# ON THE PERCEPTION OF MONOCULAR REGIONS IN BINOCULAR SCENES

# Katharina Maria Zeiner

# A Thesis Submitted for the Degree of PhD at the University of St Andrews



2012

Full metadata for this item is available in St Andrews Research Repository at: <u>http://research-repository.st-andrews.ac.uk/</u>

Identifiers to use to cite or link to this thesis:

DOI: <u>https://doi.org/10.17630/10023-3650</u> http://hdl.handle.net/10023/3650

This item is protected by original copyright

# On the perception of monocular regions in binocular scenes.

Katharina Maria Zeiner



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

March, 2012

#### 1. Candidate's declarations:

I, Katharina Maria Zeiner, hereby certify that this thesis, which is approximately 50795 words in length, has been written by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for a higher degree.

I was admitted as a research student in September 2008 and as a candidate for the degree of Doctor of Philosophy in June 2009; the higher study for which this is a record was carried out in the University of St Andrews between 2008 and 2011.

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 electronic publication:

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 electronic publication of this thesis:

(iii) Embargo on both all of printed copy and electronic copy for the same fixed period of 3 years on the following ground(s):

publication would preclude future publication

Date signature of candidate .....

signature of supervisor .....

### On the perception of monocular regions in binocular scenes

Our two forward facing eyes supply the visual system with slightly different versions of the world around us. These slight differences in the two retinal images are used to obtain depth information about a visual scene. Conventionally, it has been assumed that points in one retinal image that do not have corresponding points in the other image are treated as noise. However, if one object is partially occluded by another, monocular points and regions occur. While it is, by now, established that we do not ignore geometrically plausible monocular points and regions, we know much less about our percept of such regions. So far it has not been studied how our percept of monocular regions compares to that of binocular regions and how well we are able to use the information contained in them.

In this thesis I explore our percept of such monocular regions and ask how we treat them in relation to their binocular surround. Using classical psychophysical methods as well as reaction-times and eye-tracking data, I find that we are perfectly capable of using monocular regions to interact with them and to arrive at a stable, seemingly complete percept of them. Sometimes our performance even benefits from the presence of a monocular region. However, this breaks down when we are asked to integrate the information contained in monocular regions with information in binocular regions.

The experiments in this thesis show that even though we do not ignore monocular regions and can, if we have to, interact with them, we show a strong preference towards relying on binocular information when given the option.

# Table of Contents

Page	è
Table of ContentsI	
AcknowledgementsIX	
Copyright/LicensesX	
PublicationsXI	
1. Introduction1	
1.1 Aims1	
1.2. Stereopsis2	
1.2.1. Two Sources of Information2	
1.2.2. The Correspondence Problem6	j
1.2.3. Unmatched Points7	
1.3 Half-Occlusions9	
1.3.1. Half-Occlusions – Definition9	
1.3.2. Perceiving depth from monocular items11	
1.3.2.1. Interpreting monocular occlusions12	2
1.3.2.2. Arriving at a depth percept based on	
monocular information23	
1.4. Amodal completion25	
1.4.1. Amodal completion and texture segmentation26	
1.4.1.1. Effects on performance	
1.4.1.2. Effects on reaction-times	
1.4.2. Amodal completion of subjective contours	1
1.4.3. Amodal completion of missing information	
1.4.4. Amodal completion of information in monocular scenes	
1.4.4.1. Amodal completion with monocular	
regions	,
1.5. Rivalry	
1.6. Extreme half-occlusions	

1.6.2. Stability of the percept of extreme half-occlusions	
1.7. Our percept of monocular regions40	
1.7.1. Integration of monocular information into the overall scene	
1.7.2. Perceptual compression of space in monocular regions41	
1.7.3. Integration of monocular regions into a unified percept42	
1.7.4. Do we form a cyclopean percept?44	
1.8. Comparing monocular and binocular regions	
1.9. Structure of this thesis45	
Part I – Perceiving monocular regions in binocular scenes	
2.1. Aims	
2.2. Numerosity Perception49	
2.2.1. A Sense of Number	
2.2.2. Relative Numerosities51	
2.2.3. A sense of number vs. texture density	
2.3. General Themes Motivating the Experiments	
2.3.1. Do we perceive information in monocular regions the same as information in binocular regions?	
2.3.2. Is this any different from closing one eye?	
2.3.3. Is there a benefit from using two eyes?	
2.4. Stimulus Geometry & General Setup58	
2.4.1. Apparatus58	
2.4.1.1. Calibration59	
2.4.1.2. Stimulus Generation60	
2.4.2. Stimuli61	
2.5. Procedure61	
2.6. Presented Numerosities63	
2.6.1 A scale for discriminating numerosities	
2.6.2. Varying baseline numerosities70	
2.7. Participant Screening71	

2.8. Ethical A	Approval	73
2.9. Analysis		74
2.9.1. to the	How was the Psychometric Function fitted data?	75
	2.9.1.1. The psychometric function	76
	2.9.1.2. Confidence intervals	80
	2.9.1.3. Sensitivity analysis – adjusting the confidence intervals	81
	2.9.1.4. Goodness-of-fit	84
3. Sensitivity to diff	erences between numerosities in monocular regio	ons86
3.1. Aims		86
3.2. Experim Maximal Cor	ent 1 – Numerosity Discrimination at htrast	87
3.2.1.	Aims	
3.2.2.	Methods	87
	3.2.2.1. Stimuli	
	3.2.2.2. Participants	91
	3.2.2.3. Procedure	91
3.2.3.	Predictions	92
	3.2.3.1. Vertically Occluded vs. Binocular	92
	3.2.3.2. Monocular Unoccluded vs. Binocular	93
	3.2.3.3.a. Vertically Occluded vs. Monocular Occluded	94
	3.2.3.3.b. Vertically Occluded vs. Monocular Unoccluded	94
3.2.4.	Analysis	95
3.2.5.	Results	95
	3.2.5.1. Differences between the various baseline-numerosities	96
	3.2.5.2. Differences between left and right eye monocular conditions	98
	3.2.5.3. Differences between the 4 remaining conditions	100
3.2.6.	Discussion	101

	3.2.9.1. Differences between Vertically Occluded and Binocular Unoccluded Conditions	101
	3.2.6.2. Differences between Monocular and Binocular Unoccluded Stimuli	102
	3.2.6.3. Differences between Vertically Occluded and Monocular Occluded Stimuli	103
	3.2.6.4. Differences between Monocular Occluded and Monocular Unoccluded Stimuli	104
3.2.6.	5. General Discussion	104
3.3. Experim	nent 2	106
3.3.1.	Aims	106
3.3.2.	New Stimuli	107
3.3.3.	Hypotheses	109
	3.3.3.1. Purely Binocular vs Horizontally Occluded	110
	3.3.3.2. Horizontally Occluded vs Vertically Occluded	111
3.3.4.	Participants	111
3.3.5.	Setup	111
3.3.6.	Analysis and Results	112
	3.3.6.1. Investigating non-stationarities	112
	3.3.6.2. Differences between left and right eye monocular conditions	116
	3.3.6.2.1. Presentation-eye as a between-subjects factor	116
	3.3.6.2.1.a. Occluded conditions	117
	3.3.6.2.1.b. Unoccluded conditions	118
	3.3.6.2.2. Eye-dominance as a 'between-subjects' factor	118
	3.3.6.2.2.a. Occluded conditions	119
	3.3.6.2.2.b. Unoccluded conditions	119
	3.3.6.3. Differences between the remaining 5 conditions	120
3.3.7.	Discussion	121
	3.3.7.1. Purely Binocular vs Horizontally Occluded	121

	3.3.7.2. Vertically Occluded vs Horizontally Occluded122
	3.3.7.3. Vertically Occluded vs. Monocular123
3.3.8.	Issues with the stimuli used in experiment 2124
3.4. Experim	125 nent 2 – part 2
3.4.1.	Changes in the stimuli125
3.4.3.	Questions127
3.4.4.	Participants127
3.4.5.	Analysis127
3.4.6.	Results128
3.4.7.	Discussion129
	3.4.7.1. Can we replicate the results of experiment 2 with the changed luminances?129
	3.4.7.2. Can we replicate the results of experiment 2 after changing the baseline, stimulus levels and presentation order?130
	3.4.7.3. Is there a difference in performance between the monocular conditions with noise and those with a grey background?130
3.5. Overall discus	sion131
4. Experiment 3: The effe	ct of varying texture densities on our
percept of half-occluded	regions134
4.1. Aims	
4.2. Methods	
4.2.1. Stimu	li/Procedure135
4.2.2. Quest	tions139
4.2.3. Partic	ipants142
4.2.4. Analy	sis142
4.3. Results	142
4.3.1. Effect	s of horizontal occlusion143
4.3.2. Effect	s of vertical occlusion146
4.3.3. Do pa	articipants ignore one eye's view?148
4.4 Discussion	
5. Experiment 4: The integration	gration of monocular and binocular numerosities.156

5.1. Aims156
5.2. Methods157
5.2.1. Stimuli/Procedure157
5.2.2. Questions162
5.2.3. Participants165
5.2.4. Analysis165
5.3. Results165
5.3.1. Are monocular and binocular regions integrated?166
5.3.2. The effect of differing dot densities
5.4 Discussion171
Part II – Searching for monocular targets176
6. Introduction176
6.1. Aims176
6.2. What do we use information in monocular regions for?177
6.2.1. Recap: Monocular Regions & Depth Perception178
6.2.2. Recap: Monocular Regions & Amodal Completion179
6.2.3. Recap: Our Percept of Monocular Regions179
6.3. How is information in monocular regions processed? - Computational models181
6.4. Searching for monocular items188
6.4.1. Searching for ocular singletons and visual saliency188
6.4.2. The saliency hypothesis193
6.5. Methods195
6.5.1. Stimuli195
6.5.2. Procedure197
6.5.3. Analysis197
7. Visual search for monocular items – Are we faster at detecting monocular targets?199
7.1. Aims
7.2. Experiment 5 – searching for monocular items in binocular scenes199
7.2.1. Aims200
7.2.2. Methods

7.2.2.1. Stimuli/Procedure	200
7.2.2.2. Questions	204
7.2.2.3. Participants	207
7.2.2.4. Analysis	207
7.2.3. Results	208
7.2.3.1. How does this stimulus compare to the one used by Zhaoping (2008)?	208
7.2.3.2. Do monocular items and ocular singletons affect our visual search behaviour in the same fashion?	209
7.2.3.3. How does a larger number of monocular items/'singletons' affect our performance?	210
7.2.4 Discussion	211
7.3. Experiment 6	213
7.3.1. Methods	213
7.3.1.1. Conditions	214
7.3.1.2. Questions	214
7.3.1.3. Participants	214
7.3.1.4. Analysis	215
7.3.2. Results and discussion	215
7.3.2.1. How are reaction-times and the number of distractors related?	215
7.3.2.2. Differences between monocular/binocular target conditions	216
7.4. Overall Discussion	217
8. Experiment 7 – Do monocular regions affect our eye-movements?	219
8.1. Aims	219
8.2. Background	219
8.3. Methods	220
8.3.1. Stimuli/Procedure	220
8.3.2. Research Questions	224
8.3.3. Participants	226
8.3.4. Analysis	226
8.4. Results	227
8.4.1. Reaction-Times	228

8.4.2. Eye-Movements	229
8.5. Discussion	232
9. Overall discussion	235
9.1. Experiment 1 (chapter 3.2)	235
9.2. Experiments 2 and 2.2	236
9.2.1. Experiment 2 (chapter 3.3)	236
9.2.2. Experiment 2.2 (chapter 3.4)	237
9.2.3. Monocular Regions so far	238
9.3. Experiment 3 – or what happens if the two eyes see different things?	239
9.3.1. Experiment 3 (chapter 4)	239
9.3.2. Our percept of monocular regions – summary I	240
9.4. Experiment 4 – Are we able to integrate information from monocular and binocular regions?	242
9.4.1. Experiment 4 (chapter 5)	242
9.4.2.Our percept of monocular regions – summary II	243
9.5. Experiment 5 (chapter 7.2)	246
9.6. Experiment 6 – Interacting with increasing numbers of monocular regions	248
9.6.1. Experiment 6 (chapter 7.3)	248
9.6.2. Visual-search – what do the reaction-times tell us?	249
9.7. Experiment 7 (chapter 8)	250
9.8. Discussion	251
10. References	256
Appendices	A1

## Acknowledgements

They say it takes a village to raise a child. It felt like it took more than that to finish my PhD.

I would like to thank all those people without whom I would never have gotten this far. Without whom I would probably never have ended up in a lab, would not have found my own voice in terms of research, and who never stopped telling me that there are more important things in life than work.

My supervisor Julie Harris who gave me the best support imaginable when I needed it while allowing me the freedom I needed to develop my ideas. I don't think I could have had a better supervisor. I hope I didn't go off on too many tangents....I probably never would have realised how interesting visual perception is if you hadn't given me the opportunity to work in the lab all those years ago and I probably wouldn't have made it this far without your support.

My parents and grandparents who taught me that it doesn't matter what you do as long as you give it your best, who supported many of my crazy whims and who forced me off the phone with the comment 'you know you are just procrastinating at the moment, write a couple of pages and then we'll talk again.'.

My two RAs Manuel and Kotryna who kept me sane in crazy times, were happy to brainstorm and who are just fabulous in general!

Zhaoping Li and Keith May who allowed me to work in their lab (and the ACCN Internship Scheme who supported the collaboration) and taught me so much.

Anna Blanch who believed in my writing when I didn't.

All those people who believed in my programming even though I still can't optimise code to save my life.

The people who put up with way too many blonde questions about psychometric function fitting and maths in general.

The friends who were there for me both when I had good days and when I had not so good ones, who lived through the various series of scones, pancakes, and muffins, and who forced me to have a life every once in a while :)

The amazing people who proofread parts of of this thesis – Ashley Dennis, Valentin Haenel, Louise O'Hare, Ingo Fründ, Verena Kersken and Neil Rabinowitz.

Verena Kersken who submitted the final version of this thesis for me.

Marianne Maertens who gave me a job even though I was still writing up.

And probably most importantly, Verena Kersken, yet again for being the best friend imaginable, I wouldn't have gotten through the last few months (ok, more like the last 5 years) without you around. You better start training for that triathlon in 2013!

# Copyright/Licenses

I have obtained written permission from all authors/copyright holders for figures that reproduce/strongly resemble previously published figures.

## Reproductions

For figure 6.1, the copyright holder is the Association for Research in Vision and Ophthalmology (ARVO), I have the permission to reproduce this figure (figure 1 in Zhaoping, 2008) and the permission from Zhaoping Li.

# Figures strongly resembling published material

For figures 1.8-1.10, the copyright holder is Elsevier Ltd. While my figures are no direct reproductions of figures 3b, 4, and 6 in Nakayama and Shimojo (1990), I have License number 2863571014107 from Elsevier Ltd to reproduce 3 figures.

For figure 1.13, the copyright holder is Nature Publishing Group. While my figure is no direct reproductions of figure 1 in He & Nakayama (1992), I have License number 2863571469510 from Nature Publishing group to reproduce figure 1 and permission from Zijiang He.

For figure 1.15, the copyright holder is Elsevier. While my figure is no direct reproduction of the stimuli shown in figure 1 in Gold et al (2000), I have License number 2863570666469 from Elsevier Ltd to reproduce figure 1 and permission from Jason Gold.

For figure 1.16, the copyright holder is the American Psychological Association (APA). The reproduction of up to 3 figures from any given APA publication dies not breach copyright. While my figure is no direct reproduction of the stimuli shown in figures 5 and 7 in Bruno et al. (1997), I have permission from Nicola Bruno.

For figures 1.19 and 1.20, the copyright holder is the American Psychological Association (APA). The reproduction of up to 3 figures from any given APA publication dies not breach copyright. While my figures are no direct reproduction of the stimuli shown in figures 3 and 5 in Ono et al. (2003), I have permission from Hiroshi Ono.

# Publications

The work in this thesis has been included in the following publications:

# Upcoming

Zeiner, K.M., Grinkeviciute, K.K. & Harris, J.M. (2012). Integrating monocular and binocular regions in a visual scene. (in preparation).

Zeiner, K.M. & Harris, J.M. (2012). Perceptual integration across natural monocular regions. (in preparation).

Zeiner, K.M., Spitschan, M. & Harris, J.M. (2012). Visual search for monocular items in binocular scenes. (in preparation).

Zeiner, K.M., Spitschan, M., Grinkeviciute, K.K., & Harris, J.M. (2012). Effects of monocular distractors on monocular and binocular visual search. [Abstract]. Journal of Vision.

# 2012

Zeiner, K.M. & Harris, J.M. (2012). Perceiving monocular regions in binocular scenes. [Extended Abstract]. Predicting Perceptions 2012 conference proceedings

# 2011

Zeiner, K.M., Grinkeviciute, K.K., Spitschan, M., & Harris, J.M. (2011). Combining monocularly and binocularly visible regions of a scene for dot density perception. [Abstract]. Neuroscience Meeting Planner. Washington, DC: Society for Neuroscience.

Zeiner, K.M., Spitschan, M., May, K.A., Zhaoping, L., & Harris, J.M. (2011). Eye movements and reaction times for detecting monocular regions in binocularly viewed scenes. [Abstract]. Journal of Vision.

Zeiner, K.M., Spitschan, M., & Harris, J.M. (2011). Combination of left and right eye information for texture/density perception. [Abstract]. *Perception* 40(1), p.109.

# 2010

Zeiner, K.M., Spitschan, M., & Harris, J.M. (2010). Using monocular regions in binocular scenes to guide our attention in a visual search task. [Abstract]. Neuroscience Meeting Planner. San Diego, CA: Society for Neuroscience.

Zeiner, K.M. & Harris, J.M. (2010). What is happening behind my neighbour's fence? - Investigating our ability to use unmatched regions in a binocular scene. [Abstract]. *Perception* 39(2), pp.269-270.

Zeiner, K.M., Spitschan, M., & Harris, J.M.(2010) Using numerosity to explore monocular regions in binocular scenes. [Abstract], *Journal of Vision* 10(7):362.

#### 1. Introduction

#### 1.1 Aims

One 'puzzle' of human vision is how we perceive the world surrounding us as three dimensional while the image that each eye passes on to the visual system is two dimensional.

Unlike a camera, our visual system receives input from two sources, each eye seeing a slightly different version of any scene around us. Also, unlike many other animals, humans have two forward facing eyes, arranged such that the input from the two eyes delivers a considerable overlap. This overlap, and the slight differences in position of image points between the views belonging to the two eyes, are used by our visual system to recover depth from a scene. However, in natural scenes there is also a host of image points that cannot be matched with image points from the input from the other eye, because they are only present in one eye's view.

This literature review will discuss previous research on the perception of these monocular points, and regions, in binocular scenes and their geometry. I will discuss how monocular regions compare to binocular regions, and how our percept of monocular regions might be related to that produced by amodally completed or rivalrous stimuli. This literature review is by no means intended to be a complete review of the amodal completion and rivalry literature, so these sections will focus on the literature relevant to our understanding of the perception of monocular regions. I will also describe how the work in this thesis follows from the literature on monocular regions and how our percept of information in monocular regions relates to how we perceive depth from said monocular regions.

#### 1.2. Stereopsis

In this section I will describe how the two eyes' views are combined to arrive at a three-dimensional percept of a scene around us, the difficulties we face in this process (how is this combination accomplished?) and the limitations (what if a point is only seen by one eye?).

#### 1.2.1. Two Sources of Information

We receive separate visual input from our two horizontally offset eyes. This means we see the visual scene surrounding us from two slightly different directions.

If we are looking at two objects located at different distances from us they can fall onto different locations on the retinae in our left and right eyes. This is illustrated in figure 1.1.



Figure 1.1: Points A and B lie at different locations in space. Based on each of the two retinal images they could be located anywhere along each line of sight. However, once the two images are combined our visual systems are able to deduce that point B is located further away from us than point A.

The visual system combines the two eyes' views of points A and B and is able to use the slight differences between them to deduce the location in depth of the two points in relation to each other and the observer. I will now explain how the distance between the two objects can also be expressed as the angular difference between the image locations of the two points in the left eye ( $\Phi_l$ ) minus the angular difference between the image locations of the two points in the right eye ( $\Phi_l$ ).

Let the distance between the two eyes be i, D is the distance between our observer and the point closest to the observer (in the case of figure 1.2 this is point A),  $\Delta d$  is the distance along the depth dimension between the two points.  $\Phi_{I}$  is the visual angle between points A and B in the left eye  $\Phi_{r}$  is the visual angle between the two points in the right eye.  $\theta_A$  is the binocular subtense of A and  $\theta_B$  is the binocular subtense of B. The relative angular disparity  $\eta = \Phi_I - \Phi_r = \theta_A - \theta_B$ . Furthermore, let e1 be the distance of A from the midline and  $e_2$ the distance of B from the midline. This is shown in figure 1.2.



Figure 1.2: Illustration of the different terms used in the disparity calclation.

Let us assume our observer is fixating on point A. The relative binocular disparity can be derived as follows (Cormack & Fox, 1985, cited in Howard & Rogers II, 2002):

The convergence angle  $\theta_A$  is

$$\theta_{A} = \arctan\left[\frac{\frac{i}{2} + e_{I}}{D}\right] + \arctan\left[\frac{\frac{i}{2} - e_{I}}{D}\right]$$
(1.1)

and the binocular subtense of  $\,\theta_{\scriptscriptstyle B}$  of B (which lies  $\,\Delta d$  behind our fixation point A is

$$\theta_{B} = \arctan\left[\frac{\frac{i}{2} + e_{2}}{D + \Delta d}\right] + \arctan\left[\frac{\frac{i}{2} - e_{2}}{D + \Delta d}\right]$$
(1.2)

this means the relative angular disparity  $\eta$  of the two points is  $\theta_{A}$ -  $\theta_{B}$ , i.e.:

$$\eta = \left( \arctan\left[\frac{\frac{i}{2} + e_1}{D}\right] + \arctan\left[\frac{\frac{i}{2} - e_1}{D}\right] \right) - \left( \arctan\left[\frac{\frac{i}{2} + e_2}{D + \Delta d}\right] + \arctan\left[\frac{\frac{i}{2} - e_2}{D + \Delta d}\right] \right)$$
(1.3)

If we assume that the two points fall on the median plane then, since  $\Phi_{I} = \Phi_{r}$ ,  $\eta = 2\Phi^{\circ}$  and  $\Phi = \theta_{P} - \theta_{P}$  the above formula can be reduced to:

$$\tan\frac{\theta_A}{2} = \frac{i}{2D} \quad \text{and} \quad \tan\frac{\theta_B}{2} = \frac{i}{2(D + \Delta d)} \quad . \tag{1.4}, (1.5)$$

Since, when we are talking about small angles, the tangent of an angle is approximately the same as the angle in radians, it follows that

$$\eta = 2\Phi = \frac{i}{D} - \frac{i}{D + \varDelta d} \quad \text{or} \quad \eta = \frac{i\varDelta d}{D^2 + D\varDelta d} \tag{1.6}, (1.7)$$

and since D $\Delta$ d tends to be negligible compared to the size of D<sup>2</sup> (if  $\Delta$ d is small compared with D) this can be rewritten as

$$\eta = \frac{i\Delta d}{D^2}$$
 (in radians). (1.8)

In order to be able calculate the disparity between two points our visual system has to work out which point in the right retinal image corresponds to which point in the left retinal image and match them correctly. I will explain why it is important to correctly match corresponding points in the two eyes' views below.

#### 1.2.2. The Correspondence Problem

Much research has been conducted trying to establish how the visual system matches corresponding points between the two eyes, since any point could theoretically be matched to a large number of other points in the other retinal image (e.g. Marr and Poggio, 1976, 1979). The issue is that there is a large number of points in each retinal image that might correspond to more than one other point in the other retinal image. This is called the correspondence problem. How does the visual system match the right two points with each other? Figure 1.3 shows a basic version of this problem.



Figure 1.3: The dots that are intersected by one diagonal line fall onto the same point on the retina. This figure illustrates how we have to correctly match the correct retinal images in order to arrive at a correct representation of the world.

Figure 1.3 illustrates the importance of correctly matching points from

the two retinal images. Each eye sees a horizontal row of 3 circles. However, there are different ways the points could be matched. The circles in figure 1.3 are all possible matches between a circle see by the right eye and one seen by the left eye. The points that lie on the same radial lines (the lines from the eyes to the circles) will appear in the same location in one eye. Let us consider the grey circles as the 'real' circles that should be correctly matched. If we incorrectly matched two points in the retinal images that correspond to one of the unfilled circles this would lead to us perceiving this dot to lie at a different distance from us in relation to the other circles. Julesz (1964) showed that we are able to perceive depth in a Random Dot Stereogram (a display filled with dots that is presented stereoscopically). This means the visual system readily matches corresponding points in the two retinal images even if there is only very limited information (the only information present in a random dot stereogram is the colour of the dot and its location in the two eyes) present.

While this works for dots, one might object that we are surrounded by objects of which parts are occluded by other objects. This occlusion causes regions of these objects to be visible to one eye but have no match in the other eye. How does this affect the correspondence problem?

#### 1.2.3. Unmatched Points

Again, let us reduce the visible world to two dots that are at various distances from an observer. We can then consider how we are able to perceive them as lying at different depths in space.



Figure 1.4: Consider pints A and B - Disparity =  $\alpha$  -  $\beta$ . Now consider points A and C. We cannot define the disparity for those two points as C is only visible to the right eye. This point is called an unmatched point.

Figure 1.4 shows a top down view of points that are at different distances from the observer. Points A and B are at different locations in space and, since the two eyes are horizontally offset, lead to different retinal images. While the question for points A and B was how the visual system manages to correctly match these two views and to combine them to a meaningful percept using the difference in the visual angle of the distance between the two circles in the two eyes' views, we face a different question for point C. Notice that C is only seen by one eye as it is hidden behind the image of A in the left eye's view. If we were to try to calculate the relative disparity between dots A and C, this would not work as the unmatched point C can only be seen by the right

eye and a visual angle between point C and the other two points in the left eye view can therefore not be calculated.

Julesz (1971) suggests that these monocular points are treated as noise in our percept of the stimulus (in the case of a random dot display this could be caused by the visual system falsely matching two points in the retinal images, leaving the 'true' matches of the two matched points unmatched) and thus ignored by the visual system. However, as figure 1.4 shows, unmatched points do not always act as noise but might simply have no match in the other eye's retinal image. I will describe this further in the following section.

#### 1.3 Half-Occlusions

#### 1.3.1. Half-Occlusions - Definition

Regions in a binocular scene that can only be seen by one eye occur in virtually every visual scene. They occur when one object is occluded by another object, in relation to the observer, which leads to some parts of the object in the background to be seen only by one eye. Figure 1.5 shows a top-down view of such a situation.

The mid-grey rectangle is a background plane which contains information, for example there could be a picture on that plane or it might just be your kitchen cabinet. The dark grey rectangle is an occluder that is placed between you and the background. This could be anything opaque, a rectangular column in your kitchen or a coffee mug that you are holding up in





Figure 1.5: Birds-eye view of an occlusion situation. Note how different parts of the background object are visible to the two eyes.

Because the occluder is in a slightly different location in each of your eyes, it is in front of slightly different regions on the background plane. Note how there are regions on the background to the right and left side of the occluder that only one eye can view.

These regions have been referred to as 'half- occlusions' (e.g. Egnal & Wildes, 2002), 'unmatched' (e.g. Brooks & Gillam, 2006), 'unpaired' (e.g. Nakayama & Shimojo, 1990), 'monocular areas/regions' (e.g. Ono et a.l, 2002; Grove et al., 2006) or 'partially occluded' (e.g. Nakayama, Shimojo & Silverman,

1989). I will be referring to them throughout this thesis as monocular regions.

The two eyes see different parts of the target behind the occluder. Note how the region that can be depicted with a single viewpoint (if there were a single eye, above the nose) is significantly smaller than the region accessible to both eyes combined. Using the input from both eyes gives the observer more information than a single viewpoint would allow for. This is because each of the eyes sees a small area behind the occluder that the other eye cannot access. For example, if I pick up the coffee mug on my table, and hold it up between my eyes and the computer screen, some of the text on the screen will be occluded in each eye's view. If my visual system is able to use the monocular information of these regions this increases the amount of text I can consciously perceive and interact with dramatically (e.g. See Ono et al., 2003).

Even though there is no disparity information available in monocular regions, we also seem to be able to extract a sense of depth from them. I will now discuss some of the research on perceived depth from monocular regions.

#### 1.3.2. Perceiving depth from monocular items

Nakayama and Shimojo (1990) were some of the first to start exploring perceived depth from monocular regions in a series of experiments using random dot stereograms. They found that unpaired points were perceived as further away than binocular points when seen adjacent to said binocular points. They suggested that this showed that the visual system interprets monocular areas as being produced by an invisible occluder.

These results stand in sharp contrast to the discounting of monocular image points as noise, which would be suggested by the more classical approach taken, for example, by Julesz (1971). I will now describe some of the research on perceived depth from monocular regions.

#### 1.3.2.1. Interpreting monocular occlusions

In natural scenes, monocular regions coincide with large depth discontinuities (Wilcox & Lakra, 2007) and it seems that we interpret monocular regions based on either an innate or learned knowledge about geometrical occlusion constraints (Nakayama & Shimojo, 1990). This means we react to monocular regions as if they were caused by an occluder even if there is no visible occluder. This was first investigated by Lawson and Gulick (1967) who demonstrated that monocular cues signalling occlusion are interpreted as indicative of a depth step.

Motivated by drawings by daVinci, Nakayama and Shimojo investigated this in a series of experiments (Shimojo & Nakayama, 1990; Nakayama & Shimojo, 1990). I will now describe these experiments and discuss their significance for our knowledge on what occlusion constraints we employ to interpret our visual environment.

We find monocular regions along all natural vertical edges, but not along horizontal edges. Along vertical edges there are geometrically plausible and

12

implausible occlusion situations. If a monocular region is visible in the left eye, it can only be caused by a plausible occluder if it occurs on the left side of said occluder and vice versa.

While it is possible to create a situation in the lab in which a right eye monocular region occurs to the left of an occluder, this stimulus will not correspond to a natural situation and will thus be interpreted as invalid and appear rivalrous (for more details on rivalry, see section 1.5) because of interocular suppression.

Shimojo and Nakayama (1990) asked whether our percept of monocular regions differs if the monocular region is valid/invalid. After observing that a monocular region that is geometrically plausible is perceived as lying behind an object that seems to be causing the occlusion while a monocular region that is geometrically not plausible is perceived as lying on the same plane as the occluding surface and appearing to fade away (a sign that the region is being suppressed interocularly) Shimojo and Nakayama (1990) devised two experiments that asked under which conditions we are suppressing these monocular regions.

First they devised a stimulus which is shown as a cartoon in figure 1.6. This stimulus was presented under 8 (2x2x2 design) conditions:

- the disparities could either be crossed or uncrossed

- the blue monocular region could be presented in the left or the right eye

- the blue monocular region could be presented on the left or right side of the

13

white square.



Figure 1.6: Cartoon of simulus used by Shimojo & Nakayama (1990). The yellow and white regions where binocularly visible. The blue region was monocular. The colours were chosen to make any interocular suppression more noticable, the texture was added to facilitate binocular fusion.

While the combination of the monocular region being located to the right side of the square is a plausible stimulus if it is presented to the right eye, it is not if it is presented to the left eye.

Participants were asked to report whether the blue region appeared to fade (which is taken to be a sign of interocular suppression). The invalid occlusions were significantly more likely to lead to a fading of the blue area than the valid occlusion conditions. This shows that monocular regions are only fully integrated into the overall percept if they can be interpreted as being cause by a geometrically plausible occlusion.

Next Shimojo and Nakayama (1990) asked whether the interocular suppression for the invalid regions was caused by observers applying an 'occlusion constraint' (i.e. that monocular regions are only 'valid' if they can be caused by a geometrically plausible occlusion) or whether it was caused by what they call 'mutual correlation'.

Mutual correlation had been proposed as an alternative means of explaining the observed suppression and assumes that monocular regions can be unmatched to differing extents. If a region has no match in the other eye, then the amount of suppression depends on what information is presented in the retinotopic counterpart of this region. If the retinotopic counterpart can be binocularly matched (this corresponds to a geometrically plausible occlusion) then the monocular region is integrated. If the retinotopic counterpart is also unmatched then this leads to competition between the two eyes' views in this region.

While this alternative explanation leads to the same predictions for the horizontal disparities used in the previous experiment, it leads to very different predictions for vertical disparities. As I mentioned before, monocular regions only occur naturally at vertical edges while the regions around horizontal edges are fully binocular. This means if we assume that we apply an occlusion constraint when interpreting monocular regions, it would not matter whether a monocular region was above or below a binocular region and whether the retinotopic counterpart could be binocularly matched or not. Since this kind of monocular region could not be produced by a natural occlusion it would be treated as invalid. If we assume the interocular suppression is caused by mutual correlation then, monocular regions above and below binocular regions would lead to interocular suppression if their retinotipic counterpart was

15

monocular, too.

A cartoon of the stimulus Shimojo and Nakayama (1990) used to investigate this question is shown in figure 1.7.



Figure 1.7: Example cartoon of simulus used by Shimojo & Nakayama (1990). The red regions where binocularly visible. The green regions were monocular. The colours were chosen to make any interocular suppression more noticable, the texture and the thick dark stripes in the centre and at the top and bottom were added to facilitate binocular fusion.

The condition shown in figure 1.7 contains vertical disparities, the central region would have been rotated by 90° to contain horizontal disparities.

Participants were always presented with two monocular regions on both sides of the central rectangle and were asked to indicate which one of the two 'unpaired stripes' appeared more rivalrous in the 9 seconds the stimulus was presented for.

The results for the horizontal disparities mirrored those for the previous experiment, leading to interocular suppression of the invalid monocular regions. For the vertical disparities, the results point strongly towards the idea that monocular regions are interpreted based on occlusion constraints. Participants responded close to chance, indicating that neither of the two monocular lines appeared more suppressed than the other.

The results of these two experiments support the hypothesis that our interpretation of monocular regions depends critically on eye-of-origin information. This means that we interpret the monocularity of a region in relation to the local depth signals before arriving at the decision that a monocular region is valid or invalid.

While Shimojo and Nakayama (1990) focused on under which conditions monocular regions are interpreted as valid/invalid based on occlusion constraints, they then took these findings (Nakayama and Shimojo, 1990) and asked how these occlusion constraints are interpreted by the visual system when building a 3D percept of a scene.

When viewing a binocular scene that contains a monocular region that is caused by occlusions there is, due to the overall scene geometry, a range of distances at which the scene elements in this region can lie at. Figure 1.8 shows this 'depth constraint zone'. Since we have no disparity information about points in this zone the first question we have to ask whether we arrive at a consistent qualitative depth percept from those points.

17





To do so Nakayama and Shimojo (1990) showed participants, in a first experiment, stimuli that contained both valid and invalid occlusion situations and asked which one of two monocular bars appeared further from a binocular square. The stimulus they used is shown in figure 1.9.



Figure 1.9: Illustration of stimulus used by Nakayama & Shimojo, 1990. The top stimulus is an example of a valid stimulus, the bottom of an invalid stimulus. The distance between the binocular rectangle and monocular bar was varied.

Participants were asked in which one of the two simultaneously presented stimuli the bar 'appears to be more in front' (Nakayama & Shimojo, 1990). Valid unpaired bars appeared further away than invalid unpaired bars.

The question remaining is whether we are able to compare the depth percept we have of the monocular regions to a binocular region. Nakayama and Shimojo (1990) tested this in their second experiment by using a stimulus similar to the one used in experiment 1. This is shown in figure 1.10. Participants were asked to 'adjust the disparity until [the] perceived depth [of the binocular bar] matches that of the monocular bar' (Nakayama & Shimojo, 1990).



Figure 1.10: Illustration of stimulus used by Nakayama & Shimojo, 1990. Participants could adjust the disparity of the binocular bar until it matched that of the monocular bar.

For the invalid monocular situation, participants' disparity matching functions remained flat at 0 suggesting that not only do these invalid monocular lead to a very limited depth percept (see their previous experiment) but that this depth percept (if we can even call it a depth percept) is also not comparable to a binocular depth percept. For the valid monocular situation, on the other hand, the perceived distance increased corresponding to the closest possible theoretical distance that is compatible with the occlusion constraint, demonstrating that we use our knowledge of occlusions to estimate the perceived depth of the monocular regions.

However, the depth percept from monocular regions does not rely on a visible binocular occluder. In a final experiment they showed sparse dot stereograms in which monocular dots bounded a surface that contained no disparity information. Figure 1.11 shows such a situation.


Figure 1.11: Schematic illustration of the stimuli used in Nakayama & Shimojo (1990). The rectangular blank region (which is described by the dashed line in this schematic) is bounded by valid monocular regions.

After demonstrating that monocular dots can lead to a depth percept even if there is no disparity information about the actual occluders present, Nakayama and Shimojo (1990) showed participants stimuli similar to the one presented in figure 1.11. Two conditions were presented. In the first one a monocular region indicated an edge (in this condition the monocular region was bounded by binocular regions on both sides) in the second one participants viewed a stimulus just like the one schematised in figure 1.11, only that the blank region was triangular. In both conditions the background was black and the dots were white. Participants were asked to describe any depth perceived when viewing these stimuli. All participants saw subjective contours at least in one of the displays. This means we are able to use the presence and location of monocular regions to deduce the shape and size of an occluder that would have to be present to cause the monocular regions themselves.

Having access to eye-of-origin information is crucial for all of these tasks. Starting with the distinction between valid and invalid monocular regions

and continuing with the location of a potential occluder, we need to know which eye is providing the monocular region to be able to decide whether to use the information contained in a monocular region and at which side of an occluder it lies.

In these two papers Nakayama and Shimojo showed not only that we have access to monocular regions but also, and more importantly, that we are able to distinguish between valid monocular regions that could have been caused by a natural occlusion and invalid occlusions that could not have been caused by a natural occlusion. This means we use either innate or learned knowledge about visual scenes (Nakayama & Shimojo, 1990) to decide whether we will continue processing a monocular region or whether we will treat it as noise and try to suppress it in our percept.

This way of interpreting monocular regions has far reaching effects. In the plausible situation that a monocular region is placed next to a foreground and a background (both of which are identifiable as such due to shading, colour or texture differences (Harris & Wilcox, 2009) then the monocular region is interpreted as lying at the same depth in space as the background plane (Anderson & Nakayama, 1994; Collett, 1985; Julesz, 1971; Nakayama & Shimojo, 1990). In geometrically plausible situations, depth perception is faster (e.g. Gillam & Borsting, 1988; Saye & Frisby, 1975; Wilcox & Lakra, 2007). Divergently, if the 3D-scene geometry of the placement of a monocular region is less plausible, such as when the monocular region shares the texture of the foreground, we perceive less depth in a display (Grove, Gillam, & Ono, 2002)

and depth perception is slower (Grove & Ono, 1999).

Monocular regions clearly affect our percept of the binocular regions around them, but we can also perceive depth based on the monocular regions themselves. I will now describe some of the research that has investigated our depth percept based on monocular information.

## 1.3.2.2. Arriving at a depth percept based on monocular information

As the experiments by Nakayama & Shimojo have shown, valid monocular regions lead to a strong qualitative depth percept (see also e.g. Gillam, Cook & Blackburn, 2003). There is, however, a debate as to whether we can also arrive at a quantitative depth percept based on information contained in monocular regions alone.

Liu et al. (1994) suggested they had found evidence for quantitative depth from monocular regions, but Gillam (1995) and Liu et al (1997) find that the quantitative depth perceived in Liu et al. (1994) could be explained by stereoscopic matching of the binocular regions. Gillam & Nakayama (1999) used a stimulus that eliminated all these possible stereoscopic matches and found that even when we discount conventional stereopsis, there still seems to be a binocular process that allows us to perceive quantitative depth by being sensitive to the pattern of contours in the two eyes. They argued that the size of the monocular region (i.e. the size of the gap between two matchable binocular regions) was used as a cue. Gillam et al (1999) found that this process works even when no contours can be matched and that the perceived depth is comparable to that perceived from conventional stereopsis. Gillam and colleagues argue that this is a binocular process because we can use the size of the monocular gap as a cue.

Tsirlin and colleagues (e.g. Tsirlin et al., 2010, personal communication), on the other hand, argue in so far unpublished work (other than the aforementioned conference abstract) that the visual system is able to use the monocular information to estimate the magnitude of the depth.

While it seems clear that the visual system uses monocular regions to supplement our depth percept, and might even able to arrive at a depth percept independently of binocular information, it is as yet unclear how the visual system uses this monocular information in other situations. More specifically, if I am viewing a surface that is partly visible binocularly and partly only visible to one eye, what evidence is there that my visual system is using information from monocular regions to perceive more of the surface? Or, does the visual system attempt to 'fill-in' the information that is not accessible binocularly? In this context filling-in describes how the visual system perceptually interpolates information about the surface (Durgin, Tripathy & Levi, 1995) around such regions to make an educated guess about what the region is most likely to contain. This process is called amodal completion.

I will now review some of the research on on amodal completion and highlight how this could be how the visual system treats monocular regions.

#### 1.4. Amodal completion

In general, visual systems are very good at filling-in what is hidden (or occluded) in a visual scene and forming a representation of it. For example, we complete partly occluded objects in our percept (this process is called amodal completion (Michotte, Thinès, Costall, & Butterworth, 1991) which means even though we have no visual input as to what these occluded regions actually 'look' like, we perceive a coherent, complete object. This does not mean we consciously 'see' the amodally completed region but we can react to it as if it were visible. For example, I can pick up a cup of coffee that is standing behind my computer screen even if I can only see half the cup and the other half is occluded. This stands in contrast to modal completion where we perceive an actual measurable change in our percept (e.g. of luminance when making judgements about Kanizsa figures - see figure 1.14). Amodal completion of objects can affect texture segmentation (e.g. He & Nakayama 1994), and improve pattern discrimination (e.g. Gold, Murray, Bennet & Sekuler, 2000) which I will describe in more detail in the following sections. But, more importantly for the question of how we perceive monocular regions, amodal completion highlights that there is a marked dissociation between the two retinal images and the percept we form based on this visual input.

### 1.4.1. Amodal completion and texture segmentation

#### 1.4.1.1. Effects on performance

He and Nakayama (1994) asked participants to rapidly discriminate between two textured surfaces of a visual display. One surface was defined by a number of black, textured, elements, the other was defined by a number of white bar or L shapes. The task was to judge the orientation of the texture that was defined by the white L shaped elements. They manipulated the binocular disparity of white bars and L shapes in relation to bordering (black, textured) squares. The two possible spatial locations the black squares and white elements could be located in are shown using the L shapes as an example in figure 1.12. If the disparity was uncrossed (figure 1.12.b), which is consistent with the textured black square being located in front of a white square, participants had more trouble discriminating the different textured elements in the scene than when the disparities were crossed and a white L was perceived in front of the square (Figure 1.12.a).



White 'L' in front

White square behind

Figure 1.12: The two eyes' views can be cross-fused. After He & Nakayama, 1994.

This difficulty seems to be caused by the visual system interpreting the elements as part of an occluded object. For example an L-shaped texture

located in the background would be interpreted as the visible part of a rectangular object. But more importantly, the perceived difference between the plane of the black squares and the white Ls appears less distinct.

This difficulty also extents to participants' reaction-times. I will discuss this in the next section.

#### 1.4.1.2. Effects on reaction-times

In a a study very similar to the one described in their 1994 paper, He and Nakayama (1992) asked how amodal completion affects our ability to search for a specific shape. They used a stimulus in which a white L-shape (either L or mirror reversed L) bordered on a black square which contained several smaller white squares. Participants viewed several of these stimuli on a computer screen through shutter goggles and were asked to find a target L-shape. Disparity was added so the L-shape could either appear in front or behind the black square. This manipulation lead to observers perceiving either an L-shape occluding part of the adjacent square (in front – figure 1.13a) or a white square being partially-occluded by the black square (behind – figure 1.13b). The third (control – figure 1.13c) condition is set up so the stimulus cannot be interpreted as an occlusion situation and the L-shape could either be presented in front or behind the plane of the black square.



Figure 1.13: The two eyes' views can be cross-fused. After He & Nakayama, 1992.

Visual search latencies increased, and accuracy was reduced, when participants were searching for a target L-shape that appeared to lie behind a black square. There was no significant difference in response latencies when an L-shape that could not be interpreted as part of a partially-occluded larger surface was either presented in front of or behind the black square (third condition). This suggests that the difference between the first two conditions is not due to differences in disparities. The amodal completion of the L-shape in the second condition (b), which leads to it being perceived as part of a white background square, makes the visual search harder. Note, that monocular regions do not always affect our performance negatively, but with this task where participants are asked to search for an L-shape, perceiving a white rectangle instead of the L is a disadvantage. What these results suggest is that by the time that we start searching for the L-shape, amodal completion has already occurred.

This kind of filling-in of information about the shape of an object (that the

L that is shown is interpreted as being part of a rectangle) does not only occur with objects that appear to be occluded, but can also be observed when there is no actual object present that can be occluded.

### 1.4.2. Amodal completion of subjective contours

In the case of Kanizsa Figures (Kanizsa, 1979), such as the Kanizsa Square in figure 1.14, the object that is perceived (a white square in front of 4 black circles) is only defined by the surrounding cut-out circles.





Figure 1.14: Example of a Kanizsa Square

Gold et al. (2000) used the subjective contours found in Kanizsa Squares to investigate the effect of amodal completion on the perceived shape of a given subjective contour. Subjective contours are defined by the surrounding contextual information, not by actual contours. The neurons in visual cortical areas do, however, from V2 onward, respond to subjective contours as if they were real contours (von der Heydt et al. 1984; Bakin et al., 2000). If the angle of the cut-outs of the circles in a Kanizsa Square is varied, the perceived subjective contours corresponds to either a square that appears to be squeezed together in the middle (this situation could also be called thin) or a called fat).

Gold and colleagues (2000) manipulated the presentation of the Kanizsa Figures in order to investigate the effect of amodal completion on perceived shape. Figure 1.15 shows the different conditions the stimuli could be displayed as (from left to right): the Kanizsa Squares could be displayed as 'real' objects (the inducing circles were present as well as lines in the location where we observe the subjective contours), as defined by the subjective contours (only the inducing circles were presented), as 'occluded' objects (the inducing circles were presented but with a line completing the circle, as if looking at the inducers through cardboard cutouts), as 'textured, occluded' objects (same setup as in the occluded condition but with a stripey pattern added to the entire surround and the region that is normally interpreted as the square between the inducers), or in a 'fragmented' display (the inducing circles were presented, however the light wedges that induce the subjective contours were not aligned thus leading to no percept of the contours).



Figure 1.15: The different conditions used, after Gold et al. (2000).

Participants were asked to indicate whether a presented stimulus was 'thin' or 'fat'. Their performance indicates that, other than for the fragmented instances, their response is based on the subjective contours that were induced by the cut-out circles. The fact that participants do not only respond to the subjective contours when they are presented in the 'classical' fashion of a Kanizsa square, but also to the occluded versions (both 'occluded' and 'textured occluded') shows that subjective contours are perceived not only in cases when we are presented with part of an object but also when the object or shape causing the subjective contour is not a physical stimulus but itself has been amodally completed.

We are not consciously aware of amodal completion in everyday life. One of the most striking examples of this is the blind spot which I will discuss in the next section.

### 1.4.3. Amodal completion of missing information

While I have so far discussed how our visual system perceptually completes objects that appear occluded by other objects (whether they are 'real' objects or are defined by subjective contours), there are situations when the visual system receives no input about a specific region on the retina. We have no photo-receptors on the retina where the optic nerve attaches to the retina. This blind spot is perceptually filled-in such that we do not notice the lack of information in this region (e.g. Ramachandran, 1992; Walls, 1954) in our visual field. The same seems to happen in the case of scotomas (pathological blind spots) (Ramachandran & Gregory, 1991). Patients very rarely notice the visual impairment unless they are specifically tested for their visual deficits (Ramachandran, 1992).

Amodal completion raises the issue that there is a marked dissociation between the retinal images and the percept we form based on the visual input. This is crucial because it means that even if we are not presenting a complete object in a stimulus, an observer might be reacting to one. So if we were to completely ignore monocular information, we might not notice a difference in our percept if we were simply completing the ignored regions amodally.

In real occlusion situations, this problem becomes even more complex. Consider the following example – if you are holding up a coffee cup in front of you a part of the scene will be occluded from your view by the coffee cup (similar to the occluder in figure 1.5). The most striking difference between the cup and an amodal completion stimulus like the one used by Gold et al. (2000) (see figure 1.15), is that the cup is located at a different distance from us than the background we are interested in. So rather than having a dichoptic occlusion (both eyes receive the same input), the occlusion is in a slightly different region in the two eyes' views. This is illustrated in figure 1.5 which shows a pair of eyes, a foreground occluding object, and a background. Most studies of amodal completion use stimuli where the two eyes view identical scenes. As figure 1.5 demonstrates, this often does not occur in natural occlusion situations.

So what happens in a scene that corresponds to a natural occlusion situation?

1.4.4. Amodal completion of information in monocular scenes When we view a binocular scene that is consistent with naturally occurring

occlusions (the two eyes views are slightly different), amodal completion appears to be faster than when the two eyes views are identical (Bruno, Bertamini & Domini, 1997). I will discuss this in more detail below.

### 1.4.4.1. Amodal completion with monocular regions.

Bruno and colleagues (1997) asked whether amodal completion only occurs when the two eyes view the same input, and the occlusion is thus only defined by pictorial cues, or whether amodal completion could also be observed in stimuli that correspond to what they called an ecologically valid 3 dimensional scene. To do so they investigated the priming effects of visual scenes. The prime could be either a prime polygon that resembled the test stimulus, a prime that extended into a monocular region, or a 'mosaic' prime that had the same shape as the visible part of the occluded prime but could not be interpreted as an occlusion. Figure 1.16a) shows the different prime conditions while figure 1.16b) shows the different test patterns used in the experiment.



Figure 1.16: a) Types of primes used. The primes could either be presented in their pictoral form (as shown here) or with disparity added to the plygon. b) The test patterns used. Both after Bruno et al. (1997).

Participants were asked whether the two test patterns were either the same or different. Bruno and colleagues (1997) found that adding disparity information (including monocular regions) led to participants response latencies showing no significant difference for the complete and occluded primes - neither when the prime was presented for 600ms, nor when it was only presented for 100ms. This indicates a representation of the occluded polygon is comparable to that of the complete polygon, even at very short presentation times.

These results neither support nor refute the idea that, rather than using monocular regions to build the overall percept, participants might simply fill-in the region 'behind' the occluder. However, what they show us is that we can expect the percept of a monocular region in a binocular scene to appear comparable to that of a binocular surface, even at stimulus presentation times as short as 100ms suggesting that their integration into the overall percept

occurs at the early stages of visual processing. This is consistent with the suggestion of Nakayama and Shimojo (Nakayama & Shimojo, 1990; Shimojo & Nakayama, 1990).

While I have discussed how monocular regions affect and relate to amodal completion, there is another issue to consider. When we view a display containing monocular regions, the two retinal images will be completely different for some parts of the image. How is the visual system able to match the correct binocular regions instead of incorrectly matching a monocular region with an adjacent binocular region? Under laboratory conditions we find that we often have trouble in doing so. In the next section I will discuss what happens if the two eyes receive completely different input.

#### 1.5. Rivalry

When the two eyes view totally different items, our percept tends to be very different from that of a natural scene. We observe a switching between the two eyes' views called binocular rivalry (see e.g. Alais, O'Shea, Mesana, Alais & Wilson, 2000, Blake, Lee & Heeger, 2009, Blake & Logothetis, 2003). Usually, one image is perceived (described as dominant) and the other is described as being suppressed.

However, we can think of a situation when the two eyes' views are consistent with a binocular occlusion scene, yet there is no binocular information about the object in the background. Imagine looking through a fence that is placed at a specific distance in front of a background. Figure 1.17 shows such a scene. The placement of the fence leads to all visible background regions being fully monocular. Would we expect to find a rivalrous percept or do we perceive a continuous surface in the background? Or, more general, how is our percept of any monocular region different from that of a region that is amodally completed or from one that leads to a rivalrous percept? Is there even a difference?

In the following section I will discuss an experiment that poses the question whether a stimulus in which the entire background visible behind a binocular fence is monocular leads to a rivalrous percept.

#### 1.6. Extreme half-occlusions

Forte and colleagues (2002) asked if, when the two eyes are shown two stimuli in which each point in the background can be seen only by one eye, one sees a stable percept as opposed to experiencing binocular rivalry.

#### 1.6.1. Geometry

Taking visual occlusions to the extreme, by spacing their occluders so that no part of their background stimulus could be seen by both eyes, Forte et al. (2002) presented vertical strips of filtered 2 dimensional white noise to their participants. Figure 1.17 shows left and right eye views of a scene, illustrating the principle of the stimulus setup used by Forte and colleagues. Note how the

entire Christmas scene in the background is visible behind the occluders, though all in monocular sections.



Figure 1.17: An example of a set of fence-like occluders that lead to only monocular regions being visible. The two images are set up to be cross-fused.

This stimulus setup represents input that could only occur if a set of occluders was very specifically spaced out in front of a background. Figure 1.18 illustrates how real world occluders would have to be spaced out so that one eye would see half of the background, and the other eye would see the other half. This scene is a view from above of such a setup. The solid red lines mark what is visible to the left eye, the dotted blue lines mark what is visible to the left eye.



Figure 1.18: This figure illustrates how only monocular regions of a stimulus can be presented using fence-like foreground occluders.

Note that this is a highly specific type of stimulus. The monocular regions are only spaced like this if the display and the occluders are at a specific distance from each other and the eyes. If any of the distances were to be changed independently of each other, there would either be binocular areas visible as well, or some completely occluded regions.

## 1.6.2. Stability of the percept of extreme half-occlusions

Forte and colleagues specifically chose this stimulus setup so that they could study monocular regions in isolation from binocular ones. They found that observers could distinguish a background stimulus with a discontinuous texture from stimuli with a continuous texture in the background. Their results, they argue, suggest that the visual system pieces together a version of the 'original stimulus' from the two unmatched, monocular, views. In other words there is a clear representation of the whole scene, almost as if both eyes had viewed the whole background without occluders. However, the details of whether the representation is actually as complete as that for a full binocular view is not known, because neither these authors or anyone else has investigated these findings quantitatively rather than qualitatively.

Forte et al also showed that a stable scene was only perceived when the stimuli were consistent with natural occlusions. For example, if the two eyes' views are swapped, there is no real occluder that would be consistent with what is viewed. For such conditions, the scene was not perceived as stable. This suggests that we are able to extract information from such a scene (be it disparity information or based on some learned constraints about occlusions) that indicates to us whether we are viewing an occlusion situation or whether the two eyes view unrelated images. While our percept of monocular regions might be based on input that is not very different from input that causes a rivalrous percept, at some point during processing we distinguish between plausible and implausible input and the actual percept seems to be quite a different one.

While I have so far discussed whether monocular regions are consciously perceived and how they might be processed, what I have not focused on is how our percept of monocular regions differs from that of

binocular information surrounding them. I will now discuss previous research that has investigated what monocular regions 'look like'.

### 1.7. Our percept of monocular regions

Ono and colleagues (2003) considered an observation made by Leonardo Da Vinci (Kemp, 1989, cited by Ono et al,2003), namely that a painter is unable to depict two objects behind each other via a single viewpoint, and yet we are capable of perceiving them as being located behind each other when viewed with two eyes. This observation raises the question of how we are able to obtain a unified view of the world from the two (non-unified) views our eyes supply us with.

#### 1.7.1. Integration of monocular information into the overall scene

First, Ono et al (2003) evaluated how a small rod obstructs areas of a relatively distant background (text in their case) behind it, and how the visual system integrates the monocular information with the surrounding binocular information in these cases. They did so by asking participants to describe their perception of the text and the rod when either fixating the rod or the text (see figure 1.19). This demonstration can be reproduced by holding up a pen in front of a written text (for example as you read this chapter now). As can be observed when looking at the text and not fixating on the pen, the pen was 'seen as double and blurry' (Ono et al., 2003).



Figure 1.19: Stimulus setup for first demonstration in Ono et al. (2003).

Their next experiment, however, led to more striking results. If the distance between the rod and the background was reduced the percept of the stimulus changed.

### 1.7.2. Perceptual compression of space in monocular regions

When participants were asked to align a binocular comparison line with either a monocular line (see figure 1.20a) or the edge of an occluder (see figure 1.20b), the monocular regions were perceived as compressed in comparison to the binocular regions around them. At 'near fixation' when participants were fixating on the black occluder, when comparing a monocular line with the binocular comparison line, the monocular line appeared 1.31 min arc displaced outward (compared to when fixating on the background) whereas when the occluder edge was compared with the binocular comparison line it appeared 0.81 min arc displaced inward towards the middle of the screen (again, compared to when fixating on the background). Note that the compression is

very small and just holding up the previously mentioned pen will not allow you to detect the compression.



Figure 1.20: Participants were asked to allign a binocular comparison line with either a monocular line (a) or the edge of the occluder (b). Both situations were conducted at near- and far-fixation. After Ono et al. (2003).

From this we can again conclude that monocular information is incorporated into the overall percept. Ono and colleagues then asked whether the displacement they observed was the effect of the visual system trying to construct a more unified view of the world.

# 1.7.3. Integration of monocular regions into a unified percept

In their second experiment, Ono and colleagues presented participants with four thin vertical lines between two rectangles that served as occluders for the outside lines. The two inner lines were binocular. On top of the two occluding rectangles was a bridge (i.e. another occluding rectangle connecting the two rectangles on their top ends) on top of which a binocular comparison line was presented. The comparison line was randomly located above any of the lines below the bridge or the two inner edges of the occluder rectangles. Participants were asked to judge whether the comparison line was to the left or the right of the standard line. The set-up is depicted in figure 1.21.



Figure 1.21: Participants were asked to allign a binocular comparison line with either a monocular line (a) or the edge of the occluder (b). Both situations were conducted at near- and far-fixation. After Ono et al. (2003).

Monocular and binocular lines, as well as the occluding edges, were perceived at a slightly displaced position form their real position. More specifically, both the binocular and monocular lines appeared displaced inward toward the middle of the screen while the occluder edges appeared displaced outward. These findings are consistent with our visual system trying to achieve a more centred view of the world *without* losing the information our binocular visual system supplies us with.

## 1.7.4. Do we form a cyclopean percept?

Ono and colleagues (2003) suggest that compressing and displacing the background slightly allows us to construct a unified view of the world that includes both binocular and monocular information. Erkelens and van Ee (2002) argue that a unified (or cyclopean) view of the world is not necessary for us to perceive visual directions. There is a current debate about whether, and if so, how the visual system combines the two eyes' views into a unified cyclopean view before interacting with it. For more information on this debate refer to: Mapp & Ono (1999); Erkelens (2000); Erkelens & van Ee (2002); Ono, Mapp & Howard (2002 in Howard & Rogers, 2002); Ono, Wade & Lillakas (2009).

### 1.8. Comparing monocular and binocular regions

While Forte et al. (2002) investigated the processing of monocular information produced by extreme half-occlusions; Ono et al. (2003) were concerned with smaller monocular regions. Little research has been conducted, however, on the differences in processing of these two types of information and how they relate to binocular regions. In one of the many studies investigating the role partial occlusions play in depth perception processes, Pianta and Gillam (2003) suggest involvement of a common mechanism in the processing of monocularly occluded/binocular information for depth perception. However, their experiment investigated the possible differences in processing of depth signals generated by regions containing these different types of information. This means they were looking at whether the depth signals provided by the two types of regions were processed differently rather than the appearance of information in monocular regions compared to information in binocular regions.

No other research seems to have been conducted on the matter of to what extent monocular regions are processed differently than binocular regions, and whether it matters if these monocular regions are embedded in binocular ones or are of the extreme form used by Forte and colleagues (2002). Because the research literature is scant, many questions are ripe for exploration. One of the unanswered questions to date is, whether monocular and binocular regions are used in the same way when we are building our final percept. While it seems fairly clear that monocular regions are not ignored, Ono and colleagues' (2003) work suggests that they are not processed in exactly the same fashion as binocularly viewed regions. The experiments described in this thesis explore the question of how monocular information is processed and represented.

#### 1.9. Structure of this thesis

In this thesis I will be asking how our representation of visual information differs depending on whether it is presented binocularly or in a monocular region. The experimental chapters are split into two parts. Each part has its separate

introduction in which I will go into the theoretical and experimental background for each section and will explain the general methods used in the experimental chapter of each section.

The first three experimental chapters investigate potential effects of halfocclusions on relative numerosity/texture density judgements. In these chapters I will ask the following questions:

- Chapter 3: Are we as sensitive to differences between numerosities that are presented in monocular regions as we are to differences that are presented binocularly?
- Chapter 4: Do numerosities in monocular regions appear less numerous than those viewed binocularly?
- Chapter 5: When both monocular and binocular information is present, how do monocular and binocular regions contribute to our overall percept of a stimulus?

Related to the general question of how our percept of monocular regions differs from that of binocular regions, is the question of why we use monocular information in the first place. It has been suggested that using information in monocular regions when forming our overall percept might have been a useful development for detecting predators (Changizi & Shimojo, 2008). If this is the case then information in monocular regions will influence our visual search behaviour.

The last two experimental chapters will describe two visual search studies

designed to investigate the effect of half-occlusions on our performance and behaviour during a visual search task. These chapters will ask the following questions:

- Chapter 6: When performing visual search, are we able to use information in monocular regions as an additional cue to cues along other dimension (e.g form)?
- Chapter 7: Does the presence of monocular items change our eyemovements during a visual search task?

# Part I – Perceiving monocular regions in binocular scenes.

#### 2. Introduction

In this first group of experimental chapters I will focus on the perception of monocular regions in binocular scenes. I will ask whether we perceive monocular information differently than binocular information, whether we are able to integrate the two types of information, and whether we might weight them differently if they are conflicting. This chapter serves as an introductory chapter to the concept of numerosity discrimination in general, as well as to the following experimental chapters.

#### 2.1. Aims

This chapter will describe the general methods for the relative numerosity experiments in chapters 3-5. I will describe the stimuli, the setup, and will explain how I analysed the data from these experiments using a generic example.

Before I go into the stimulus-specific methods, I will introduce the topic of numerosity perception in more detail. While the question of how numerosities are processed and perceived is an interesting one in itself, I decided to use a task in which participants compare two sets of dots to decide which has more: the larger numerosity (a relative numerosity task) to investigate how our visual system deals with monocular regions that are embedded in binocular scenes. There is a debate on how relative numerosities are perceived and processed. I will mention this debate, but I am agnostic to this issue for the purpose of the experiments described here. The experiments in chapters 3-5 were not designed with the intention of investigating how numerosities are perceived but rather use relative numerosities to ask questions about how monocular regions are perceived.

I will now give an overview of previous research on relative numerosities that are relevant to our question at hand – how information contained in monocular regions is processed and represented by the visual system when it is embedded in a binocular scene. I will then discuss how relative numerosity judgements can be used to investigate how the visual system deals with monocularly presented information that is embedded in binocular information.

## 2.2. Numerosity Perception

#### 2.2.1. A Sense of Number

Perceptionists have long been interested in how we interact with sets of objects. Early work by Jevons (1871) focused on how many different items observers were able to simultaneously attend to. This was further elaborated on by Kaufman et al. (1949) who introduced the distinction between subitising (a small number of items can be immediately assessed) and estimating of a numerosity.



Figure 2.1: Example of a numerosity within the subitising range.

Within the subitising range (see figure 2.1 for an example) observers are able to consistently and accurately enumerate one to three items. Between four and nine items the error rate linearly increases while reaction times slow significantly (e.g.Mandler & Shebo, 1982). The lower limit of the subitising range coincides with the capacity of the visuospatial sketchpad and the episodic buffer (e.g. Cowan, 2005) in the context of the multi-component model of working memory (Baddeley & Hitch, 1974; Baddeley, 2000). In this context Luck and colleagues (Luck & Vogel, 1997; Vogel, Woodman & Luck, 2001) observed that features (such as simple geometrical shapes) of a visual display can automatically be bound into objects if they appear to relate to each other (for example, due to proximity or similarity caused by size, shape, orientation, or colour). They found an upper limit of four objects that can be immediately assessed and enumerated irrespective of how many features had been bound into one of these objects. This pattern of performance is also found for numerosity judgements (van Oeffelen & Vos, 1982).

Above the subitising range our ability to accurately assess a single

numerosity drops exponentially or linearly along the log-scale (van Oeffelen & Vos, 1982).

### 2.2.2. Relative Numerosities

If you ever tried to judge the number of marbles in a jar at a fundraiser you will have noticed how bad we are at this task. Similarly, if we are asked to estimate the number of dots in a display such as in figure 2.2 we struggle greatly with this task.



Figure 2.2: Example of a numerosity above the subitizing range.

However, we are a lot better at comparing different numerosities. For example, we are able to reliably discriminate differences as small as that between 25 and 28 dots (van Oeffelen & Vos, 1982).



Figure 2.3: Example of a numerosity discimination above the subitizing range.

Note how easy it is to work out the difference between the two sets of dots in figure 2.3 (right side has more dots) compared to working out the absolute number (97 dots) displayed in figure 2.2.

So how does the visual system discriminate between two numerosities? There is a debate about whether we have direct access to the number of elements within stimuli or whether we use the texture density of said stimuli to make a judgement. I will describe some of the key aspects of the debate in the next section. Note, however, that the question of whether we use texture density or numerosity directly does not affect the task I am using such a stimulus for.

#### 2.2.3. A sense of number vs. texture density

When comparing different numerosities there are several putative mechanisms. Durgin (1995) suggested that we use the 'texture density' in this kind of display to discriminate between two stimuli. This would mean that we

use the number of dots per unit area rather than processing the overall number to distinguish between two displays. Durgin (1995) found significant adaptation to the density of a presented cloud of dots which means that density of a comparison stimulus was perceived as less dense if participants had adapted to dense displays of dots.

Durgin's (1995) experiments showed that if we adapt to a dense texture, this affects our numerosity judgements of a test numerosity – we underestimate how many elements are present. This effect is stronger at larger numerosities, but is even present at a level just above the subitising range. Durgin (1995) suggests that relative numerosity judgements are made using texture density.

This texture density adaptation is binocular (Durgin, 2001) which means there is significant interocular transfer (if we adapt to a dense texture in one eye, we will underestimate a test numerosity that is presented to the other eye). Interestingly, this stands in contrast to texture contrast adaptation where Durgin (2001) observes only monocular aftereffects. These findings suggest that the comparison between two numerosities is made at a stage when the two eyes' views have already been fully combined. Along the visual pathway this is the case once information has been processed by complex cells (in V2, note that there are complex cells that respond to binocular information in V1 as well but this information does not seem to get passed on to the extent that information is passed on from V2 (see Cumming & Parker, 1997) – for example when we view anti-correlated random dot stereograms, complex cells in V1 are

activated by the disparity information contained in the random dot stereogram, but this does not lead to a percept of depth, whereas complex cells in V2 respond to the disparity signal in correlated random dot stereograms which also lead to a depth percept (see e.g. Cogan et al. 1995; Cumming et al. 1998)). This lead Durgin (2001) to propose that texture contrast judgements could be made at the stage of simple cells, whereas texture density coding had to be located at the stage of complex cells or even later in the pathway.

Burr and Ross (2008), on the other hand, argue that they have evidence for adaptation to number rather than density. They suggest that the adaptation to Durgin's dense textures is actually adaptation to the numerosity itself. They view our percept of a numerosity as a very direct one, comparable to qualia (Jackson, 1982). Burr and Ross (2008) tested adaptation to varying sized dots and found no effect of the element size on the overall adaptation effects which, they suggest, shows that observers cannot be reacting to the texture density.

However, as Durgin (2008) also points out in a response, Burr and Ross did not change the overall size of the stimulus, so changing the size of single elements changes the overall density of the texture as well. He tests this in a follow-up experiment in which observers adapt to two stimuli, one of which is more numerous, while the other is more dense. Figure 2.4 is a demonstration of the experiment.



Figure 2.4: A demonstration of the stimulus used by Durgin (2008). Fixate on the cross in the top row for 30 seconds, then fixate on the cross in the bottom row. The top left stimuls is more numerous than the top right stimulus, which is more dense, after adaptation, the bottom left stimulus appears more numerous than the bottom right simulus. The two bottom stimuli contain the same number of dots. After Durgin (2008).

This paper lead to another response from Ross and Burr (2008) who suggest that this adaptation only occurs in the adapted region and does not affect judgements of the overall numerosity, i.e. if we were to adapt to a density within a very small area we would not observe the underestimation of numerosity/density we find with the bottom right stimulus in figure 2.4.

Work in the Dakin lab (personal communication, but see Dakin et al. 2010) suggests that a middle ground seems more likely than the two extremes. More specifically, they find that numerosity and density are not judged independently and that the manipulation of one biases our percept of the other.

This debate, while interesting and important for the question of how we perceive numerosities, does not directly affect the question I am trying to answer using a relative numerosity task. The experiments described in chapters 3-5 investigate the differences in the spatial representation of patterns between binocular and monocular regions by utilising relative numerosity tasks. I will measure our ability to compare two sets of elements and then decide which set contains more, and will be specifically interested in changes in our sensitivity and a potential bias in our percept.

Durgin (2001) suggests that numerosity judgements are made at a stage when the two eyes' views have already been combined. I decided to investigate our percept of monocularly occluded regions using relative numerosities for the following reason - if monocular regions are not completely integrated with binocular regions, this should lead to a disruption in our performance when performing a numerosity task. By completely integrated, I mean that observers have a stable percept of monocular regions that does not appear different from that of binocular regions but one that also leads to the same behaviour (e.g. same accuracy and precision when making judgements about their content).

Based on the previously discussed research, I asked how our representation of numerosity differs depending on whether it is presented binocularly, monocularly or in a partially occluded form in a monocular region.
## 2.3. General Themes Motivating the Experiments

Before describing the basic experimental design in more detail I will discuss some general questions and themes that motivated the experiments in chapters 3-5.

# 2.3.1. Do we perceive information in monocular regions the same as information in binocular regions?

The central question in this thesis is whether information in monocular regions is perceptually treated the same as binocular information. In chapters 3-5 I ask this question using relative numerosities - are numerosities that are presented in monocular regions treated the same as numerosities that are presented in binocular regions?

Monocular regions do, theoretically, provide the visual system with the same information that a purely binocular region would. However, since they are presented only to one eye, and because we have to link spatially separate regions to form our percept, they might be treated as less reliable. If monocular regions do not have the same status as binocular regions in our representation of objects and space, we may be less sensitive to these differences in a stimulus containing information in these monocular regions.

# 2.3.2. Is this any different from closing one eye?

Monocular regions are embedded in a binocular scene. Therefore, the

question is whether monocular regions and purely monocular stimuli lead to a similar percept of numerosities presented in them. One way that monocular regions could be treated as purely monocular is if the visual system deals with the occlusions and the conflict between the two eyes' views by ignoring one of the two retinal images. This would be akin to rivalry.

# 2.3.3. Is there a benefit from using two eyes?

In a situation without any occlusions, a binocular stimulus provides the visual system with the same information as a monocular stimulus. It may, however, give the visual system a 'second go' at the presented scene, each eye providing us with an independent input. But, this is mainly useful in scenes with a lot of noise or low contrasts between different components. The prediction is that participants will show no, or very limited, improvement in performance when viewing binocular stimuli compared to monocular stimuli.

I will now introduce the general setup of the following numerosity experiments and the stimulus geometry used for the displays.

# 2.4. Stimulus Geometry & General Setup

# 2.4.1. Apparatus

The stimuli were presented on a liyama 22in Vision-Master-Pro monitor which had a resolution of 1280 x 1024 pixels and had a refresh rate of 100Hz.

The screen was viewed through a Modified Wheatstone Stereoscope (see figure 2.5).



Figure 2.5: Top-down view of the Modified Wheatstone Steroscope setup used for the experiments.

The two sets of mirrors that are placed between a monitor and an observer allow for two views of a stimulus to be presented side by side on the screen, but each visible to only one eye. The viewing distance was 100cm. The head position was stabilised using a chin rest.

### 2.4.1.1. Calibration

I calibrated the stereoscope at the beginning of each experiment and checked the calibration after every second particiant. In this setup, the stereoscope is calibrated as follows:

A 'fixation cross' is displayed in the centre of each half of the screen. In figure 2.5 this corresponds to the actual locations on the screen. Three plumblines are hung from the ceiling above the screen. Two of these reach the left and right eye image fixation crosses and are lined up with them. The third plumb-line ends just above the top of the screen and is lined up with the centre of the screen (this corresponds to the point marked as 'perceived location' in figure 2.5).

The mirrors are calibrated with the screen by viewing the screen through the stereoscope. First the right set of mirrors is adjusted by closing the left eye and by then moving the mirrors until when one looks at the plumb-line starting at the fixation cross moving upwards, one sees a continuous line extending past the top edge of the screen. This means that the 'actual location on the screen' is perceived as lying in the middle of the screen at the 'perceived location'. This is then repeated for the left eye. Finally, the overall percept needs to be checked by viewing the plumb-line with both eyes. If the calibration was successful the plumb-line will appear to be continuous and in the middle of the screen. There will be no changes in the depth of the line and it will not appear slightly shifted to one side once we move past the top of the screen. If any of these occur the calibration has to be repeated to ensure that the perceived location of the presented stimuli is correct.

## 2.4.1.2. Stimulus Generation

The stimuli were generated and presented using Matlab and the Psychophysics Toolbox 3 (Brainard, 1997; Pelli, 1997) on a PC workstation.

#### 2.4.2. Stimuli

While the general setup was the same across all experiments in chapters 3-5, there were subtle differences between the stimuli (especially where the luminances are concerned). I will thus describe the specific stimuli in the methods section of each experiment.

What all of these experiments had in common was that observers viewed two squares (size 6.84°by 6.84°) which were framed by a line (width: 0.42 arcmin) and contained a cloud of dots (0.2 arcmin diameter). The two squares were separated by a gap of 4.56° so, once viewed through the stereoscope, each eye was presented with only one of the squares. The clouds of dots had a disparity of 41 arcmin with respect to the plane of the screen. This corresponds to them being located 18.5 cm behind the plane of the screen.

### 2.5. Procedure

All relative numerosity experiments in this thesis used a 2 Interval Forced Choice (2IFC) paradigm coupled with method of constant stimuli and difference thresholds for the data collection. Participants were presented with a series of pairs of stimuli (trials) that were presented sequentially.

In a given trial participants saw a white fixation cross on the screen for 1 second. This was followed by the first stimulus interval, which was displayed for 0.4 seconds (for different possible stimuli refer to the methods sections of

the individual experiments). This was followed by a second fixation cross for another second and a second stimulus interval for another 0.4 seconds. After this, a third fixation cross was presented until participants made a choice as to 'which of the two stimulus intervals was more numerous'. Figure 2.6 shows this in the form of a timeline.





In each trial, one interval contained a 'standard' in which a baseline dot density was presented and a 'test' in which the number of dots was varied. In experiments 1 and 2 the standard stimulus was presented first and the baseline (see section 2.6.2) was changed at regular intervals. In later experiments the standard could either be presented first or second.

I will now describe the different numerosities used in the experiments in chapters 3-5.

#### 2.6. Presented Numerosities

In order to be able to ask whether we are more or less sensitive to differences between numerosities, given changes in the stimuli, we need a scale that allows us to describe the difference between two numerosities. I will now describe the scale I used in chapters 3-5.

#### 2.6.1 A scale for discriminating numerosities

While there are several different approaches to describing the difference between two numerosities, I have chosen to use the approach suggested by van Oeffelen and Vos (1982). This is based on the Weber threshold, which is defined as the point at which an observer correctly identifies the difference between two numerosities 75% of the time. Van Oeffelen and Vos (1982) do this for a 2 alternative forced-choice (single interval) situation (2AFC). I will first discuss the reasoning behind the scale advocated by Van Oeffelen and Vos (1982) and then explain how this applies to the 2 interval forced-choice (2IFC) task I am using in this thesis.

In a 2AFC task an observer is presented with one test stimulus per trial and is asked to make one of two responses about this stimulus. I will call this stimulus T from now on. When making a response about T we are doing so based on an internal representation of T which we will refer to as t. Van Oeffelen and Vos propose that the discriminability of two stimuli obeys Thurstone's law of comparative judgement (Thurstone, 1927, cited in Van Oeffelen and Vos, 1982) and that we represent the numerosity of a stimulus using a normally distributed probability distribution (the mean of which we will call  $q_t$  and with a standard deviation  $\sigma$ , which is assumed to remain constant across different numbers of dots (n) in T) which explain the pattern of our responses to a given t. Let x be the random variable that represents the internal noise factors acting upon our representation of number. This can be described as follows:

$$t \mapsto \frac{1}{\sigma_t \sqrt{2\pi}} \exp \left(\frac{(q_t - x)^2}{2\sigma^2}\right)$$
(2.1)

Now let us assume that a participant is presented with a stimulus T that contains n dots. The probability that our participant will respond that the number of dots was n is

$$P_{I}(n|n) = \frac{1}{\sigma_{n}\sqrt{2\pi}} \int_{C(n;n-1)}^{C(n;n+1)} \exp(-\frac{(q_{n}-x)^{2}}{2\sigma_{n}^{2}} dx$$
(2.2)

I denotes that any integer is a possible response in this case (i.e. we are not operating in a forced-choice environment, yet) and the limens of the integral denote the category bounds for our responses. Van Oeffelen and Vos point out that our representation of numerosities appears to be on a logarithmic scale so, following Parducci (1963) (cited in Van Oeffelen, and Vos, 1982) the category bounds are placed 'halfway between the internal representation of n and n-1 for C(n;n-1) and n and n+1 for C(n;n+1). This means we can reformulate equation 2 as

$$P_{I}(n|n) = \frac{1}{\sigma_{n}\sqrt{2\pi}} \int_{1/2[\ln(n) + \ln(n-1)]}^{1/2[\ln(n) + \ln(n-1)]} \exp{-\frac{[\ln(n) - x]^{2}}{2\sigma_{n}^{2}}} dx$$
(2.3)

and if we substitute y for

$$\ln(n) - \frac{x}{\sigma_n}, \tag{2.4}$$

we arrive at:

$$P_{I}(n|n) = \frac{1}{\sqrt{2\pi}} \int_{\frac{1}{2\sigma_{n}} \ln\left(\frac{n+1}{n}\right)}^{\frac{1}{2\sigma_{n}} \ln\left(\frac{n+1}{n}\right)} \exp\left(-\frac{y^{2}}{2}\right) dy.$$
(2.5)

Since we made the assumption that an observer's internal representation of a stimulus can be described using a standardised normal distribution, the integral limits are actually z scores. Here Van Oeffelen and Vos (1982) use the fact that we know that  $P_1(7|7)=0.5$  (Averbach, 1963 and Hunter & Sigler, 1940, cited in Van Oeffelen & Vos, 1982), from which we can calculate a value for  $\sigma$  and we find  $\sigma=0.1080$ . Based on this, since we have assumed  $\sigma$  to be constant, we could calculate  $P_1(n|n)$  for all n. In this situation, however, a participant is allowed to choose their response freely. In a situation when a participant is forced to choose between two alternatives, the probability that they will respond with n (the alternative being another response 'm') is:

$$P_{n,m}(n|n) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{1}{2\sigma} \ln\left(\frac{m}{n}\right)} \exp\left(-\frac{y^2}{2}\right) dy, form > n$$
(2.6)

$$P_{n,m}(n|n) = \frac{1}{\sqrt{2\pi}} \int_{\frac{1}{2\sigma}\ln\left(\frac{m}{n}\right)}^{\infty} \exp\left(-\frac{y^2}{2}\right) dy_{s} form < n.$$
(2.7)

If we then define the larger of the two values as 'max' and the smaller of the two as 'min', the conditional probability of the response n given the two alternatives m and n can be expressed as the following general expression:

$$P_{n,m}(n|n) = \frac{1}{\sqrt{2\pi}} \int_{\frac{1}{2\sigma} \ln\left(\frac{min}{max}\right)}^{\infty} \exp\left(-\frac{y^2}{2}\right) dy.$$
(2.8)

Technically, the Weber threshold is defined as the point at which we can discriminate between two stimuli and be 50% correct in our response. Practically, this corresponds to the point at which we respond that one of two stimuli is larger 75% of the time. So, by looking for

$$P_{n,m} = 0.75$$
 (2.9)

in (8), we can thus derive

$$\left(\frac{1}{2\sigma}\right)\ln\left(\frac{max}{min}\right),$$
 (2.10)

from which we can calculate the Weber fraction needed to discriminate between two numbers:

$$W_{b} = \frac{(max - min)}{min}$$
(2.11)

Keep in mind that because we are dividing through the smaller number,

and

the difference between two presented numbers, if m < n is smaller than the difference if m > n for the same n (for more details on this refer to formulae 9-13 in van Oeffelen and Vos (1982) or, for example, to Crossman (1956)).

Since I have so far discussed the 2AFC situation, let us look at the situation when two stimuli are presented and observers are asked to discriminate between the two. In an 2 interval forced-choice (2IFC) task an observer is presented with two stimuli. Let us call them S and T. The observer then has two possible responses they can give to a stimulus. In our case the response would be whether the standard (S) or test (T) interval appeared more numerous. Van Oeffelen and Vos (1982) suggest that our response to a given stimulus, based on our internal representation of said stimulus follows a normal distribution. So, individually, the representation of each stimulus (s and t) can be described as follows:

$$t = N(q_t, \sigma 2) \tag{2.12}$$

$$s=N(q_s,\sigma 2) \tag{2.13}$$

therefore, our internal representation of the difference between the two stimuli is:

$$t - s = N(q_t - q_s, 2\sigma^2)$$
 (2.14)

We are interested in the probability that a participant responds that T is more numerous than S given the two stimuli and if T can be either larger or smaller than S this means we want to know when t is larger than s. Since we are making the assumption that our internal representation is normally distributed with a mean q and a variance of  $\sigma^2$ , we know that the probability of a given x can be described as:

$$P(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2\sigma^2} (x-q)^2\right] .$$
 (2.15)

We are interested in the difference between t and s, i.e., the subject's approximation of the probability that the stimulus T was greater than the stimulus S

$$P(T-S>0) = \int_{0}^{\infty} \left(\frac{1}{\sqrt{4\pi\sigma^{2}}}\right) \exp\left[\frac{-1}{4\sigma^{2}} \left(x - \left(q_{t} - q_{s}\right)\right)^{2}\right] dx \quad .$$
(2.16)

Remember that our representation of q appears to be on a logarithmic scale (Van Oeffelen & Vos, 1982). This means

$$q_t = \ln(t) \tag{2.17}$$

and

$$q_s = \ln(s)$$
 . (2.18)

We are interested in the point at which we are able to reliably judge T as being more numerous than S if it, indeed, is more numerous than S. As before, we will use the difference between T and S at which we respond that t is more numerous in 75% of the time as our threshold. In other words, we are looking for the point at which the probability that a participant responds that t>s is 0.75 and at which the probability that the participant responds that s>t is 0.25.

Since we are operating on a standardised normal distribution, we can look up this point and it is 0.675 standard deviations away from where t = s.

This means:

$$q_n - q_m = 0.675 \cdot \sqrt{2\sigma^2}$$
 (2.19)

and based on the previous equations we arrive at:

$$\ln\left(\frac{n}{m}\right) = \left(0.675\sqrt{2}\right)\sigma\tag{2.20}$$

Van Oeffelen and Vos (1982) argue that the noise in our internal representation leads to  $\sigma = 0.1080$  for most participants so if we assume this number to be accurate then

$$\frac{n}{m} = e^{(0.675\sqrt{2}\sigma)} \tag{2.21}$$

leads to

$$\frac{n}{m} = 1.109$$
 (2.22)

and since we are interested in the proportion difference, i.e.

$$\frac{n-m}{m}$$
(2.23)

which corresponds to

$$\frac{n-m}{m} = \frac{n}{m} - 1 \tag{2.24}$$

our Weber Fraction is

$$W_{b} = 0.109$$
 . (2.25)

Keep in mind, though, that this number is based on the assumption that  $\sigma = 0.1080$ . If a participant's estimate of the two stimuli is more noisy and they thus have a higher  $\sigma$ , this number will obviously differ.

Note that this description has a directionality inherent to it. If the test stimulus is smaller than the standard, the Weber Fraction is negative and vice versa. This allows us to investigate whether there is an effect of comparing a small test with a large standard or a large test with a small standard (see experiments 1 and 2, chapter 3).

Rather than using the underlying numbers of dots (e.g. the difference between 80 and 93 and the difference between 50 and 58), the Weber fraction allows me to compare performance across several baseline numerosities (both number pairs above correspond to the same Weber fraction of approximately 0.16). Since Durgin's (1995) findings suggested that there is a difference in the effect of adaptation depending on the adaptor numerosity, I used 5 different baseline numerosities in my first set of experiments (chapter 3).

# 2.6.2. Varying baseline numerosities

The stimuli contained varying numbers of dots that were randomly distributed within the square frame. In experiments 1 and 2 (chapter 3), the number of dots for the standard stimulus in a trial could be 50, 65, 80, 95, or

110 dots, in later experiments (experiments 2.2 - 4, chapters 3-5) the baselinenumerosity was held constant at 80. The test stimulus could vary by 0, ±16, ±25, ±50 or ±75 percent of the numerosity (rounded to the nearest integer) of the standard display.

While the majority of work on relative numerosities has used a single baseline numerosity, some groups (e.g. Durgin (1995); Dakin Lab at UCL (John Greenwood, personal communication, 2010)) have found that for a lot of situations participants' behaviour varies as the baseline numerosity is increased.

Therefore, I initially decided to test a number of baseline numerosities and to then combine them for analysis *only* if my participants' performance indicated that there was no significant difference between the different baseline numerosities. I will discuss this aspect of the analysis at the beginning of the analysis section for the first two experiments (chapter 3).

## 2.7. Participant Screening

As a pre-test, participants were screened in regard to whether they had normal vision. We were interested in two things – visual acuity and stereo vision. Visual acuity was assessed in order to make sure that all participants had normal vision. Furthermore, we assessed whether participants had normal stereo vision.

Visual acuity was assessed using the Snellen Eye-Chart. All participants

used in this thesis had 25/20 vision or better and could at least identify 50% of letters at 20/20. Stereo vision was assessed using the Butterfly and TNO tests.

The Butterfly test consists of a set of displays that are presented with stereo glasses. The first of the displays consists of displays of 4 rings in which one was given a crossed binocular disparity whereas the others were on the plane of the page (see figure 2.7).



Figure 2.7: Mock-up of first part of the Butterfly test, one of the rings has a different disparity from the rest of the display. Participants are asked to point out the ring.

Participants were asked to point out which one was protruding. This first part of the tests assesses whether observers can detect disparities of 800 to 400 arcsec. All participants were able to correctly identify all protruding rings.

The second part of the Butterfly Test that potential participants completed showed a butterfly and they were asked to describe the shape of it. This second part tests ranges of 2500 to 1200 arcsec disparity. All participants were able to point out the different aspects of the butterfly and differentiate between the different levels of depth.

The TNO Test uses red/green glasses and consists of pacman-like shapes that seem to float on top of a noise texture (see figure 2.8).



Figure 2.8: Approximate appearance of pacman-like shape presented in TNO Test. The shaded shape protrudes from the surrounding background noise.

Participants were asked to point out which way the 'mouth' of the 'pacman' was facing. Through the course of the displays the disparity between the noise texture and the 'pacman' is reduced from 480 arcsec to 15 arcsec. All participants were able to correctly identify the direction of the opening down to 60 arcsec, or better.

# 2.8. Ethical Approval

All numerosity experiments were approved by the School of Psychology Ethics Committee of the University of St Andrews Teaching and Research Ethics Committee. Participants were informed of their right to withdraw from the studies and participants initials were used to code their data. They were given the opportunity to either volunteer for the experiments or to participate for a reimbursement of £5.00 per hour.

### 2.9. Analysis

I will now describe how I analysed the data in chapters 3-5. Let us take a step back, first. The data was collected using method of constant stimuli and difference thresholds were collected.

This means that the difference between the standard and test numerosities was varied in a systematic way (9 levels of difference  $\div$  0,  $\pm$ 0.16,  $\pm$ 0.25,  $\pm$ 0.5,  $\pm$ 0.75 (proportion difference from standard, see section 2.5) and each level was repeated several times. For each condition we thus have a proportion of responses in which the test interval was perceived as more numerous for each stimulus level. Data was analysed by fitting a cumulative probability function (such as in figure 2.9) to the data and then basing further comparisons on the fitted function.



Figure 2.9: Example Psychometric function.

In this plot of the psychometric function the distance between the 50% point which corresponds to the point of subjective equality (PSE) and the 75% point is defined here as the point of just noticeable difference (JND). These

points allow us to make statements about the shape of the psychometric function. The overall shape of the psychometric function depends on the sigmoid and the linear transformation (the 'core', see section 2.9.1.1) that is used to fit the function to the data. However, since we will be fitting the same sigmoid (a cumulative gaussian sigmoid) to our data for all numerosity experiments, being able to compare the PSE and the JND will suffice for most of our data analysis. I will discuss the shape of the psychometric function used in section 2.9.1.1.

For experiments 1, 2, and 2.2 we are interested in the point at which participants reliably choose the test interval as containing more dots when it actually contained more dots and vice versa. For this I will use the Weber threshold discussed in section 2.6.1. For later experiments we will also be interested in whether participants show a bias towards perceiving certain stimulus variations as more numerous than others. In this case we will be interested in the bias of the 50% point (i.e. Point of Objective Equality – Point of Subjective Equality).

# 2.9.1. How was the Psychometric Function fitted to the data?

I will now explain how the data was analysed using a sample dataset. Let us assume that a person participated in a relative numerosity experiment and was asked to indicate whether the test in a given interval was more numerous than a standard. Our hypothetical participant's responses were recorded for 360 trials and we then counted the number of trials they responded that the test interval appeared more numerous for each level of difference between the test and the standard (see table 2.1 below)

Proportion Difference	-0.75	-0.50	-0.25	-0.16	0.00	0.16	0.25	0.50	0.75
Count (out of 40) 'test more'	0	0	8	19	31	36	36	40	40

Table 2.1: Example Data-Set

We could simply plot this as a line-chart (figure 2.10).



Figure 2.10: Line-Graph of Example Data-Set.

We could, however, also fit a function to the data to be able to describe participant behaviour.

# 2.9.1.1. The psychometric function

I used Psignifit 3.0 (Fründ et al., 2011, Wichmann & Hill, 2001 a&b), a Python library, which allows us to fit sigmoidal functions to psychophysical data.

The function we are fitting to the data has the general form:

$$\Psi(\mathbf{x};\boldsymbol{\theta}) = \gamma + (1 - \gamma - \lambda) F(\mathbf{x};\boldsymbol{\theta})$$
(2.26)

Where F is a sigmoid function, with a parameter vector  $\theta = (\alpha, \beta, \gamma, \lambda)$ where alpha and beta describe the inflection point and the slope of the psychometric function, respectively, and  $\gamma$  and  $\lambda$  are the guess and lapse rate respectively. In all experiments in this thesis the parameters were regularised based on the same assumptions. Since we are making certain assumptions about the different parameters (specifically: the slope, and lapse- and guessrates of our psychometric function are positive) we could either hard-code this by constraining the possible values to be either in the desired range or to be discarded; or, we could 'regularise' the parameters by applying so called soft constraints. What this does is make certain values (the values we assume are most likely) more likely without imposing a stark cut-off point for the less likely values. Regularising parameters has, in this context, the same effect as the assumed prior distribution in a Bayesian framework. Since the description using assumptions about the prior distribution seems more intuitive I will be using this terminology for the non-Bayesian context as well.

For  $\alpha$  a broad Gaussian distribution centred at 0 was assumed. For  $\beta$  a Gamma distribution was assumed. A Gamma distribution is only defined for values greater than 0. This means by using a Gamma distribution we are

making the assumption that the slope of our function will be positive. In our case it makes sense to assume this because we expect participants to be more likely to respond that the test interval contained more dots when it actually contained more dots and to respond that it contained less dots when the standard was more numerous.

Beta distributions were assumed for the guess and lapse rate (by choosing fairly broad beta priors we are not making any specific assumptions about participant performance but avoid some lapse/guess rates having a negative probability, a problem we would encounter with a Gaussian distribution for the priors and without having the steep cut-off we would encounter by using a uniform distribution – this is the case for all the priors I used for the fits in this thesis).

F can be decomposed into two functions, g, a simple scalar function and f, a higher dimensional function (which injects a nonlinearity) such that

$$F(\mathbf{x};\boldsymbol{\alpha},\boldsymbol{\beta}) := f(\mathbf{g}(\mathbf{x},\boldsymbol{\alpha},\boldsymbol{\beta})).$$
(2.27)

g is also called the core and f is also referred to as the sigmoid. The psychometric functions in this thesis used a linear core of the form

$$g(x,a,b) = (x-a)/b$$
 (2.28)

and a Gaussian cumulative density function as the sigmoid (Note: the Gaussian

cumulative density function has no closed form).

When we fit this psychometric function to our data using constrained Maximum Likelihood Estimation (MLE), the following parameter values give us the best fit to the data:

 $\alpha = -0.1121, \quad \beta = 0.1912, \quad \gamma = 0.0153, \quad \lambda = 0.0093$ 

This leads to the following sigmoid that describes the data (figure 2.11):



Figure 2.11: Psychometric Function for Example Data.

From this fitted psychometric function (the error bars are the 95% confidence intervals, see section 2.9.1.2.) we can now find the value that corresponds to our threshold, which, in this case, is 0.1290 (this means this participant needs the test interval to have 12.9% more dots than the standard interval to reliably say that the test interval is more numerous).

#### 2.9.1.2. Confidence intervals

Natural processes, including participants' performance in such a task, are noisy. This means, rather than relying on a point estimate of the threshold we need a measure of variability. Ideally we are looking for a measure of precision of the point estimate we have of our threshold. Since we assume that the variability in the parameters is caused by variability in the data, a good way of doing so is assessing how much varying the data affects the parameter estimates. One technique that allows us to do this in a systematic fashion is the bootstrap. The bootstrap (for an introduction to the bootstrap see e.g. Efron & Tibshirani, 1993) allows us to obtain confidence intervals for a fitted value using Monte Carlo simulations of the fit. In this thesis the bootstrap used was a bias corrected accelerated bootstrap (BCa) (Effron, 1987).

When we bootstrap the 50% point (this is an example, we also bootstrap the parameters and the deviance in the following section), we assume that the fitted model is a good description of the underlying pattern that describes our participants' behaviour (we are using a parametric bootstrap here). We know the x value of the 50% point on the fitted function. We then generate new samples for this specific x value by drawing from a Binomial distribution with the parameters n and Psi(x). n is the number of trials at the x value we are bootstrapping and Psi(x) is the predicted probability (under the maximum likelihood estimator) at said x value. We then repeat this 2000 times. The 2.5 and 97.5 percentiles of this bootstrap distribution are then used as the limits of the 95% confidence interval. This, however, does not necessarily give

us a correct estimate of the true confidence intervals.

If we were to run a simulation of how well an algorithm fits a model with confidence intervals (M) to a known dataset (N) we would be interested in the proportion M/N i.e. the proportion that our confidence intervals contain the actual dataset N. A general problem of bootstrapped confidence intervals when using constrained MLE, is that the confidence intervals tend to underestimate the spread of the data (e.g. Wichmann & Hill, 2001b; Fründ et al., 2011). This means our 95% confidence intervals contain the true parameters in less than 95% of cases. We are thus more likely to conclude that two parameters or threshold measures are significantly different from each other when they are not.

### 2.9.1.3. Sensitivity analysis – adjusting the confidence intervals

To asses whether the bootstrapped confidence intervals are underestimating the true spread of the data we run a sensitivity analysis (or bootstrap bridging assumption,(Wichmann & Hill, 2001b)). This is visualised for our example data-set in figure 2.12 by showing the joint probability of parameters a and b of the fitted model.



Figure 2.12: Confidence intervals (red elypsoid line) and probable data points (cloud of blue dots).

The dark shading shows the density of the joint distribution estimated from the bootstrapped parameters. The red dot in the middle is the point estimate of the two parameters that was determined by MLE. The red diamonds that are connected by a line are points at which additional bootstrap samples were drawn. This lets us evaluate how well our 95% confidence intervals cover probable data points. Based on this sensitivity analysis we can expand the 95% confidence intervals by using the widest confidence intervals along the connecting line and thus have more realistic confidence intervals (Wichmann & Hill, 2001b). All confidence intervals that are reported in the following chapters have already been corrected based on the results of such a sensitivity analysis. For our example dataset the sensitivity analysis leads to an adjustment of our confidence intervals for the threshold (which we previously found to be at 0.1290) from 0.0779 - 0.1845 to -0.3327 - 0.1935.

When we bootstrap our thresholds this gives us a measure of accuracy of the point estimate, what it does not provide us with is a measure of goodness-of-fit of our psychometric function in relation to the data. This is important because the confidence intervals are commonly used as a measure of goodness-of-fit, and generally small confidence intervals go hand in hand with a good fit, but, only generally, so we need another measure of goodnessof-fit to assess how well our fitted function actually describes the underlying data. For example, we might have fitted a model to the data that does not describe the data very well, but there is one constellation of the parameters that gives us an acceptable (but still not very good) fit. In most iterations of the bootstrap we will therefore end up with this constellation of parameters and thus very small confidence intervals. If we were to rely on the confidence intervals in this case we would assume that the fitted model describes the data really well, when the truth is that it does not. We thus need a measure of goodness-of-fit other than the confidence intervals. At the other extreme there is the situation in which we have not collected enough data but have found a model that fits the data perfectly. In this case we would end up with extremely large confidence intervals. Again, we need a measure of goodness-of-fit of the

model for the data.

# 2.9.1.4. Goodness-of-fit

There are various measures that we could use to assess goodness-offit.  $X^2$ , for example or the deviance (in a nutshell, this measure is based on the sum of squares error metric). I will be using the deviance because Psignifit, which I am using for this analysis, uses the deviance as a measure of goodness-off-fit. For a discussion of the advantages of using the deviance over  $X^2$  as a measure of goodness-of fit see Wichmann & Hill (2001a).

In order to assess whether a fitted function is a good fit for the data, we assume that the model is a perfect fit for the data and compare the 'expected deviance' with the actual deviance using Monte Carlo Simulations. If the observed deviance is outside the 95% confidence intervals of the deviance distribution (critical deviance) we treat the fitted model as a 'bad fit' to the data.

Figure 2.13 shows a visualisation of this for the sample data set:



Figure 2.13: Deviance Distribution. Solid line: Observed deviance, dotted line: critical deviance.

In this thesis I will not be explicitly reporting the deviances for all fitted functions, but they will be included in the figures depicting single fitted sigmoids for each participant and for each experiment in this part of the thesis I will mention what percentage of fitted functions was deemed to have a good fit. I used the deviance as a means of finding the sigmoid and priors that would give me the best fit for the highest proportion of data-sets. So while this means that some data-sets might be better described by a higher or lower lapse rate, or by a different sigmoid, all data-sets were fitted with the same sigmoid (Gaussian), core (linear transformation), and prior assumptions about the parameters. This was done to be able to compare thresholds and biases across all experiments.

In the following chapters I will describe the experiments in which I used relative numerosities to investigate how monocular zones in binocular regions are perceived. I will refer back to this chapter where appropriate but will explain experiment-specific aspects in more detail in the methods of each chapter.

# 3. Sensitivity to differences between numerosities in monocular regions

This chapter explores our sensitivity to differences in numerosity under monocular, occluded and binocular viewing situations.

In two experiments (Experiments 1 & 2) participants were sequentially presented with two dot displays and were asked to decide which one of the two displays was more numerous. Experiment 1 uses maximal contrast, experiment 2 is an improved version of experiment 1, with stimulus changes that address issues discussed in the discussion section of experiment 1.

# 3.1. Aims

While the overall aim of the numerosity experiments in this thesis was to examine if/how our representation of patterns differs depending on whether they are presented in monocular zones; the aim of the experiments in this chapter was to explore our sensitivity to differences between two numerosities.

More specifically, I asked the following questions:

When performing a numerosity discrimination task,

- does our sensitivity to information in monocular zones differ from that to binocular information? And, if so,

- is our performance such that it would suggest we are using only the input from one eye?

I will now describe the first out of 5 sensitivity experiments (2 of which

are presented in this chapter) in more detail.

# 3.2. Experiment 1 – Numerosity Discrimination at Maximal Contrast

#### 3.2.1. Aims

This experiment was designed to explore how our numerosity discrimination thresholds differ between binocular conditions and when numerosities are seen in monocular zones of binocular scenes (akin to those used by Forte et al, see figures 1.17 and 1.18, section 1.6).

More specifically I asked the following questions:

- Do observers deliver different thresholds when asked to discriminate numerosity in monocular zones than numerosities in binocular scenes?

- Does our sensitivity differ between a fully binocular stimulus and a fully monocular (i.e. one eye closed) stimulus?

- Does our sensitivity to numerosities in monocular zones suggest that we are ignoring one eye's view?

### 3.2.2. Methods

## 3.2.2.1. Stimuli

Participants were sequentially presented with two dot displays and were asked to decide which of the two displays was more numerous. The displays

could be either binocular or monocular. In each configuration the stimulus could either be occluded or unoccluded (see figure 3.1). Thus the stimulus was presented in one of 4 configurations:

(a) Vertically Occluded – A cloud of dots was presented behind binocular, vertical occluders. The occluders were spaced so that when viewed binocularly, all dots in the cloud were visible, though each dot only monocularly. See figure 3.1.a.

(b) Binocular (Unoccluded) – A cloud of dots was presented binocularly.See figure 3.1.b.

(c) Monocular Occluded – One eye was presented with a cloud of dots behind vertical occluders. As opposed to the vertically occluded condition, this means 50% of the dots were occluded, the density of the cloud, however, remained constant. Note, however, that since we are comparing Weber-Fractions this reduction in visible dots should not affect performance in relation to the Weber-Fraction itself. See figure 3.1.c.

(d) Monocular Unoccluded – One eye was presented with a cloud of dots. See figure 3.1d.





3.1.d) Monocular, Unoccluded

Figure 3.1: The different stimulus configurations for experiment 1.

The following 4 conditions were used, measuring performance for each of the 4 stimuli separately:

(a) Vertically Occluded – The binocular version of the occluder stimulus (figure 3.1a) was shown in both intervals.

(b) Binocular (Unoccluded) – The binocular version of the unoccluded stimulus (figure 3.1b) was shown in both intervals.

(c) Monocular Occluded – A monocular version of the occluder stimulus

(figure 3.1c) was shown in both intervals.

(d) Monocular Unoccluded – A monocular version of the unoccluded stimulus (figure 3.1d) was shown in both intervals.

Conditions (c) and (d) were collected as left and right eye conditions. In figure 3.1.c you can see the version presented to the left eye, whereas figure 3.1.d. shows a right eye stimulus. These sub-conditions were later compared and after no difference between left and right eye sub-conditions was found

combined into the two general monocular conditions – occluded and unoccluded. I will discuss this in more detail in section 3.2.8.2.

The stimulus square (size:  $6.79^{\circ}$  by  $6.71^{\circ}$ ) was framed by a white line (width:  $0.14^{\circ}$ , luminance: 23.4 cd/m<sup>2</sup>) and contained  $0.07^{\circ}$  white dots, and was displayed on a black background (effectively 0 cd/m<sup>2</sup>, below the limit of our luminance meter).

The monocular conditions could be presented either to the right or the left eye. The different conditions were run in separate trials. For example, if the first of two stimuli in a trial was monocular unoccluded, the second one was monocular unoccluded as well.

#### 3.2.2.2. Participants

As a pre-test all 8 participants completed both the Butterfly and the TNO stereo vision tests and did visual acuity tests at 3 metres (see general methods at the beginning of chapter 2). All participants had normal or corrected to normal acuity. One participant had to be excluded due to impaired stereo vision (TNO test: needed a disparity of 4 arcmin). Thus 7 participants participated in the experiment.

## 3.2.2.3. Procedure

For a description of the within-trial timeline see section 2.5. Each participant completed 40 sessions of 54 trials each. This took just under 3

hours for most participants.

The stimuli contained varying numbers of dots that were randomly distributed within the square frame. The number of dots for the standard stimulus in a trial could be 50, 65, 80, 95, or 110 dots, whereas the test stimulus could vary by 0,  $\pm 16$ ,  $\pm 25$ ,  $\pm 50$  or  $\pm 75$  percent of the numerosity (rounded to the nearest integer) of the standard display. Each participant performed 8 sessions at each baseline, the order of which was randomised for each participant.

## 3.2.3. Predictions

#### 3.2.3.1. Vertically Occluded vs. Binocular

In the vertically occluded condition, which is the condition that contains monocular zones, the visual system potentially has access to all the dots that it would have access to in the unoccluded condition. In the former, each dot is presented to one or the other eye, and in the latter to both eyes. If we were to assume that there is a difference in numerosity sensitivity between binocular and monocular regions, we would expect thresholds to be higher for vertically occluded stimuli than for fully binocular stimuli. An extreme prediction would be that information in monocular zones is treated as noise and therefore ignored. We know, however, that this is very unlikely. We can expect the monocular zones not to be ignored, since the very similar setup used by Forte and colleagues (2002) showed that monocular regions that are embedded in
binocular scenes appear to form part of a stable representation.

Another possibility is that binocular regions are treated as more reliable than monocular regions. This should be reflected in differences in thresholds in this experiment. More specifically, if monocular regions are treated the same as binocular regions there should be no significant difference between the threshold levels of the two conditions. On the other hand, if information in monocular zones is treated as less reliable, the threshold levels of the vertically-occluded condition might be significantly higher than the binocular unoccluded condition.

#### 3.2.3.2. Monocular Unoccluded vs. Binocular

There are conditions, such as when the contrast in a stimulus is reduced, under which we gain extra information from the input of our two eyes that adds to the information gained from only one eye. In the case of low contrast this leads to an improvement of contrast sensitivity performance by a factor of  $\sqrt{2}$  (e.g. Campbell & Green, 1965). This improvement of performance is, however, mostly observed under conditions that contain a lot of noise. The present stimulus contained no noise in the cloud of dots and had very high contrast differences (white dots at 23.4 cd/m<sup>2</sup> on an effectively black background), which would suggest that the performance gain from having a binocular stimulus might be very small. If the monocular version of this stimulus is treated as less reliable, this would show up in differences in participants' thresholds.

## 3.2.3.3.a. Vertically Occluded vs. Monocular Occluded

If participants perform the task by ignoring one eye, one might expect performance in these two conditions to be similar. At the same time, if the presented number of dots does have an effect this should not only affect the pattern between the different baselines within conditions but should also lead to similarities between the lower baselines in the vertically occluded condition and the higher baselines in the monocular occluded condition (because in the later condition participants are comparing smaller numerosities than in the former).

## 3.2.3.3.b. Vertically Occluded vs. Monocular Unoccluded

This is a control. The monocular occluded stimuli contain only half the number of visible dots of the unoccluded display (the density of dots remains constant, though). If there are significantly different thresholds in the two conditions this would warrant further investigation in a later experiment to investigate whether participants utilize a texture density or a basic numerosity approach. This is because if participants rely solely on the number of presented dots and are thus responding to the raw number, the monocularly occluded condition might lead to different results because of the smaller numbers of visible dots.

#### 3.2.4. Analysis

Participants' data was recorded in relation to whether they had responded that the test interval was more numerous. As described in chapter 2, all data-sets were fitted to the same sigmoid, core and prior expectations for the parameters. This lead to 195 out of the original 204 fits to be deemed a good fit, based on the deviance of the fits.

The confidence intervals of the fits for each baseline-numerosity for each condition for each participant were compared (see figure 3.2 and appendices A1.1-6). Then the combined data sets were re-fitted.

The parameters for the monocular conditions were compared using a repeated measures ANOVA, then the left and right eye conditions were combined and the sigmoids re-fitted (see figure 3.3 and appendices A1.7-12).

The thresholds for the remaining 4 conditions were then compared using a One-Way ANOVA. The planned comparisons were Sidak corrected (see figure 3.4).

## 3.2.5. Results

I will now discuss the results for the various tests in the order that they were performed. This means the differences between the conditions are explained in section 3.2.5.3.

#### 3.2.5.1. Differences between the various baseline-numerosities

Some numerosity discrimination tasks (John Greenwood, personal communication) have indicated that our performance differs depending on the overall numerosities used in an experiment. This led to me using 5 different baseline-numerosities for this experiment. Figure 3.2 shows the fitted functions and confidence intervals for the 5 different baseline-numerosities for each condition for participant gk. The fitted functions for the remaining participants can be found in appendices A1.1-6. Note how the confidence intervals overlap greatly for the 5 baseline numerosities. Also, note how there is no specific trend in the differences between the sigmoids in each graph – i.e. the fit for the 50 dot baseline does *not* consistently have a shallower slope than the fit for the 110 dot baseline. Hence the 4 sigmoids for the sub-plots in appendix A1.1.



**Proportion Difference** 

Figure 3.2: The fitted functions to the 5 baseline numerosities for the 6 conditions for participant gk. The lightest sigmoid is the 50 dot baseline, the other sigmoids darken with ascending numerosity.

Figure 3.2.b visualises the differences or lack thereof by plotting participants thresholds against the baseline-numerosities. If at all, participants perform



slightly better at higher baselines. This trend is not significant, though.

Figure 3.2.b - The thresholds of all participants plotted against the baseline numerosities.

Because there was no significant difference between the results found with different baseline-numerosities across the different conditions, I combined the different data-sets so that each condition now had only 1 sigmoid fitted to a larger data-set.

# 3.2.5.2. Differences between left and right eye monocular conditions.

I compared performance for the left and right eye versions of the monocular stimuli. The goal was to establish whether they could be combined into one data set.

The fitted sigmoids for the 6 conditions, but this time collapsed across all baseline numerosities, are shown separately for participant gk in figure 3.3. The fitted sigmoids for the remaining participants can be found in appendices A1.7-12. The sigmoids for the monocular conditions are plotted in the same sub-plots (c and d, the open circles were chosen for better legibility) to make a comparison between the left and the right eye conditions easier. Fits with a deviance (D) above the critical level are marked with a \*.



**Proportion Difference** 

Figure 3.3: The fitted functions for the 6 conditions for participant gk.

Note the large overlap of the confidence intervals of the left and right eye versions of the monocular stimuli (subplots c) and d)). This means that participants do not treat information from the two eyes differentially (which might have indicated a participant showing strong dominance of one eye). I thus combined the data sets for the two monocular, occluded (figure 3.3c), and the two monocular, unoccluded (figure 3.3d) conditions and fit new sigmoids.

For completeness, the new fits for the monocular conditions for all participants are shown in appendices A1.13-14. This leaves us with fits for 4 conditions for each participant.

## 3.2.5.3. Differences between the 4 remaining conditions.

A summary of the different participants' performance for the 4 conditions are shown in figure 3.4, where thresholds are plotted, one sub-graph per condition.



Figure 3.4: Thresholds across the 4 different conditions for experiment 1.

Overall, the differences between the groups were highly significant

(ANOVA F(3,24) = 5.112, p = 0.007). The threshold levels for the vertically occluded condition were not significantly higher than those for the fully binocular condition (p = 0.464). The monocular unoccluded condition was not significantly different from the binocular condition (p = 0.99996) and the monocular occluded condition led to threshold levels significantly different from the two unoccluded conditions (binocular: p = 0.015; monocular: p = 0.01). The two occluded conditions, however, where not significantly different from each other (p = 0.462).

#### 3.2.6. Discussion

I will now go through our hypotheses and discuss the results in relation to the hypotheses.

# 3.2.6.1. Differences between Vertically Occluded and Binocular Unoccluded Conditions.

We are interested in the differences in thresholds between the vertically occluded and binocular unoccluded conditions because while we are presented with the same number of dots in both conditions, the dots are only visible monocularly, embedded in a binocular scene, in the first while they are fully binocular in the second. If we are less sensitive to information in monocular zones compared to binocular information, then we would expect differences in participants' thresholds. There was no significant difference in thresholds between the vertically occluded and fully binocular conditions. This indicates that we are able to integrate the dots in monocular zones fully and that they are treated as just as reliable as dots that are presented binocularly. One might argue that participants could have performed the task by suppressing one eye's view and relying on the visible density since the texture density remained constant across the two eyes' views. There was also no significant difference in threshold between the two occluded stimuli (see section 3.3) which suggests that while participants did not simply rely on one eye's view, they might have used a filling in mechanism to compensate.

# 3.2.6.2. Differences between Monocular and Binocular Unoccluded Stimuli.

Participants do, if there is a lot of uncertainty in a scene, treat a monocular stimulus as less reliable than a binocular stimulus. As if the binocular stimulus allows them a 'second go' at seeing the stimulus. Here we are interested in whether the specific stimulus used here, led to different thresholds for the monocular and binocular versions of the unoccluded stimuli, a lack of which would suggest that the two stimuli are treated as equally reliable where numerosity information is concerned.

There was no significant difference between the two unoccluded stimuli, suggesting that, as predicted, at maximum contrast we do not gain from having the two eyes' views present. However, the maximum contrast is also a

limitation. This setup allows us to say that if the contrast is sufficiently high, we do not seem to gain much in relation to sensitivity if we have input from the two eyes. What we are not able to say is whether this is the case under different contrast situations as well.

# 3.2.6.3. Differences between Vertically Occluded and Monocular Occluded Stimuli.

In section 3.2.6.1. I mentioned that since the density of the dots remained constant across the stimulus, participants could have ignored one eye's view to perform the discrimination task. Doing so, would have rendered the vertically occluded stimulus equivalent to the monocular occluded stimulus. Here I ask whether participants' sensitivity to differences suggests that one eye's view is ignored to interpret the vertically occluded stimulus.

The thresholds for the vertically occluded condition were lower than the thresholds for the monocular occluded condition. While this difference was not significant, the trend supports the results discussed in section 3.2.6.1. As opposed to ignoring one eye's view and performing the task by turning the vertically occluded situation into one that resembles the monocular occluded stimulus, participants seem to integrate the two eyes' views and perform the task in a fashion that delivers similar thresholds to when they are dealing with fully binocular stimuli. However, since performance is somewhere between the fully binocular and monocular occluded conditions, the integration might come at a cost or the difference might be caused by the presence of the occluders.

# 3.2.6.4. Differences between Monocular Occluded and Monocular Unoccluded Stimuli.

Participants' thresholds in the monocular unoccluded situation were significantly lower than in the monocular occluded condition. One might think that this could be caused by participants only having access to 50% of the dots they have access to in the unoccluded situation, but this is partly why I am using the proportion difference as a measure – yes, participants will have access to less dots overall, but the proportion difference remains constant. So this should not be the cause of the difference in sensitivity.

The main difference between the two types of stimuli is the addition of the white noise occluders in the occluded situation. One thing that this experiment does not allow us to look at is whether the occluders themselves have an effect on participants' thresholds. For the binocular situations (both unoccluded and vertically occluded) the presence of the occluders leads to higher (though not significantly) thresholds as well. This indicates that the lowered sensitivity in the occluded conditions might be due to the occluders rather than the monocularity/binocularity of the dots.

# 3.2.6.5. General Discussion

Overall, it seems to be the case that our representation of the presented numerosities does not differ between when they are presented in monocular zones, in the vertically occluded stimulus, or fully binocularly. However, there are a number of issues this experiment did not address.

Firstly, we have to ask whether participants would have been able to perform the task using only one eye (i.e. ignoring the input from one eye, be it their non-dominant eye or in a random fashion). While it seems that this is not the case since participants show no raised thresholds in the vertically occluded condition compared to the fully binocular condition, there was also no significant difference between the two occluded conditions suggesting that information in monocular regions is not interpreted exactly the same as binocular information.

This lack of difference between the two occluded conditions and the general pattern, that, in both the monocular and binocular cases, the occluded conditions lead to higher thresholds than the unoccluded conditions (though not both significant) might suggest that there is a general effect of the occluders on our performance. This could be making the task harder by attracting our attention, another possibility might be that the white noise in the occluders is treated as dots (in the white noise patterns each pixel is randomly assigned to be either black or white, the dots in the background plane are 3 pixels in diameter) or as part of the overall texture that is to be compared. This would make the difference between the two displays that is added by the background plane insignificant compared to the vast number of dots/texture that remains constant across the two intervals. However, in its extreme form, this hypothesis would make participant performance much worse than what we observe here.

Additionally, while using maximal contrast was a good starting point for a first experiment, the large differences in luminance between the black dots and the white background raise another issue – could participants have used the contrast as a cue rather than having to use numerosity/density cues to perform the task? While this would not necessarily have to change participant performance, the stimuli used in this experiment do not allow us to make a statement about how participants might have approached the task itself.

To deal with the issues raised in the section above, namely:

- the possibility that participants could have ignored one eye's view in the vertically condition and still have performed just as well,
- the fact that we cannot disambiguate the effect of the monocularity and of the presence of the occluder itself,
- the possibility that the contrast between the black background and the white dots was used as a cue to numerosity,

I decided to run a second experiment with the same general setup but with stimuli that addressed the perceived shortcomings of the stimuli in the present experiment. I shall describe the second experiment in the following section.

#### 3.3. Experiment 2

## 3.3.1. Aims

This experiment was designed to explore how our numerosity

discrimination thresholds differ under binocular conditions or when numerosities are seen in monocular zones of binocular scenes. I tried to take into account the issues raised with experiment 1 and thus expanded the list of questions we were interested in to the following list:

Under conditions that give participants no luminance information to perform the task (this will be explained in more detail in the method section below),

- Do numerosities in monocular zones lead to different thresholds than numerosities in binocular scenes?

- Is there an effect on our performance that is mediated by the presence of the occluders?

- Does our sensitivity differ between a fully binocular stimulus and a fully monocular stimulus?

- Does our sensitivity to numerosities in monocular zones suggest that we are ignoring one eye's view?

### 3.3.2. New Stimuli

As before participants were sequentially presented with two dot displays and were asked to decide which of the two displays was more numerous. Also, as before, the displays could either be binocular or monocular, as for the previously used conditions (figure 3.1). A 5<sup>th</sup> condition was added to investigate the potential effects of the occluders. Figure 3.5 shows the different conditions used in the present experiment. The overall luminance was changed to a midlevel-grey to avoid the potential confounding effects of luminance (participants might have used the overall contrast as a cue to numerosity rather than performing the task I asked them to do). In order to eliminate the effects of luminance further, half the dots were black and the other half white – this means 80 dots would have approximately the same overall luminance as 140.



Figure 3.5: The different stimulus configurations for experiment 2.

The stimulus square (the size was kept constant at 6.79° by 6.71°) was now framed by a black line (width: 0.14°, luminance: 0.01 cd/m<sup>2</sup>) and contained 0.07° sized black and white (now 66.54 cd/m<sup>2</sup>) dots (50% of the dots were white, 50% were black, the colour was assigned randomly to each dot) displayed on a midlevel grey background (28.12 cd/m<sup>2</sup>). The white noise surround and occluders now had an overall luminance of 33.40 cd/m<sup>2</sup>. I will discuss the differences in luminance in section 3.3.8.

In the newly added Horizontally Occluded condition the white noise occluders were rotated by 90° (6.79° by 0.67°) this lead to a stimulus in which only 50% of the dots were accessible, and all binocularly (yet the density remained constant). This condition was intended to allow us to establish whether the differences between stimuli containing monocular zones and the purely binocular stimuli were due to the presence of the occluders or truly an effect of the information being contained in monocular zones.

### 3.3.3. Hypotheses

There were two types of hypotheses for this experiment, the first kind are about the effects of the general changes made to the stimuli, the second about how the newly introduced horizontally occluded condition compares to the previously used conditions.

As for the changes to the stimuli, the changed luminance was intended to stop participants from being able to use the change in luminance as a cue. If this was the case then we might expect a change in the results. However, it is not clear whether this would improve participants' performance when judging monocular stimuli or stimuli containing monocular zones or whether their performance would suffer.

I will now reframe the hypotheses to include the horizontally occluded condition.

## 3.3.3.1. Purely Binocular vs Horizontally Occluded

If participants were just relying on the information on the background plane to make their judgements then there should be no difference in performance between the binocular and horizontally occluded conditions. The proportion difference between the standard and test stimuli would remain constant, and should thus not be affected by the fact that we have halved the visible area of the background plane. If there is an effect of the occluding stripes, however, this could lead to a drop in sensitivity. If the occluding stripes are (for example) treated as part of the overall numerosity/density, any differences in the number of dots on the background plane will become insignificant compared to the dominating number of 'dots' that are present in the occluders. Note, however, that this is a very extreme hypothesis that would lead to participants having much larger thresholds than what we have previously observed.

# 3.3.3.2. Horizontally Occluded vs Vertically Occluded

Potential differences between horizontally occluded and vertically occluded conditions follow the same logic as for 3.3.3.1 above. If there is an effect of the presence of the white-noise occluders, then this should lead to similar effects in both conditions, if there is, however, both an effect of the occluders and of the monocular zones then this should be visible in this comparison. If the monocular zones impair our performance then the vertically occluded stimuli should lead to even worse thresholds than for the horizontally occluded stimuli.

## 3.3.4. Participants

11 participants completed the experiment. Just as in the previous experiment, participants were screened for potential issues with their stereo vision. Two additional participants were excluded during this screening process due to very limited stereo vision (cut-off criterion, as before, no stereo vision below 4 arcmin in TNO test).

# 3.3.5. Setup

Each participant completed 40 sessions of 63 trials each. The timeline of each block of trials remained the same. This took just under 4 hours for most participants. Again, there were 5 different baseline numbers for the dots in the standard stimulus (50, 65, 80, 95, 110). This baseline number remained constant for 8 sessions and was then changed in a random order that was generated separately for each participant.

## 3.3.6. Analysis and Results

#### 3.3.6.1. Investigating non-stationarities

Non-stationarity of a participant's performance describes changes in their performance (and thus the shape of the psychometric function) across time. These non-stationarities could be caused by the participant still learning a task, becoming tired, being distracted, or adopting a different strategy while performing the task.

As there was no significant difference between the confidence intervals between the different baseline numerosities in the previous experiment, for the analysis of this experiment the different baseline-numerosities were treated as different blocks within the same experiment. This allows us to not only fit all the data to one psychometric function but at the same time allows us to investigate whether participants are still learning at the beginning of the experiment or if there happens to be a difference between the baselines that did not show up with the stimulus configuration used in experiment 1. To assess whether there are any non-stationarities in participant behaviour we decompose the overall deviance into residuals. Deviance residuals are no different from standard residuals in a generalized linear model (e.g. Dobson & Barnett, 2008) but are a generalisation that also applies to binomial data such as the data in our datasets. We can then calculate the correlation between the deviance residuals and the sequence the blocks were presented in. If an observer were learning we would expect a positive correlation, on the other hand, a negative correlation would indicate fatigue. In the same way the overall deviance is bootstrapped to assess whether it is excessive (outside the 95% confidence intervals of the bootstrapped distribution) the correlation coefficients (Rkd) were bootstrapped as well. I will indicate cases where the Rkd indicates non-stationary behaviour.

All conditions were fitted with the same cumulative Gaussian function used in experiment 1. Figure 3.6 shows the fitted functions for the different conditions for participant js. Each 'block' of trials is depicted by a dot along the psychometric function. Note how there is very little variance across blocks in the binocular unoccluded condition (figure 3.6.a), most dots for a stimulus level lie on top of each other. Conversely, the monocular occluded conditions lead to a larger variance across blocks (figures 3.6.d and e).



Figure 3.6: The fitted functions to the 4 baseline numerosities for the 7 conditions for participant js.

The fitted functions for all other participants can be found in appendices A1.15-24.

5 out of the original 77 fits show a deviance higher than the critical level suggested by the bootstrapped distribution

As for the non-stationarity, 8 out of the original 77 fits show a Rkd that is outside the 95% confidence intervals of the bootstrapped distribution, 6 out of which are in the monocular conditions. However, note that other than participant sh who shows learning for 4 of their 7 fits (all positive correlation), no participant, including participant js (see figure 3.6) showed a nonstationarity for more than 1 condition.

This suggests, that if there is any learning happening during the duration of the experiment the effect is negligible for most participants.

The reader might notice the much larger deviances compared to the previous experiment (see figure 3.3, for example), yet only 5 fits have a deviance that is higher than the critical level. Keep in mind that the 'size' of the observed deviance itself depends on several factors. In the present case, the number of blocks used has increased from 9 (the different stimulus levels) to 45 (the different stimulus levels for 5 different baselines). If I were to combine the different baselines and go back to the 9-block approach used for experiment 1, the deviance drops back to the levels we observed for experiment 1. So even if the deviances appear large, unless it is indicated that an observed deviance is above the 95% confidence interval suggested by the bootstrapped deviances, the fitted model is a good approximation of the data.

Compare figure 3.7a with figure 3.7b, which shows the fit for participant Ih based on 45 blocks whereas figure 3.7b shows the fit for participant Ih

based on 9 blocks in which the 5 baseline-numerosities had already been combined.



**Proportion Difference** 

Figure 3.7: The fitted functions for the binocular condition for participant lh. A) the different baselines were used as separate blocks in the experiment, b) the responses for the different baselines were averaged.

Note that the fitted sigmoid does not differ between the 45 and the 9 block approach. I will therefore continue using the functions fitted to the data combined in 9 blocks for the remainder of this chapter.

3.3.6.2. Differences between left and right eye monocular conditions.

3.3.6.2.1. Presentation-eye as a between-subjects factor.

As in experiment 1, the monocular stimuli could be presented to either

the left or the right eye. In order to assess whether the different monocular conditions differed from each other, or whether they could be treated as one condition for this analysis (and could be collapsed into one condition for future experiments), I tested whether  $\theta$  differs significantly for the two eyes. To do so I ran a two-way mixed ANOVA and treated the eye as a between-subjects factor. This means you could think of this as us treating the two conditions as two separate sessions during which the same variables (the parameters  $\alpha$ , $\beta$ , $\lambda$ , and  $\gamma$ ) are assessed.

We can then test whether there is a main effect of the eye the stimulus was presented to and whether there might have been an interaction between the eye and the different parameters.

This was done twice, once for the two unoccluded monocular conditions and once for the two occluded monocular conditions.

I will now go through the results for the two monocular occluded conditions first.

#### 3.3.6.2.1.a. Occluded conditions

For the occluded conditions, Maulchy's test of sphericity was significant (W = 0.07, df = 5, p < 0.01), thus the results for the interaction of the parameters with the eye were adjusted using a Greenhouse-Geisser correction (epsilon was estimated at 0.498). The interaction of the parameters with the eye were non-significant at p= 0.12, F(1.49, 29.86) = 2.47. There was also no

significant main effect for the eye at p = 0.09, F(1,20) = 3.17. This suggests that, as is experiment 1, there was no difference between the left and right eye monocular occluded conditions.

For the two monocular unoccluded conditions the results were similar.

## 3.3.6.2.1.b. Unoccluded conditions

For the unoccluded conditions, Maulchy's test of sphericity was significant (W = 0.05, df = 5, p < 0.01), thus the results for the interaction of the parameters with the eye were adjusted using a Greenhouse-Geisser correction (epsilon was estimated at 0.45). The interaction of the parameters with the eye were non-significant at p= 0.23, F(1.34, 26.84) = 1.59. There was also no significant main effect for the eye at p = 0.38, F(1,20) = 0.81.

One might argue that this does not allow us to claim that there was no difference for the individual participants because they might show different eye-dominance patterns. This potential criticism is dealt with in the next section.

## 3.3.6.2.2. Eye-dominance as a 'between-subjects' factor.

I created an additional variable, that we shall call dominance, which codes a participant's eye dominance. Rather than relying on eye-dominance tests which can lead to very different results depending on the test used, I used the  $\beta$  parameter to drive this classification. Whichever eye led to a

steeper slope of the psychometric function was treated as the dominant eye for a given participant.

I then repeated the two-way mixed ANOVA, this time using the dominance variable as the between-subjects factor.

### 3.3.6.2.2.a. Occluded conditions

Let us look at the results for the two monocular occluded conditions first. Maulchy's test of sphericity was significant (W = 0.06, df = 5, p < 0.01), thus the results for the interaction of the parameters with the eye were adjusted using a Greenhouse-Geisser correction (epsilon was estimated at 0.49). The interaction of the parameters with the dominance variable were non-significant at p= 0.19, F(1.45, 28.96) = 1.80 There was also no significant main effect for the eye at p = 0.19, F(1,20) = 1.87.

#### 3.3.6.2.2.b. Unoccluded conditions

For the two unoccluded monocular conditions, the results were similar. Maulchy's test of sphericity was significant (W = 0.06, df = 5, p < 0.01), thus the results for the interaction of the parameters with the eye were adjusted using a Greenhouse-Geisser correction (epsilon was estimated at 0.44). The interaction of the parameters with the eye-dominance variable were nonsignificant at p= 0.09, F(1.31, 26.14) = 2.89. There was also no significant main effect for the eye at p = 0.87, F(1,20) = 0.03. It is thus safe to say that there is no significant effect of which eye these stimuli are presented to. I therefore combined the two conditions in each pair, treated them as one, and refit the psychometric functions for these conditions.

## 3.3.6.3. Differences between the remaining 5 conditions.

A summary of the different participants' performance for the 5 remaining conditions are shown in figure 3.8, where thresholds are plotted for each condition. A One-Way ANOVA was performed and the planned pairwise comparisons were Sidak corrected.



Figure 3.8: Thresholds across the 5 conditions for experiment 2.

Overall, the differences between the groups were highly significant (ANOVA F(4,55) = 13.86, p < 0.01). The threshold levels for the horizontally occluded condition were significantly higher than those for the fully binocular condition (p < 0.01). However, they were not significantly different from the

vertically occluded condition (p = 0.43). The monocular unoccluded condition was significantly different from the horizontally occluded condition (p < 0.01).

Interestingly, the vertically occluded condition was not significantly different from the monocular unoccluded condition (p = 0.18), but was significantly different from the monocular occluded condition (p = 0.01).

### 3.3.7. Discussion

I will now discuss these results in regard to the hypotheses for this experiment and will then continue with some further observations.

## 3.3.7.1. Purely Binocular vs Horizontally Occluded

The horizontally occluded condition was added in this experiment to assess the effect of the presence of the white-noise occluders on our sensitivity to differences between two numerosities. Unlike the vertically occluded condition, all visible dots in the stimulus were fully binocular. Thus, differences between the purely binocular and horizontally occluded conditions would indicate an effect of the occluders on our sensitivity to differences in numerosity.

There was a significant difference between the purely binocular and the horizontally occluded conditions. This suggests that there might indeed be an effect of the occluders on participant performance. This could, for example, be caused by the occluders attracting our attention, thus making the numerosity

discrimination task harder to accomplish because we have less resources to devote towards performing the discrimination. Alternatively, we could also hypothesise that the white-noise texture in the occluders is interpreted as a very large number of very small dots. If this is the case, then any changes in numerosity in the background plane would be dwarfed by the sheer number of dots on the occluder plane. However, this is a very extreme hypothesis that would lead to a much worse performance than what we observe here. The bottom line of this is that, performance is worse when occluders are present.

This means rather than comparing vertically occluded stimuli with purely binocular stimuli I will be comparing the vertically and horizontally occluded conditions next - comparing two conditions that both contain occluders and thus cancelling out the effects of the occluders. This will allow me to focus on the effect of monocularity/binocularity.

# 3.3.7.2. Vertically Occluded vs Horizontally Occluded

If we then compare the vertically occluded conditions, any differences between the two should be caused by differences between how information in monocular regions is processed compared to fully binocular information rather than effects caused by the presence of the occluders.

Threshold levels for the vertically and horizontally occluded conditions were not significantly different. If we assume that the occluders led to a reduction in sensitivity for both stimuli, then the lack of difference between the two conditions suggests that, at least from a sensitivity perspective, we are able to interact with vertically occluded information as well as with binocular information (in the horizontally occluded condition). However, we have to ask – how can we be sure that participants cannot simply perform this task for the vertically occluded condition by ignoring one eye's view? While participants reported that the stimuli appeared stable and non-rivalrous, this does not exclude the possibility that one eye's view is suppressed. The following comparison between the monocular conditions and the vertically occluded conditions is intended as a first step towards this issue.

#### 3.3.7.3. Vertically Occluded vs. Monocular

The monocular conditions had been intended to be compared to the vertically occluded conditions, to investigate whether the participants' thresholds for the vertically occluded condition might be due to suppression of one eye's view.

There was no significant difference between the vertically occluded and the monocular unoccluded conditions, whereas there was a significant difference between the vertically occluded and monocular occluded conditions. This suggests that the performance in the vertically occluded condition is unlikely to be generated by a process in which participants are simply ignoring one eye's view. However, due to the constant density of the background numerosity, we cannot discount the possibility that participants were simply ignoring one eye's view when judging vertically occluded stimuli. A good way to investigate this question would be to present the two eyes with different density displays in the vertically occluded condition, which would lead to a biased percept if one eye was suppressed.

I will discuss this and several other issues with the stimuli used in this experiment further in the following section.

#### 3.3.8. Issues with the stimuli used in experiment 2

An issue with the monocular stimuli used in experiment 2 is that the background for the non-stimulus eye was filled with white noise. It is thus hard to decide whether the difference between the vertically occluded and monocularly occluded conditions was due to the monocularity of the later or whether the difference was a side-effect of the added white noise in the monocular conditions.

This means the observed differences are not fully interpretable. Before I discuss how this was changed for the following experiment, let us return to the presented luminances, which suggest another design issue.

When the experiment was designed the midlevel-grey background was supposed to have a luminance that was 50% of that of the white dots used in the experiments. The maximum luminance used was 66.54 cd/m<sup>2</sup> so the midlevel-grey should have been approximately 33.28 cd/m<sup>2</sup>. However, because of a mistake in the code used, the grey background had a luminance of 28.12 cd/m<sup>2</sup>. Unfortunately, there were some other issues with the stimulus setup.

After analysing the main experiment, I noticed that rather than

calculating the proportion difference used for the different stimulus levels based on the following equation

$$Wb = (test - standard)/min$$
 (3.1)

the proportions had been calculated the following way

$$Wb = (test - standard) / standard$$
 (3.2)

This means the difference between two numerosities at the negative side of our scale had been larger than they should have been.

Finally, in the previous experiments the standard stimulus had always been presented first. While this does not affect the threshold itself, it can affect the PSE (for example, if a participant tends to continuously respond by pressing one button if they are guessing). This can easily be controlled for by presenting the standard stimulus first on half the trials and second for the other half of trials.

I thus repeated the experiment after making the changes described in the following section.

# 3.4. Experiment 2 - part 2

#### 3.4.1. Changes in the stimuli

The overall stimulus setup used in this re-run were the same as in the main experiment.

The changes made to the stimuli were the following:

The luminance of the midlevel-grey background was 33.23 cd/m<sup>2</sup>.

Furthermore, because experiment 1 and the first version of experiment 2 had neither suggested a significant difference between the different baseline numerosities (if at all, some participants showed issues with the lowest baseline indicating a different, less effective, strategy for the numerosity judgements), one baseline numerosity (80 dots) was used. Because only one baseline-numerosity was used, the number of trials for each data-set was reduced from 360 to 270 per condition (that is each data point on a fitted psychometric function is based on 30 observations. This means each participant completed 1890 trials.

A revised set of levels of difference between the standard and test numerosities was used. The new numerosities used for the test were: 45, 53, 64, 69, 80, 93, 100, 120, 140.

Finally, since the difference between the left and right eye versions of the monocular conditions have not been excessive (section 3.3.6.2. above), I combined the monocular conditions, presenting the stimulus to the right eye 50% of the time and to the left eye 50% of the time. While this reduces the number of monocular conditions to two, I decided to add two more monocular conditions as controls. In the main experiment, the 'non-stimulus eye' had been presented with a continuous white-noise background with a black line framing a square the same size as the background the dots were presented on in the 'stimulus eye' to aid fixation for the monocular conditions. This control experiment also included monocular conditions that had a midlevel-grey

square surrounded by a black line to control for effects of the white-noise background on thresholds.

# 3.4.3. Questions

This experiment was thus intended as a control for the following questions:

1. Do the results of experiment 2 replicate with a midlevel-grey background rather than a slightly darker background?

2. Do we observe differences in thresholds between the binocular condition and other conditions after adjusting the baseline, the stimulus levels and the presentation order?

3. Are the differences we have observed between the monocular and binocular conditions in experiment 2 due to the white-noise in the background of the monocular conditions?

#### 3.4.4. Participants

5 observers participated in the experiment. All observers had normal or corrected to normal accuracy (Snellen Eye-Chart) and normal stereo vision (TNO-Test).

# 3.4.5. Analysis

The data for this follow-up was analysed in the same way as the data for

the main experiment. However, since only one baseline-numerosity had been used, one sigmoid was fitted for each stimulus level. The fitted functions can be found in appendices A1. 25-28.

Then a One-Way-ANOVA was performed. The a priori pairwise comparisons were Sidak corrected.

### 3.4.6. Results

Figure 3.9 shows the mean thresholds for the 7 different conditions for all participants.

Overall, the differences between the groups were significant (ANOVA F(6,21) = 4.46, p < 0.01). However, after being corrected for multiple comparisons (I used a Sidak correction), none of the pairwise comparisons that were performed were significant. As opposed to before, the threshold levels for the horizontally occluded condition were not significantly higher than those for the fully binocular condition (p = 0.22). Neither was the vertically occluded condition different from the fully binocular condition (p = 0.22). The monocular occluded and the grey monocular occluded conditions were not significantly different (p = 0.99), neither were the monocular unoccluded and the grey monocular unoccluded conditions (p = 0.95). Unlike in the main experiment, the vertically occluded condition was not significantly different from the grey monocular occluded condition the grey monocular occluded condition the grey monocular occluded condition (p = 0.39).
The trends we previously observed in the data, namely, that the thresholds for the binocular occluded conditions were higher than the thresholds for the binocular unoccluded condition, that the monocular occluded conditions lead to much higher thresholds than the monocular unoccluded or the vertically occluded condition, did not change.

#### 3.4.7. Discussion

I will now discuss the results in relation to the questions I asked at the beginning of this experiment.

### 3.4.7.1. Can we replicate the results of experiment 2 with the changed luminances?

While the differences between the different conditions are nonsignificant in this experiment, the trends we observed in experiment 2 remained the same. More importantly, they are the same for all participants. What we see, however, is more consistent individual differences. For example, participant js (2<sup>nd</sup> participant from the left in the bar-charts in figure 3.9) has consistently low thresholds whereas participant ms (2<sup>nd</sup> from the right in figure 3.9) has consistently high thresholds. Yet, both participants show the same pattern in their differences between thresholds for the different conditions. This suggests that the underlying differences between our percept of information in monocular regions and of fully binocular information also has an effect in the present experiment. We can thus assume, that we can replicate the results of experiment 2 with the changed luminances.

# 3.4.7.2. Can we replicate the results of experiment 2 after changing the baseline, stimulus levels and presentation order?

When comparing the fully binocular conditions in experiment 2 and 2.2 the thresholds between the two conditions are not significantly different (bin 2: 0.146; confidence intervals: lower bound: 0.124 upper bound: 0.169; bin 2.2: 0.121; confidence intervals: lower bound: 0.065 upper bound: 0.177). For this comparison we are, however, also interested in the bias of the PSE. If presenting the standard stimulus always first led to a biased response pattern, then presenting the standard stimulus in either interval should lead to a bias that is not significantly different from 0. Experiment 2 led to a mean bias of -0.165 with confidence intervals that were not significantly different from 0 (lower: -0.382; upper: 0.051). Experiment 2.2 lead to a mean bias of -0.046 with confidence intervals that were not significantly different from 0 either (lower: -0.168; upper: 0.076). Thus while the original bias had not been significantly different from 0, it was reduced by changing the presentation order.

# 3.4.7.3. Is there a difference in performance between the monocular conditions with noise and those with a grey background?

The monocular stimuli in experiment 2 had contained white-noise within

the stimulus square in the non-presentation eye. Here I tested whether the differences between the monocular stimuli and the vertically occluded stimuli were caused by the additional white-noise. I thus had two versions of the occluded and unoccluded monocular stimuli: one in which the stimulus square was filled with white noise and one in which the square was filled with a midlevel-grey. There was no significant difference between the white-noise and grey background versions, suggesting, just like the lack of difference between the horizontally occluded and fully binocular conditions, that in this experiment the effect of the occluders was negligible.

#### 3.5. Overall discussion

The experiments discussed in this chapter intended to ask whether information in monocular regions leads to a percept that is comparable to that of binocular regions. Let us summarise what we have learned in this chapter.

Information in monocular regions seems to lead to a percept that is very similar to that of binocular information. In experiment 1 I found no difference between the vertically occluded and fully binocular conditions. There appears to be a slight effect of the textured occluders though. In experiment 2 there was a significant difference between the newly introduced horizontally occluded condition and the fully binocular one, a pattern that was repeated in experiment 2.2. This effect is small but seems to relate to the occluders rather than the textured white-noise pattern itself. There was no significant difference between the monocular conditions white-noise in the stimulus

square and those in which the white-noise was replaced by midlevel-grey. Once we discount the effect of the occluders though, we find no significant differences between the vertically occluded conditions and the horizontally occluded conditions in experiments 2 and 2.2. This suggests that information in monocular regions leads to a percept that is not very different from that of binocular information, a relationship that is further supported by the finding that there is a significant difference between the vertically occluded and monocular occluded conditions in experiment 2. This difference would suggest that we are very unlikely to simply ignore one eye's view when processing the vertically occluded stimuli. However, the stimuli used in the experiments in this chapter do not allow us to come to this conclusion.

So far the different conditions had a density that was more or less constant across the whole stimulus and the same in both eyes. If we use a texture density mechanism (see e.g. Durgin et al. 1995) to differentiate between numerosities and if the density remains the same, participants would, technically, be able to ignore one eye's view in the vertically occluded conditions and to simply extrapolate the density of the remaining region that is placed behind the occluding bars. A stimulus that has different densities across the stimulus plane would allow the exclusion of this scenario.

A 'stripy' binocular stimulus can, once occluded by vertical occluders, be turned into a stimulus that has one eye view containing a higher density than the other eye. If participants use only one eye then they should be responding to the wrong density all of the time (either too high by the

proportion the higher density eye was assigned or too low by the proportion the lower density eye was assigned) and their threshold levels should skyrocket for this condition because they would need a much higher difference between the two densities to make a reliably correct response.

In the following chapter I will describe an experiment that uses such a stimulus to ask whether differences in the stimulus density lead to a biased percept and whether these changes have an effect on our sensitivity to differences between two displays.

## 4. Experiment 3: The effect of varying texture densities on our percept of half-occluded regions

Our visual systems are able to represent an occluded object by integrating information from spatially distant regions at the edges of the occluder. In the previous chapters, I have shown that at constant densities we are as sensitive to differences between numerosities that are presented binocularly, as to those that are presented in monocular regions. In this chapter I ask how the two eyes' views are *combined* to represent a background scene that is composed of only monocular regions. This can be done using a stimulus that does not allow participants to extrapolate the overall numerosity/density based on one eye's view. The details of this manipulation will be described below. The perceived texture density of the background was explored by asking observers to compare the number of dots present, with that in a fully visible binocular image. In sum, there was no evidence that participants were less accurate than for an equivalent binocular stimulus, and they were only slightly less *precise*. This suggests that a background scene can be fully perceptually represented, as least by some observers, despite each region of the whole scene being visible to only one eye.

#### 4.1. Aims

The present experiment was intended to further explore the question of whether information that is presented in monocular regions is represented in a similar fashion as information that is presented binocularly. The previous

experiments did not allow us to make any statements about whether numerosities in monocular regions appear as numerous as when they are presented binocularly, nor allowed us to disambiguate whether participants were making judgements about the vertically occluded stimuli by ignoring one eye's view. Here, I made changes to both the stimuli and the procedure to allow these issues to be explored.

The stimuli used in this experiment were designed to disambiguate the later issue, and a slightly changed procedure was intended to allow us to make a statement about whether monocularly presented numerosities appear less numerous.

#### 4.2. Methods

#### 4.2.1. Stimuli/Procedure

As in the previous experiments (2 and 2.2), I studied two occlusion situations, both natural, one containing monocular zones, and compared them to a stimulus with a fully binocular background. Figure 4.1 shows cartoons of the stimuli used.



Figure 4.1: Cartoons of the stimuli used. a) Fully binocular stimulus; b) Horizontally occluded stimulus; c) Vertically occluded stimulus.

I compared density discrimination for three types of scene:

(1) A fully binocular scene where observers viewed a pattern of dots on a grey background (figure 4.1a);

(2) A horizontally occluded scene, where observers viewed a pattern of dots as if behind a foreground horizontally slatted fence (figure 4.1b).

Here, the fence and dotted background were both fully binocular, but half the dots were not visible to either eye as they were completely occluded by the fence;

(3) A vertically occluded scene (as used by Forte et al, 2002, figure 4.1b) in which each slice of background was only visible to one eye, but which was constructed so that all of the background was visible to one eye, or the other, as if behind a foreground vertical slatted fence (figure 4.1c).

The baseline stimulus was a grey square of size  $6.84 \times 6.84$  degrees (luminance  $32.47 \text{ cd/m}^2$ ), upon which a random pattern of black (luminance

0.01 cd/m<sup>2</sup>) and white (luminance 66.54 cd/m<sup>2</sup>) dots of size 0.18 x 0.18 min arc were superimposed. Half the dots were black and the other half white. The grey square was surrounded by a binary white noise texture (the luminance of each pixel, 0.06x0.06 min arc, was allocated randomly) that filled the remainder of the screen. The 'standard' stimulus contained 80 dots (and was always an instance of the binocular stimulus (1) described below), while the 'test' stimulus could contain 45, 53, 64, 69, 80, 93, 100, 120, or 140 dots. As in the experiments in chapter 3, the difference between number of dots in the standard and test stimuli will be expressed in terms of the proportion difference between them (refer to chapter 2.6.1. for more details on how this was calculated).

The test stimulus was one of the following three stimulus types:

(1) Fully binocular: both eyes viewed an identical pattern of dots (figure 4.1a). The dots were located on a background grey plane displayed at a disparity of 41 min arc with respect to the plane of the screen, this corresponds to them being located 18.5 cm behind the white noise surround which was presented in the plane of the screen.

(2) Horizontal occluded: both eyes viewed an identical pattern of random dots, half of which were hidden behind horizontally oriented foreground occluders (figure 4.1b). The dots had a disparity of 41 min arc with respect to the plane of the screen, whereas the white noise surround and occluders were at zero disparity. The 5 horizontal occluders (binary white noise, surround luminance: 34.69 cd/m<sup>2</sup>] size: 6.84°x0.68°) were

spaced so that 41 min arc wide strips of the background pattern were visible between them.

(3) Vertical occluded: I used the stimulus setup first described by Forte et al. (2002) to develop a stimulus that was consistent with a real 3D scene, but where each part of a background plane was only viewed by one eye. Observers viewed a foreground 'fence' (binary white noise occluders, luminance: 34.80 cd/m<sup>2</sup>, size: 0.68°x6.84°) with strips of background (each of width 41 min arc) only visible to one eye, or the other (figure 4.1c). This was achieved by generating an identical background image for each eye, then shifting each eye's view of the background pattern of dots by 20.5 min arc away from the centre in opposite directions, to deliver a relative disparity of 41 min arc. All the dots in the background were visible, but all were visible to only one eye.

In a 2 Interval Forced Choice (2IFC) task, participants were asked to indicate which one of two intervals contained the stimulus with more dots.

Two stimuli were presented in each trial. The standard stimulus was always of type (1), the test stimulus could be any of the 3 stimulus types. This led to 3 conditions:

(1) Binocular (in standard interval) – Binocular (in test interval)

(2) Binocular – Horizontal Occluded

(3) Binocular – Vertical Occluded

Observers were presented with a fixation cross for 1s, this was followed by the first stimulus interval which was displayed for 0.4s, a second fixation cross (1s), and the second stimulus interval (0.4s). Then a third fixation cross appeared, that stayed on the screen until participants made their response. Participants completed a total of 1350 trials each (60 trials for each stimulus level for conditions (1) and (3), 30 trials for each stimulus level for condition (2). Responses were given using one of two keys on a standard computer keyboard.

The stimuli were presented on a liyama 22in Vision-Master-Pro monitor (resolution: 1280 x 1024 pixels, refresh rate: 100Hz) and viewed through a Modified Wheatstone Stereoscope. The distance between the screen and observer was 100cm. The head position was stabilised using a chin rest. Stimuli were generated and presented using Matlab and the Psychophysics Toolbox 3 (Brainard, 1997; Pelli, 1997) on a PC workstation.

#### 4.2.2. Questions

There were a number of research questions this experiment tried to answer.

1. Do scenes appear less numerous when dots are presented in monocular regions that are embedded in binocular scenes?

Since experiments 2 and 2.2 suggested that there might be an effect of the occluders on participants' performance, rather than comparing the bias of

the point of subjective equality (PSE) between the binocular condition and the vertically occluded condition, I will be comparing the vertically occluded condition (3) with the horizontally occluded condition (2). Simultaneously, I have to ask whether there is an effect of the presence of the occluders on participants' performance. This leads us to the related (control) question 2.

#### 2. Is there an effect of occlusion on participant performance?

In order to be able to answer question 1 we have to first ask whether there is an overall effect of the occluders on participant performance. To do so we will look at differences in precision and accuracy of participants' judgements between conditions 1 and 2. A difference in precision (i.e. a higher JND) will indicate that when comparing stimuli in one of the two conditions we need a larger difference between the stimuli to perceive it. A difference in accuracy (i.e. a bias in the PSE) will indicate that one stimulus type appears less or more numerous than the other. This, in turn, would suggest that there is a difference in our percept of occluded information.

### 3. Is one eye's view ignored when making judgements about vertically occluded numerosities?

This question is related to question 1, but is more specific. It is possible that for stimulus (3), visual information is not seamlessly integrated from the two eyes. One possibility is that observers primarily use their dominant eye

and ignore the dots presented to the other eye. To explore whether observers used primarily one eye, or the other, when viewing stimulus (3), I used different dot densities in different regions of the display for all three stimulus types. The background stimulus was divided into ten 0.68° wide strips. Half of the strips were assigned 1/3 of the dots; alternate strips were assigned the remaining 2/3 of the dots. For stimulus (3), one eye was therefore presented with only the lower density strips, the other eye was presented with only the higher density strips. The aim of this manipulation was to test whether observers used information from only one eye when viewing stimulus (3). For example, if they used only the eye with 2/3 of the dots and assumed the density in the background to be constant, we would expect the perceived number of dots to be 50% larger when compared with the number of dots presented in stimulus (1). The higher density was presented to the right eye on 50% of stimulus presentations and to the left eye on the other 50% of presentations.

For stimuli (1) and (2) the arrangement meant that the local density was different in different regions of the stimuli. This manipulation did not affect observers' ability to perform the task compared to the previous experiments (compare participant thresholds for condition (1) in figure 4.3 and participant thresholds for the binocular conditions in experiments 1, 2, and 2.2 in figures 3.4, 3.8, and 3.9). Participants also did not notice a difference in the local densities when asked after completing the experiment. Note that this stimulus was perceived as flat and participants did not report perceiving the background plane as slanted or corrugated.

#### 4.2.3. Participants

8 participants, students aged 20-29, completed the study. All participants had normal or corrected to normal vision and normal stereo vision (TNO Stereo Test & Snellen EyeChart). Participants were paid expenses for their participation or volunteered. The experiment was approved by the University Teaching and Research Ethics Committee (UTREC) of the University of St Andrews. All participants gave written informed consent.

#### 4.2.4. Analysis

The proportion of trials for which the test intervals were perceived as containing more dots was recorded as a function of the proportion difference between test and standard. Psychometric functions were fitted as described in chapter 2.9.1. To compare the threshold levels (distance along the x-axis between the 50% and 75% points of the psychometric function) and PSE's (50% point on the psychometric function) between the different conditions, a 1-way-ANOVA was run on the values obtained for participants' thresholds and PSE's. To adjust the significance levels for multiple comparisons, the planned pairwise comparisons were Sidak corrected (see e.g. Miller, 1981, pp.254-255).

#### 4.3. Results

The 1-Way-ANOVAs for the thresholds and PSEs revealed that participants showed significant differences in PSEs when comparing between the different stimuli (F(2,21)=4.13, p=0.03) (see figures 4.2.b and 4.3.b). There

was also a significant difference in sensitivity between the different conditions (F(2,12)=9.77, p<0.01) (see figures 4.2.c and 4.3.c).

I will now go through the results in relation to the questions I asked in section 4.2.2., and present them graphically. When looking at the two graphs for each comparison, keep in mind that condition 2 appears in both graphs. Also, the 95% confidence intervals in the graphs are based on the bootstrapped fits of the data.

#### 4.3.1. Effects of horizontal occlusion

The main purpose here was to explore how information is integrated across regions that are partially occluded in one eye, due to occlusion by vertically oriented objects at different depths (used in condition 3). However, we first have to establish if there were any effects of horizontal occlusion (used in condition 2) on thresholds, or whether horizontal occlusion produced a perceptual bias. Figure 4.2 compares results for conditions 1 (binocular) and 2 (horizontal occlusion). Example psychometric functions for observer rp are shown in figure 4.2a, and summary data showing PSE's (figure 4.2b) and thresholds (figure 4.2c) for each observer.



Figure 4.2: a) comparison of fitted sigmoids for stimuli 1) (black, dashed) and 2) (red) for participant RP. PSEs and thresholds of all participants are compared in b) and c) respectively. Error-bars indicate bootstrapped 95% confidence intervals of the point estimates.

Figure 4.2b shows that PSEs were consistently higher for condition 2 than condition 1. This difference was significant (p = 0.02). This demonstrates that the pattern presented in stimulus 2 appears less dense than the pattern presented in stimulus 1. There was also a significant difference between thresholds found for conditions 1 and 2 (p = 0.03). However, while the point estimates of the threshold for the two conditions were different (see figure 4.2c), there was a large overlap in the bootstrapped confidence intervals. Unfortunately there is no simple way of integrating the bootstrapped confidence intervals into the analysis while testing for significant differences between the conditions. What the large overlap of the confidence intervals suggests is that the difference between the two conditions might not be as pronounced as it appears. The difference in the point estimates suggests that Participants were less sensitive to density differences when horizontal occluders hid 50% of the dots than when discriminating between density in fully binocular stimuli.

Note the large positive bias exhibited by participant ms for condition 2. This bias is consistent with this participant responding only to the visible dots on the background plane for condition 2, and comparing this to the number seen in total in condition 1, rather than assuming a constant density of dots behind the horizontal occluders. The large bootstrapped confidence intervals for participant ms are slightly misleading in this case. Since this participant shows such a large bias, this means the upper 25% of the psychometric function exceeded the proportion-difference range sampled from. This, in return, means the bootstrapped error bars might be too big because we used an inappropriate stimulus range for this participant.

What participant ms's performance suggests, is that while the majority of participants seemed to assume a constant density behind the horizontal occluders, the task could be interpreted differently and that this participant could be responding only to the *visible* dots. However, participant ms was an experienced observer who was also aware of the experimental design and stimulus generation. While this should not have affected their performance, they could have responded only to the dots they *knew* were being presented. All other participants were naïve in terms of the specifics of the experimental design and the stimulus generation which suggests that in general, when faced with occluders which cause parts of a stimulus to be completely occluded, people assume the background stimulus continues in a continuous fashion behind the occluders. This is consistent with the majority of people responding to texture density rather than raw numerosity.

I will now look at the effects of the vertical occluders on our perception.

#### 4.3.2. Effects of vertical occlusion

I next compared the results found for condition 2, with those for condition 3, to test specifically for an impact of delivering regions of the background to only one eye. Since the previous comparison shows that there is a significant difference between the two binocular conditions, I decided to compare the two occluded conditions with each other. The aim of comparing the two occluded conditions with each other, rather than the fully binocular condition with the vertically occluded condition, is to avoid interpreting changes in performance that are caused by the presence of the occluder as changes caused by the monocular regions.

Thresholds for basic measures such as contrast detection are as good in monocular zones as binocular regions, when monocular zones have a binocularly visible boundary (Su et al, 2009). Here I wanted to test whether the integration of information across several of these bounded regions was as precise and accurate as when the observer had a binocular view.



Figure 4.3: a) comparison of fitted sigmoids for stimuli 2) (red) and 3) (blue, dashed) for participant RP. PSEs and thresholds of all participants are compared in b) and c) respectively. Error-bars indicate bootstrapped 95% confidence intervals of the point estimates.

Figure 4.3a shows example psychometric functions for participant rp comparing the results for conditions (2) and (3). Figures 4.3b and 4.3c show PSE's and thresholds for all observers, respectively. Here, thresholds for the two conditions (figure 4.3c) were not significantly different (p= 0.19), suggesting that we are equally sensitive to numerosity in vertically occluded and horizontally occluded stimuli. There was also no significant difference between PSEs for the two stimuli (p= 0.14). However, I found large individual differences between participants' PSEs. I will discuss the individual differences for condition 3 in detail in section 4.3.3.

Overall, the lack of difference between the PSEs and thresholds suggests that, although we have very different content in the right and left eye for stimulus (3), and although the visual system must integrate information from across monocular zones to perform this task, statistically, there is no consistent evidence that the monocularly presented displays (stimulus 3) appear as less numerous than the binocularly presented displays (stimulus 2).

Stimulus type (3) was designed to deliver background dots to only one eye, or the other, yet be consistent with a real world situation where the background was hidden behind a vertical slatted fence, at a different depth (same distance as screen). The issue of interest is the extent to which the visual system can integrate information from across such a series of monocular regions. I tested for this by exploring any potential biases or lowered precision in the psychometric function. For example, the visual system might suppress one eye's view during viewing. As described in the Methods, one eye's view

contained 1/3 of the dots the other 2/3. Had suppression due to rivalry occurred, the non-dominant eye's pattern would likely have been suppressed (see Blake et al, 1980).

I found large individual differences in participant performance for condition 3. Participant sc showed a large negative bias which is consistent with them ignoring the lower density eye and assuming a constant density across the entire background plane. Participants sr, and ta, on the other hand, showed large positive biases, consistent with the vertically occluded test stimulus appearing as less numerous than the binocular standard. This could be caused by participants ignoring one eye's view either assuming a constant density across the background plane or only responding to the dots that are visible to the non-ignored eye. I will discuss this in more detail in the following section. If this pattern of results is caused by participants' consistently ignoring one eye's view this would lead to large differences in performance depending on which eye the higher density of dots was presented to.

### 4.3.3. Do participants ignore one eye's view?

As a control measure, I analysed condition 3 separately for the situations when the higher density had been presented to the right eye and vice versa. This comparison is shown in figure 4.4 for all participants. The psychometric function fitted to the data can be described using 4 parameters (Fründ et al. 2011, see also general methods, chapter 2.9.1.1). I used these parameters for a repeated measures ANOVA. The interaction of the

parameters with the eye the higher density was presented to were nonsignificant, F(1.40, 19.66) = 0.08, p= 0.86 (Greenhouse-Geisser corrected). There was also no significant main effect for the presentation eye, F(1,14) = 0.44, p = 0.84. This suggests that, overall, observers do not use the high or low density eye differentially. Figure 4.4 shows the fitted functions for all participants for the cases when the left or the right eye was presented with the higher density of dots.



Figure 4.4: Condition3 fitted separately for cases when the right eye (red, dashed) or the left eye (black) was presented with a higher density of dots.

No single participant shows significant differences between the trials when the higher density was presented to the left and when it was presented to the right eye (see overlap of confidence intervals in figure 4.4). This means that even though participants sr, and ta show a strong positive bias, consistent with the vertically occluded test stimulus appearing as less numerous than the binocular standard, these participants show this bias both when the higher density is presented to the left and when it is presented to the right eye. Participant sc, on the other hand, shows a negative bias consistent with only using information from the higher density eye and assuming this density remains constant in the entire background. However, as with participants sr, and ta, sc exhibits this pattern both for the left and the right eye situations.

These results suggest that the bias exhibited by these participants is not simply caused by participants ignoring the non-dominant eye's view. However, these participants might simply be ignoring the lower density eye's view, irrespective to which eye it is presented to. This means we cannot be sure that the vertically occluded numerosities appear less numerous than the binocular numerosities as a property of the percept.

#### 4.4 Discussion

In the present experiment I found, unlike in experiment 2.2 (chapter 3.4) but in line with the findings for the main experiment 2 (chapter 3.3.6.3), that there is an effect of binocular occlusion of background information on our sensitivity to dot density (conditions 1 and 2, compare figure 4.3 with figures 3.8 (exp. 2) and 3.9 (exp.9)). The size of the difference between the two conditions remains relatively constant across all 3 experiments. There is no

significant effect of making the occlusion monocular (by using a vertical occluder, condition 3) rather than binocular (condition 2). This is in line with the findings in experiments 2 and 2.2.

I have demonstrated that information from a number of separate monocular regions can be integrated into our overall percept of dot density. On a population-level, we are able to integrate regions of different densities that have no spatial connection into a meaningful percept and compare it to a binocular comparison stimulus with no loss of accuracy. While the two eyes receive the same input in typical amodal completion studies (see chapter 1.4), our occluded stimuli in condition 3 show different parts of the background plane to the two eyes. We do not find this affects observer performance.

In a study where observers viewed a scene containing objects (such as polygons and squares), Bruno and colleagues' (1997) found that an amodally completed 3D (geometrically plausible) occlusion situation (with monocular regions) leads to a percept that is comparable to that of a fully binocular view of the same scene. Our stimuli contained no objects and yet observers report seeing a continuous surface on the background plane. This is consistent with the findings of Bruno et al. (1997).

In contrast to other situations when the two eyes are presented with different input, there is no consistent support for the idea that the monocular zones that are behind the occluders generally lead to a rivalrous percept. Occasionally, an observer seemed to be responding to only one eye's view, but

most participants did not and could combine the information almost seamlessly. Further, all participants reported a stable percept.

Our results are also consistent with some other findings from the literature. Amodal completion occurs not only in binocularly occluded regions but also when a region is only occluded monocularly such as is the case for the blind spot (this occurs both in the natural blindspot, Durgin et al., 1995, and pathological scotomas, Tripathy & Levi, 1999). Unlike for information that is filled-in at the blind spot, my work in this chapter shows that the majority of observers do not ignore one eye's view of a vertically occluded stimulus.

The binocular occluders that created the monocular zones in stimulus 3 can be considered as forming boundary contours. He, Ooi and colleagues (e.g. van Bogaert, Ooi & He, 2008; Ooi & He, 2006) used stimuli that were constructed such that there was only one element in the stimulus that could be interpreted as partial occlusion (i.e. there was only one monocular region). In contrast, the stimuli used here contained several monocular regions which, due to the scene geometry (see fig. 1b), could be in conflict with each other when we are trying to match the correct foreground regions with each other – while the binocular occluder regions were in the same location in the two retinal images, the monocular regions in the left eye were in the same locations as the monocular regions in the right eye. If our visual system were to try to match two non-corresponding, conflicting, monocular regions with each other this could be expected to lead to a rivalrous percept (see chapter 1.5). Yet, participants' thresholds suggest that the monocular regions in the present

stimuli are integrated into our percept and are perceived as part of a stable 3 dimensional scene. These results also show that the patterns in the monocular regions are used fully to make a numerosity judgement. This suggests that the white noise occluders act as an anchor for a number of monocular regions (in our stimuli) just as monocular boundaries do for the single partially occluded region in the stimuli used by Ooi & He, 2006. You could think of this as follows: even though the left and right eye monocular regions are in the same location in the two retinal images, they are attached to the binocular contours (the edges of the occluders). This connection allows us to 'anchor' them within the 3D scene and interpret them as lying behind the binocular occluders. It seems that this allows us to use information in monocular regions in the same fashion as binocularly presented information.

In all the experiments I have discussed so far, the dots were either presented binocularly or monocularly, participants did not have to integrate both monocular and binocular dots to be able to make a judgement about the overall numerosity. This means, while the occluders were intended to ask the question of how monocular and binocular regions are integrated, the actual numerosity judgements did not require any integration of the two types of regions to be performed. In other words, while participants had to integrate monocular and binocular regions to arrive at a stable percept, the numerosities themselves were always either monocular or binocular. I had no situation in which the dots themselves were split across monocular and binocular regions. This opens up the question – are we actually able to integrate information in

monocular and binocular regions to form an overall percept of density or number?

In the following experimental chapter I will introduce a stimulus that allows us to ask this specific question – are observers able to integrate monocularly and binocularly presented numerosities to make a texture density judgement or do they rely on one type of region?

# 5. Experiment 4: The integration of monocular and binocular numerosities.

When we look around, there are a vast number of monocular regions embedded in our binocular view of the world. While the experiments in the last two chapters have showed that we can make judgements about information presented in these monocular regions, as well as about information that is presented binocularly, so far I have not asked how well monocular and binocular information is integrated between the two types of regions. In this Chapter I ask: are observers able to integrate monocularly and binocularly presented information not only to form a stable percept, but also to make judgements about the content of the regions?

#### 5.1. Aims

The aim of this experiment was to investigate whether observers are able to integrate numerosities across both monocular and binocular regions. The previous experiments suggested that monocular regions are as reliable and thus as useful as binocular regions.

The present experiment asks whether we are able to integrate information from the two types of regions and if so, whether this affects our performance.

#### 5.2. Methods

#### 5.2.1. Stimuli/Procedure

A scene very similar to the vertical occluder condition used in experiment 3 (chapter 4, section 4.2.1) was used. However, the background plane was moved closer towards the occluder plane, which led to both monocular and binocular regions being visible. This was done in the following fashion:

The background plane contained a grey square of size  $6.84 \times 6.84$  degrees (luminance  $32.47 \text{ cd/m}^2$ ), which had black (luminance  $0.01 \text{ cd/m}^2$ ) and white (luminance  $66.54 \text{ cd/m}^2$ ) dots ( $4.11 \times 4.11$  min arc) distributed across it. 50% of the dots were black, 50% white. The dots were arranged in 'stripes' of dots that would be visible monocularly or binocularly. The remainder of the screen was filled by a binary white noise texture (the luminance of each pixel,  $1.37 \times 1.37$  min arc, was allocated randomly).

This was overlaid with fence-like vertical occluder stripes (made up of binary white noise, luminance: 34.80 cd/m<sup>2</sup>, size: 0.68°x6.84°). The occluders were spaced 0.68° apart. While the occluders were placed in the same location in both eyes' views, the background plane was shifted outward by 0.23° which meant between the occluders there was a binocular region (width: 0.46°) as well as a monocular region (width: 0.23°) visible to each eye. This placement of the two stimulus planes means the background plane appeared to lie 6.18cm behind the occluder plane. Figure 5.1 shows a schematic of the scene and how the monocular and binocular strips on the background were arranged.

In this setup, unlike in the previous vertically occluded stimuli, there are regions of the background plane that are completely occluded from view by either eye. For the purposes of this experiment I will be referring to the *visible* dots (both binocular and monocular) only.



Figure 5.1: A schematic top down view of the stimulus arrangement used for experiment 4. The solid lines bound the area visible to the left eye and mark the regions of the background plane at the top visible to the left eye. The dashed lines mark the same for the right eye. The areas on the background plane where the solid and dashed lines overlap are binocularly visible. Note that this is a schematic, the distances do not correspond to the actual stimulus used.

This stimulus was presented in the following 5 configurations:

(1) Baseline - The monocular and binocular regions had the same density of dots, i.e. ½ of dots were presented binocularly, ½ monocularly.

(2) Binocular High 1 - 3/3 of dots were presented binocularly, 1/3 monocularly.

(3) Binocular High 2 - <sup>3</sup>/<sub>4</sub> of dots were presented binocularly, <sup>1</sup>/<sub>4</sub> monocularly.

(4) Monocular High 1 - <sup>2</sup>/<sub>3</sub> of dots were presented monocularly, <sup>1</sup>/<sub>3</sub> binocularly.

(5) Monocular High 2 - <sup>3</sup>/<sub>4</sub> of dots were presented monocularly, <sup>1</sup>/<sub>4</sub> binocularly.

Figures 5.2.i and 5.2.ii show screenshots of the 5 stimulus configurations used in the experiment (the left square was viewed by the left eye, the right square by the right eye).



1) Baseline



2) Binocular High 1



#### 3) Binocular High 2

Figure 5.2i: The first 3 stimulus configurations. The images are uncrossed (but not cross!) fusable.



4) Monocular High 1



5) Monocular High 2 Figure 5.2.ii: The last 2 stimulus configurations. The images are uncrossed (but not cross!) fusable.

Using a 2IFC task, participants were asked to indicate which one of two intervals contained the stimulus with more dots. Two stimuli were presented in each trial. The standard stimulus was always of type (1), the test stimulus could be any of the 5 stimulus configurations. This led to 5 conditions:

- (1) Baseline Baseline
- (2) Baseline Binocular High 1

- (3) Baseline Binocular High 2
- (4) Baseline Monocular High 1
- (5) Baseline Monocular High 2

As in the experiments described in the previous chapters, stimuli were presented on a liyama 22in Vision-Master-Pro monitor (resolution: 1280 x 1024 pixels, refresh rate: 100Hz) and viewed through a Modified Wheatstone Stereoscope. The distance between the screen and observer was 100cm. The head position was stabilised using a chin rest. Stimuli were generated and presented using Matlab and the Psychophysics Toolbox 3 (Brainard, 1997) on a PC workstation.

Observers were presented with a fixation cross for 1s, this was followed by the first stimulus interval which was displayed for 0.4s, a second fixation cross (1s), and the second stimulus interval (0.4s). Then a third fixation cross appeared, that stayed on the screen until participants made their response. Participants completed a total of 1800 trials each (40 trials for each stimulus level for each condition). Responses were given using one of two keys on a standard computer keyboard.

#### 5.2.2. Questions

There were a number of questions this experiment tried to answer. 1. Are we able to integrate monocular and binocular regions to form an overall percept of a stimulus that can be compared to another stimulus? Put more simply, I am interested in whether the thresholds for the baseline condition 1 are much higher than for the stimuli used in previous experiments and whether participants report a stable perception of the stimulus. If this stimulus leads to a rivalrous percept if monocular and binocular regions are not integrated, this would show up in a heightened threshold for this condition (compared to the thresholds for comparable conditions in previous experiments).

### 2. If the texture density in the presented monocular and binocular regions is different, do we start ignoring one type of region?

In conditions 2-5 the monocular and binocular regions contain proportionally different numbers of dots. If participants were to ignore one type of region this would lead to a biased percept of the overall numerosity and would lead to a bias in their responses. From the results in experiment 3 (chapter 4, section 4.3.4) we have learned that there are two ways for participants to deal with regions that they cannot access (be it because it is fully occluded or ignored). They could either assume a constant density of dots across the entire background plane or they can base their judgement solely on the visible/not ignored regions. If participants were to ignore the monocular regions and rely solely on the input from the binocular regions, we would arrive at two sets of extreme predictions. Let us start with the scenario in which participants base their numerosity judgement only on the binocularly visible dots. In this scenario the perceived numerosity of stimulus 2 would be 16.7% higher than the presented baseline numerosity (66.7% of dots binocular compared to 50% binocular), stimulus 3 would have numerosity overestimated by 25%, stimulus 4 underestimated by 16.7%, and stimulus 5 underestimated by 25%.

In a second possible scenario, participants assume that the density of dots remains constant across the entire display, including in monocular areas, that are ignored (i.e. if the binocular region contains a large number of dots the ignored monocular regions are assumed to contain the same number of dots in the same amount of space). This means that the perceived numerosity of stimulus 2 would be 33.3% higher than the presented numerosity, in stimulus 3 the perceived numerosity would be overestimated by 50%, in stimulus 4 underestimated by 33.3% and in stimulus 5 underestimated by 50%.

I will mark both sets of predictions on the graph depicting the biases so we can compare them to the results.

3. Does participants' sensitivity to differences between two numerosities drop when the density in the monocular and binocular regions is varied?

Is there a rise in threshold when participants are asked to integrate more different densities (conditions 3 and 5) compared to more similar densities (2 and 4)? If participants are asked to integrate monocular and binocular regions that are very different we might expect a change in performance, caused by
the increased difficulty of the integration, as the difference between the numerosities is increased.

### 5.2.3. Participants

4 participants, students aged 22-25, completed the study. All participants had normal or corrected to normal vision and normal stereo vision (TNO Stereo Test & Snellen EyeChart). Participants were paid expenses for their participation. The experiment was approved by the University Teaching and Research Ethics Committee (UTREC) of the University of St Andrews. All participants gave written informed consent.

#### 5.2.4. Analysis

Psychometric functions were fitted as described in the methods chapter (2) and the previous experimental chapters (3 and 4). Two 1-Way-ANOVAs compared the threshold levels and PSE's for the different conditions. The planned pairwise comparisons were Sidak corrected (see e.g. Miller, 1981, pp.254-255).

## 5.3. Results

There was a significant difference between participants' bias for the different conditions (F(4,19)=36.81, p<0.01). There was, however no significant difference in sensitivity between the different conditions (F(4,19)=1.34, p=0.30).

I will now go through the results in relation to the questions I asked in section 5.2.2.

## 5.3.1. Are monocular and binocular regions integrated?

The main aim of this experiment was to investigate whether monocularly presented numerosities were used to the same extent as binocularly presented ones when making texture density comparisons. However, to do so we have to first ask whether the stimulus used in this experiment is integrated seamlessly into a stable percept and whether participants' sensitivity to differences between stimuli remains comparable.

Participants reported a stable percept and were just as sensitive to differences between two stimuli as participants were in the previous experiments. For example, the vertically occluded condition in experiment 3 lead to a mean threshold of 0.28 which is comparable to the threshold of 0.30 for the baseline condition 1 in the present experiment. Figure 5.3 compares the three vertically occluded conditions.



Figure 5.3: Comparison of vertically occluded conditions across experiments 2.2, 3, and 4. The error-bars depict the 95% confidence intervals.

Keep in mind that I excluded participants in experiment 3 for having excessively high thresholds; I did not have to exclude any participants for this experiment since all participants showed relatively comparable performance.

Let us now consider how well participants perform when the density is varied in the monocular and binocular regions.

# 5.3.2. The effect of differing dot densities

The aim of this manipulation was to investigate whether participants are able to integrate both monocular and binocular regions to form a stable percept that uses information from both types of regions. The baseline stimulus has a constant dot density across the stimulus plane. Participants could, in theory, completely ignore either monocular or binocular regions altogether in this baseline condition, and still show no bias in their performance. If a participant were to use this strategy for conditions 2-5, however, this would lead to a highly biased percept. For example, if monocular regions were completely ignored and the numerosity comparison was made based solely on the binocular regions this would lead to a bias of -0.16 in condition 2, of -0.25 in condition 3, of +0.16 in condition 4, and of +0.25 in condition 5.

Figure 5.4 shows participant bs's psychometric functions. The vertical lines in 5.4.a) mark the predicted bias based on scenario 1 from section 5.2.2, the vertical lines in 5.4.b) mark the predicted bias based on scenario 2. This observer is clearly biased for conditions 2-5 and scenario 1 seems a good fit.



**Proportion Difference** 

Figure 5.4: Fitted functions for participant bs for all conditions. The vertical lines in a) are the predicted biases if the monocular regions were completely ignored and only the binocular regions compared across intervalls. The vertical lines in b) are the predicted biases if the monocular regions were completely ignored and a constant dot density was assumed across the entire background plane

The remaining participants performed similarly, all showing considerably biased PSEs. Figure 5.5 shows the PSEs for the different conditions for all participants.



Figure 5.5: The PSEs for the different conditions for all participants. Mean PSEs are displayed by the solid lines, predictions based on suppression of monocular regions (section 5.2.2) are displayed by evenly (only binocular regions) and unevenly (constant density assumed) dashed lines.

These large biases are significantly different from the baseline. These biases are also consistent with monocularly presented numerosities being ignored (scenario 1 above). While biases are not as large as one would expect if a constant density of dots was assumed across the entire background plane, they are consistent with the scenario in which numerosity judgements are based solely on the binocular regions. The observed biases are a strong indicator that information presented in these monocular regions is not integrated into the overall percept in the same fashion as information in binocular regions is. The fitted psychometric functions with data for all participants can be found in appendix A5.1.

There were no systematic effects on sensitivity, there was no significant difference between the different conditions. Figure 5.6 shows the thresholds for

all participants.



Mean thresholds are displayed by the solid lines.

When we look at the thresholds across the different participants, there is no systematic difference. Note how there is not even a trend between the different conditions.

# 5.4 Discussion

The aim of this experiment was to investigate whether and how well observers are able to integrate monocular and binocular information to form a stable percept and make judgements about its contents.

I have shown that there is a distinct change in bias when the density of monocular and binocular regions in the stimulus differs. Varying the density (i.e. the number of dots in a monocular or binocular region) across the two types of regions allows us to ask whether we use both types of regions or whether we rely on binocular regions when comparing two sequentially presented stimuli containing both monocularly and binocularly visible dots. The observed bias is consistent with the monocular regions being ignored.

On the other hand, there is no drop in sensitivity across the different conditions and we are not just as sensitive to differences between the stimuli in the baseline condition as we are to vertically occluded stimuli in experiments 2.2 and 3.

While this does not contradict the results of the experiments in the previous chapters and with findings from depth perception (see chapter 1), one might have expected that we are able to utilise the monocularly presented information better. In particular, experiment 3 had suggested that a vertically occluded stimulus containing only monocular regions in the background appeared as equally numerous as a comparable binocular stimulus. This means when only integrating across monocular regions we are capable of using monocular information to make relative numerosity judgments. When we are required to integrate this monocular information with binocular information (Experiment 4), this process seems to break down and we rely solely on the binocularly presented information. Why is this the case? In experiment 3 we were able to use monocular information. Similarly, monocular information is used when making judgements about the 2 and 3 dimensional shape of an object, whether this object is visible (e.g. Wilcox & Lakra, 2007)) or amodally completed (e.g. Bruno et al. 1997). So why ignore the monocular regions for this stimulus?

I can only speculate on this point, but it seems likely that this is due to

the difference between the tasks. With the present stimulus, participants report a stable percept just as in my own previous experiments (experiments, 1-3, chapters 3-4) and the discussed experiments on amodal completion (e.g. Bruno et al. 1997) and on depth perception (see chapter 1). The similar thresholds when the density is varied between monocular and binocular regions (figure 5.6) suggests that as the difference between the monocular and binocular regions becomes accentuated, the monocular regions do not interfere more or less with the binocular regions. It appears as if participants fully ignore the monocular regions when making the numerosity judgements. The bias suggests that the dots used in this experiment appear significantly less numerous. Unlike in other experiments (such as in Forte et al. 2002) where participants were asked to make a judgement about the overall stimulus, this experiment (while using the same task) required participants to assess the content of the monocular regions and then integrate it with the content of the binocular regions to arrive at a non-biased percept.

Why go through the effort of integrating monocular information with binocular information yet use only binocular regions for numerosity judgements? Wouldn't it be much simpler to either completely use monocular information (if we are already integrating it into our percept) or to treat monocular information as noise and ignore it completely? It could be that the way we use monocular information is more computationally expensive (since we have to match the binocular regions between the two eyes views and integrate the monocular regions into this percept) while not giving us any

additional information (since we are not using the monocularly visible dots to make the numerosity judgements) than either alternative.

One suggestion for why we do not completely ignore information presented in monocular regions, is that using the information is advantageous from a survival point of view (e.g. Changizi & Shimojo, 2008).

For example, imagine standing in a forest, there will be a large number of trees and shrubs at varying distances. If you were foraging for berries, it would be clearly beneficial to be able to assess the number of berries on a shrub, even if parts of it were 'hidden' behind another. More of the berries are potentially visible if information from both eyes were used, so at least theoretically it would be really useful if we were able to assess the overall number of berries on a tree accurately to compare between different trees so we could choose the tree with the highest number of berries. The present experiment suggests that we are not very good at this task when some berries are visible to only one eye and some are visible to both.

Do we actually use these monocular regions as anything other than supporting information? More specifically, do we use the monocular regions simply to support a conclusion we have already arrived at based on binocular regions? If there is some truth to the evolutionary argument that information that is present in monocular regions allows us to detect predators that might be somewhat hidden from sight by trees and bushes, it would seem likely that we use monocular regions as more than just supporting information. Perhaps monocular regions attract our attention towards them, to allow us to interact

with binocular regions in their surround. To investigate this question I decided to employ a visual search experiment and study both the effect of the monocularity of a visual search target on our overall visual search performance and on our eye-movements during said visual search.

In the following three chapters I will introduce the paradigm used and then discuss first two reaction-time versions of it, followed by an eye-tracking version.

# Part II – Searching for monocular targets

#### 6. Introduction

The experiments in chapters 3, 4, and 5 have shown that we do not ignore monocular regions when making judgements about a binocular scene. We have also learned that we use this monocular information nearly as efficiently as we use binocular information. What I have, so far, not asked is whether we are able to exploit the additional information in the monocular regions and interact with it. This is the purpose of the work in the following chapters (7 and 8).

Previous research on amodal completion has shown that the presence of monocular regions affects our processing of a visual scene, changing our reaction times and accuracy at visual search tasks (see section 1.4). None of these experiments investigated whether the content of monocular regions, rather than the mere presence of monocular regions, affects our visual search performance. In this second set of experimental chapters I will ask just this question.

### 6.1. Aims

This chapter will introduce the research that motivated the experiments in chapters 7 and 8. I will discuss some of the computational models that led me to ask how far our visual search behaviour differs for monocular and binocular items, and I will conclude by introducing the stimuli that inspired my experimental design, in more detail. Before doing any of this, though, let us go back to the basic question of what we might use the information in monocular regions for.

#### 6.2. What do we use information in monocular regions for?

Imagine standing in the jungle and seeing a tiger, which is hidden behind several trees. Since a tiger is camouflaged, it might, depending on the light, be quite hard to pick out from grass and foliage in the foreground, and from more trees in the background. In such a situation it would be prudent to use as much of the information provided by the environment as possible. Having two eyes allows more of the tiger to be seen, but the extra visual information is contained within monocular regions. Also, note that the scene you would see contains a whole host of monocular regions all over the scene due to the large number of trees, leaves and foliage both in the foreground and background.

While one might argue that attempting to use the monocular information increases the initial computational load, being able to use the added information accessible through viewing such a scene binocularly seems advantageous. Let us return to figure 1.4 (p.8), where only a small proportion of the background object can be seen in both retinal images. If this were the case for the tiger in the jungle, relying only on binocular information would mean ignoring the majority of the tiger. Previous research and the results of the previous experimental chapters suggest that we would definitely not ignore the tiger completely. At the same time we might not be not quite as good at making judgements about the appearance of the tiger.

Let us review what we know about how we process monocular information in binocular regions so far.

## 6.2.1. Recap: Monocular Regions & Depth Perception

Where our usage of monocular regions for depth perception is concerned, we know that monocular regions coincide with large depth discontinuities (Wilcox & Lakra, 2007). It has been suggested that we interpret monocular regions based on either an innate or learned knowledge about geometrical occlusion constraints (Nakayama & Shimojo, 1990). This means, in the plausible situation that a monocular region is placed next to a foreground and a background (both of which are identifiable as such due to shading, colour or texture differences, Harris & Wilcox, 2009), then the monocular region is interpreted as lying at the same depth in space as the background plane (Anderson & Nakayama, 1994; Collett, 1985; Julesz, 1971; Shimojo & Nakayama, 1990). In such geometrically plausible situations, depth perception is faster (e.g. Gillam & Borstig, 1988; Saye & Frisby, 1975; Wilcox & Lakra, 2007). Divergently, if the 3D-scene geometry of the placement of a monocular region is less plausible, such as when the monocular region shares the texture of the foreground, we perceive less depth in a display (Grove, Gillam, & Ono, 2002).

The findings for the effect of monocular regions on amodal completion follow along similar lines.

## 6.2.2. Recap: Monocular Regions & Amodal Completion

Amodal completion appears to be faster when we view a binocular scene that is consistent with naturally occurring occlusions than when the two eyes views are identical (Bruno, Bertamini & Domini, 1997). We respond to an amodally completed object that is defined by a monocular region in the same way as we would respond to the whole object (He & Nakayama, 1992), showing appropriate changes in reaction times and accuracy.

This suggests that we are able to integrate monocular information into our overall percept of a scene. In how far is this reflected by our conscious percept of such monocular information?

## 6.2.3. Recap: Our Percept of Monocular Regions

In the previous experimental chapters I have investigated our percept of monocular regions. So what have we learned so far?

We have a stable percept of monocular regions. This is equally the case when we view a binocular scene that has only one, geometrically plausible monocular item (e.g. Nakayama & Shimojo, 1990), when we are integrating larger numbers of monocular and binocular zones (experiment 4), and in the extreme case when we are only integrating across spatially distant monocular regions (experiments 1-3, Forte et al., 2002).

However, it seems that in order to accommodate the additional information provided by monocular regions, information contained in them can

appear slightly compressed and displaced (Ono et al., 2003).

When we have to integrate information across spatially distant monocular regions, we are just as accurate at making judgements about said information as when judging information contained in binocular regions (experiments 1-3). We are just as accurate and precise when integrating across monocular and binocular regions that are similar (experiment 4). When the two types of regions contain conflicting information, however, we ignore the monocular information and rely on the information provided binocularly (experiment 4).

Changizi & Shimojo (2008) suggest that leafy environments such as the previously mentioned jungle scene made forward facing eyes an evolutionary advantage. They argue that a cluttered environment (whether an environment is cluttered depends on the animals size, i.e. their example is that a forest will appear cluttered for a bigger animal but uncluttered for a mouse due to its size) means that the loss of the extra information from a wider visual field (as for example horses have) is made up by the extra information gained by being able to see past the first layer of leaves (which is something two forward facing eyes allow us to do). If we are able to integrate the two eyes' views, this allows us to see past this first layer of occluders (remember the added regions that using the information from the two eyes gives us in figure 1.4).

We now know that we can do this quite well. However, the experiments in the previous chapters have focused on how well we are able to integrate monocular and binocular regions to make judgements about the information

contained in them. If the monocularity of some information means forward facing eyes are advantageous from an evolutionary perspective, one might wonder whether we interact with these monocular regions in a slightly different fashion as with binocular regions. More specifically – do they affect our visual search behaviour? If the speed-up observed for depth perception is any indication, we might expect an improvement in our overall visual search performance.

Before I go further into this question, let us step back and review some of what we know about how we might process monocular regions. I will now discuss some of the computational models describing how the visual system might process monocular regions.

# 6.3. How is information in monocular regions processed? - Computational models.

While there is no model of how the brain processes monocular information specifically, there are a number of models of stereo vision that take monocular information into account, some of which might allow us to shed some light on the question of how rivalry, amodal completion and our percept of monocular regions might relate to each other. I will discuss some of these models in this section.

Marr and Poggio (1976) developed an algorithm to extract binocular disparity, that was able to match the correct image points from the left and the

right eye to each other. They introduced constraints that the visual system might use, based on properties of the world. For example, one image point can only occupy one point in space at a time; that is, every point has a unique location (Marr, 1982). At the same time Marr (1982) used the observation that objects tend to have a relatively smooth surface. Making this assumption about a scene eases the correspondence problem significantly because it reduces the possible number of image points that can be realistically matched with each other. There have been many improved models developed since Marr's work, but the majority of theories and algorithms that support the view that the smoothness constraint is central to stereo matching work on the assumption that unmatched image points are noise and that the visual system treats them as such.

There are several models of disparity extraction that incorporate monocular information from monocular regions, like the unmatched point illustrated in figure 1.3 (p.6). These will not be discussed in detail as the focus of this thesis is not a computational one. However, see Egnal and Wildes (2002) for a review of the different approaches utilised by several computational models. What this section is intended to do is draw your attention to *some* of the models of stereo-matching that take monocular regions into account with a focus on those models that might allow us to better understand how monocular regions might relate to amodal completion, rivalry, and differences in visual search performance.

Let us return to Marr and Poggio's constraints. Marr & Poggio (1979)

make two assertions in their cooperative stereo-algorithm – uniqueness of a point and smoothness of surfaces. The algorithm matches points in the two 'retinal' images in the following (simplified) fashion:

1. the images are filtered with bar masks (Wilson & Giese (1977) suggest that receptive fields of simple cells can be modelled using linear units, called bar masks, that have bar-shaped receptive fields) at different orientations.

2. then the algorithm looks for zero-crossings in the second directional derivative in the filtered images as they are indicative of edges in objects (this implements the smoothness constraint).

3. next zero-crossings of the same sign are matched between the two images (uniqueness constraint)

The algorithm allows us to arrive at the 2½-D sketch (Marr&Poggio, 1979).

Watanabe and Fukushima (1999) extend this model with detectors for unmatched regions. So until zero-crossings are matched between the two images, the Watanabe and Fukushima algorithm is essentially the same as Marr and Poggio's cooperative stereoalgorithm. At the end of step 3) in the Marr and Poggio's algorithm, zero-crossings in the two images can be classified as either having a match in the other eye or not having a match (this can occur either in the left or right image). Watanabe and Fukushima apply what they call the 'occlusion constraint' at this stage, this means that based on the occlusion geometry discussed earlier, and depicted in Figure 1.4 (p.8), an

unmatched region in the right eye will only occur naturally if there is an occluding object between the left eye and the background. This means the zero-crossing which occurs in the left eye will have no match in the filtered right eye image.

This occlusion constraint is applied in the following fashion:

Two layers of 'cells' in the model perform two very different computations at the same time (iteratively in this case). One layer reconstructs the 3D surface based on the matched zero-crossings, the other layer 'discriminates' (Watanabe & Fukushima, 1999) the unmatched zero-crossings. The output from the each layer is used as input for the other layer in the next iteration, thus allowing the layer performing the 3D surface reconstruction to use the unmatched zero-crossings as a cue to large depth steps.

However, the 'receptive fields' used in this model are of a binary nature. This means they respond to one disparity and show no response to all other disparities. While this proves very useful for this model and, allows us to reconstruct a 3D surface, this is not physiologically plausible. Qian (1997) points out that receptive fields of V1 cells are tuned to a range of disparities. He suggests that even when a cell is incredibly sharply tuned to one disparity, this will still mean that it is tuned to disparity range of at least 0.2° (and most cells will have a much broader tuning curve). This means that a model which approximates receptive fields with a filter that leads to a binary response is not physiologically plausible. While such models might allow us to arrive at a similar 3D representation of a scene as our visual system does, they will not

allow us to make any statements about how the visual system might arrive at the percept.

Hayashi and colleagues (2004) extended the model proposed by Watanabe and Fukushima (1999) while taking V1 physiology into account. Their model uses a disparity energy model (Ohzawa et al. 1990). The disparity energy model is based on V1 physiology. To model visual processing by simple and complex cells, the algorithm convolves input from each retinal image with a Gabor function that describes the receptive field shape of a simple cell. Then the input from two receptive fields with a 90° phase difference in the two eyes are summed, and the binocular sum for each receptive field is then squared and summed with the corresponding receptive field's sum. This leads to a good approximation of V1 complex cells.

The model proposed by Hayashi and colleagues (2004) uses the output of the disparity energy model and adds an additional layer, which detects monocular regions by examining the population response of the disparity energy neurons. While binocular regions lead to excitation of neurons within a very specific disparity range, a monocular region leads to a much broader activation at various disparities. They then add an occlusion constraint. When the two retinal images are lined up there can only be one monocular region in a specific location. This means this model does not only allow us to describe processing of monocular zones in binocular scenes but also of rivalrous stimuli, suggesting rivalry and our perception of monocular zones might rely on the same binocular suppression mechanism.

While Hayashi and colleagues (2004) claim the first stage of their model is based on a physiologically plausible mechanism, this is not quite the case. Although they apply receptive fields to the input stimuli, they then convert the responses of the receptive fields into a binary response, thus the main computation is then done on responses that are not very different from those generated by the Watanabe and Fukushima (1999) model (Assee & Qian, 2007).

So where does this leave us? Assee and Qian (2007) suggest that V2 disparity edge selective cells allow us to model how the visual system detects monocular occlusions. They argue that monocular (V1) cells cannot always detect occlusions, but if the monocular cells feed forward into binocular (V2) cells via feedforward connections, then the population response of the binocular cells allows us to reliably detect monocular occlusions and which eye they occur in.

While this model is definitely more physiologically plausible than previous models, we have to ask how much the physiological plausibility adds to our understanding of monocularly occluded regions. None of the discussed models take into account the question of what happens to our percept of information *in* these regions. Zhaoping (e.g. 2002, 2006) proposes a mechanism that might allow us to detect monocular regions rapidly and suggests how we then direct our attention towards such regions. I will discuss this model in the next section. Watanabe and Fukushima (1999) use a simple occlusion constraint in combination with a detector for unmatched zero-

crossings to detect monocularly occluded regions, which is a surprisingly computationally efficient way of correctly identifying unmatched regions. Hayashi et al (2004) suggest a tie to rivalry mechanisms. This is interesting, especially given that geometrically implausible occlusions lead to rivalrous percepts (Forte et al. 2002). All these different models highlight how the inclusion of monocular regions into the overall percept can aid our depth perception.

If we assume that our visual system applies an occlusion constraint such as the one suggested by Hayashi and colleagues (2004), then this would suggest that processing does not have to occur in higher visual areas but could already be accomplished at V1 or, if we require the model to be more physiologically plausible, such as the one proposed by Assee and Qian (2007), then our visual system could well be able to distinguish between naturally possible and impossible occlusions based on the population response of disparity edge selective cells.

It has been suggested that visual attention is driven by bottom-up processes (e.g. Zhaoping, 2002). If this is the case and if monocular occlusions can be detected as early as V1 or V2, these occlusions might directly attract our attention and thus affect our visual search behaviour. I will now describe this idea in more detail.

### 6.4. Searching for monocular items

Zhaoping (2008) asks whether we show improved visual search performance when searching for items in a stimulus that are visible to one eye while the distractors are presented to the other eye (she defined such targets as ocular 'singletons'). Her question seems to be related to the question I have been asking – whether monocular items lead to changes in our visual search behaviour from what we find with binocular targets. I will now introduce the experiment used by Zhaoping (2008) and discuss its implications for my question.

# 6.4.1. Searching for ocular singletons and visual saliency

In a series of 4 experiments, Zhaoping (2008) investigated whether ocular singletons can attract our attention and thus improve our visual search performance.

To do so, she first asked whether we have an added benefit from eye-oforigin information when searching for a feature singleton (in this case a bar that was rotated by  $\pm 20^{\circ}$  compared to horizontal distractor bars). Participants viewed a black background with a grid of 660 bars. 659 of these bars were oriented horizontally whereas the target was the feature singleton. The target could be located at one of 28 locations, which were between 12° and 15° shifted outward from the centre of the screen. This stimulus was presented in one of 3 conditions. (1) 'Dichoptic Congruent' in which participants viewed all distractors in one eye and the target in the other, (2) 'Dichopic Incongruent' in which the target and 658 of the distractor bars were presented to one eye and one distractor was presented to the other eye, and (3) 'Monocular' in which all 660 bars were presented to the same eye. Figure 6.1 shows the stimuli used in this experiment.



c) Monocular

Figure 6.1: The stimuli used in experiment 1a in Zhaoping (2008). © ARVO, Reproduced with permission from the author and the copyright owner.

On a given trial, this stimulus was presented for 200ms, followed by a mask containing stars made up of several differently oriented bars similar to the ones used in the actual stimulus which remained on the screen until a participant responded as to which direction the target was tilted in. This response was not timed. Participants were significantly better at detecting a dichoptic congruent target than a monocular or dichoptic incongruent target. Zhaoping then asked whether the eye-of-origin information (available in the dichoptic congruent condition, but not the other two) was consciously accessible. She therefore ran a modified version of the experiment in which all bars were horizontal. The target bar was presented to one eye, all distractor bars were presented to the other eye. The target was present in 50% of trials and participants were asked to indicate whether a target was present on a given trial. When the luminance of all bars was constant, participants were significantly better than chance at detecting the ocular singleton target. However, if the luminance of the bars was non-uniform performance was at chance. This suggests that the eye-of-origin information is not consciously accessible but forms a very strong cue to attention (since in this part of the experiment the target was solely defined by it being an ocular singleton). Indeed, it provides a stronger cue than the feature-singleton (the oriented bar) used in the first part of this experiment.

In the later 3 experiments, reaction-times were measured for a stimulus very similar to the one used in figure 6.1. The orientation contrast between the target and the distractors was increased to  $50^{\circ}$  (either +25° or -25° rotated from

horizontal), furthermore, the luminance across all bars was uniform. Participants were asked to report the location of either an orientation singleton or a texture border (where the orientation of a group of bars suddenly changed from one column of bars to the next). Unlike in experiment 1, reaction times were recorded and the stimuli remained visible until participants responded rather than being masked.

In experiment 2 the stimuli could either be (1) binocular (the two eyes viewed exactly the same stimulus), (2) monocular (all bars were presented to one eye), or (3) 'dichoptic congruent' (the orientation singleton/texture change coincided with the ocular singleton/change in presentation eye). Participants were not informed of the different conditions.

In experiment 3 the stimuli could be any of the 3 conditions in experiment 2, or (4) 'dichoptic incongruent' (the orientation singleton/texture change did *not* coincide with the ocular singleton/change in presentation eye. Participants were informed of the different conditions and of the possibility that the ocular singleton distractor in condition (4) might appear more salient than the actual target and that they should try to ignore this distractor as best as possible.

In experiment 4 the conditions used were (2), (3), and (4) and, unlike in experiment 3, participants were not warned about the distractor.

An example stimulus for experiments 2-4 is shown in figure 6.2.



Figure 6.2: Example stimulus of the stimuli used in experiments 2-4 in Zhaoping (2008). © ARVO, Reproduced with permission from the author and the copyright owner.

In experiments 2-4 the reaction times mirrored the error rates in experiment 1. When an ocular singleton or a change in presentation eye was present at the location of an orientation singleton/ a texture border, then participant were significantly faster at detecting the singleton/border. When the eye-of-origin information was incongruent with the orientation changes, then visual search latencies were significantly slower. This was the case irrespective of participants' awareness of the presence of the ocular singletons. In experiment 2 participants did not report noticing differences between the conditions, yet their reaction times suggest they are responding to the eye-oforigin information as a very strong cue. This is in line with the findings of experiment 1b where participants were not aware of the presence of the ocular singleton, yet were performing significantly better than chance when asked to detect it.

Similarly, in experiment 3 participants had been instructed to try to ignore the singleton distractor as far as possible when searching for the differently oriented target or texture yet they were unable to do so effectively, leading to significantly slower reaction times.

When participants were not informed about the presence of a distractor (in experiment 4) and the only cue to the target location was the difference in eye-of-origin this pattern of results remained the same. Most observers did not notice the ocular singleton distractor/the distractor texture yet performed significantly better than at chance when a singleton target was present. This further supports the notion that eye-of-origin information is a very strong cue for attention, but not one that is processed at a high level, thus remaining unconscious. Zhaoping argues that eye-of-origin information is a strong bottom-up saliency factor which influences attention (and thus reaction-times and error rates) irrespective of whether we try to exert (top-down) control over our attention. This exemplifies the 'saliency hypothesis', a model of visual saliency which Zhaoping proposed. In the next section I will describe the saliency hypothesis in more detail. Note that I am describing this model to form a background for this set of experiments, but that I chose to use these experiments for their implications for monocular regions rather than the saliency-hypothesis itself.

# 6.4.2. The saliency hypothesis

Zhaoping (1998a, 1998b, 2002, 2006) proposes a model of bottom-up visual saliency which computes saliency maps by modelling pyramidal cells and interneurons in V1. The pyramidal cells receive 'visual' input from the region of the stimulus that is located in their 'receptive field'. A receptive field of such a cell is modelled by applying a filter to the input image that roughly

corresponds to a Gabor patch with a Gaussian applied to it. However, there are both monosynaptic excitatory connections between the different pyramidal cells, as well as disynaptic inhibitory connections that are mediated by the interneurons. This means while the initial response of a pyramidal cell is determined by its direct input, the response is then modulated by the surrounding cells.

This model makes a specific prediction about how the visual system should deal with eye-of-origin information. Based on the visual saliency model, differences in the eye-of-origin lead to increased saliency of an item in a visual scene, even if the item is not distinguishable based on other feature information.

Figure 6.3 is a schematic of the setup used in the different experiments in Zhaoping (2008) (compare to the stimuli shown in figures 6.1 and 6.2). The dotted bar in figure 6.3 is presented to one eye (e.g. left), and all other bars are presented to the other eye (e.g. right). The V1 saliency model predicts that the bar will be more salient both when it is also distinguishable by a feature (a) and when there is no difference in feature present (b).



Figure 6.3: The dotted bar is presented to the other eye than all other bars, this coincides with a difference in orientation in a) and is the only difference between the bars in b).

Unmatched points drive our attention on a both conscious and subconscious level (Zhaoping,2008). That is, we do not have to be consciously aware of an unmatched point for it to affect our performance in a visual search task. Moreover this explains why an ocular singleton can effectively attract our attention even when orientation singletons are present.

If the eye-of-origin has such large effects on visual saliency, is this only the case for situations when every item in a stimulus is monocular (as in Zhaoping's studies), or is this also the case for situations when the majority of items are binocular but a number of items are monocular? In other words, does monocularity drive visual saliency in the same fashion that eye-of-origin does?

I thus decided to adapt the stimuli used by Zhaoping (2008) to investigate whether monocular items affect our attention the same way ocular singletons do.

#### 6.5. Methods

#### 6.5.1. Stimuli

The experiments in chapters 7 and 8 are based on the singleton search task used in experiments 2-4 in the paper that I introduced in section 6.4.1 (Zhaoping 2008).

I chose a stimulus that was intended to require slightly more processing than the stimulus used by Zhaoping. While her stimulus had contained angled bars (see figure 6.2) and participants were asked to find the 'odd one out' I presented a C amongst O's and asked participants to find the C.

What all of these experiments had in common was that observers viewed two squares (the size and luminance was varied between chapters 7 and 8 because two different setups were used for the experiments in these chapters. I will discuss these aspects of the stimulus separately in the two chapters) which were framed by a line and contained tiles with letters. The two squares were separated by a gap so, once viewed through the stereoscope, each eye was presented with only one of the squares. Within the squares 225 tiles were arranged in a 15x15 square. On each tile a letter was displayed. The letters were located centrally on the tile but were jittered both vertically and horizontally to avoid false matches (the jitter was the same in the two eyes for each tile). Figure 6.4 shows an example of the stimulus used in experiment 5 (chapter 7). Note how the letters are rarely completely centred in their tile but how the shift is the same in both eyes' views.

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	
	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	
	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	
	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	000000000000000000	
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	000000000000000000000000000000000000000	
000000000000000	000000000000000	
0000000000000000000	000000000000000000	
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	
0.0.0.0.0.0.0.0.0.0.0.0.0	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	
0.0.0.0.0.0.0.0.0.0.0.0.0.0	0.0.0.0.0.0.0.0.0.0.0.0.0	
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	

Figure 6.4: Example stimulus from experiment 5. Note the jitter added to the letters.

#### 6.5.2. Procedure

In all experiments in chapters 7 and 8 participants were instructed to find the target C amongst the Os' and report the side (left or right) of the square it was located in by pressing a button. To avoid higher-order cognitive effects on the reaction-times, participants were asked to respond with the same digit of their left or right hand when the target was located in the left or right side of the square respectively.

In a given trial, participants viewed the square with a fixation cross at the centre. They were instructed to fixate the cross. They then initiated the stimulus presentation by pressing any button. In chapter 8 stimulus presentation was also fixation contingent. The stimulus remained visible until a response was made at which point it was replaced by the square and fixation cross until the next trial was initiated.

## 6.5.3. Analysis

In all experiments in chapters 7 and 8 I first calculated the harmonic mean for each participant for each condition. The harmonic mean is the reciprocal (multiplicative inverse) of the arithmetic mean of the reciprocals. Unlike the arithmetic mean, the harmonic mean has a strong tendency towards the smaller values in the list it is calculated from. This is very useful for reaction time data. Reaction times are positively skewed and thus simply calculating an arithmetic mean is inappropriate – if a participant simply forgets to respond then this outlier would dramatically increase the arithmetic mean. There are

several ways of calculating a measure of central tendency for reaction time data, some advocate the use of the mode, some suggest calculating the arithmetic mean and the standard deviation from it, followed by the removal of all data-points that are either two or three standard deviations larger than the mean and then to recalculate the arithmetic mean. While these are all valid ways of dealing with the skew in the data, I chose to use the harmonic mean as advocated by Ratcliff (1993) because it places less weight on the higher values (which are most likely outliers) while retaining the sample size. The first and third quartiles were used as a measure of spread.

In the following two chapters I will first describe three visual search experiments, two reaction-time experiments and one eye-tracking experiment.

# 7. Visual search for monocular items – Are we faster at detecting monocular targets?

The work on ocular singletons suggests that eye-of-origin information can attract our attention in a bottom-up fashion. If monocular regions in a binocular scene were processed similarly such a bottom-up effect on our visual attention might allow us to use said monocular regions to locate edges of objects more quickly or find the predator behind a tree that Changizi and Shimojo (2008) hypothesised about.

#### 7.1. Aims

The aim of experiments 5 and 6 was to investigate whether and how monocular items affect our reaction times during visual search. More specifically, by adapting the stimulus used by Zhaoping (e.g. 2008) using monocular targets/distractors in a binocular scene rather than only monocular items these experiments also ask whether we can draw parallels between the work on ocular singletons and what we know about monocular regions in binocular scenes.

# 7.2. Experiment 5 – searching for monocular items in binocular scenes.

This experiment was motivated by the question how environmentally plausible the stimuli used by Zhaoping (e.g.2008) are. She uses stimuli

containing ocular singletons to test her hypothesis of how visual attention is driven by a low-level bottom-up processing and these stimuli are a good way of testing the saliency hypothesis and the neural network model she proposes.

However, the issue with this is the fact that unless we are sitting in a lab looking at a screen through a stereoscope or through shutter-goggles, we are not likely to see a ocular singletons when we look around us. As I discussed in chapter 6, there are parallels between ocular singletons and monocular items in binocular scenes. This raises the question whether the findings for ocular singletons, especially about how they affect our reaction times, hold true for monocular items as well.

#### 7.2.1. Aims

The aim of this experiment was to investigate whether stimuli that followed the same principles as the stimuli used by Zhaoping (2008) would lead to the same pattern of results when it was presented with ocular singletons and when it was presented with monocular items.

## 7.2.2. Methods

## 7.2.2.1. Stimuli/Procedure

Participants viewed two black squares (luminance: 0.01 cd/m<sup>2</sup>, size:  $6.84^{\circ}$ by  $6.84^{\circ}$ ) which were framed by a white line each (luminance 66.54 cd/m<sup>2</sup>; width: 8 arcmin) on a black screen (luminance 0.01 cd/m<sup>2</sup>). The two squares
were separated by a gap of 4.49° so, once viewed through the stereoscope, each eye was presented with only one of the squares. Within each square 225 tiles (size: 27x27 arcmin) were arranged in a 15x15 grid. On each tile a white letter was displayed (size: 16x20 arcmin, luminance 66.54 cd/m<sup>2</sup>). The letters were located centrally on the tile but were jittered by up to 3.5 arcmin both vertically and horizontally to avoid false matches (the jitter was the same in the two eyes for each tile).

This stimulus was presented in one of 7 conditions.

(1) monocular target, all distractors binocular

(2) binocular target, 1 monocular distractor, all other distractors binocular

(3) monocular target, 4 monocular distractors, all other distractors binocular

(4) all binocular

(5) singleton target, all distractors in other eye

(6) 1 singleton distractor, all other distractors and target in other eye

(7) 1 'singleton' target, 4 'singleton' distractors, all other distractors in other eye

Even though neither the target nor the distractors are technically singletons in condition 7, I will continue using the term singletons for the items because this condition is an extension of conditions 5 and 6 and I am using the term to highlight their connection.

Figures 7.1-7 show crossed/uncrossed fusable screenshots of the different conditions.



Figure 7.1: Condition1 - Monocular target, all distractors binocular.



Figure 7.2: Condition 2 - binocular target, 1 monocular distractor, all other distractors binocular.



Figure 7.3: Condition 3 - monocular target, 4 monocular distractors, all other distractors binocular.



Figure 7.4: Condition 4 - all binocular.



Figure 7.5: Condition 5 - singleton target, all distractors in other eye.



Figure 7.6: Condition 6 - 1 singleton distractor, all other distractors and target in other eye.



Figure 7.7: Condition 7 - 1 'singleton' target, 4 'singleton' distractors, all other distractors in other eye.

When the target was monocular or a singleton, it was presented to right eye 50% of trials and vice versa.

In a given trial, participants viewed the square and bounding line with a fixation cross (size: 27x27 arcmin) at the centre. They were instructed to fixate the cross. Stimulus presentation was initiated by pressing any button. The stimulus remained visible until a response was made at which point it was replaced by the square and fixation cross until the next trial was initiated.

Participants were instructed to find the target C amongst the Os' and report the side (left or right) of the square it was located in by pressing a button.

# 7.2.2.2. Questions

There was a number of questions this experiment tried to answer.

1. Are monocular items in a binocular scene more salient than items that are feature-singletons?

If monocular items are more salient than feature-singletons (the C) then this supports the notion that they attract our attention to regions in a visual scene that are likely to contain valuable information (whether this is the edge of an object or a tiger behind a tree). If this is the case condition 1 should lead to significantly faster reaction-times than conditions 2 and 3. To be able to interpret the results for conditions 1-3 (in which monocular items are embedded in a binocular scene) we have to ask two related questions first.

#### 2. Is the search task chosen for this experiment appropriate?

More specifically, do the 'singleton' conditions (5-7) lead to results comparable to those found for the stimuli containing oriented bars (Zhaoping, 2008)?

From this the third question follows:

#### 3. How do monocular items and ocular singletons compare?

Are there parallels between the pattern of results for the 'singleton' conditions (5-7) and the monocular item conditions (1-3)? Part of this question is also how both sets compare to a purely binocular stimulus (condition 4).

I added another two conditions (3 and 7 for the monocular and singleton conditions respectively) in which the number of monocular/'singleton' distractors was increased which leads us to the final question.

4. Our environments are cluttered and tend to contain a host of monocular regions – how is our performance when there is more than one monocular item?

This question was born out of a conversation with Zhaoping Li. She suggested that one of the reasons why ocular singletons are more salient is because they are very easy to find in the 'set' of items we are searching in.

Suppose you are searching for a feature-singleton in a binocular display such as the one in this experiment. Since the difference in features between our target C and distractor Os is not very large all items will be relatively equal in their salience. This means you have to search the display in a fairly serial manner. If the target C is the only item presented to one eye and all distractor Os are presented to the other eye, Zhaoping suggests that the reason we are so much faster because suddenly we are not searching for an O amongst 224 other items but for the only item in this eye which is therefore not suppressed by any surrounding items (this is caused by disynaptic inhibition in Zhaoping's model (see section 6.4.2). Let us come back to the actual question - how is our performance when there is more than one monocular item? When we look around us there is a number of monocular regions rather than just one. Therefore, if any drop in reaction-times is due to the smaller number of items having to be searched in a set, the question is whether this drop is still observable if there are several monocular/'singleton' items in the stimulus. The prediction here is that, since the maximum number of these additional items is 4, reaction-times should still be lower than in a fully binocular or incongruent

stimulus (conditions 2 and 6).

# 7.2.2.3. Participants

7 participants, students aged 23-28, completed the study. All participants had normal or corrected to normal vision and normal stereo vision (TNO Stereo Test & Snellen EyeChart). Participants volunteered to participate. The experiment was approved by the University Teaching and Research Ethics Committee (UTREC) of the University of St Andrews. All participants gave written informed consent.

# 7.2.2.4. Analysis

There was a 1.27s variance for the binocular condition (4) which can be thought of as a baseline condition (the only cue to the target location in this condition is the feature difference between the C and the distractor Os) and the reaction-times for the remaining conditions showed equally large variances across participant reaction times. I thus normalised the reaction-times for all other conditions by the reaction-times for the binocular condition 4. For condition this means all reaction-times were 1.

I used a 1-Way-ANOVA to compare the normalised harmonic means for the 7 conditions (I will continue referring to the conditions as described in section 7.2.2.1). The planned pairwise comparisons were Sidak corrected (see e.g. Miller, 1981, pp.254-255).

### 7.2.3. Results

There was a significant difference between the mean reaction times for the different conditions (F(6,42)=36.98, p<0.01).

I will now go through the results in relation to the questions I asked in section 7.2.2.2.

# 7.2.3.1. How does this stimulus compare to the one used by Zhaoping (2008)?

Before asking whether monocular items affect our reaction-times the same way ocular singletons do, we have to ask whether the stimulus I chose for this experiment leads to results comparable to those found by Zhaoping (2008).

Conditions 5 and 6 use the same stimulus configurations (in regard to the eye-of-origin of the target/distractors) as the 'dichoptic congruent' and 'dichoptic incongruent' conditions used in experiments 2 and 3 in Zhaoping (2008). Participants' reaction-times follow the same pattern and were significantly faster (p<0.01) in condition 5 in which the target coincided with the ocular-singleton than in condition 6 in which the ocular singleton was a distractor O. One might wonder whether the singleton distractor actually 'distracted' participants' attention from the C. This was clearly the case, participants were significantly slower when searching for the target in condition 6 than when all items were binocular in condition 4 (p<0.01). These results directly reflect the pattern of results used by Zhaoping (2008).

# 7.2.3.2. Do monocular items and ocular singletons affect our visual search behaviour in the same fashion?

For the singleton conditions this experiment repeated the reaction-time pattern of condition 1 < condition 4 < condition 2. Do the monocular conditions follow the same pattern? Figure 7.8 shows the mean normalised reaction times for the monocular and singleton conditions.



Figure 7.8: Mean normalised reaction times for conditions 1, 2, 3, 5, 6, and 7. The horizontal line corresponds to the binocular baseline condition 4. The error bars are 95% CIs.

Condition 1, in which the target C is the only monocular item in the stimulus, leads to participants detecting the target significantly faster than in the binocular condition 4 (p<0.01) however, while the monocular distractor condition leads to even higher reaction-times, performance is not significantly worse than in condition 4 (p=0.15). It seems that while the pattern of results

mirrors that of the singleton conditions, the effect of the monocular items is less pronounced. The error rates for these conditions follows the same pattern. While too low (the mean error rates do not rise above 2.57% for any of the conditions) to be truly interpretable, participants perform worse in the monocular/singleton distractor conditions 2 and 6 (2.14% and 2.57% error rates respectively) than in the monocular/singleton target conditions 1 and 5 (0.79% and 0.5% respectively). Just as for the reaction times, monocular/singleton items affect our error rates.

# 7.2.3.3. How does a larger number of monocular items/'singletons' affect our performance?

Our environment is cluttered and contains a host of monocular regions. In all the conditions I have so far looked at, the maximum number of monocular/singleton items was one. But how do we perform when there is more than one monocular item? Conditions 3 and 7 were intended to answer this question because they include 4 monocular/'singleton' distractors. So how does performance differ between the conditions with our one monocular/singleton target and conditions 3 and 7? Let us start with the singleton conditions because the effect of the singleton has so far been more pronounced than the effect of the monocular item. Reaction times for condition 7 are slower than for condition 5 (though not significantly, p=0.06) but still significantly faster than for condition 6 (p<0.01). For the monocular conditions the pattern is comparable. Condition 3 is significantly slower than condition 1

(p <0.01). But note how, as with the differences between the conditions in section 7.2.3.2. the differences between the monocular conditions lead to much less pronounced changes in reaction-times.

# 7.2.4 Discussion

Similarly to ocular-singletons monocular items that are embedded in binocular scenes give us different eye-of-origin information from their surround. Ocular-singletons attract our attention and are more salient than featuresingletons (Zhaoping 2008). The aim of this experiment was to investigate whether there are parallels between the effect of ocular-singletons and monocular items on our visual-search behaviour. This is indeed the case. Our reaction-times for the different monocular conditions follow the same pattern the reaction-times for the singleton conditions do. Just like a singleton target, a monocular target appears more salient than a target that only differs from the remaining scene because it is a feature-singleton. At the same time, if a monocular/singleton distractor was present participants were significantly slower when searching for the target C. This suggests that our low-level access to eye-of-origin information can orient our attention towards monocular regions and their surround and content. In the context of the tiger in the jungle, the monocular regions we might be seeing some of the tiger in could be the part that attracts our attention towards the tiger.

We will most likely see more than one monocular region when we look around us. I thus investigated whether the attentional benefit of the monocular/

singleton targets remains if there is more than one monocular/singleton item in a scene. This was the case for the ocular singletons - while reaction-times were slower than for condition 5 in which the target was also a true ocularsingleton, they where still much faster than for the binocular condition 4. For the monocular items the beneficial effect of the monocular target was still present compared to the monocular distractor condition 2, but participants performed more or less the same as when viewing the fully binocular stimulus in condition 4. In conditions 1 and 2 we have already seen that monocular targets and distractors are less salient than singleton distractors and targets. It appears as if the monocular target and four distractors are no more salient than the surrounding binocular distractors and participants are essentially performing a serial search task just as in the fully binocular condition 4. This result would suggest that any potential benefit of the monocularity of a region disappears once there are several of them, leading us back to the original question – can we use the monocularity of a region as a cue to large depth steps or to potentially important information in the region itself? It appears we might not be very good at this.

However, in this experiment the number of monocular or 'singleton' distractor items did not exceed 4 and the target was always monocular or a 'singleton, too. While having participants search for a C amongst Os will never be truly comparable to perceiving a natural scene around us, a definite limitation of the current experiment is that there were only a maximum of 4 monocular or 'singleton' distractors. Natural scenes tend to contain a whole

host of monocular regions and the question is whether the slowing in reactiontimes in conditions 3 and 7 compared to conditions 1 and 5 continues linearly as the number of such distractors increases or whether it follows a different pattern.

Experiment 6 investigates this question further by comparing reaction times across a larger number of monocular distractors.

# 7.3. Experiment 6

Experiment 5 showed that monocular items affect our visual-search performance similarly to ocular-singletons. This effect is strongly diminished when there is a number of monocular distractors. What happens if the number of distractors is further increased? The aim of this experiment was to investigate the effect of an increased number of distractors that are monocular if the overall scene is binocular. To do so I used the same stimuli as in experiment 5 but changed the number of possible monocular distractors.

# 7.3.1. Methods

The overall stimulus setup and procedure for this experiment was the same as in experiment 5. The only difference was the conditions the stimuli were presented under. The singleton conditions were abandoned for this experiment because they are not pertinent to the question of how we deal with a large number of monocular distractors in a binocular scene.

# 7.3.1.1. Conditions

This stimulus was presented in one of two conditions. The target C could either be monocular or binocular. For each of these target conditions 6 'distractor number' levels were tested – each stimulus contained either 0, 1, 4, 8, 16, or 32 monocular distractors. Larger numbers of distractors led to participants reporting trouble fusing the two eyes' views in a pilot experiment, hence the number of distractors stops at 32. The distractors were randomly placed within the 15x15 grid. All other distractors were binocular.

# 7.3.1.2. Questions

This experiment was intended to answer the following two questions:

1. How does the number of distractors relate to participants' performance? Does it scale linearly?

2. How does performance compare when the target is monocular or binocular?

# 7.3.1.3. Participants

4 participants, students aged 22-24, completed the study. All participants had normal or corrected to normal vision and normal stereo vision (TNO Stereo Test & Snellen EyeChart). Participants volunteered to participate. The experiment was approved by the University Teaching and Research Ethics Committee (UTREC) of the University of St Andrews. All participants gave written informed consent.

# 7.3.1.4. Analysis

I used a two-way mixed ANOVA to compare the harmonic means for the different distractor levels for the monocular and binocular target conditions.

# 7.3.2. Results and discussion

I will now discuss the results in relation to the two questions that motivated this experiment. Maulchy's test of sphericity was significant (W < 0.01, df = 14, p < 0.01), thus the results for the within-subjects effects were adjusted using a Greenhouse-Geisser correction (epsilon was estimated at 0.218).

# 7.3.2.1. How are reaction-times and the number of distractors related?

Within subjects the was a significant within-subjects effect of the different distractor numbers at p = 0.001, F(1.09, 6.55) = 27.63.

Figure 7.9. illustrates the increase in reaction-times for all participants for monocular and binocular targets respectively.



Figure 7.9: Reaction times for all participants for the different distractor numbers for monocular targets(a) and binocular targets(b).

So, higher numbers of distractors lead to higher reaction-times.

When compare figures 7.9.a and b it also appears as if this effect is mediated by the target condition.

# 7.3.2.2. Differences between monocular/binocular target conditions.

The interaction of the number of distractors with the condition was significant at p= 0.01, F(1.09, 6.55) = 11.83. This means that while reaction-times increase simply as the number of distractors increased, they increased less if the target was binocular than when the target was monocular. Note how reaction-times for the binocular target remain relatively stable across distractor levels – as if the majority of participants is able to partially ignore the monocular distractors – whereas reaction-times keep increasing up to 16 distractors (it appears as if reaction-times might be asymptotic for higher numbers of distractors were to be presented in

a suitable stimulus) which suggests that whatever allowed participants to partially ignore the distractors, it is not working when the target is monocular like the added distractors.

# 7.4. Overall Discussion

So what have we learned in the past two experiments? Let us go back to experiment 5. I found that monocular items affect our visual-search performance similarly to ocular-singletons. When an item we are looking for in a binocular scene is monocular we are much more likely to find it quickly than when we are looking for a binocular item and there is a monocular item present in the scene. These results would suggest that we are indeed able to use monocular regions to direct our attention; a tempting idea because monocular regions are indicative of large depth-steps in natural scenes (Wilcox & Lakra, 2007).

So how would a monocular item guide out attention and thus lower our reaction-times? Zhaoping (2010) found that ocular-singletons automatically attract our gaze (even when we are not aware of the singleton). If the previous parallels between ocular-singletons and monocular items hold, then we might expect monocular items to attract our gaze and thus improve our visual-search performance. I thus decided to test whether monocular items attract our gaze the same way ocular-singletons do.

However, since I found that the effect of the monocular items is smaller

than for ocular singletons and is strongly diminished as the number of monocular regions in a scene increases in experiment 6 it seems unlikely that the stimuli used in experiment 6 will lead to easily interpretable results. While they are more realistic stimuli in terms of the number of monocular regions, they are too complex for a first exploration of whether our gaze is attracted by monocular items (due to eye-tracker calibration accuracy). The simpler stimuli used in experiment 5 allow us to ask whether a participant's gaze shifts towards a target . The following experiment 7 is a replication of experiment 5 with the addition of the measurement of participants' eye-movements.

As part of the ACCN '09 Internship Programme I was given the opportunity to ask whether monocular items attract our gaze the same way ocular-singletons do. In the following chapter (8) I will describe experiment 7 which I ran in collaboration with Zhaoping Li and Keith May using the same setup as in Zhaoping (2010).

8. Experiment 7 – Do monocular regions affect our eye-movements?

# 8.1. Aims

The experiments described in the previous chapters have suggested that we are able to incorporate monocular information into our overall percept of a scene. It seems that we are able to use the information that is presented in monocular regions to nearly the same extent that we can use binocular information and that monocular regions affect our visual-search performance. The aim of this experiment was to explore how monocular regions compete for our attention with feature singletons and how they affect our eye movements during visual search.

# 8.2. Background

In experiment 5 in chapter 7 I found that, similarly to the ocular-singletons described by Zhaoping (e.g. 2008, 2010), monocular items affect our visual search performance. Zhaoping had found that if we are searching for a ocular-singleton target amongst non-singleton distractors. I found that we are much faster at finding a monocular target than when we are searching for a binocular target amongst binocular distractors. We are even slower when we are searching for a binocular target amongst binocular distractors if there is a monocular distractor present. This would suggest that monocular items do, indeed, drive our attention towards them. Since Zhaoping (2008) found that ocular-singletons automatically attract our gaze, the question is whether the previously found parallels between ocular-singletons and monocular items hold. If this is the case then we might expect monocular items to attract our gaze and thus improve our visual-search performance. I decided to test whether monocular items attract our gaze as ocular-singletons do using the same stimuli as in experiment 5.

While the stimuli I used in experiment 6 are more realistic stimuli in terms of the number of monocular regions, they are too complex for the first exploration in the present experiment. The small number of monocular regions in the stimuli used in experiment 5 allow us to ask whether a participant's gaze shifts towards a target. In the present experiment, I replicate experiment 5 using a different method: the measurement of participants' eye-movements.

### 8.3. Methods

# 8.3.1. Stimuli/Procedure

Participants viewed two white squares (luminance: 110 cd/m<sup>2</sup>, size: 14.23°by 14.16°) which were framed by a black line each (luminance: 0.01 cd/m<sup>2</sup>; width: 16 arcmin) on a white screen (luminance: 110 cd/m<sup>2</sup>). The two squares were separated by a gap of  $8.98^{\circ}$  so, once viewed through the stereoscope, each eye was presented with only one of the squares. Within each square 225 white tiles (size: 54x54 arcmin; luminance: 110 cd/m<sup>2</sup>) were arranged in a 15x15 grid. On each tile a black letter was displayed (size: 32x40

arcmin, luminance 0.01 cd/m<sup>2</sup>). The letters were located centrally on the tile but were jittered by up to 8 arcmin both vertically and horizontally to avoid false matches (the jitter was the same in the two eyes for each tile).

This stimulus was presented in one of 3 conditions.

(1) monocular target, all distractors binocular

0.0.0.0.0.0.0.0.0.0.0.0.0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0.0.0.0.0.0.0.0.0.0.0.0
0.0.0.0.0.0.0.0.0.0.0.0.0
0.0.0.0.0.0.0.0.0.0.0.0.0.0
0.0.0.0.0.0.0.0.0.0.0.0
0.0.0.0.0.0.0.0.0.0.0.0.0.0
0.0.0.0.0.0.0.0.0.0.0.0.0.0
000000000000000000000000000000000000000



Figure 8.1: Condition1 - Monocular target, all distractors binocular.

(2) binocular target, 1 monocular distractor, all other distractors binocular

$\begin{array}{c} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 $
0.0.0.0.0.0.0.0.0.0.0.0.0.0
0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0
0.0.0.0.0.0.0.0.0.0.0.0.0.0
0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0
0.0.0.0.0.0.0.0.0.0.0.0.0.0.0
0.0.0.0.0.0.0.0.0.0.0.0.0.0.0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

Figure 8.2: Condition 2 - binocular target, 1 monocular distractor, all other distractors binocular.

# (3) all binocular

0 0
0.0.0.0.0.0.0.0.0.0.0.0.0.0.0
0.0.0.0.0.0.0.0.0.0.0.0.0.0.0
0.0.0.0.0.0.0.0.0.0.0.0.0.0.0
0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0
0.0.0.0.0.0.0.0.0.0.0.0.0.0.0
0.0.0.0.0.0.0.0.0.0.0.0.0.0
0.0.0.0.0.0.0.0.0.0.0.0.0.0



Figure 8.3: Condition 3 - all binocular.

Figures 8.1-3 above show crossed/uncrossed fuseable screenshots of the three conditions.

Participants' eye movements were recorded (at 250Hz) using a infrared video eye-tracker from Cambridge Research Systems which was set on top of a modified Wheatstone Stereoscope while performing the same reaction time task as in experiment 5, with the only change in the timeline being that stimulus presentation was fixation contingent. This is shown in figure 8.4.



Figure 8.4: Within-trial timeline.

The experiment was split into 32 blocks of trials, at the beginning of each block the eye tracker was calibrated, then participants initiated the first trial by pressing any button, then the programme paused until participants fixated on the fixation dot (the fixation criterion used was the following: If the participant's pupil was recorded to be within the 54x54 arcmin square surrounding the fixation dot for 50 ms this was classed as fixation for this purpose).

Participants were instructed to 'Search for the C amongst the O's and respond by pressing one of two keys on a computer keyboard that to indicate which side of the display square the C was located in.' They were instructed to respond as quickly and as accurately as possible. Furthermore they were instructed that it was not necessary to maintain fixation after the stimulus appeared and they were encouraged to move their eyes while searching for the stimuli. Each stimulus was presented until participants indicated their choice on the computer keyboard, then an empty frame and a fixation dot were shown until participants initiated the next trial.

Each participant completed a set of 2 practice blocks at the beginning of each session. Most participants split the 32 blocks over 2 1.5 hour sessions.

# 8.3.2. Research Questions

#### 1. How does the presence of monocular regions direct our attention?

As in experiment 5, the main question I wanted to explore with this experiment was that of how a number of monocular regions affects our attention and our performance. However, while I used changes in reaction-times as a proxy for effects on attention in Experiment 5, in this experiment I asked whether early eye-movements during visual search show a similar pattern. This leads us to two related questions.

#### 2. Are eye-movements driven by the presence of monocular regions?

Zhaoping (2008) found that ocular-singletons significantly affect the first saccade during visual-search (we are significantly more likely to look towards an ocular-singleton than anywhere else, even if we are not consciously aware of the presence of the singleton) in a fashion that mirrors the effects of ocular-singletons on our reaction-times during visual-search. In experiment 5, I showed that monocular regions affect our reaction-times just like ocular-singletons do (if to a slightly lesser extent) so the question is whether monocular regions affect our first saccade direction during visual-search similarly to how ocular-singletons affect this first saccade.

Zhaoping argues that we process the information about ocular-singletons in V1 (see section 6.4.2) and that this information can directly affect our eye movements through a pathway that connects V1 with the Frontal Eye Field through the Superior Colliculus (for more details on this see e.g. Fecteau, Bell & Munoz (2004); Tehovnik, Slocum & Schille (2003); Zhaoping (2008)). We can thus add a third, though tentative question.

# 3. Is the monocularity of a region processed as early as the identity of an ocular-singleton?

This question relates back to section 6.3 and the different suggestions as to how early monocular regions could be identified as such and used in the overall processing of a visual scene. While there appears to be consensus that processing of monocular regions happens early on during visual processing, opinions differ in regard to whether this can be achieved in V1 alone (e.g. Hayashi et al. 2004) or whether part of the processing still happens in V2 (e.g. Assee & Qian, 2007). If early saccades during visual-search are directly driven by information from V1 then this experiment also allows us to ask whether we are able to assess the monocularity of a region as early as V1. Keep in mind that this experiment was not explicitly designed to test this question and that any results should be interpreted with caution. However, if participants' eye movements are not driven by monocular regions as systematically as they are by ocular-singletons then this would suggest that the monocularity of a region is not fully processed in V1 and it could be that further processing in V2 happens before we can correctly identify a monocular region.

# 8.3.3. Participants

4 Participants completed the experiment. All participants had normal or corrected to normal vision and had good stereo vision (TNO-Test). Participants were reimbursed for their participation or volunteered. The experiment was approved by the University Teaching and Research Ethics Committee (UTREC) of the University of St Andrews. The experiment was performed in Zhaoping's lab at UCL and was covered by a blanket ethical approval from the Department of Computer Science at the University College London. All participants gave written informed consent.

# 8.3.4. Analysis

I analysed the reaction-time data for this experiment the same way I had analysed the reaction-times in experiments 5 and 6, using the harmonic mean as the measure of central tendency, and the first and third quartiles as the measure of dispersion. However there was a further constraint on the reactiontime data that was applied beforehand – only trials in which participants eyemovements had been successfully recorded were included in the harmonic mean and quartiles. While this reduces the number of trials used to calculate these measures, it allows us to better compare the reaction-time and eyemovement data. Furthermore, trials with reaction-time under 200ms were excluded (there was a distinct gap in the latency distribution between these trials, which coincided with trials in which participants mentioned that they had accidentally pressed a button in response to the stimulus appearing on the screen and the lower-bound for the remaining trials which occurred between 380 and 500ms depending on the participant).

For the eye-movement data, I explored the direction of the first saccade (to the left or right from the fixation dot). I recorded the entire visual-search path for each trial, but after having large difficulties with the fixation criterion (the data fed back from the eye-tracker suggested participants were not fixating at the beginning of each trial yet the video recording that was provided at the same time showed that they were indeed fixating (two of the four participants were very experienced psychophysical observers)) I decided to only use the measure of the general direction rather than a finer measure.

# 8.4. Results

I will now discuss the reaction-time and eye-movement data separately.

# 8.4.1. Reaction-Times

On average, participants' reaction-times replicate the trends found in experiment 5. However, not all participants showed the same pattern of results. Figure 8.5 depicts the reaction-times for the 4 participants for the three conditions. The dashed lines depict the mean reaction-times for each condition.





None of the 3 conditions led to significantly different reaction-times. If we normalise the first two conditions by the binocular one (and thus treat the binocular condition as the baseline) the underlying baseline differences between participants disappear. Figure 8.6 shows the reaction-time differences between the first two conditions normalised by the binocular condition. Note how for all participants other than ps, reaction-times are slightly faster in the monocular target condition than in the monocular distractor condition. This is consistent with the findings in experiment 5 (see figure 7.8).



Figure 8.6: Reaction-Times in the monocular target/distractor conditions normalised by the reaction times in the binocular condition. The dashed lines depict the mean reaction-times for each condition.

As for the proportion of trials that were correct responses, there was no significant difference between the conditions, with between 96.5 and 100% of responses being correct and there was no significant difference between conditions. This indicates that this specific group of participants was performing at ceiling, this was also perceived as such by the participants who reported the task itself being an easy one.

# 8.4.2. Eye-Movements

The first saccades, however, paint a different picture. While we could observe a slight difference between the conditions for the reaction-times, there was no systematic difference between conditions where eye-movements are concerned (F(1,3) = .48; p = .54). Figure 8.7 shows the proportion of 'correct'

eye-movements towards the target C (again, normalised by the performance in the binocular condition).



Figure 8.7: Proportion correct eye movements normalized by the binocular condition. The dashed lines depict the mean proportion correct eye-movements for each condition.

Even though participants' eye-movements and eye-movements do not seem to be following the same pattern when we are simply looking at figures 8.6 and 8.7, when we look at the correlation between the reaction-times and eye-movements we see that across the two normalised conditions they are highly related (Pearson's Correlation: -.85; p > .01; adjusted  $R^2 = .67$ ; F(1,6) = 15.358). However, this picture becomes more differentiated when we look at the two normalised conditions separately. Figure 8.8 shows the regression function for all participants.



Figure 8.8: Regression function for all participants for the monocular target condition (a), and the monocular distractor condition (b), both normalised by the binocular condition.

While performance in the psychophysical task and eye movements are highly related for the monocular target condition (Pearson's Correlation: -.96; p = .04; adjusted  $R^2$  = .89; F(1,2)=24.48), they do not follow the same pattern for the monocular distractor condition (Pearson's Correlation: -.28; p = .72; adjusted  $R^2$  = -.38; F(1,2)= .17). This suggests that while it appears that reaction-times can be used as a predictor for whether a participant looked towards the target and vice versa, this breaks down when we are looking at participant performance when the target is binocular and there is a monocular distractor. This does, however, not mean that some participants are looking towards the monocular distractor. Note how the 'proportion correct eye-movement' hovers around 1 (i.e. no difference from the binocular condition) for the monocular distractor times a between participants.

# 8.5. Discussion

In summary, participants' reaction-times roughly mirrored those in experiment 5, but the difference between the monocular target and monocular distractor conditions was not significant. This is most likely due to the small sample size and the fact that the experienced participants were performing at ceiling. Ideally, I would have liked to test more participants but since I was collecting the data in London and was thus constrained by both the distance to St Andrews and the fact that I had no ready participant-pool to draw from, this was not possible. As for the eye-movements, there was little difference between the two conditions. The two measures were significantly correlated, though, with a high correlation between the reaction-times and eye-movements for the monocular target condition but no real correlation between them for the monocular distractor condition.

Let us return to the questions from section 8.3.2.

#### 1. How does the presence of monocular regions direct our attention?

As in experiment 5 participants found the target C faster if it was monocular than when it was binocular and the presence of a monocular distractor slowed this process even further (figure 8.6). This effect was not as pronounced as in experiment 5. Participants in this experiment were more experienced psychophysical observers than in past experiments which might explain why the effect was not as pronounced as in the past. As experiment 5 did, these results suggests that the presence of monocular regions can direct our attention, but that the answer might not be as clear-cut. It seems that this effect might not be as large with experienced psychophysical observers.

#### 2. Are eye-movements driven by the presence of monocular regions?

Unlike ocular-singletons, monocular regions do not seem to affect our eye-movements. While we are slightly more likely to look towards a monocular target (figure 8.7), this effect is nowhere near significant and the differences between participants are fairly large. Our reaction-times and eye-movements do, at least to a certain extent, go hand in hand but compared to the correlation between the two for ocular-singletons (Zhaoping, 2008), the relation between the two is smaller for monocular regions and nearly non-existent for the monocular distractor condition (figure 8.8b). This suggests that while there might be parallels between the processing of monocular regions and ocular-singletons, monocular regions do not appear to affect our eye-movements the way ocular-singletons do.

#### This leads us to the last question.

3. Is the monocularity of a region processed as early as the identity of an ocular-singleton?

Since monocular regions do not seem to drive our first saccade during visual-search search the way ocular-singletons do, we might ask whether this

means that the two are processed differently. One possibility is the following: if ocular-singletons do, indeed, affect our eye-movements during visual search because the eye-of-origin information is processed in V1 and ocular-singletons are recognised as very salient by the superior colliculus, then the other side of this argument is that if the processing of a specific aspect of an item (in this case the monocularity of a region) has not occurred in V1, it should not be able to affect our eye-movements via the superior colliculus route. This is in line with the proposal of Qian and Assee (2007) that monocular regions are not completely processed in V1 and that further processing in V2 is needed to determine whether a monocular region is caused by an occlusion (and thus valid) or not (and thus most likely noise). In this experiment the monocularity of a region did not affect our eye-movements significantly. Thus, since monocular regions do not drive our eye-movements the way ocular singletons do, if the above argument holds, this suggests that we do not process monocular regions the same way we process ocular singletons. This also suggests that in order to process the monocularity of a region we need information about the binocularity of the surrounding regions.

In the following chapter I will discuss how the different experiments relate to each other and close with an overall discussion of the results.

# 9. Overall discussion

In the experimental chapters of this thesis I have investigated our percept of monocular regions in binocular scenes and how their presence affects our visual-search performance. This chapter will give us a chance to review the findings of the experiments in the past chapters and to explore trends in the results. I will now review what I have found in the different experiments and discuss how the results of each experiment relate to those of the previous experiments.

# 9.1. Experiment 1 (chapter 3.2)

I started this thesis with the question of whether monocular regions in binocular scenes are perceived the same way that the binocularly visible regions surrounding them are. To do so, I asked in experiment 1 whether numerosities presented in monocular regions (caused by vertical occluding bars that were made up of white-noise) lead to similar discrimination thresholds as numerosities that are presented fully binocularly or fully monocularly. In this experiment, I started with a stimulus in which the contrast between the dots and the background was maximal. There was no significant difference between our discrimination thresholds for stimuli containing monocular regions and those containing fully binocular information. At the same time, the vertically occluded condition (which contained the monocular regions) was not significantly different from the two monocular control conditions either. While not significant, the two occluded conditions (monocular

and binocular) lead to higher thresholds than the two unoccluded conditions (monocular and binocular). It was thus unclear whether there was an effect of the occluders I had used to create the monocular regions. I thus decided to repeat this experiment using slightly altered stimuli.

### 9.2. Experiments 2 and 2.2

# 9.2.1. Experiment 2 (chapter 3.3)

Experiment 2 was an improved version of experiment 1. Rather than presenting white dots on a black background I now presented black and white dots on a midlevel-grey background. Since it was unclear in experiment 1 whether there was an effect of the white noise occluders used to produce the monocular regions, I introduced a condition that would allow me to test this. In experiment 1 I had used a fully binocular condition and a vertically occluded condition that gave rise to monocular regions. In experiment 2 I introduced a horizontally occluded condition. In this condition binocular white-noise occluders were placed horizontally in front of the cloud of dots thus leading to only binocular dots being visible behind the occluders. If there was an effect of the white-noise occluders themselves, this effect should be present both in the horizontally and vertically occluded conditions. Using this set-up I found a significant difference between the horizontally occluded and binocular conditions, yet no significant difference between the horizontally and vertically occluded conditions. This suggests that, while there seems to be a detrimental effect of the occluders on our performance, from a sensitivity perspective, we
are able to interact with vertically occluded information as well as with binocular information (in the horizontally occluded condition).

#### 9.2.2. Experiment 2.2 (chapter 3.4)

After finishing experiment 2 I noticed several issues with the experimental setup. First of all, the midlevel-grey used in experiment 2 was darker than what was intended. Furthermore, the Weber-Fractions used had been calculated wrong for the negative range, and the standard stimulus had always been presented first. I thus reran experiment 2 with a slightly changed experimental setup in experiment 2.2, amongst others, changing the luminance of the grey background to lying at 50% between the black and white luminances used, revising the levels of difference between the standard and test numerosities. Additionally, I randomised the standard-test presentation order to account for any effects of the presentation order, and added two monocular conditions in which the white-noise background used in experiment 2 was replaced with a midlevel-grey background. This last change was made to test whether the white-noise in the occluders was causing the observed slight drop in sensitivity or whether the occlusion itself was causing this drop.

The trends I previously observed in the data, namely, that the thresholds for the occluded conditions were higher than the thresholds for the fully binocular condition, that the monocular occluded conditions led to much higher thresholds than the monocular unoccluded or the vertically occluded condition, did not change. I found no difference between the monocular

conditions that had a white-noise background and those that had a midlevelgrey background. This suggests that the drop in sensitivity we observed for the occluded conditions was caused by the spatial arrangement of the stimulus rather than the white-noise texture used in the occluders.

### 9.2.3. Monocular Regions so far.

Information in monocular regions seems to lead to a percept that is very similar to that of binocular information. Remember, in experiment 1, I found no difference between the vertically occluded and fully binocular conditions. There appears to be a slight effect of the textured occluders though. In experiment 2, there was a significant difference between the newly introduced horizontally occluded condition and the fully binocular one, a pattern that was repeated in experiment 2.2. This effect is small but seems to relate to the occluders rather than the textured white-noise pattern itself. Once we discount the effect of the occluded conditions and the horizontally occluded conditions in experiments 2 and 2.2.

Overall, this indicates that information in monocular regions leads to a percept that is not very different from of the percept obtained when viewing binocular information, a relationship that is further supported by the finding that there is a significant difference between the vertically occluded and monocular occluded conditions in experiment 2. This difference would suggest that we are very unlikely to simply ignore one eye's view when processing the vertically occluded stimuli. However, the stimuli used in the experiments up to this point

do not allow us to come to this conclusion because I used clouds of dots with a constant density in the background plane. Thus if participants were to ignore one eye's view and rely on the density of the dots visible to the non-ignored eye they could still arrive at an unbiased estimate of the number of dots. I therefore decided to ask what would happen if there was a different density of dots visible to the two eyes.

# 9.3. Experiment 3 – or what happens if the two eyes see different things?

### 9.3.1. Experiment 3 (chapter 4)

Since the constant density of dots across the stimuli and eyes in the previous experiments did not allow me to be sure that we were not ignoring one eye's view, I used a new set of stimuli in the following experiment. I designed 'stripy' stimuli, that contained different densities in different regions of the stimulus. This stimulus, once a set of vertical occluders is added, leads to different densities being delivered to the two eyes. If participants were to use only one eye, their responses would be heavily skewed because each eye alone delivers a wrong estimate of number (either too high by the proportion the higher density eye was assigned).

In experiment 3, I found that there is a significant effect of binocular occlusion of background information on our sensitivity to dot density. The size of the difference between the two conditions remains relatively constant across

all 3 experiments (the thresholds for the fully binocular conditions ranged between 0.12 and 0.17 while the thresholds for the binocularly occluded conditions ranged between 0.23 and 0.26). There is no significant effect of making the occlusion monocular (by using a vertical occluder) rather than binocular. This is in line with the findings in experiments 2 and 2.2.

## 9.3.2. Our percept of monocular regions – summary I.

With experiment 3 I showed that information from a number of separate monocular regions can be integrated into our overall percept of number. In general, we are able to integrate regions of different densities that have no spatial connection into a meaningful percept and compare it to a binocular comparison stimulus with no loss of accuracy. This is comparable to work on amodal completion in 3 dimensional scenes (Bruno et al., 1997) where a 3D occlusion situation leads to a percept that is comparable to that of a fully binocular view of the same scene. The stimuli used here contained no objects and yet observers report seeing a continuous surface on the background plane.

In contrast to other situations when the two eyes are presented with different input (work on binocular rivalry, e.g. Alais, O'Shea, Mesana, Alais & Wilson, 2000, Blake, Lee, & Heeger, 2009, Blake & Logothetis, 2003), the results of experiment 3 do not support the idea that geometrically plausible monocular regions (caused by occluders) generally lead to a rivalrous percept, because participants appear to be using information from both eyes to arrive at

an estimate of the number of dots.

This pattern of results is also consistent with some other findings from the literature. Amodal completion occurs not only in binocularly occluded regions but also when a region is only occluded monocularly, such as is the case for the blind spot (this occurs both in the natural blind-spot, Durgin et al., 1995, and in pathological scotomas, Tripathy & Levi, 1999). Unlike for information that is filled-in at the blind spot, my work in the experiments so far suggest that the majority of observers do not ignore one eye's view of a vertically occluded stimulus. If this had been the case, we would have expected participants to overestimate the number of dots if they had only used the higher density eye and underestimate the number of dots if they had only used the lower density eye. For a more detailed discussion of this, refer back to chapter 4.3.3. A group of observers, however does seem to ignore the lower density eye's information. I will discuss this further in section 9.8.

In all the experiments I have discussed so far, the dots were either presented binocularly or monocularly, participants did not have to integrate both monocular and binocular dots to be able to make a judgement about the overall numerosity. This means, while participants had to integrate monocular and binocular regions to arrive at a stable percept, the dots themselves were always either monocular or binocular. I presented no situation in which the dots themselves were split across monocular and binocular regions.

I thus introduced a stimulus in the next experiment that allows us to ask

whether observers are able to integrate monocularly and binocularly presented dots to make a texture density judgement.

# 9.4. Experiment 4 – Are we able to integrate information from monocular and binocular regions?

#### 9.4.1. Experiment 4 (chapter 5)

The aim of experiment 4 was to investigate whether observers are able to integrate numerosities across both monocular and binocular regions. The previous experiments suggested that monocular regions are as reliable, and thus as useful as, binocular regions.

Here I introduced a new stimulus, one for which participants had to integrate both monocular and binocular regions to arrive at a non-biased percept. For this, I moved the background plane with the dots closer towards the occluder plane. This means both monocular and binocular regions became visible. These regions could either have a consistent texture density or the monocular or binocular regions could have a higher density. The reasoning behind this was that if participants were to ignore one type of region this would lead to markedly biased responses.

I found that there is a distinct change in the perceived texture density when the density of monocular and binocular regions in the stimulus differs. Varying the density (i.e. the number of dots in a monocular or binocular region) across the two types of regions allows us to ask whether we use both types of

regions or whether we rely on binocular regions when comparing two sequentially presented stimuli containing both monocularly and binocularly visible dots. The observed bias is consistent with the monocular regions being ignored.

Interestingly, while there is no drop in sensitivity across the different conditions across this experiment, we are not as sensitive to differences between the stimuli in the baseline condition here as we are to vertically occluded stimuli in experiments 2.2 and 3.

# 9.4.2.Our percept of monocular regions – summary II

At first glance, one might be tempted to think that the results of experiment 4 stand in contrast with the results of experiments 1-3. The first three experiments had all consistently suggested that we are very capable of using information in monocular regions to make judgements about dot densities. In particular, experiment 3 had suggested that a vertically occluded stimulus containing only monocular regions in the background appeared as equally numerous as a comparable binocular stimulus. Even when the monocular regions had a varying texture density. Experiment 4, on the other hand suggests that we completely ignore monocular information. So do these results really contradict one another? I do not think so. While the task in experiment 4 is the same as in the previous experiments, monocular and binocular regions have to be combined slightly differently by our visual system in the two situations. This difference is crucial here. When only integrating

across monocular regions, we are capable of using monocular information to make relative numerosity judgments (e.g. in experiment 3). When we are required to integrate this monocular information with binocular information (in experiment 4), this process seems to break down and we rely solely on the binocularly presented information.

With the new stimulus, participants report a stable percept just as in my own previous experiments (experiments, 1-3, chapters 3.2, 3.3, 3.4, and 4) and the discussed experiments on amodal completion (e.g. Bruno et al. 1997) and on depth perception (e.g. Gillam & Borsting, 1988; Gillam, Cook & Blackburn, 2003; Nakayama & Shimojo, 1990; Shimojo & Nakayama, 1990). The similar thresholds across conditions in experiment 4, when the density is varied between monocular and binocular regions, suggest that, as the difference between the monocular and binocular regions is varied, there is no change in our response to the binocular regions that is mediated by the monocular regions. It appears as if participants suppress the monocular regions when binocular information about the same plane is present (here both monocular and binocular regions are located behind the occluders on the same background plane; refer to figure 1.18 (p.37) for the relationship behind the foreground occluders and the background plane, some of which was visible monocularly and some of which was visible binocularly in this experiment). The bias (which mirrors the proportion of the overall number of dots that was presented monocularly) suggests that the dots used in this experiment appear significantly less numerous. Unlike in experiments 1-3, where participants were

asked to make a judgement about either a fully binocular cloud of dots or monocular dots that were spread across a number of spatially distant (but all equally monocular) regions, this experiment (while using the same task) required participants to assess the content of the monocular regions and then integrate it with the content of the binocular regions to arrive at a non-biased percept. This raises the question why, if we are able to use information in monocular regions, we do not seem to use it if binocular information is also present? Perhaps binocular information is treated as more reliable by the visual system. None of the experiments in this thesis allowed me to explicitly ask this question, but the results of experiment 4 do definitely hint towards this conclusion. I will go into more detail on this in section 9.8.

One reason for participants ignoring information from monocular regions might be that while we are able to consciously perceive information in monocular regions, we do tend to use them for reasons that do not require us to have explicit access to their content. For example, we know monocular regions are indicative of large depth steps in a scene (Wilcox & Lakra, 2007) and as such point to regions of interest.

Changizi and Shimojo (2008) suggested that we can use information in monocular regions in so far as that they could guide our attention towards them. If this is the case then, while we might not be able to integrate monocular information with binocular information, the presence of monocular regions might affect our attention during visual-search and direct it either towards the monocular region (by shifting our head, for example, we can bring a monocular region into our field of view for the occluded eye as well) or to the binocular regions directly adjacent to it.

To investigate how monocular regions affect our attention I decided to use a visual search experiment and study both the effect of the monocularity of a visual search target on our overall visual search performance and on our eyemovements during said visual search.

### 9.5. Experiment 5 (chapter 7.2)

The aim of experiment 5 was to investigate whether information in monocular regions affects our visual-search performance. I used a stimulus similar to the one used by Zhaoping (2008) because of the similarities between monocular regions and her group's studies of ocular-singletons (single monocularly presented elements, usually presented in a field of binocular items). Similarly to ocular-singletons, monocular items that are embedded in binocular scenes give us different eye-of-origin information from their surround. Ocular-singletons attract our attention and are more salient than feature-singletons: binocularly presented items that differ by some featural aspect (shape, size, colour) (Zhaoping 2008). I thus wondered whether there are parallels between the effect of ocular-singletons and monocular items on our visual-search behaviour. This is indeed the case. In experiment 5, participants searched for a target C amongst distractor Os. This target C could either be monocular in a binocular scene or a singleton (for a detailed description of the different stimulus conditions refer back to chapter 7.2.2.1 and specifically to

figures 7.1-7 (pp. 202-204). Our reaction-times for the different monocular conditions follow the same pattern that the reaction-times for the singleton conditions do. Just like a singleton target, a monocular target appears more salient than a target that only differs from the remaining scene because it is a feature-singleton. At the same time, if a monocular/singleton distractor was present in chapter 5, participants were significantly slower when searching for a target C amongst distractor Os. If we assume that these changes in reaction times are indicative of a bottom-up process affecting our attention, this suggests that just like for ocular singletons, monocular items attract our attention because they are more salient. However, if we do, indeed, use monocular regions to find either the large depth-steps adjacent to them (Wilcox & Lakra, 2007) or because they might contain important information about predators (Changizi & Shimojo, 2008), then we have to wonder what happens if we are confronted with more than one monocular region at once.

We will most likely see more than one monocular region when we look around us. In experiment 5, reaction-times were slower when people were judging monocular targets in binocular scenes than when the target was an ocularsingleton. While there still seems to be a beneficial effect on our visual search performance if we are searching for a monocular item, the number of monocular or 'singleton' distractor items did not exceed 4 and the target was always monocular or a 'singleton', too. Having participants search for a C amongst Os will never be truly comparable to perceiving a natural scene around us, but a definite limitation of experiment 5 is that there were only a

maximum of 4 monocular or 'singleton' distractors. Natural scenes tend to contain a whole host of monocular regions and the question is whether the slowing in reaction-times as the number of monocular items increases continues linearly or whether it follows a different pattern.

# 9.6. Experiment 6 – Interacting with increasing numbers of monocular regions

### 9.6.1. Experiment 6 (chapter 7.3)

This experiment followed out of experiment 5 and its aim was to investigate the effect of an increased number of distractors that are monocular if the overall scene is binocular. To do this I used the same stimuli as in experiment 5 but changed the number of possible monocular distractors to vary between 0 and 32. As expected, higher numbers of distractors led to higher reaction-times. I had varied the number of distractors in two different conditions and here an interesting effect emerged. There had been two possible target conditions – monocular and binocular. So, while reaction-times increased simply as the number of distractors increased, they increased less if the target was binocular than when the target was monocular. Reaction-times for the binocular target (see figure 7.9, p.216) – as if the majority of participants was able to partially suppress the monocular distractors – whereas reaction-times for the monocular target condition keep increasing up to 16 distractors (it appears as if reaction-times might be asymptotic for higher numbers of distractors if higher

numbers of distractors were to be presented in a suitable stimulus) which suggests that whatever allowed participants to partially ignore the distractors, it is not working when the target is monocular like the distractors.

#### 9.6.2. Visual-search – what do the reaction-times tell us?

#### So what have we learned in the past two experiments?

Let us go back to experiment 5. I found that monocular items affect our visualsearch performance similarly to ocular-singletons. When an item we are looking for in a binocular scene is monocular we are much more likely to find it quickly than when we are looking for a binocular item and there is a monocular item present in the scene. These results would suggest that we are indeed able to use monocular regions to direct our attention; a tempting idea because monocular regions are indicative of large depth-steps in natural scenes (Wilcox & Lakra, 2007 who also found a reaction-time advantage for plausible monocular regions). So how would a monocular item guide out attention and thus lower our reaction-times? Zhaoping (2010) found that ocular-singletons automatically attract our gaze (even when we are not aware of the singleton). If the previous parallels between ocular-singletons and monocular items hold, then we might expect monocular items to attract our gaze and thus improve our visual-search performance. I thus decided to test whether monocular items attract our gaze the same way ocular-singletons do.

### 9.7. Experiment 7 (chapter 8)

The aim of this experiment was to explore how monocular regions compete for our attention with feature singletons and how they affect our eye movements during visual search. I used the same stimuli as in experiment 5 (except that the contrast polarity was reversed to improve eye-tracker performance). For this experiment, participants' reaction-times roughly mirrored those in experiment 5, but the difference between the monocular target and monocular distractor conditions was not significant. This was most likely due to the small sample size and the fact that the experienced participants appeared to be performing at ceiling. As for the eye-movements, there was little difference between the two conditions for the first saccade (I measured whether the first saccade was directed towards the target C or away from it). The eye-movements and reaction-times were significantly correlated, though, with a high correlation between the reaction-times and eye-movements for the monocular target condition but no real correlation between them for the monocular distractor condition.

As in experiment 5, participants found the target C faster if it was monocular than when it was binocular and the presence of a monocular distractor slowed this process even further. This effect was not as pronounced as in experiment 5 (compare figures 7.8 (p.209) and 8.6 (p. 229)). This could have been because participants in experiment 7 were more experienced psychophysical observers than in past experiments, which might explain why the effect was not as pronounced as in the past. These results support the notion that the presence of

monocular regions can direct our attention, but unlike the results for experiment 5, they also suggest that the answer might not be as clear-cut. It seems that this effect might not be as large with experienced psychophysical observers.

Unlike ocular-singletons, monocular regions do not seem to affect our eyemovements. While we are slightly more likely to look towards a monocular target (figure 8.7 (p.230)), this effect is nowhere near significant and the differences between participants are fairly large. Our reaction-times and eye-movements do, at least to a certain extent, go hand in hand but compared to the correlation between the two for ocular-singletons (Zhaoping, 2008), the relation between the two is smaller for monocular regions and nearly non-existent for the monocular distractor condition (figure 8.8b (p. 231)). This suggests that while there might be parallels between the processing of monocular regions and ocular-singletons, I do not believe the group of observers in this experiment allow me to draw a specific conclusion. To be able to find out whether monocular regions affect our eye-movements the same way ocular-singletons do or not, this experiment would have to be repeated with a group of observers that is not as experienced with psychophysical experiments.

#### 9.8. Discussion

So what have learned overall from the experiments presented in this thesis? Let us return to the 'bigger question' I had started with in chapter 1. I set out asking whether monocular regions in binocular scenes are perceived the same

way the binocular regions surrounding them are. We saw, that this seems to be the case if we are only integrating information across monocular regions (experiments 1-3). At the same time we saw that once we have to integrate across monocular and binocular regions a seemingly very different picture emerges (experiment 4). Here, we are more or less unable to use monocular regions and suppress their information, relying solely on binocular information. At the same time participants did not report a bistable percept. This might seem surprising because due to the monocular regions, the two retinal images are very different in the areas that contain the dots participants are examining.

Arnold (2011) suggests there are several reasons why monocular regions do not cause binocular rivalry. In most situations we will have differential occlusion cues for the two eyes (e.g. a monocular region will occur on the temporal side of the retinal image in relation to the occluder). But even when we have no such occlusion cues available we still only rarely experience binocular rivalry. For example, if you hold up a pencil between your eyes and either your computer screen or this page (like we did in chapter 1) and read some text, you will be able to read the text while still being marginally aware of the pencil in front of it. Arnold (2011) argues this is achieved by active suppression rather than binocular rivalry. The findings in this thesis are in line with this suggestion. Not a single participant reported binocular rivalry. Yet in experiment 4, where participants had to integrate monocular and binocular information to arrive at an unbiased percept, integration of the two types of information did not happen either. Participants were consistently suppressing

the information provided in monocular regions. So what does the visual system suppress in which situations?

Signal strength (such as blur, luminance and chromatic contrast), seems to be an important factor, see Arnold (2011). In binocular rivalry, when two conflicting images are presented to the two eyes, the image with a higher signal strength can consistently suppress the weaker signal even when the two signals keep switching between the two eyes (e.g. Arnold et al. 2007, 2008; Logothetis et al. 1996; Kovács et al. 1996). This indicates that we are not consistently suppressing one eye's view. We are, instead, able to consistently suppress the weaker signal across the two eyes. If we make the assumption that a monocular dot is not as 'strong' as a binocular dot then the results of experiment 4 directly follow from said assumption. At the same time, if we assume that a region with a higher density of dots is 'stronger' than a region with a lower density, then the individual differences we observed in experiment 3 (where several participants seemed to be consistently ignoring the lowerdensity eye's input) might be explained. The participants who appeared to be ignoring the lower density eye's input did so irrespective of whether the higher density was presented to the left or the right eye which is what would be expected if the assumption that a higher density of dots is treated as a stronger signal than the lower density of dots holds. While this is conjecture at this point, such a suppression process could explain these interesting patterns of results.

So how does this relate to how we treat monocular regions during

visual-search? If we are searching for a monocular target amongst binocular distractors then the target itself is very easy to find if there are no, or very few other monocular items or regions around. At the same time, if we are searching for a binocular target this task is fairly hard, but our performance does not change drastically if I start adding a number of monocular items to the same scene. This suggests that, similarly to what I found in experiment 4, where participants seemed to suppress the information from monocular regions, we are quite adept at ignoring information from monocular scenes. What this also shows is that once we have to attend to the monocular items (if the target is monocular) we are unable to suppress the monocular distractors.

While I started out asking about our percept of monocular regions, it seems there is no single answer. We can use monocular regions and are quite sensitive to the information contained in them, but at the same time, we do seem to rely on binocular information if we have to integrate the two types of information. Monocular items heavily affect our visual-search patterns when we are searching for them, yet if they are distractors and we are searching for a binocular target we are able to more or less completely ignore them. One might be tempted to say that this is a rather irregular pattern, but I do not believe this is the case. What all those differences in how we treat monocular regions show is the following: monocular regions are not the same as binocular regions.

We are able to use them and often, such as is the case with amodal completion (Bruno et al. 1997), with texture discrimination (He & Nakayama, 1994) or my own visual-search experiments (experiments 5 & 6, chapter 7), we

even seem to be able to improve our performance at the tasks at hand by using monocular regions. But we do not equate monocular and binocular regions. This could be caused by us treating monocular information as less reliable than binocular information, but it could also be caused by the fact that, due to some computation during visual processing, binocular information gets passed on further than monocular information. How this difference occurs was not the focus of this thesis. But there is a distinct difference between how we use the two types of regions. Interestingly, our conscious percept of monocular regions is not that different from that of binocular regions. I mentioned that participants do not report a bistable percept and when given enough time to look at the dots in the numerosity stimuli we are able to count all dots in the display. This tells us we can use the monocular information. But when binocular information is present this does not matter. It is as if when our visual system is presented with both monocular and binocular information, rather than treating the two kinds of information as equal, it shows a preference for the binocular information.

10. References

- Alais, D., Cass, J., O'Shea, R. P., & Blake, R. (2010). Visual sensititity underlying changes in visual consciousness. Current Biology, 20, 1362–1367.
- Anderson, B. L., & Nakayama, K. (1994). Toward a General Theory of Stereopsis: Binocular Matching, Occluding Contours and Fusion.
  Psychological Review, 101(3), 414–445.
- Arnold, D. H. (2011). Why is binocular rivalry uncommon? Discrepant monocular images in the real world, 1–7.
- Arnold, D. H., Grove, P. M., & Wallis, T. S. A. (2007). Staying focused: A functional account of perceptual suppression during binocular rivalry. JOV, 7(7), 1–8.
- Arnold, D. H., Law, P., & Wallis, T. S. A. (2008). Binocular switch sup- pression:a new method for persistently rendering the visible "invisible". VisionRes., 48, 994–1001.
- Assee, A., & QIAN, N. (2007). Solving da Vinci stereopsis with depth-edgeselective V2 cells. Vision Research, 47, 2585–2602.
- Baddeley, A. D. (2000). The episodic buffer: a new component of working memory? Trends in Cognitive Scienes, 4(11), 417–423.
- Baddeley, A. D., & Hitch, G. J. L. (n.d.). Working Memory. In G. A. Bower (Ed.), The psychology of learning and motivation: advances in research and theory (Vol. 8). New York: Academic Press.

Bakin, J. S., Nakayama, K., & Gilbert, C. D. (2000). Visual responses in monkey

areas V1 and V2 to three-dimensional surface configurations. Journal of Neuroscience, 20, 8188–8198.

- Blake, R., & Logothetis, N. K. (2002). Visual competition. Nature Reviews Neuroscience, 3, 13–21.
- Blake, R., Lee, S.-H., & Heeger, D. (2009). Binocular Rivalry and Neural Dynamics. Phys Rev Lett, 102(25), 1–15.
- Brainard, D. H. (1997). The Psychophysics Toolbox. Spatial Vision, 10, 433– 436.
- Brooks, K. R., & Gillam, B. J. (2006). Quantitative perceived depth from sequential monocular decamouflage. Vision Research, 46, 605–613.
- Bruno, N., Bertamini, M., & Domini, F. (1997). Amodal Completion of Partly Occluded Surfaces: Is There a Mosaic Stage? Journal of Experimental Psychology: Human Perception and Performance, 23(5), 1412–1426.
- Burr, D., & Ross, J. (2008). Response: Visual number. Current Biology, 18(18), R857–R858.
- Campbell, F. W., & Green, D. G. (1965). Optical and Retinal Factors Afecting Visual Resolution. J Physiol, 576–593.
- Changizi, M. A., & Shimojo, S. (2008). "'X-ray vision'" and the evolution of forward-facing eyes. Journal of Theoretical Biology, 254, 756–767.
- Cogan, A. I., Kontsevich, L. L., Lomakin, A. J., Halpern, D. L., and Blake, R. (1995). Binocular disparity processing with opposite-contrast stimuli.

Perception 24, 33–47.

- Collett, T. S. (1985). Extrapolating and interpolating surfaces in depth. Proceedings of the Royal Society of London Series B – Biological Sciences, 224, 43–56.
- Cowan, N. (2005). Working Memory Capacity. Essays in Cognitive Psychology. Hove: Psychology Press.
- Crossman, E. R. F. W. (1956). The measurement of discriminability. Quarterly Journal of Experimental Psychology, 7, 176–195.
- Cumming, B. G. & Parker, A. J. (1997). Responses of primary visual cortical neurons to binocular disparity without depth perception. Nature, 389, 280–283.
- Cumming, B. G., Shapiro, S. E., & Parker, A. J. (1998). Disparity detection in anticorrelated stereograms. *Perception*, 27, 1367–1377.
- Dakin, S. C., Cass, J., Greenwood, J. A., & Bex, P. J. (2010). Probabilistic, positional averaging predicts object-level crowding effects with letter-l ike stimuli. Journal of Vision, 10(10), 14–14.
- Dobson, A. J., & Barnett, A. (2008). An Introduction to Generalized Linear Models. Chapman & Hall/CRC Texts in Statistical Science (3rd ed.). Boca Raton, FL: Chapman and Hall/CRC.
- Durgin, F. H. (1995). Texture Density Adaptation and the Perceived Numerosity and Distribution of Texture. Journal of Experimental Psychology: Human Perception and Performance, 21(1), 1–21.

- Durgin, F. H. (2001). Texture contrast aftereffects are monocular; texture density aftereffects are binocular. Vision Research, 41, 2619–2630.
- Durgin, F. H. (2008). Texture density adaptation and visual number revisited. Current Biology, 18(18), 855–856.
- Durgin, F. H., Tripathy, S. P., & Levi, D. M. (1995). On the filling in of the visual blind spot: some rules of thumb. Perception, 24, 827–840.
- Efron, B. (1987). Better bootstrap confidence intervals. Journal of the American Statistical Association, 82, 171–200.
- Efron, B., & Tibshirani, R. J. (1993). An introduction to the bootstrap. New York: Chapman and Hall.
- Egnal, G., & Wildes, R. P. (2002). Detecting Binocular Half-Occlusions: Empirical Comparisons of Five Approaches. IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(8), 1127–1133.
- Fecteau, J. H., Bell, A. H., & Munoz, D. P. (2004). Neural correlates of automatic and goal-driven biases in orienting spatial attention. Journal of Neurophysiology, 912, 1728–1737.
- Forte, J., Peirce, J. W., & Lennie, P. (2002). Binocular integration of partially occluded surfaces. Vision Research, 42, 1225–1235.
- Fründ, I., Haenel, N. V., & Wichmann, F. A. (2011). Inference for psychometric functions in the presence of nonstationary behavior. JOV, 11(6), 1–19.

Gillam, B., & Borsting, E. (1988). The role of monocular regions in stereoscopic

displays. Perception, 17, 603-608.

- Gillam, B., & Nakayama, K. (1999). Quantitative depth for a phantom surface can be based on cyclopean occlusion cues alone. Vision Research, 39, 109–112.
- Gillam, B., Blackburn, S., & Nakayama, K. (1999). Stereopsis based on monocular gaps: Metrical encoding of depth and slant without matching contours. Vision Research, 39, 493–502.
- Gillam, B., Cook, M., & Blackburn, S. (2003). Monocular discs in the occlusion zones of binocular surfaces do not have quantitative depth – A comparison with Panum's limiting case. Perception, 32, 1009–1019.
- Gold, J. M., Murray, R. F., Bennett, P. J., & Sekuler, A. B. (2000). Deriving behavioural receptive fields for visually completed contours. Current Biology, 10(11), 663–666.
- Grove, P. M., & Ono, H. (1999). Ecologically invalid monocular texture leads to longer perceptual latencies in random-dot stereograms. Perception, 28, 627–639.
- Grove, P. M., Brooks, K. R., Anderson, B. L., & Gillam, B. J. (2006). Monocular transparency and unpaired stereopsis. Vision Research, 46(1695-1705), 1–11.
- Grove, P. M., Gillam, B., & Ono, H. (2002). Content and context of monocular regions determine perceived depth in random dot, unpaired background and phantom stereograms. Vision Research, 42, 1859–1870.

- Harris, J. M., & Wilcox, L. M. (2009). The role of monocularly visible regions in depth and surface perception. Vision Research, 1–20.
- Hayashi, R., Maeda, T., Shimojo, S., & Tachi, S. (2004). An integrative model of binocular vision: a stereo model utilizing interocularly unpaired points produces both depth and binocular rivalry. Vision Research, 44, 2367– 2380.
- He, Zijland J, & Nakayama, K. (1992). Surfaces Versus Features in Visual Search. Nature, 359, 231–233.
- He, Zijiang J., & Nakayama, K. (1994). Perceiving Textures: Beyond Filtering. Vision Research, 34(2), 151–162.
- Heydt, von der, R., Peterhans, E., & Baumgartner, G. (1984). Illusory contours and cortical neuron responses. Science, 224, 1260–1262.
- Howard, I. P., & Rogers, B. J. (2002). Seeing in depth. II. Porteous Press.
- Jackson, F. (1982). Epiphenomenal Qualia. Philosophical Quarterly, 127–136.
- Jevons, W. S. (1871). The power of numerical discrimination. Nature, 3, 363– 372.
- Julesz, B. (1964). Binocular depth perception without familiarity cues. Science, 145, 356–362.
- Julesz, B. (1971). Foundations of Cyclopean Perception. Chicago: University of Chicago Press.

Kanizsa, G. (1989). Organization in vision. New York: Praeger. (1979) .(1989).

Leonardo on painting. (M. Kemp, Ed.). New Haven, CT: Yale University Press.

- Kaufman, E. L., Lord, M. W., Reese, T. W., & Volkman, J. (1949). The discrimination of visual number. American Journal of Psychology, 62, 498–525.
- Kovács, I., Papathomas, T. V., Yang, M., & Fehér, Á. (1996). When the brain changes its mind: Interocular grouping during binocular rivalry. Proc. Natl. Acad. Sci. USA, 93, 15508–15511.
- Lawson, R. B., & Gulick, W. L. (1967). Stereopsis and anomalous contour. Vision Research, 7, 271–297.
- Liu, L., Stevenson, S. B., & Schor, C. M. (1997). Binocular Matching of Dissimilar Features in Phantom Stereopsis. Vision Research, 37(5), 633– 644.
- Liu, L., Stevenson, S. B., & Schor, C. M. (1994). Quantitative stereoscopic depth without binocular correspondence. Nature, 367, 66–69.
- Logothetis, N. K., Leopold, D. A., & Sheinberg, D. L. (1996). What is rivalling during binocular rivalry? Nature, 380, 621–624.
- Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for f eatures and conjunctions. Nature, 279–281.
- Mandler, G., & Shebo, B. J. (1982). Subitizing: An analysis of its component processes. Journal of Experimental Psychology: General, 111, 1–22.

Marr, D. (2010). Vision: A Computational Investigation into the Human

Representation and Processing of Visual Information (1982). Cambridge, MA: MIT Press.

- Marr, D., & Poggio, T. (1976). Cooperative Computation of Stereo Disparity. Science, 194(4262), 283–287.
- Marr, D., & Poggio, T. (1979). A computational theory of human stereo vision. Proceedings of the Royal Society of London. Series B.
- Michotte, A., Thinès, G., Costall, A., & Butterworth, G. (1991). Michotte's experimental phenomenology of perception. Hillsdale, NJ: Erlbaum.
- Miller, R. G. (1981). Simultaneous Statistical Inference (2nd ed.). New York: Springer Verlag.
- 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, 55–68.
- Ohzawa, I., DeAngelis, G. C., & Freeman, R. D. (1990). Stereoscopic depth discrimination in the visual cortex: Neurons ideally suited as disparity detectors. Science, 249, 1037–1041.
- Ono, H., Lillakas, L., Grove, P. M., & Suzuki, M. (2003). Leonardo's constraint: Two opaque objects cannot be seen in the same direction. Journal of Experimental Psychology: General, 132(2), 253–265.

- Ono, H., Wade, N. J., & Lillakas, L. (2002). The pursuit of Leonardo's constraint. Perception, 31(1), 83–102.
- Ooi, T. L., & He, Z. J. (2006). Binocular rivalry and surface-boundary processing. Perception, 35(5), 581–603.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. Spatial Vision, 10, 437–442.
- Pianta, M. J., & Gillam, B. J. (2003). Paired and unpaired features can be equally effective in human depth perception. Vision Research, 43, 1–6.
- QIAN, N. (1997). Binocular disparity and the perception of depth. Neuron, 359– 368.
- Ramachandran, V. S. (1992). Blind spots. Sci. Am., 266, 89–91.
- Ramachandran, V. S., & Gregory, R. L. (1991). Perceptual filling in of artificially induced scotomas in human vision. Nature, 350, 699–70.
- Ratcliff, R. (1993). Methods for dealing with reaction time outliers. Psychological Bulletin, 114, 510–532.
- Saye, A., & Frisby, J. P. (1975). Role of monocularly conspicuous features in facilitating stereopsis from random-dot stereograms. Perception, 4, 159–171.
- Shimojo, S., & Nakayama, K. (1990). Real World Occlusion Constraints and Binocular Rivalry. Vision Research, 30(1), 69–80.
- Su, Y., He, Z. J., & Ooi, T. L. (2009). Coexistence of binocular integration and

suppression determined by surface border information. PNAS, 106(37), 15990–15995.

- Tehovnik, E. J., Slocum, W. M., & Schiller, P. H. (2003). Saccadic eye movements evoked by microstimulation of striate cortex. European Journal of Neuroscience, 17, 870–878.
- Tripathy, S. P., & Levi, D. M. (1999). Looking behind a pathological blind spot in human retina. Vision Research, 39, 1917–1925.
- Tsirlin, I., Wilcox, L. M., & Allison, R. S. (2010). Monocular occlusions determine the perceived shape and depth of occluding surfaces. JOV, 10(6), 1–12.
- van Bogaert, E. A., Ooi, T. L., & He, Z. J. (2008). The monocular boundary contour mechanism in binocular surface representation and suppression. Perception, 37(8), 1197–1215.
- van Ee, R., & Erkelens, C. J. (2000). Is there an interaction between perceived direction and perceived aspect ratio in stereoscopic vision? Perception & Psychophysics, 62, 910–926.
- van Oeffelen, M. P., & Vos, P. G. (1982). A probabilistic model for the discrimination of visual number. Perception & Psychophysics, 32(2), 163–170.
- Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions, and objects in visual working memory. Journal of Experimental Psychology: Human Perception and Performance, 27, 92– 114.

- Walls, G. L. (1954). The filling-in process. Am. J. Optom. Arch. Am. Acad. Optom., 31, 329–341.
- Watanabe, O., & Fukushima, K. (1999). Stereo algorithm that extracts a depth cue from interocularly unpaired points. Neural Networks, 12, 569–578.
- Wichmann, F. A., & Hill, N. J. (2001a). The psychometric function: I. Fitting, sampling, and goodness of fit. Perception & Psychophysics, 63(8), 1293–1313.
- Wichmann, F. A., & Hill, N. J. (2001b). The psychometric function: II. Bootstrapbased confidence intervals and sampling. Perception & Psychophysics, 63(8), 1314–1329.
- Wilcox, L. M., & Lakra, D. C. (2007). Depth from binocular half-occlusions in stereoscopic images of natural scenes. Perception, 36(6), 830–839.
- Wilson, H. R., & Giese, S. C. (1977). Threshold visibility of frequency gradient patterns. Vision Res., 17, 1177–1190.
- Zhaoping, L. (2002). A saliency map in primary visual cortex. Trends in Cognitive Sciences, 6(1), 9–16.
- Zhaoping, L. (2008). Attention capture by eye of origin singletons even without awareness--A hallmark of a bottom-up saliency map in the primary visual cortex. Journal of Vision, 8(5), 1–1.
- Zhaoping, L., & Guyader, N. (2010). Interference with Bottom-Up Feature Detection by Higher-Level Object Recognition. Current Biology, 17, 26– 31.

Zhaoping, L., & May, K. A. (2007). Psychophysical Tests of the Hypothesis of a Bottom-Up Saliency Map in Primary Visual Cortex. PLoS Computational Biology, 3(4), e62.



**Proportion Difference** 

A1.1: The fitted functions to the 4 baseline numerosities for the 6 conditions for participant ad.



**Proportion Difference** 

A1.2: The fitted functions to the 5 baseline numerosities for the 6 conditions for participant dg.

Proportion Responses 'test appears more numerous'



Proportion Responses' test appears more numerous'



A1.3: The fitted functions to the 5 baseline numerosities for the 6 conditions for participant kz.



**Proportion Difference** 

A1.4: The fitted functions to the 5 baseline numerosities for the 6 conditions for participant lb.

Proportion Responses 'test appears more numerous'



Proportion Responses 'test appears more numerous'

**Proportion Difference** 

A1.5: The fitted functions to the 5 baseline numerosities for the 6 conditions for participant ml.




A1.6: The fitted functions to the 5 baseline numerosities for the 6 conditions for participant vk.

Proportion Responses 'test appears more numerous'



Proportion Difference A1.7: The fitted functions for the 6 conditions for participant ad.



Proportion Difference A1.8: The fitted functions for the 6 conditions for participant dg.







Proportion Difference A1.10: The fitted functions for the 6 conditions for participant lb.



Proportion Difference

A1.11: The fitted functions for the 6 conditions for participant ml.



Proportion Difference

A1.12: The fitted functions for the 6 conditions for participant vk.



Proportion Responses' test appears more numerous'



Proportion Responses 'test appears more numerous'







A1.16: The fitted functions to the 4 baseline numerosities for the 7 conditions for participant dl.







A1.18: The fitted functions to the 4 baseline numerosities for the 7 conditions for participant jl.



A1.19: The fitted functions to the 4 baseline numerosities for the 7 conditions for participant kz.



## Proportion Difference

A1.20: The fitted functions to the 4 baseline numerosities for the 7 conditions for participant lb.



A1.21: The fitted functions to the 4 baseline numerosities for the 7 conditions for participant lh.



A1.22: The fitted functions to the 4 baseline numerosities for the 7 conditions for participant lo.











