055

Geisler, W. S., & Diehl, R. L. (2003). A Bayesian approach to the evolution of perceptual and cognitive systems. *Cognitive Science*, *27*(3), 379–402. http://doi.org/http://dx.doi.org/10.1016/S0364-0213(03)00009-0

Gendron, R. P., & Staddon, J. E. R. (1983). Searching for Cryptic Prey: The Effect of Search Rate. *The American Naturalist*, *121*(2), 172–186. Retrieved from http://www.jstor.org/stable/2461121

Georgeson, M. A., Yates, T. A., & Schofield, A. J. (2009). Depth propagation and surface construction in 3-D vision. *Vision Research*, *49*(1), 84–95. http://doi.org/10.1016/j.visres.2008.09.030

Georgieva, S., Peeters, R., Kolster, H., Todd, J. T., & Orban, G. a. (2009). The processing of three-dimensional shape from disparity in the human brain. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *29*(3), 727–742. http://doi.org/10.1523/JNEUROSCI.4753-08.2009

Gilchrist, I. D., & Harvey, M. (2006). Evidence for a systematic component within scan paths in visual search. *Visual Cognition*, *14*(4–8), 704–715. http://doi.org/10.1080/13506280500193719

Gilchrist, I. D., North, A., & Hood, B. (2001). Is visual search really like foraging? *Perception*, *30*(12), 1459–1464. http://doi.org/10.1068/p3249

Gillam, B. (1968). Perception of slant when perspective and stereopsis conflict: Experiments with aniseikonic lenses. *Journal of Experimental Psychology*, *78*, 299–305.

Gillam, B., Blackburn, S., & Brooks, K. R. (2007). Hinge versus twist: The effects of "reference surfaces" and discontinuities on stereoscopic slant perception. *Perception*, *36*(4), 596–616.

Gillam, B., & Ryan, C. (1992). Perspective, orientation disparity, and anisotropy in stereoscopic slant perception. *Perception*, *21*(4), 427–39. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1437462

Glennerster, A., & McKee, S. P. (1999). Bias and sensitivity of stereo judgements in the presence of a slanted reference plane. *Vision Research*, *39*(18), 3057–69. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10664804

Glennerster, A., Rogers, B. J., & Bradshaw, M. F. (1996). Stereoscopic Depth Constancy Depends on the Subject's Task. *Vision Research*, *36*(21), 3441–3456.

Godwin, H. J., Menneer, T., Riggs, C. a., Cave, K. R., & Donnelly, N. (2014). Perceptual failures in the selection and identification of low-prevalence targets in relative prevalence visual search. *Attention, Perception, & Psychophysics*, *77*(1), 150–159. http://doi.org/10.3758/s13414-014-0762-8

Goutcher, R., & Hibbard, P. B. (2010). Evidence for relative disparity matching in the perception of an ambiguous stereogram. *Journal of Vision*, *10*, 1–16. http://doi.org/10.1167/10.12.35.Introduction

Goutcher, R., & Hibbard, P. B. (2014). Mechanisms for similarity matching in disparity measurement. *Frontiers in Psychology*, *4*, 1–11. http://doi.org/10.3389/fpsyg.2013.01014

Goutcher, R., & Mamassian, P. (2002). A ground plane preference for stereoscopic slant. *Perception ECVP Abstract*, *31*.

Goutcher, R., & Mamassian, P. (2005). Selective biasing of stereo correspondence in an ambiguous stereogram. *Vision Research*, *45*(4), 469–483. http://doi.org/10.1016/j.visres.2004.08.025

Griffiths, T. L., Lieder, F., & Goodman, N. D. (2015). Rational use of cognitive resources: Levels of analysis between the computational and the algorithmic. *Topics in Cognitive Science*, *7*(2), 217–229. http://doi.org/10.1111/tops.12142

Harris, J. M. (2004). Binocular vision : moving closer to reality. *Philosophical Transactions of the Royal Society London A*, *10*(1098).

Harris, J. M. (2010). Monocular zones in stereoscopic scenes: a useful source of information for human binocular vision? *Proc. SPIE*, *7524*. http://doi.org/10.1117/12.837465

Harris, J. M. (2014). Volume perception: disparity extraction and depth representation in complex three-dimensional enviroments. *Journal of Vision*, *14*(12), 1–25.

Harris, J. M., Chopin, A., Zeiner, K., & Hibbard, P. B. (2012). Perception of relative depth interval: systematic biases in perceived depth. *Quarterly*

Journal of Experimental Psychology, *65*(1), 73–91. http://doi.org/10.1080/17470218.2011.589520

Harris, J. M., McKee, S. P., & Smallman, H. S. (1997). Fine-scale processing in human binocular stereopsis. *Journal of the Optical Society of America A*, *14*(8), 1673–1683.

Harris, J. M., & Parker, A. J. (1992). Efficiency of stereopsis in random-dot stereograms. *Journal of the Optical Society of America A*, *9*(1), 14–24. http://doi.org/10.1364/JOSAA.9.000014

Harris, J. M., & Wilcox, L. M. (2009). The role of monocularly visible regions in depth and surface perception. *Vision Research*, *49*(22), 2666–2685. http://doi.org/10.1016/j.visres.2009.06.021

Harris, J. M., & Willis, A. (2001). A binocular site for contrast-modulated masking. *Vision Research*, *41*(7), 873–881. http://doi.org/10.1016/S0042-6989(00)00292-3

Harris, J. P., & Gregory, R. L. (1973). Fusion and rivalry of illusory contours. *Perception*, *2*(2), 235–247. http://doi.org/10.1068/p020235

Harrison, J. J. (2008). Altered from Asiatic hybrid lilium stereogram. *CC BY SA 3.0*. Retrieved from https://commons.wikimedia.org/wiki/File:Asiatic_hybrid_lilium_stereogram.jpg

Heesy, C. P. (2009). Seeing in stereo: The ecology and evolution of primate binocular vision and stereopsis. *Evolutionary Anthropology*, *18*(1), 21–35. http://doi.org/10.1002/evan.20195

Heiling, A. M., Herberstein, M. E., & Chittka, L. (2003). Pollinator attraction: Crab-spiders manipulate flower signals. *Nature*, *421*(6921), 334. http://doi.org/10.1038/421334a

Heron, S., & Lages, M. (2012). Screening and sampling in studies of binocular vision. *Vision Research*, *62*, 228–234. http://doi.org/10.1016/j.visres.2012.04.012

Hibbard, P. B., & Bouzit, S. (2005). Stereoscopic correspondence for ambiguous targets is affected by elevation and fixation distance. *Spatial Vision2*, *18*(4), 399–411.

Hillis, J. M., Watt, S. J., Landy, M. S., & Banks, M. S. (2004). Slant from texture and disparity cues: optimal cue combination. *Journal of Vision*, *4*(12), 967–992. http://doi.org/10.1167/4.12.1

Hobgood, N. (2005). Soft coral crab. *CC BY SA 3.0*. Retrieved from https://commons.wikimedia.org/wiki/File:Hoplophrys_oatesi.jpg

How, M. J., & Zanker, J. M. (2013). Motion camouflage induced by zebra stripes. *Zoology*, 1–8. http://doi.org/10.1016/j.zool.2013.10.004

Howard, I. P. (2002). *Seeing in depth. Vol. 1 Basic Mechanisms*.

Howard, I. P., & Rogers, B. (2002). *Perciveing in Depth, Volume 2: Stereoscopic Vision*.

Huber, G. R. (2004). Blue Ridge Pastures. *CC BY SA 2.5*.

IBM SPSS Statistics for Windows. (2013). Armonk, NY: IBM Corp.

Isbell, L. a. (2006). Snakes as agents of evolutionary change in primate brains. *Journal of Human Evolution*, *51*(1), 1–35. http://doi.org/10.1016/j.jhevol.2005.12.012

Iwaniuk, A. N., & Wylie, D. R. W. (2006). The evolution of stereopsis and the Wulst in caprimulgiform birds: A comparative analysis. *Journal of Comparative Physiology A*, *192*(12), 1313–26. http://doi.org/10.1007/s00359-006-0161-2

Johannesson, K., & Ekendahl, A. (2002). Selective predation favouring cryptic individuals of marine snails (Littorina). *Biological Journal of the Linnean Society*, *76*(1), 137–144. http://doi.org/10.1111/j.1095-8312.2002.tb01720.x

Johnsen, S. (2002). Cryptic and conspicuous coloration in the pelagic environment. *Proceedings. Biological Sciences / The Royal Society*, *269*(1488), 243–56. http://doi.org/10.1098/rspb.2001.1855

Johnsen, S. (2014). Hide and seek in the open sea: pelagic camouflage and visual countermeasures. *Annual Review of Marine Science*, *6*(August 2013), 369–92. http://doi.org/10.1146/annurev-marine-010213-135018

Johnsen, S., & Sosik, H. M. (2003). Cryptic coloration and mirrored sides as camouflage strategies in near-surface pelagic habitats: Implications for foraging and predator avoidance. *American Society of Limnology and Oceanography*, *48*(3), 1277–1288.

Johnston, E. B. (1991). Systematic distortions of shape from stereopsis. *Vision Research*, *31*(7), 1351–1360.

Johnston, E. B., Cumming, B. G., & Landy, M. S. (1994). Integration of stereopsis and motion shape cues. *Vision Research*, *34*(17), 2259–2275.

Jones, J. P., & Palmer, L. A. (1987). An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex. *Journal of Neurophysiology*, *58*(6), 1233–1258. Retrieved from

http://jn.physiology.org/content/58/6/1233.abstract

Jones, R. K., & Lee, D. N. (1981). Why two eyes are better than one: The two views of binocular vision. *Journal of Experimental Psychology: Human Perception and Performance*, *7*(1), 30–40.

Joshua, D. E., & Bishop, P. O. (1970). Binocular single vision and depth discrimination. Receptive field disparities for central and peripheral vision and binocular interaction on peripheral single units in cat striate cortex. *Experimental Brain Research*, *10*(4), 389–416. http://doi.org/10.1007/BF02324766

Julesz, B. (1971). *Foundations of cyclopean perception*. The University of Chicago Press.

Kaisl, T. (2007). Herd of Zebras. *CC BY-NC-ND 2.0*. Retrieved from https://www.flickr.com/photos/scirpus/5331355068

Kane, D., Guan, P., & Banks, M. S. (2014). The limits of human stereopsis in space and time. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *34*(4), 1397–1408. http://doi.org/10.1523/JNEUROSCI.1652-13.2014

Kaufman, L., Bacon, J., & Barroso, F. (1973). Stereopsis without image segregation. *Vision Research1*, *13*(1), 137–147.

Kiltie, R. A., & Laine, A. F. (1992). Visual textures, machine vision and animal camouflage. *Trends in Ecology & Evolution*, *7*(5), 163–166.

Kim, D. I. (2013). The detection and representation of foreground vs . background objects. *Wellesley College Digital Scholarship and Archive*.

Kingdom, F. A. A., & Prins, N. (2009). *Psychophysics: A Practical Introduction* (1st ed.). Academic Press.

Kleffner, D. a, & Ramachandran, V. S. (1992). On the perception of shape from shading. *Perception & Psychophysics*, *52*(1), 18–36. http://doi.org/10.1038/331163a0

Kleiner, M., Brainard, D. H., & Pelli, D. (2007). What's new in Psychtoolbox-3? In *Perception 36 ECVP Abstract Supplement*.

Knill, D. C. (2007). Robust cue integration : A Bayesian model and evidence from cue-con fl ict studies with stereoscopic and fi gure cues to slant. *Journal of Vision*, *7*(5), 1–24. http://doi.org/10.1167/7.7.5.Introduction

Knill, D. C., & Saunders, J. A. (2003). Do humans optimally integrate stereo and texture information for judgments of surface slant? *Vision Research*, *43*(24), 2539–2558. http://doi.org/10.1016/S0042-

6989(03)00458-9

Koenderink, J., Doorn, A. Van, & Kappers, A. (1992). Surface perception in pictures. *Perception & Psychophysics*, *52*(5). Retrieved from http://link.springer.com/article/10.3758/BF03206710

Krug, K., & Parker, A. J. (2011). Neurons in Dorsal Visual Area V5/MT Signal Relative Disparity. *Journal of Neuroscience*, *31*(49), 17892–17904. http://doi.org/10.1523/JNEUROSCI.2658-11.2011

Land, M. F. (1997). Visual acuity in insects. *Annual Review of Entomology*, *42*(46), 147–177. http://doi.org/10.1146/annurev.ento.42.1.147

Landy, M. S., & Brenner, E. (2001). Motion-Disparity Interaction and the Scaling of Stereoscopic Disparity. In M. Jenkin & L. Harris (Eds.), *Vision and Attention* (pp. 129–150). New York, NY: Springer New York. http://doi.org/10.1007/978-0-387-21591-4_7

Legge, G. E., & Yuanchao, G. (1989). Stereopsis and contrast. *Vision Research*, *29*(8), 989–1004. http://doi.org/http://dx.doi.org/10.1016/0042-6989(89)90114-4

Leillinger, O. (1998). Misumena vatia with wasp. *CC BY SA 3.0*. Retrieved from https://commons.wikimedia.org/wiki/File:Misumena.vatia.beute.wespe.1771.jpg

Liu, Y., Bovik, A. C., & Cormack, L. K. (2008). Disparity Statistics in Natural Scenes. *Journal of Vision*, *8*(2008), 1–14. http://doi.org/10.1167/8.11.19.Introduction

Loffler, G. (2008). Perception of contours and shapes: low and intermediate stage mechanisms. *Vision Research*, *48*(20), 2106–27. http://doi.org/10.1016/j.visres.2008.03.006

Lovell, P. G., Bloj, M., & Harris, J. M. (2012). Optimal integration of shading and binocular disparity for depth perception. *Journal of Vision*, *12*, 1–18. http://doi.org/10.1167/12.1.1.Introduction

Lovell, P. G., Egan, J., Scott-Brown, K., & Sharman, R. J. (2015). Exploring the role of enhanced-edges in disruptive camouflage. *Perception*, *44*(10), 1236–1248.

Lovell, P. G., Gilchrist, I. D., Tolhurst, D. J., To, M., & Troscianko, T. (2008). Predicting search efficiency with a low-level visual difference model. *Journal of Vision*, *8*(6).

Lovell, P. G., Gilchrist, I. D., Tolhurst, D. J., & Troscianko, T. (2009). Search for gross illumination discrepancies in images of natural objects.

Journal of Vision, *9*(37). http://doi.org/10.1167/9.1.37.Introduction

Lovell, P. G., Scott-Brown, K., Egan, J., & Sharman, R. J. (2016). Can't see the snake for the leaves: enhanced edges disrupt depth perception. *Perception1*, *45*(3), 355–355.

Lunn, P. D., & Morgan, M. J. (1995). "The analogy between stereo depth and brightness": a reexamination. *Perception*, *24*(8), 901–904. http://doi.org/10.1068/p240901

Maloney, L. T. (2003). Statistical decision theory and evolution. *Trends in Cognitive Sciences*, *7*(11), 473–475. http://doi.org/10.1016/j.tics.2003.09.007

Marr, D. (1976). Early Processing of Visual Information. *Philosophical Transactions of the Royal Society of London B*.

Marr, D. (1982). *Vision*. *Journal of the Acoustical Society of America* (Vol. 30).

Marr, D., & Poggio, T. (1976a). Cooperative Computation of Stereo Disparity. *Science*, *194*(4262), 283–287.

Marr, D., & Poggio, T. (1976b). From Understanding Computation to Understanding Neural Circuitry. *Massachusetts Institute of Technology Artificial Intelligence Laboratory*, 1–22.

Marr, D., Poggio, T., Hildreth, E. C., & Grimson, W. E. L. (1991). A computational theory of human stereo vision. In L. Vaina (Ed.), *From the Retina to the Neocortex: Selected Papers of David Marr* (pp. 263–295). Boston, MA: Birkhäuser Boston. http://doi.org/10.1007/978-1-4684-6775-8_11

Martin, G. R. (2015). What is binocular vision for ? A birds ' eye view. *Journal of Vision*, *9*(2009), 1–19. http://doi.org/10.1167/9.11.14.Introduction

Mathematica. (n.d.). Wolfram Research, Inc.

MATLAB and Statistics Toolbox. (n.d.). The MathWorks, Inc., Natick, Massachusetts, United States.

McClure, T. (2014). Automatic Terrain Generation. Retrieved from http://uk.mathworks.com/matlabcentral/fileexchange/39559-automatic-terrain-generation

McKee, S. P. (1983). The spatial requirements for fine stereoacuity. *Vision Research*, *32*(2), 191–198.

McKee, S. P., & Taylor, D. G. (2010). The precision of binocular and monocular depth judgments in natural settings. *Journal of Vision*,

10(10), 5. http://doi.org/10.1167/10.10.5

McKee, S. P., & Verghese, P. (2002). Stereo transparency and the disparity gradient limit. *Vision Research*, *42*(16), 1963–77. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12160569

Mckee, S. P., Watamaniuk, S. N. J., Harris, J. M., Smallman, H. S., & Taylor, D. G. (1997). Is stereopsis effective in breaking camouflage for moving targets? *Vision Research*, *37*(15), 2047–2055. http://doi.org/10.1016/S0042-6989(96)00330-6

McSorley, E., & Findlay, J. M. (2001). Visual search in depth. *Vision Research*, *41*(25–26), 3487–3496. http://doi.org/10.1016/S0042-6989(01)00197-3

Merilaita, S. (2003). Visual background complexity facilitates the evolution of camouflage. *Evolution; International Journal of Organic Evolution*, *57*(6), 1248–1254. http://doi.org/10.2307/3448848

Milner, A. D., & Goodale, M. A. (1998). *The Visual Brain in Action*.

Mitchison, G. J., & McKee, S. P. (1987a). Interpolation and the detection of fine structure in stereoscopic matching. *Vision Research*, *27*(2), 295–302. http://doi.org/http://dx.doi.org/10.1016/0042-6989(87)90192-1

Mitchison, G. J., & McKee, S. P. (1987b). The resolution of ambiguous stereoscopic matches by interpolation. *Vision Research*, *27*(2), 285–294. http://doi.org/http://dx.doi.org/10.1016/0042-6989(87)90191-X

Mitchison, G. J., & Westheimer, G. (1984). The perception of depth in simple figures. *Vision Research*, *24*(9), 1063–1073. http://doi.org/http://dx.doi.org/10.1016/0042-6989(84)90084-1

Moraglia, G., & Schneider, B. (1990). Effects of direction and magnitude of horizontal disparities on binocular unmasking. *Perception*, *19*(5), 581–593. http://doi.org/10.1068/p190581

Nakayama, K., & Shimojo, S. (1990). Da vinci stereopsis: Depth and subjective occluding contours from unpaired image points. *Vision Research*, *30*(11), 1811–1825.

Nakayama, K., Shimojo, S., & Silverman, G. H. (1989). Stereoscopic depth: Its relation to image segmentation, grouping, and the recognition of occluded objects. *Perception*, *18*(1), 55–68. http://doi.org/10.1068/p180055

Nakayama, K., & Silverman, G. H. (1986). Serial and parallel processing of visual feature conjunctions. *Nature*, *320*, 264–265.

http://doi.org/10.1038/320264a0

NASA, E. and the H. H. T. (2014). Pillars of Creation. *Public Domain*. Retrieved from http://hubblesite.org/newscenter/archive/releases/2015/01/image/c/warn/

Nassi, J. J., & Callaway, E. M. (2009). Parallel processing strategies of the primate visual system. *Nature Reviews. Neuroscience*, *10*(5), 360–72. http://doi.org/10.1038/nrn2619

Neider, M. B., Ang, C. W., Voss, M. W., Carbonari, R., & Kramer, A. F. (2013). Training and transfer of training in rapid visual search for camouflaged targets. *PLoS ONE*, *8*(12). http://doi.org/10.1371/journal.pone.0083885

Neider, M. B., Boot, W. R., & Kramer, A. F. (2010). Visual search for real world targets under conditions of high target-background similarity: Exploring training and transfer in younger and older adults. *Acta Psychologica*, *134*(1), 29–39. http://doi.org/10.1016/j.actpsy.2009.12.001

Neider, M. B., Brotzen, S., & Zelinsky, G. (2010). Cutting through the clutter: Searching for targets in evolving realistic scenes. *Journal of Vision*, *7*(9), 1056–1056. http://doi.org/10.1167/7.9.1056

Neider, M. B., & Zelinsky, G. J. (2006). Searching for camouflaged targets: Effects of target-background similarity on visual search. *Vision Research*, *46*(14), 2217–2235. http://doi.org/10.1016/j.visres.2006.01.006

Neri, P. (2004). A stereoscopic look at visual cortex. *Journal of Neurophysiology*, *93*(4), 1823–6. http://doi.org/10.1152/jn.01068.2004

Nienborg, H., Bridge, H., Parker, A. J., & Cumming, B. G. (2004). Receptive field size in V1 neurons limits acuity for perceiving disparity modulation. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *24*(9), 2065–76. http://doi.org/10.1523/JNEUROSCI.3887-03.2004

Nikara, T., Bishop, P. O., & Pettigrew, J. D. (1968). Analysis of retinal correspondence by studying receptive fields of rinocular single units in cat striate cortex. *Experimental Brain Research*, *6*(4), 353–372. http://doi.org/10.1007/BF00233184

Nodine, C., & Kundel, H. (1987). Using eye movements to study visual search and to improve tumor detection. *Radiographics*, *7*(6), 1241–

1250. http://doi.org/10.1148/radiographics.7.6.3423330

Norcia, A. M., & Tyler, C. W. (1984). Temporal frequency limits for stereoscopic apparent motion processes. *Vision Research*, *24*(5), 395–401.

Norman, J. F., Lappin, J. S., & Zucker, S. W. (1991). The discriminability of smooth stereoscopic surfaces. *Perception*, *20*(6), 789–807. http://doi.org/10.1068/p200789

Norman, J. F., Norman, H. F., Craft, A. E., Walton, C. L., Bartholomew, A. N., Burton, C. L., … Crabtree, C. E. (2008). Stereopsis and aging. *Vision Research*, *48*(23–24), 2456–65. http://doi.org/10.1016/j.visres.2008.08.008

Norman, J. F., Todd, J. T., Norman, H. F., Clayton, A. M., & McBride, T. R. (2006). Visual discrimination of local surface structure: slant, tilt, and curvedness. *Vision Research*, *46*(6–7), 1057–69. http://doi.org/10.1016/j.visres.2005.09.034

Norman, J. F., & Wiesemann, E. Y. (2007). Aging and the perception of local surface orientation from optical patterns of shading and specular highlights. *Perception & Psychophysics*, *69*(1), 23–31. http://doi.org/10.3758/BF03194450

O'Brien, V. (1958). Contour perception, illusion and reality. *Journal of the Optical Society of America*, *48*(2), 112–119.

O'Toole, a J., & Walker, C. L. (1997). On the preattentive accessibility of stereoscopic disparity: evidence from visual search. *Perception & Psychophysics*, *59*(2), 202–18. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9055616

Ogle, K. N. (1952). On the limits of stereoscopic vision. *Journal of Experimental Psychology*, *44*(4), 253–259.

Ogle, K. N. (1958). Note on Stereoscopic Acuity and Observation Distance. *Journal of the Optical Society of America*, *48*(11), 794–798. http://doi.org/10.1364/JOSA.48.000794

Olsen, J. (2004). Realtime procedural terrain generation. *MIT*.

Orban, G. a., Janssen, P., & Vogels, R. (2006). Extracting 3D structure from disparity. *Trends in Neurosciences*, *29*(8), 466–473. http://doi.org/10.1016/j.tins.2006.06.012

Osorio, D., & Srinivasan, M. V. (1991). Camouflage by edge enhancement in animal coloration patterns and its implications for visual mechanisms. *Proceedings. Biological Sciences / The Royal Society*, *244*(1310), 81–85.

http://doi.org/10.1098/rspb.1991.0054

Ott, M., Schaeffel, F., & Kirmse, W. (1998). Binocular vision and accommodation in prey-catching chameleons. *Journal of Comparative Physiology A*, *182*(3), 319–330. http://doi.org/10.1007/s003590050182

Paaliaq. (2007). Zludzenie Hermana. *Public Domain*. Retrieved from https://commons.wikimedia.org/wiki/File:Zludzenie_Hermana.png

Panum, P. L. (1858). *Physiologische untersuchungen u¨über das sehen mit zwei augen.* Kiel, Germany: Schwers.

Parker, A. J. (2004). From binocular disparity to the perception of stereoscopic depth . *The Visual Neurosciences*, 779–792.

Parker, A. J. (2007). Binocular depth perception and the cerebral cortex. *Nature Reviews. Neuroscience*, *8*(5), 379–91. http://doi.org/10.1038/nrn2131

Parker, A. J., & Yang, Y. (1989). Spatial properties of disparity pooling in human sterevision. *Vision Research*, *29*(11), 1525–1538.

Parker, a J., & Cumming, B. G. (2001). Cortical mechanisms of binocular stereoscopic vision. *Progress in Brain Research*, *134*, 205–216. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11702545

Patterson, R., Cayko, R., Short, G. L., Flanagan, R., Moe, L., Taylor, E., & Day, P. (1995). Temporal integration differences between crossed and uncrossed stereoscopic mechanisms. *Perception {&} Psychophysics*, *57*(6), 891–897. http://doi.org/10.3758/BF03206803

Patterson, R., Moe, L., & Hewitt, T. (1992). Factors that affect depth perception in stereoscopic displays. *Human Factors*, *34*(6), 655–667.

Pelli. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442.

Penacchio, O., Lovell, P. G., Cuthill, I. C., Ruxton, G. D., & Harris, J. M. (2015). Three-Dimensional Camouflage: Exploiting Photons to Conceal Form. *American Naturalist*, *186*(4), 553–563. http://doi.org/10.1086/682570

Penacchio, O., Lovell, P. G., Sanghera, S., Cuthill, I. C., Ruxton, G. D., & Harris, J. M. (2016). Does countershading impede efficient visual search? *I-Perception*, *7*(1), 4. http://doi.org/10.1177/2041669515593223

Penacchio, O., Lovell, P. G., Sanghera, S., Cuthill, I. C., Ruxton, G., & H, J. M. (2015). Countershading camouflage and the efficiency of visual search. *Perception*, *44*(10), 1236–1248.

Peterhans, E., & von der Heydt, R. (1993). Functional Organization of Area V2 in the Alert Macaque. *European Journal of Neuroscience*, *5*(5), 509–524. http://doi.org/10.1111/j.1460-9568.1993.tb00517.x

Peterson, M. A., & Gibson, B. S. (1993). Shape Recognition Inputs To Figure-Ground Organization in Three-Dimensional Displays. *Cognitive Psychology*, *25*(3), 383–429. http://doi.org/http://dx.doi.org/10.1006/cogp.1993.1010

Pettigrew, J. D. (1980). Comparative Physiology of Binocular Vision. *The Australian Journal of Optometry*, *63*(5), 204–210. http://doi.org/10.1111/j.1444-0938.1980.tb02933.x

Pietrewicz, A. T., & Kamil, A. C. (1979). Search Image Formation in the Blue Jay (Cyanocitta cristata). *Science*, *204*(4399), 1332–1333. Retrieved from http://science.sciencemag.org/content/204/4399/1332.abstract

Pingstone, A. (2005). Dead leaf mantis (Deroplatys desiccata) at Bugworld in Bristol Zoo, Bristol, England. *Public Domain*. Retrieved from https://commons.wikimedia.org/wiki/File:Bristol.zoo.dead.leaf.mantis.arp.jpg

Pizlo, Z., Li, Y., & Francis, G. (2005a). A new look at binocular stereopsis. *Vision Research*, *45*(17), 2244–2255. http://doi.org/10.1016/j.visres.2005.02.011

Pizlo, Z., Li, Y., & Francis, G. (2005b). A new look at binocular stereopsis. *Vision Research*, *45*(17), 2244–2255. http://doi.org/10.1016/j.visres.2005.02.011

Pizlo, Z., Li, Y., & Steinman, R. M. (2008). Binocular disparity only comes into play when everything else fails; a finding with broader implications than one might suppose. *Spatial Vision*, *21*(6), 495–508.

Plaisted, K. C., & Mackintosh, N. J. (1995). Visual search for cryptic stimuli in pigeons: implications for the search image and search rate hypotheses. *Animal Behaviour*, *50*(5), 1219–1232. http://doi.org/http://dx.doi.org/10.1016/0003-3472(95)80039-5

Poggio, G. E. (1995). Mechanisms of stereopsis in monkey visual cortex. *Cerebral Cortex (New York, N.Y. : 1991)*, *5*(3), 193–204.

Poggio, T. (2012). Computer Science and Artificial Intelligence Laboratory Technical Report The Levels of Understanding framework. *MIT*.

Prins, N., & Kingdom, F. A. A. (2009). Palamedes: MATLAB (r) routines for analyzing psychophysical data.

Prinzmetal, W., & Gettleman, L. (1993). Vertical-horizontal illusion: one eye

is better than two. *Perception & Psychophysics*, *53*(1), 81–88. http://doi.org/10.3758/BF03211717

Qian, N., & Zhu, Y. (1997). Physiological computation of binocular disparity. *Vision Research*, *37*(13), 1811–1827. http://doi.org/10.1016/S0042-6989(96)00331-8

Qiu, F. T., & Von Der Heydt, R. (2005). Figure and ground in the visual cortex: V2 combines stereoscopic cues with Gestalt rules. *Neuron*, *47*(1), 155–156. http://doi.org/10.1016/j.neuron.2005.05.028

Ramachandran, V. S. (1986). Capture of stereopsis and apparent motion by illusory contours. *Percept Psychophys*, *39*(5), 361–373. http://doi.org/10.1038/324393a0

Read, J. C. A., & Cumming, B. G. (2003). Testing quantitative models of binocular disparity selectivity in primary visual cortex. *Journal of Neurophysiology*, *90*(5), 2795–2817. http://doi.org/10.1152/jn.01110.2002

Read, J. C. A., & Cumming, B. G. (2007). Sensors for impossible stimuli may solve the stereo correspondence problem. *Nature Neuroscience*, *10*(10), 1322–1328. http://doi.org/10.1038/nn1951

Regan, D. (n.d.). *Human Perception of Objects*.

Regan, D., & Hamstra, S. J. (1994). Shape discrimination for rectangles defined by disparity alone, by disparity plus luminance and by disparity plus motion. *Vision Research1*, *34*(17), 2277–2291. http://doi.org/10.1016/0042-6989(94)90107-4

Reid, P. J., & Shettleworth, S. J. (1992). Detection of cryptic prey: Search image or search rate? *Journal of Experimental Psychology: Animal Behavior Processes*, *18*(3), 273–286.

Rensink, R. a, & Cavanagh, P. (2004). The influence of cast shadows on visual search. *Perception*, *33*(11), 1339–1358. http://doi.org/10.1068/p5322

Richards, W. (1977). Stereopsis with and without monocular contours. *Vision Research*, *17*(8), 967–969.

Richards, W., & Foley, J. M. (1974). Effect of luminance and contrast on processing large disparities. *Journal of the Optical Society of America*, *64*(12), 1703–1705. http://doi.org/10.1364/JOSA.64.001703

Rob. (2011). Big Red Octopus (Octopus rubescens). *CC by-ND 2.0*. Retrieved from www.bbmexplorer.com

Rogers, B. J., & Graham, M. (1983). Anisotropies in the perception of three-dimensional surfaces. *Science (New York, N.Y.)*, *221*(4618), 1409–1411. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6612351

Rohaly, A. M., & Wilson, H. R. (1999). The effects of contrast on perceived depth and depth discrimination. *Vision Research*, *39*(1), 9–18. http://doi.org/10.1016/S0042-6989(98)00034-0

Rowland, H. M. (2009). From Abbott Thayer to the present day: what have we learned about the function of countershading? *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *364*(1516), 519–527. http://doi.org/10.1098/rstb.2008.0261

Ruxton, G. D., & Sherratt, T. N. (2004). *Avoiding attack: the evolutionary ecology of crypsis, warning signals and mimicry*. Oxford University Press.

Ruxton, G. D., Speed, M. P., & Kelly, D. J. (2004). What, if anything, is the adaptive function of countershading? *Animal Behaviour*, *68*(3), 445–451. http://doi.org/http://dx.doi.org/10.1016/j.anbehav.2003.12.009

Ryan, C., & Gillam, B. (1994). Cue conflict and stereoscopic surface slant about horizontal and vertical axes. *Perception*, *23*(6), 645–658.

Samonds, J. M., Potetz, B. R., & Lee, T. S. (2009). Cooperative and competitive interactions facilitate stereo computations in macaque primary visual cortex. *Journal of Neuroscience*, *29*(50), 15780–95. http://doi.org/10.1523/JNEUROSCI.2305-09.2009

Sand, H., Cederlund, G., & Danell, K. (1995). Geographical and latitudinal variation in growth patterns and adult body size of Swedish moose (Alces alces). *Oecologia*, *102*(4), 433–442. http://doi.org/10.1007/BF00341355

Sanger, T. D. (1988). Stereo disparity computation using Gabor filters. *Biological Cybernetics*, *59*(6), 405–418. http://doi.org/10.1007/BF00336114

Scharstein, D., & Szeliski, R. (2003). High-accuracy stereo depth maps using structured light. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, *1*.

Schneider, B. a, Moraglia, G., & Speranza, F. (1999). Binocular vision enhances phase discrimination by filtering the background. *Perception & Psychophysics*, *61*(3), 468–489. http://doi.org/10.3758/BF03211967

Schor, C. M., & Howarth, P. A. (1986). Suprathreshold stereo-depth matches as a function of contrast and spatial frequency. *Perception*, *15*(3), 249–

258.

Schor, C. M., & Wood, I. (1983). Disparity range for local stereopsis as a function of luminance spatial frequency. *Vision Research*, *23*(12), 1649–1654. http://doi.org/http://dx.doi.org/10.1016/0042-6989(83)90179-7

Schwark, J. D., MacDonald, J., Sandry, J., & Dolgov, I. (2013). Prevalence-based decisions undermine visual search. *Visual Cognition*, *21*(5), 541–568. http://doi.org/10.1080/13506285.2013.811135

Scinto, L., & Pillalamarri, R. (1986). Cognitive strategies for visual search. *Acta Psychologica*, *62*, 263–292.

Seuntiëns, P. J. H., Meesters, L. M. J., & IJsselsteijn, W. A. (2005). Perceptual attributes of crosstalk in 3D images. *Displays*, *26*(4–5), 177–183. http://doi.org/http://dx.doi.org/10.1016/j.displa.2005.06.005

Seymour, K. J., & Clifford, C. W. G. (2012). Decoding conjunctions of direction-of-motion and binocular disparity from human visual cortex. *Journal of Neurophysiology*, *107*(9), 2335–41. http://doi.org/10.1152/jn.01103.2011

Shiozaki, H. M., Tanabe, S., Doi, T., & Fujita, I. (2012). Neural activity in cortical area V4 underlies fine disparity discrimination. *Journal of Neuroscience*, *32*(11), 3830–41. http://doi.org/10.1523/JNEUROSCI.5083-11.2012

Simoncelli, E. P., & Olshausen, B. A. (2001). Natural image statistics and neural representation. *Annual Review of Neuroscience*, *24*, 1193–1216. http://doi.org/10.1146/annurev.neuro.24.1.1193

Smith, A. D., Hood, B., & Gilchrist, I. D. (2008). Visual search and foraging compared in a large-scale search task. *Cognitive Processing*, *9*(2), 121–6. http://doi.org/10.1007/s10339-007-0200-0

Smith, H. M. (2009). The costs of moving for a diurnally cryptic araneid spider. *Journal of Arachnology*, *37*(1), 84–91. http://doi.org/10.1636/ST07-62.1

Snowden, R. J., & Freeman, T. C. A. (2004). The visual perception of motion. *Current Biology*, *14*(19), R828-31. http://doi.org/10.1016/j.cub.2004.09.033

Srinivasan, M. V. (1995). Strategies for visual navigation, target detection and camouflage: inspirations from insect vision. *Proceedings of ICNN'95 - International Conference on Neural Networks*, *5*. http://doi.org/10.1109/ICNN.1995.487747

Steinman, S. B. (1987). Serial and parallel serach in pattern vision? *Perception1*, *16*(3), 389–398.

Stevens, K. a. (2006). Binocular vision in theropod dinosaurs. *Journal of Vertebrate Paleontology*, *26*(2), 321–330. http://doi.org/10.1671/0272-4634(2006)26[321:BVITD]2.0.CO;2

Stevens, M. (2013). *Sensory ecology, behaviour, & evolution* (1st ed.). Oxford University Press.

Stevens, M., Hopkins, E., Hinde, W., Adcock, A., Connolly, Y., Troscianko, T., & Cuthill, I. C. (2007). Field experiments on the effectiveness of "eyespots" as predator deterrents. *Animal Behaviour (Science Direct)*, *74*(5), 1215–1227.

Stevens, M., & Merilaita, S. (2009). Animal camouflage: current issues and new perspectives. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *364*(1516), 423–427. http://doi.org/10.1098/rstb.2008.0217

Stevenson, S. B., Cormack, L. K., & Schor, C. M. (1991). Depth attraction and repulsion in random dot stereograms. *Vision Research*, *31*(5), 805–813.

Suto, M., & Sykes, D. (2014). Stereophotomaker.

Sven. (2008). stlwrite. Retrieved January 1, 2016, from https://uk.mathworks.com/matlabcentral/fileexchange/20922-stlwrite-filename--varargin-?requestedDomain=www.mathworks.com

Thoen, H. H., How, M. J., Chiou, T.-H., & Marshall, J. (2014). A different form of color vision in mantis shrimp. *Science (New York, N.Y.)*, *343*(6169), 411–3. http://doi.org/10.1126/science.1245824

Thomas, O. M., Cumming, B. G., & Parker, A. J. (2002). A specialization for relative disparity in V2. *Nat Neurosci*, *5*(5), 472–478. Retrieved from http://dx.doi.org/10.1038/nn837

Tinbergen, L. (1960). The Natural Control of Insects in Pinewoods. *Archives Néerlandaises de Zoologie*, *13*(3), 265–343. http://doi.org/http://dx.doi.org/10.1163/036551660X00053

TNO Stereotest, Richmond Products. (2014). Retrieved from http://www.richmondproducts.com/shop/index.php?route=product/product&product_id=650

Todd, J. T. (2004). The visual perception of 3D shape. *Trends in Cognitive Sciences*, *8*(3), 115–121. http://doi.org/10.1016/j.tics.2004.01.006

Todd, J. T., Norman, J. F., Koenderink, J. J., & Kappers, A. M. L. (1997). Effects

of texture, illumination, and surface reflectance on stereoscopic shape perception. *Perception*, *26*, 807–822.

Torralba, A., & Oliva, A. (2003). Statistics of natural image categories. *Network: Computation in Neural Systems*, *14*(3), 391–412. http://doi.org/10.1088/0954-898X_14_3_302

Treisman, A. (1982). Perceptual grouping and attention in visual search for features and for objects. *Journal of Experimental Psychology: Human Perception and Performance*, *8*(2), 194.

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

Troscianko, T., Benton, C. P., Lovell, P. G., Tolhurst, D. J., & Pizlo, Z. (2009a). Camouflage and visual perception. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *364*(1516), 449–461. http://doi.org/10.1098/rstb.2008.0218

Troscianko, T., Benton, C. P., Lovell, P. G., Tolhurst, D. J., & Pizlo, Z. (2009b). Camouflage and visual perception. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *364*(1516), 449–461. http://doi.org/10.1098/rstb.2008.0218

Tsai, J. J., & Victor, J. D. (2003). Reading a population code: a multi-scale neural model for representing binocular disparity. *Vision Research*, *43*(4), 445–466. http://doi.org/10.1016/S0042-6989(02)00510-2

Tsirlin, I., Allison, R. S., & Wilcox, L. M. (2008). Stereoscopic transparency : Constraints on the perception of multiple surfaces. *Journal of Vision*, *8*(5), 1–10. http://doi.org/10.1167/8.5.5.Introduction

Tsirlin, I., Wilcox, L. M., & Allison, R. S. (2010). Monocular occlusions determine the perceived shape and depth of occluding surfaces. *Journal of Vision*, *10*(6), 1–11. http://doi.org/10.1167/10.6.11

Tyler, C. W. (1973). Stereoscopic Vision: Cortical Limitations and a Disparity Scaling Effect. *Science*, *181*(4096), 276–278.

Tyler, C. W. (1975). Spatial organization of binocular disparity sensitivity. *Vision Research*, *15*(5), 583–590. http://doi.org/doi:10.1016/0042-6989(75)90306-5

Tyler, C. W. (1991). The horopter and binocular fusion. *Vision and Visual Dysfunction*, *9*, 19–37.

Tyler, W. C., & Julesz, B. (1980). On the Depth of the Cyclopean Retina. *Experimental Brain Research*, *40*(2), 196–202.

Umeda, K., Tanabe, S., Fujita, I., Minini, L., Parker, A. J., Bridge, H., … Fujita, I. (2007). Representation of Stereoscopic Depth Based on Relative Disparity in Macaque Area V4 Representation of Stereoscopic Depth Based on Relative Disparity in Macaque Area V4. *Journal of Neurophysiology*, *98*(May 2007), 241–252. http://doi.org/10.1152/jn.01336.2006

Vidal-Naquet, M., & Gepshtein, S. (2012). Spatially invariant computations in stereoscopic vision. *Frontiers in Computational Neuroscience*, *6*. http://doi.org/10.3389/fncom.2012.00047

Vishwanath, D., & Hibbard, P. B. (2013). Seeing in 3-d with just one eye: stereopsis without binocular vision. *Psychological Science*, *24*(9), 1673–85. http://doi.org/10.1177/0956797613477867

Von Der Heydt, R., Zhou, H., & Friedman, H. S. (2000). Representation of stereoscopic edges in monkey visual cortex. *Vision Research*, *40*(15), 1955–1967. http://doi.org/10.1016/S0042-6989(00)00044-4

Wagemans, J. (2015). *Oxford Handbook of Perceptual Organisation*. Oxford University Press, USA.

Wallace, J. M., & Mamassian, P. (2004). The efficiency of depth discrimination for non-transparent and transparent stereoscopic surfaces. *Vision Research*, *44*(19), 2253–67. http://doi.org/10.1016/j.visres.2004.04.013

Wardle, S. G., Cass, J., Brooks, K. R., & Alais, D. (2010). Breaking camouflage : Binocular disparity reduces contrast masking in natural images. *Journal of Vision*, *10*(14), 1–12. http://doi.org/10.1167/10.14.38.Introduction

Warren, P. H., & Lawton, J. H. (1987). Invertebrate predator-prey body size relationships: an explanation for upper triangular food webs and patterns in food web structure? *Oecologia*, *74*(2), 231–235. http://doi.org/10.1007/BF00379364

Welchman, A. E., Deubelius, A., Conrad, V., Bülthoff, H. H., & Kourtzi, Z. (2005). 3D shape perception from combined depth cues in human visual cortex. *Nature Neuroscience*, *8*(6), 820–827. http://doi.org/10.1038/nn1461

Welchman, A. E., Tuck, V. L., & Harris, J. M. (2004). Human observers are biased in judging the angular approach of a projectile. *Vision Research*, *44*(17), 2027–2042. http://doi.org/10.1016/j.visres.2004.03.014

Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt. II. *Psychologische Forschung*, *4*(1), 301–350. http://doi.org/10.1007/BF00410640

Westheimer, G. (1975). Visual Acuity and Hyperacuity. *Investigative Ophthalmology*, *64*(8), 570–572. http://doi.org/10.1097/00006324-198708000-00002

Westheimer, G. (1979). The spatial sense of the eye. *Investigative Ophthalmology and Visual Science*, *18*(9), 893–912. http://doi.org/10/1979; 18(9):893-912

Westheimer, G., & McKee, S. P. (1977). Spatial configurations for visual hyperacuity. *Vision Research*, *17*(8), 941–947. http://doi.org/http://dx.doi.org/10.1016/0042-6989(77)90069-4

Westheimer, G., & Tanzman, I. J. (1956). Qualitative Depth Localization with Diplopic Images. *Journal of the Optical Society of America*, *46*(2), 116–117. http://doi.org/10.1364/JOSA.46.000116

Wheatstone, C. (1838). Contributions to the Pysiology of Vision. Part the First. On some remarkable, and hitero ubobserved, phenomena of binocular vision. *Philosophical Transactions of the Royal Society of London.*, *123*, 371–394.

Wiesel, T. N. (1968). Receptive fields and functional architecture. *J. Physiol.*, *195*, 215–243.

Wilcox, L. M. (1999). First and second-order contributions to surface interpolation. *Vision Research*, *39*(14), 2335–47. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10367055

Wilcox, L. M., & Duke, P. A. (2005). Spatial and temporal properties of stereoscopic surface interpolation. *Perception*, *34*(11), 1325–1338. http://doi.org/10.1068/p5437

Willigen, R. F. van der, Frost, B. J., & Wagner, H. (1998). Stereoscopic depth perception in the owl. *NeuroReport*, *9*(6). Retrieved from http://journals.lww.com/neuroreport/Fulltext/1998/04200/Stereoscopic_depth_perception_in_the_owl.50.aspx

Wolfe, J. M. (1994). Guided Search 2 . 0 A revised model of visual search, *1*(2), 202–238.

Woods, A. (2010). Understanding crosstalk in stereoscopic displays. In *Keynote Presentation at the Three-Dimensional Systems and Applications Conference, Tokyo, Japan* (pp. 19–21).

Woods, A. J. (2012). Crosstalk in stereoscopic displays : a review Crosstalk in stereoscopic displays : a review. *Journal of Electronic Imaging*, *21*(4), 40902. http://doi.org/10.1117/1.JEI.21.4.040902

Xiao, Y., Wang, Y., & Felleman, D. J. (2003). A spatially organized

representation of colour in macaque cortical area V2. *Letters to Nature*, *421*, 535–539. http://doi.org/10.1038/nature01239.1.

Xing, L., You, J., Ebrahimi, T., & Perkis, A. (2012). Assessment of Stereoscopic Crosstalk Perception. *IEEE Transactions on Multimedia*. http://doi.org/10.1109/TMM.2011.2172402

Yang, Y., & Blake, R. (1995). On the accuracy of surface reconstruction from disparity interpolation. *Vision Research*, *35*(7), 949–960. Retrieved from http://www.sciencedirect.com/science/article/pii/004269899400177N

Young, M. J., Landy, M. S., & Maloney, L. T. (1993). A perturbation analysis of depth perception from combinations of texture and motion cues. *Vision Research*, *33*(18), 2685–2696. http://doi.org/http://dx.doi.org/10.1016/0042-6989(93)90228-O

Zylinski, S., Osorio, D., & Shohet, a J. (2009). Perception of edges and visual texture in the camouflage of the common cuttlefish, Sepia officinalis. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *364*(1516), 439–48. http://doi.org/10.1098/rstb.2008.0264

1. Candidate's declarations:

I, Philip Cammack, hereby certify that this thesis, which is approximately 77,000 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for a higher degree.

I was admitted as a research student in September 2013 and as a candidate for the degree of Doctor of Philosophy in September 2016; the higher study for which this is a record was carried out in the University of St Andrews between 2013 and 2016.

Date signature of candidate

2. Supervisor's declaration:

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of Doctor of Philosophy in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date signature of supervisor

3. Permission for publication: (*to be signed by both candidate and supervisor*)

In submitting this thesis to the University of St Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that my thesis will be electronically accessible for personal or research use unless exempt by award of an embargo as requested below, and that the library has the right to migrate my thesis into new electronic forms as required to ensure continued access to the thesis. I have obtained any third-party copyright permissions that may be required in order to allow such access and migration, or have requested the appropriate embargo below.

The following is an agreed request by candidate and supervisor regarding the publication of this thesis:

PRINTED COPY

Embargo on all or part of print copy for a period of 2 years on the following ground(s):

- Publication would preclude future publication

ELECTRONIC COPY

Embargo on all or part of electronic copy for a period of 2 years on the following ground(s):

- Publication would preclude future publication

Date signature of candidate …… signature of supervisor ……

University of St Andrews | FOUNDED 1413

University Teaching and Research Ethics Committee

12 April 2016

Dear Philip

Thank you for submitting your amendment application which comprised the following documents:

1. Ethical Amendment Application Form
2. Advertisement
3. Participant Consent Form
4. Participant Information Sheet
5. Participant Debriefing Form
6. Data Management Plan

The School of Psychology & Neuroscience Ethics Committee is delegated to act on behalf of the University Teaching and Research Ethics Committee (UTREC) and has approved this ethical amendment application. The particulars of this approval are as follows –

Original Approval Code:	PS10641	**Approved on:**	04/12/2013
Amendment Approval Date:	12/04/2016	**Approval Expiry Date:**	04/12/2018
Project Title:	Breaking Camouflage using Depth Perception		
Researchers:	Philip Cammack, Ana Porskalaite and Andrew Chua		
Supervisor:	Professor Julie Harris		

Ethical amendment approval does not extend the originally granted approval period of five years, rather it validates the changes you have made to the originally approved ethical application. If you are unable to complete your research within the original five year validation period, you are required to write to your School Ethics Committee Convener to request a discretionary extension of no greater than 6 months or to re-apply if directed to do so, and you should inform your School Ethics Committee when your project reaches completion.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that you adhere to the 'Guidelines for Ethical Research Practice' (http://www.st-andrews.ac.uk/media/UTRECguidelines%20Feb%2008.pdf).

Yours sincerely

pp

Convener of the School Ethics Committee

cc Professor Julie Harris (Supervisor)

School of Psychology & Neuroscience, St Mary's Quad, South Street, St Andrews, Fife KY16 9JP
Email: psyethics@st-andrews.ac.uk Tel: 01334 462071