

The Impact of Monocular Self-Occlusions on Depth Perception

Domenic Au

A thesis submitted to the Faculty of Graduate Studies

in partial fulfillment of the requirements for the

Degree of Master of Science

Graduate Program in Biology

York University

Toronto, Ontario

October 2020

©Domenic Au, 2020

ABSTRACT

The amount of depth perceived from half-occlusion arrangements between two objects has been shown to depend on both the size of the occluded region and the region's texture properties (Nakayama & Shimojo, 1990; Tsirlin, Wilcox, & Allison, 2010). Depth judgements from such occlusions increase with occlusion size but can be degraded when the monocular texture information is inconsistent with the binocular regions in the scene (Grove, Gillam, & Ono, 2002). To date, depth from half-occlusions has been studied almost exclusively in the two-object configurations described above. However, monocular regions also result from *self-occlusions* and in such cases, because the monocular regions form part of a single object, the size of the monocular regions and the monocular texture properties could influence its perceived shape. In a series of experiments, I manipulated the degree to which self-occluded regions were consistent with binocular viewing geometry. I also assessed sensitivity to varying monocular texture information, perceived object coherence, and occlusion size under geometrically valid and invalid conditions. I found that monocular texture changes had little impact on perceived depth when binocular disparity was present in the stimulus suggesting that disparity in the binocular part of the object is extrapolated across monocular regions to determine object shape. There was also considerable tolerance to invalid occlusion arrangements when monocular regions were perceived as part of a coherent object. Furthermore, unlike half-occlusions involving separate objects, quantitative depth from self-occlusions do not follow predictions based on binocular viewing geometry such as the minimum depth constraint. My experiments show that the interpretation of self-occlusions depends critically on their context; heuristics used to extract depth information from such monocular regions are contingent on foreground/background segmentation.

TABLE OF CONTENTS

ABSTRACT.....	ii
TABLE OF CONTENTS.....	iii
INTRODUCTION	1
Stereopsis and monocular occlusions.....	1
Types of half-occlusions	2
Depth percepts from monocular occlusions	3
Qualitative percepts	3
Quantitative percepts	4
Objectives.....	6
GENERAL METHODS.....	8
EXPERIMENT 1	9
Introduction	9
Observers.....	10
Stimuli	10
Procedure.....	11
Results and Discussion.....	12
EXPERIMENT 2	13
Introduction	13
Observers.....	14
Stimuli	14
Procedure.....	19
Results and Discussion.....	19
EXPERIMENT 3	24
Introduction	24
Observers.....	25
Stimuli	25
Procedure.....	28
Results and Discussion.....	29
EXPERIMENT 4	32

Introduction	32
Observers	32
Stimuli	33
Procedure	33
Results and Discussion	34
EXPERIMENT 5	36
Introduction	36
Observers	37
Stimuli	37
Procedure	38
Results and Discussion	39
GENERAL DISCUSSION	41
Summary	41
Self-occlusions and half-occlusions	45
Self-occlusions in natural scenes	48
Future work	49
Conclusion	50
REFERENCES	51

INTRODUCTION

Stereopsis and monocular occlusions

Binocular stereopsis is the percept of depth that is obtained from the horizontal separation of the two eyes. This results in two slightly different images of the same scene presented to the left and right retinas. The positional difference between the two images is referred to as binocular disparity, which is used by the brain to perceive a three-dimensional representation of the world in a process known as stereopsis. Stereopsis can provide a strong and compelling percept of depth even when it is the only depth cue present in a scene (Wheatstone, 1838; Howard & Rogers, 2012). It has been shown that under ideal viewing conditions, experienced observers possess remarkably low thresholds of 2-6 seconds of arc when discriminating relative disparity (Coutant & Westheimer, 1993; Ogle, 1952; Blakemore, 1970; Westheimer, 1979; Westheimer & McKee, 1979; Badcock & Schor, 1985). A wide range of low level stimulus properties have been shown to impact the precision of stereopsis, including spatial frequency, contrast, luminance, and exposure duration (Ogle & Weil, 1958; Schor & Wood, 1983; Christopher & Rogers, 1994; Schor, Wood, & Ogawa, 1984; Howard & Rogers, 1995; Siderov & Harwerth, 1995).

Binocularity does not only produce this disparity between matched features, but also introduces content that is unique to one eye. For instance, monocular areas can arise when an object physically occludes a more distant object in a scene in one eye but not the other. Such phenomena are referred to in the literature as ‘monocular occlusions’ or ‘half-occlusions’ because these regions do not contain a corresponding match in the other eye. (Gillam & Borsting, 1988). Leonardo da Vinci illustrated this phenomenon when he described distant surfaces being occluded by vertically oriented edges of surfaces placed closer in the foreground of the scene (Da Vinci, 1877; Lawson & Gulick, 1967; Gillam & Borsting 1988; Nakayama &

Shimojo, 1990). Specular highlights are another type of monocular information, but these features do not involve occlusions. Instead, the surface reflectance properties can cause specular highlights that are unique to one eye but not the other.

Types of half-occlusions

The most commonly studied type of half-occlusion configuration, illustrated in Figure 1a, consists of an object which occludes parts of more distant surfaces or objects from one eye's view (Gillam & Borsting 1988, Nakayama & Shimojo, 1990; Harris & Wilcox, 2009; Tsirlin, Wilcox, & Allison, 2010). Similar to these traditional half-occlusions, self-occlusion occurs when a single object occludes regions within itself (see Figure 1b). Both types of occlusions are the result of nearer binocular objects or parts occluding more distant regions of a scene.

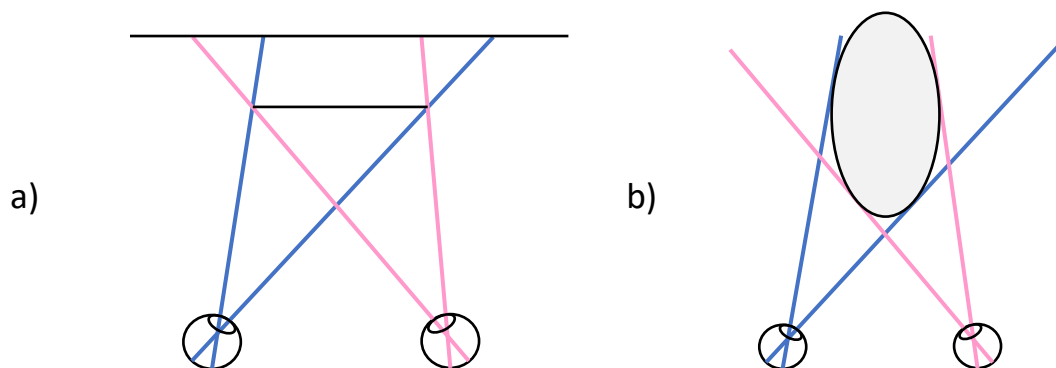


Figure 1. Diagram of viewing geometry for a half-occlusion (a) and a self-occlusion (b).

Half-occlusions involving separate objects are prevalent in the world around us because we often encounter multiple objects at different distances. For example, monocular regions of a background are present when looking through a picket fence because each eye views different portions of the background behind the fence (Forte, Peirce, & Lennie, 2002; Harris & Wilcox, 2009). Self-occlusions are less common because they require an object of a certain shape and

size be seen from a particular viewing position that lies between the eyes. However, self-occlusions do occur naturally in the real-world particularly when objects are very close. For example, faces of a cardboard box can be occluded to one of the eyes when viewed from a certain position (Wilcox & Lakra, 2007). While parsimony would suggest that all types of monocular occlusions are processed in the same way by the visual system, because self-occlusions have received less empirical attention, this assumption has not been confirmed.

Depth percepts from monocular occlusions

Qualitative Percepts

Gillam and Borsting (1988) predicted that stereoscopic processing would be facilitated by vertical depth edges created by monocular regions because most stereoscopic processing is located at depth discontinuities. They investigated this by measuring the speed at which observers indicated which of two surfaces appeared closer in depth. They presented two random-dot surfaces positioned at different disparities and evaluated two types of monocular occlusion in the region between the two disparate surfaces: 1) monocular texture consistent with the stereogram's dot density and 2) a blank textureless monocular region. The condition with the texture contained unmatched regions, which created a salient monocularly occluded region. Observers took longer to determine which surface was nearer in the untextured condition. Gillam and Borsting (1988) concluded that the presence of half-occlusions facilitates stereoscopic processing. In subsequent experiments, Grove and Ono (1999) also investigated the impact of texture properties within monocular regions on reaction times. They reported that it took longer to discriminate depth order when viewing stimuli with a substantially greater texture density in the monocular zone relative to the rest of the stimulus. However, they observed no significant

difference between textureless and consistent textured monocular regions, which was contrary to Gillam and Borsting's findings (1988).

Qualitative percepts obtained from half-occlusions have also been shown to affect the perceived shape of the occluding object. For example, an occluded region along the vertical edge between the occluder and the background can influence the perceived depth as being at the depth of the occluder, causing a change in the perceived shape of the occluder from square to rectangular. These results suggest that monocular regions can define object discontinuities which in turn constrain surface interpolation (Tsirlin, Wilcox, & Allison, 2010).

Quantitative Percepts

Although the monocular regions in half-occlusions, by definition, do not contain a corresponding match in the other eye, the distance of the occluded region from the occluder can be constrained from the viewing geometry. That is, the minimum possible distance between the two regions is constrained by the line of sight from the eye that cannot see the occluded region, to the nearest edge of the occluding surface (see Figure 2). This minimum possible distance is calculated from the size of the monocular region (Tsirlin, Wilcox, & Allison, 2010; Nakayama & Shimojo, 1990). There is a positive linear relationship between the size of the monocular region and the minimum possible amount of depth that can be perceived.

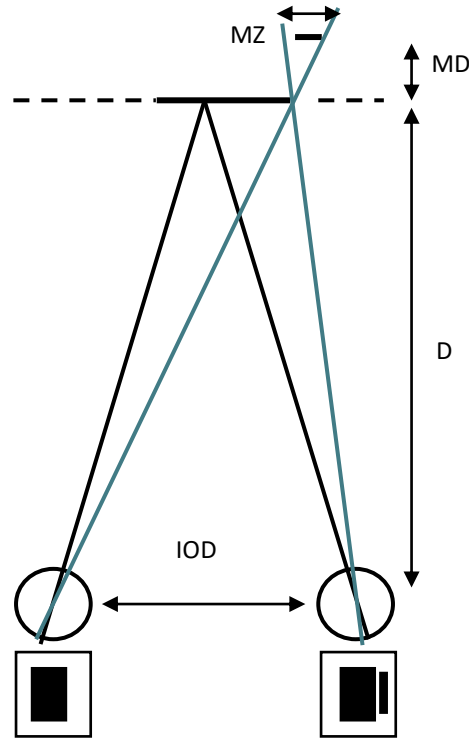


Figure 2. Diagram of viewing geometry for the minimum depth constraint. *MZ* = monocular zone size, *D* = viewing distance, *IOD* = interocular distance, *MD* = minimum depth. $MD = MZ * VD/(IOD-MZ)$

Tsirlin, Wilcox, & Allison (2010) demonstrated that depth estimates correspond to the minimum depth constraint under specific viewing conditions. However, they noted that in the so-called Da Vinci arrangements used by Gillam and Borsting (1988) and Nakayama and Shimojo (1990) the textured edges of the binocular occluder could provide stereoscopic matches with the edges of the occluded monocular region. They showed that this was a potential factor that could influence the perceived depth between the occluder and the monocular region (Tsirlin, Wilcox, & Allison, 2010; Gillam et al., 2003; Liu et al., 1997). When devising such stimuli, it is important to consider potential stereoscopic depth matches and exclude them when possible.

In addition to measuring the amount of perceived depth based on the size of a monocular region, investigators have also examined the impact of varying the similarity of the texture in the monocular and binocular regions on perceived depth. For instance, Grove, Gillam, and Ono (2002) presented half-occluded stimuli using planar surfaces like those shown in Figure 1 and manipulated the degree to which the half-occluded region was consistent with the background texture or the occluding surface. The results showed that depth was underestimated when the monocular region's texture was different than that of either the background's texture or the occluding surface's texture.

Objectives

Considerable research has evaluated the impact of half-occlusions on depth percepts for simple stimuli (Gillam & Borsting, 1988; Grove & Ono, 1999; Grove, Gillam, & Ono, 2002; Tsirlin, Wilcox, & Allison, 2010). It stands to reason that similar results would be observed for different types of occlusion. However, very few studies have examined self-occlusions or asked if depth from self-occlusions is perceived the same way as half-occlusion configurations that consist of two objects. Results of Wilcox and Lakra (2007) suggest that the presence of self-occlusions aids depth perception, however they did not evaluate if self-occluded regions contribute to depth magnitude percepts.

In the experiments reported here I investigate the qualitative and quantitative percepts obtained from monocular self-occlusions in 3D rendered objects. The first set of studies investigated how monocular texture information impacts the perceived depth of self-occluded objects. As noted above, it is not known if texture consistency (across monocular and binocular regions) is important in cases of self-occlusion. In Experiments 1 and 2, I evaluate the role of monocular texture information on the amount of perceived depth in self-occluded objects and the

results reveal that estimated depth is consistent across monocular texture manipulations when the stimulus contains familiar object shape information from texture and binocular disparity. When familiar shape is reduced, the impact of monocular texture is more consistent with previous findings (Gillam, Grove, & Ono, 2002). In Experiment 3, I evaluate the impact of self-occlusions on perceived object coherence by introducing gaps of variable size between the occluded and occluding object regions. In this study, observers report whether they perceived a single object in valid and invalid arrangements. The results of Experiment 3 reveal that observers are tolerant violations of binocular viewing geometry when assessing object coherence. In Experiments 4, quantitative depth is assessed using a subset of the stimuli in Experiment 3 and both depth magnitude estimation and disparity matching tasks. The geometric constraints of the objects in Experiment 4 encompassed a limited range of monocular region sizes, which made changes in perceived depth difficult to assess. The stimuli used in Experiment 5 were modified to make the monocular region size more salient. However, depth was overestimated and was not consistent with minimum depth constraint predictions under ecologically valid viewing conditions and estimates did not increase as monocular region size increased under invalid viewing conditions. Taken together, these experiments suggest that there are substantial individual differences in how sensitive observers are to self-occlusions. As might be expected, observers do not use the minimum depth constraint to estimate depth in invalid configurations, but they do not show some dependence on it under valid configurations. It appears that the occluding object properties such as 3D shape play a larger role in determining the impact of self-occlusions on perceived depth.

GENERAL METHODS

Observers

A total of 18 observers were recruited and participated in the experiments described here. These observers were comprised of lab members and lab volunteers. All observers had stereoacuity of at least 40 seconds of arc as assessed using the RANDOT™ stereoacuity test and self-reported normal to corrected-to-normal visual acuity. Each observer's interocular distance (IOD) was measured using an optical digital pupillary distance measuring tool. IODs ranged from 57 mm to 67 mm with an average IOD of 63 mm. The research protocol used in these experiments was approved by the York University research ethics board.

Stimuli and Apparatus

Stimuli were modeled and rendered using Blender® version 2.79b. All stimuli were rendered as stereopairs with off-axis projection. Experiments were programmed using the Psychtoolbox (Brainard, 1997; Pelli, 1997) package for Matlab® on an Intel® Xeon computer with the Windows 10 operating system. The stimuli were displayed on two calibrated HP Z Display Dream Color monitors in a mirror stereoscope arrangement. In each experiment, stimuli were displayed on a light grey background (73.8 cd/m²). Each monitor's resolution was 1920 x 1200 pixels with a refresh rate of 60 Hz. During each experiment, observers were seated facing the mirrors with their chin in a chin rest to stabilize their head position. The viewing distance was 45 cm. Depth estimates were made using a Logitech® gamepad to adjust the length of a virtual horizontal line on the display. This estimation method has been validated by Hartle and Wilcox (2016) to provide reliable depth estimates.

EXPERIMENT 1

Previous studies have investigated the role of consistent and inconsistent texture information on the facilitation of how half-occlusions impact the speed of stereoscopic depth processing as most processing is located at depth discontinuities. Gillam and Borsting (1988) reported that when the monocular texture is inconsistent with that in the binocular regions, depth discrimination was slow. However, Grove and Ono (1999) found no such effect. These experiments used simple versions of half-occluded (daVinci) stimuli like those shown in Figure 2 and it is unclear why the results are inconsistent (Gillam & Borsting, 1988; Grove & Ono, 1999).

In a follow-up study Grove, Gillam, and Ono (2002) measured the impact of the texture within the monocular regions on the perceived location in depth of the occluded region using an adjustable disparity probe. In all the conditions tested, the position in depth was underestimated when the monocular region texture was inconsistent with that of the background or the surrounding surface.

The aim of Experiment 1 was to investigate whether inconsistent monocular texture information would also cause underestimation of the depth of self-occluded objects. If the previous findings of Grove, Gillam, and Ono (2002) are generalizable to self-occlusions, then less depth should be observed for stimuli that contain a greater texture inconsistency in the monocular region. To evaluate this, 3D stimuli were rendered with correct projective geometry. The stimuli were self-occluded objects containing monocular regions on both sides.

Observers

Six observers participated in Experiment 1. Two were male and four were female. Ages ranged from 20 to 35 years with an average age of 28.2 years. All had previous experience in psychophysical experiments but were naïve to this experiment. Observers had self-reported normal or corrected-to-normal visual acuity and stereoacuity of at least 40 seconds of arc as assessed using the RANDOT™ stereoacuity test.

Stimuli

Stimuli were half-cylinders modeled and rendered using Blender (see Figure 3). The peak of the half-cylinder was located at the screen plane and the distance from the peak to the edges of the object was rendered to produce a depth of 10.1 cm (1.46 degrees). The width and height of the objects' binocular regions were 7.8 degrees and 14.4 degrees in angular size respectively. They were textured with a non-repeating marble pattern (40.1 cd/m²) that was distorted and scaled to fit the size of the object. The monocular regions presented on each side of the object were 0.55 degrees wide in angular size and spanned the full height of the object. All stimuli were physically possible objects with a minimum depth constraint of 2.88 cm for the 10.1 cm peak depth condition. The texture density in the monocular regions ranged from being consistent with the binocular region to extremely dense (5 levels were selected as shown in Figure 3-a). Monocular regions were identified on the object mesh in Blender and textured relative to the local object scale. All stimuli were displayed with the top and bottom hidden behind a viewing aperture to eliminate additional depth information provided by the edge curvature.

To help avoid response stereotyping and to confirm that observers' depth percepts varied with changing disparity, additional catch trials were included with peak-to-edge depths of 7.7 cm and 5.5 cm. Monocular regions were still present in these conditions, but were smaller due to the shallower depths.

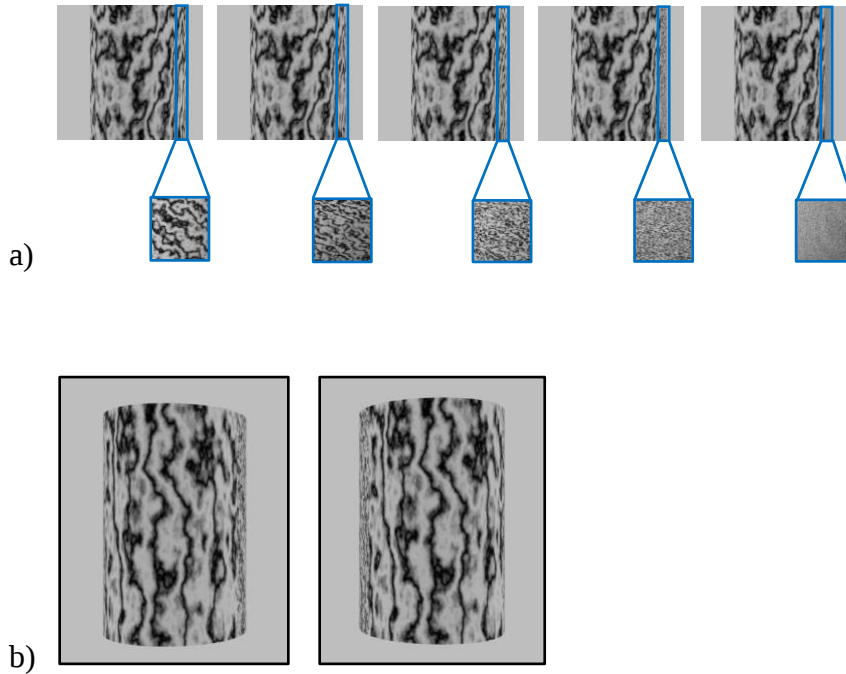


Figure 3. a) Monocular regions of the object contained five texture. b) Example of self-occluded half-cylinder arranged for cross fusion.

Procedure

On each trial, observers were asked to estimate the perceived depth of the object using an adjustable horizontal line positioned below the stimulus. The length of the line was adjusted to indicate the amount of perceived depth from the peak of the object to its furthest edge. Stimuli remained visible until a response was made using a gamepad. Peak depths of 10.1 cm, 7.7 cm, and 5.5 cm were repeated ten times for each texture density. A total of nine conditions (five experimental conditions and four control conditions) containing varying texture densities were randomly interleaved resulting in 50 trials for 10.1 cm, 20 trials for 7.7 cm, and 20 trials for 5.5 cm.

Results and Discussion

Figure 4 shows observers' depth estimates as a function of the texture density in the monocular region. While estimated depth increases with increasing object depth, there appears to be no effect of changing the texture density within the monocular region. A repeated measures ANOVA confirmed these observations. A Mauchly's test of sphericity showed that the assumption of sphericity was violated for the main effect of texture density, $p < 0.0001$ so the Greenhouse-Geisser correction was applied. As expected, there was no effect of texture density ($F(4,20) = 0.4222$, $p > 0.05$; $\eta_G^2 = 0.005$) and there was a significant effect of peak depth, which corresponded to the different monocular region sizes ($F(2,22) = 11.7109$, $p = 0.00009$). Simple main effects of condition (with a Bonferroni adjustment) did not reveal any significant interactions.

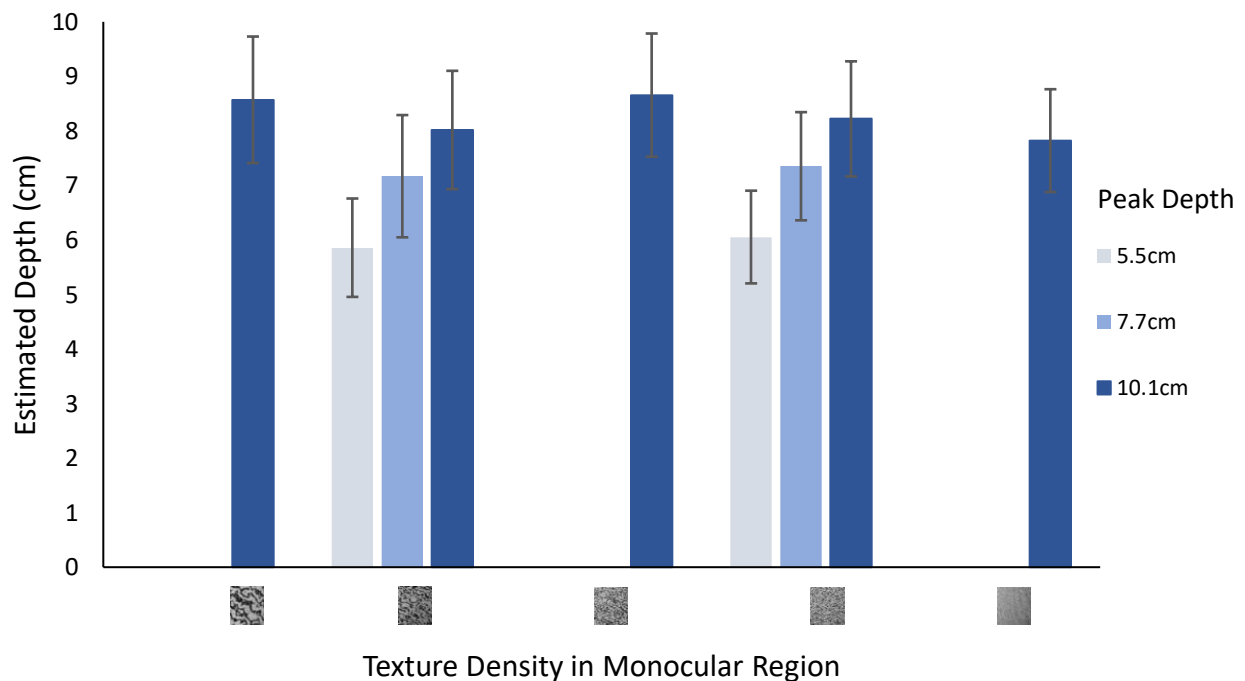


Figure 4. Estimated depth is plotted as a function of the texture density in the monocular region for three peak depth conditions: 5.5 cm and 7.7 cm peak depths are shown for the second and

fourth texture densities. The lowest texture density on the far left is identical to that used in the binocular region. Error bars show standard error of the mean.

It was predicted that the introduction of a change of texture density within the monocular region's texture content relative to the rest of the object would result in depth underestimates (Grove, Gillam, & Ono, 2002). However, depth estimates do not vary substantially across texture density conditions. Instead, it appears that depth percepts remain robust even in the presence of relatively large self-occluded regions as well as texture differences across conditions and between surface regions.

A potential explanation for the effect of peak depth shown here, is that observers were able to extrapolate the volumetric shape of the stimuli from the binocularly defined surface peak. In all stimuli the surface was a smooth object of the same elliptical shape. To evaluate the role of 3D shape on depth estimates in Experiment 2, the stimulus was modified to make this information less informative.

EXPERIMENT 2

It was hypothesized that the monocular texture density manipulation in Experiment 1 did not impact observers' depth estimates because of the potential influence of 3D shape information present along the curved front face of the object. If this is true, then reducing the amount of binocular disparity within the object and increasing the size of the monocular regions will cause observers to rely more on monocular texture information such as linear perspective as opposed to interpolating the shape from binocular disparity.

In Experiment 2, we varied the amount of shape information from binocular disparity by truncating the curved peak of the half-cylinder at different points along the front face of the object. This resulted in a set of stimuli with different amounts of binocular disparity, monocular region sizes, and texture information (see Figure 5 and 6-a). If familiar shape information was responsible for consistent depth estimates across monocular texture differences in Experiment 1, then we would expect to find that the presence of 3D shape from binocular disparity should aid depth estimation in Experiment 2. That is, monocular texture manipulations should not impact depth estimates when there is binocular disparity present along the nearest edges of the object sides. However, when little or no binocular disparity is present then we should see an effect of texture on depth estimation; depth estimates will be degraded as monocular texture density increases (and perspective information become less useful).

Observers

Eight observers participated in Experiment 2. Two were male and six were female. Ages ranged from 19 to 38 years with an average age of 26.6 years. All had previous experience in psychophysical experiments and five had participated in Experiment 1, none were aware of the experimental hypothesis. Observers had self-reported normal or corrected-to-normal visual acuity and stereoacuity of at least 40 seconds of arc assessed using the RANDOT™ stereoacuity test.

Stimuli

The stimuli were truncated versions of the those used in Experiment 1. In the first condition, the tip of the cylinder was cut off resulting in a 3 cm (3.81 degree) front face width. This condition was similar to Experiment 1 in that it contained binocular disparity along the

sides of the object, but the peak was removed. This reduced the amount of curvature at the peak but still provided some 3D shape from binocular disparity. For a comparison, a condition was created with minimal binocular disparity along the sides. This stimulus had a front face width of 6.2 cm (7.85 degrees in angular size) and it contained larger monocular regions (see Table 1). To help avoid response stereotyping, we generated an intermediate condition with a 4.5 cm (5.71 degrees in angular size) front face width that contained binocular disparity and monocular region sizes between that of the two other conditions (see Table 1). Each condition was tested at three peak depths of 5.5 cm (0.87 degrees), 7.7 cm (1.2 degrees), and 10.1 cm (1.5 degrees) as in Experiment 1. Consistent object depth was achieved by scaling the object after it was truncated. The texture mapping was consistent throughout each condition and did not scale with the object. The lowest, middle, and highest density from Experiment 1 was chosen (see Figure 6-b). As in Experiment 1, all stimuli were displayed with the top and bottom hidden behind an aperture to eliminate additional depth cues provided by the object's boundaries.

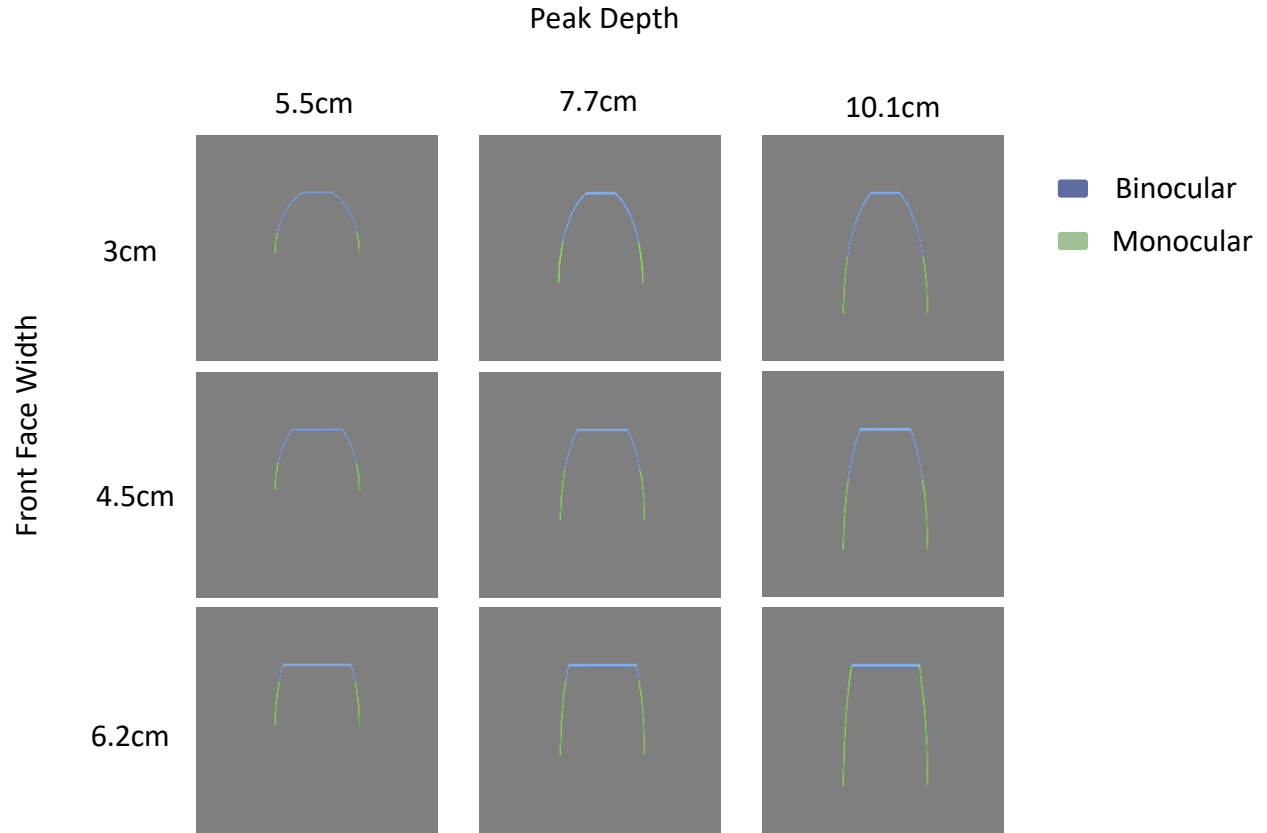


Figure 5. All conditions in Experiment 2 depicting top-down view the of object structure.

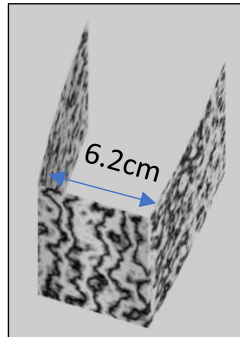
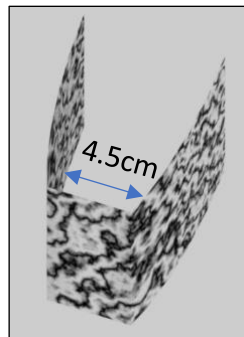
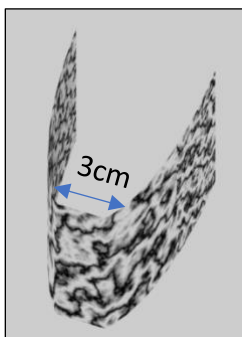
Binocular (blue) and monocular (green) regions are highlighted.

Table 1: Overview of monocular and binocular region widths, binocular disparity, and total depth for all conditions in Experiment 2.

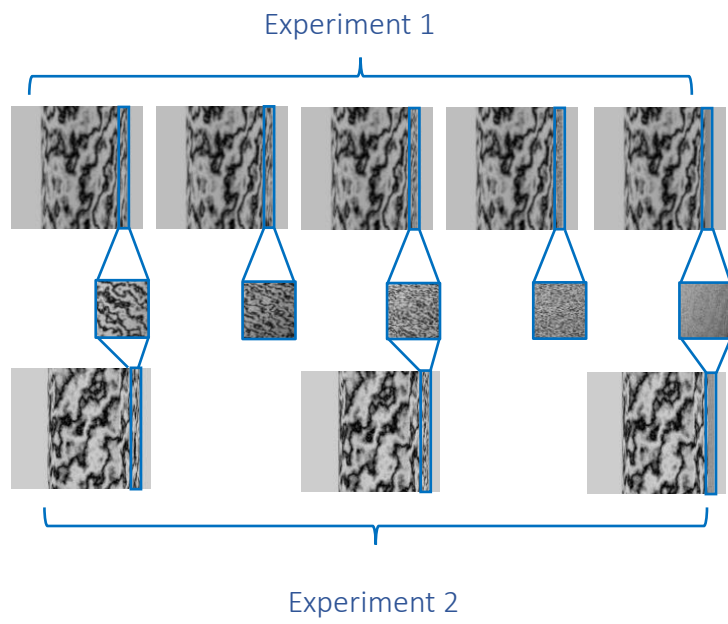
3cm Front Face Width			
Peak Depth	5.5 cm	7.7 cm	10.1 cm
Monocular Region Size	0.24 cm (0.31 deg)	0.32 cm (0.41 deg)	0.38 cm (0.48 deg)
Binocular Disparity	3.60 cm (0.59 deg)	4.10 cm (0.67 deg)	5.30 cm (0.84 deg)
Binocular Region Size	3.80 cm (4.83deg)	3.38 cm (4.30 deg)	3.22 cm (4.10 deg)

4.5cm Front Face Width			
Peak Depth	5.5 cm	7.7 cm	10.1 cm
Monocular Region Size	0.27 cm (0.34 deg)	0.38 cm (0.48 deg)	0.43 cm (0.55 deg)
Binocular Disparity	2.90 cm (0.48 deg)	3.30 cm (0.55 deg)	4.20 cm (0.68 deg)
Binocular Region Size	2.24 cm (2.85 deg)	1.84 cm (2.34 deg)	1.68 cm (2.14deg)
6.2cm Front Face Width			
Peak Depth	5.5 cm	7.7 cm	10.1 cm
Monocular Region Size	0.38 cm (0.48 deg)	0.51 cm (0.65 deg)	0.73 cm (0.93 deg)
Binocular Disparity	1.50 cm (0.26 deg)	1.30 cm (0.22 deg)	0 cm (0 deg)
Binocular Region Size	0.57 cm (0.73 deg)	0.16 cm (0.20 deg)	0 cm (0 deg)

a)



b)



c)

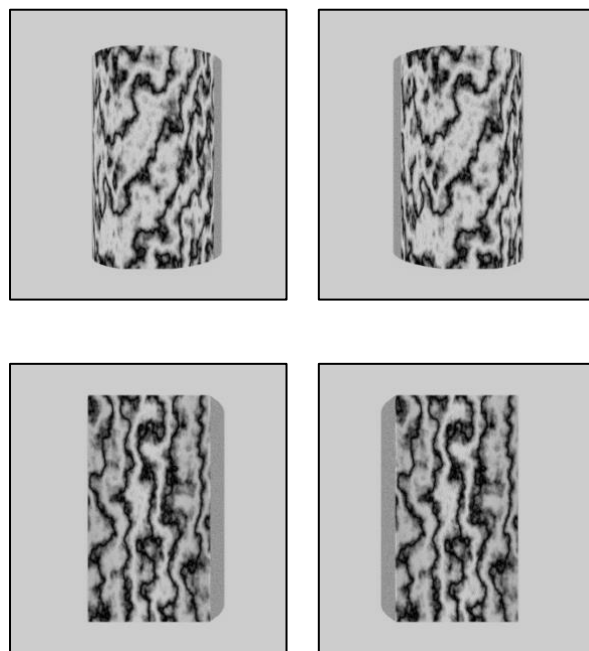


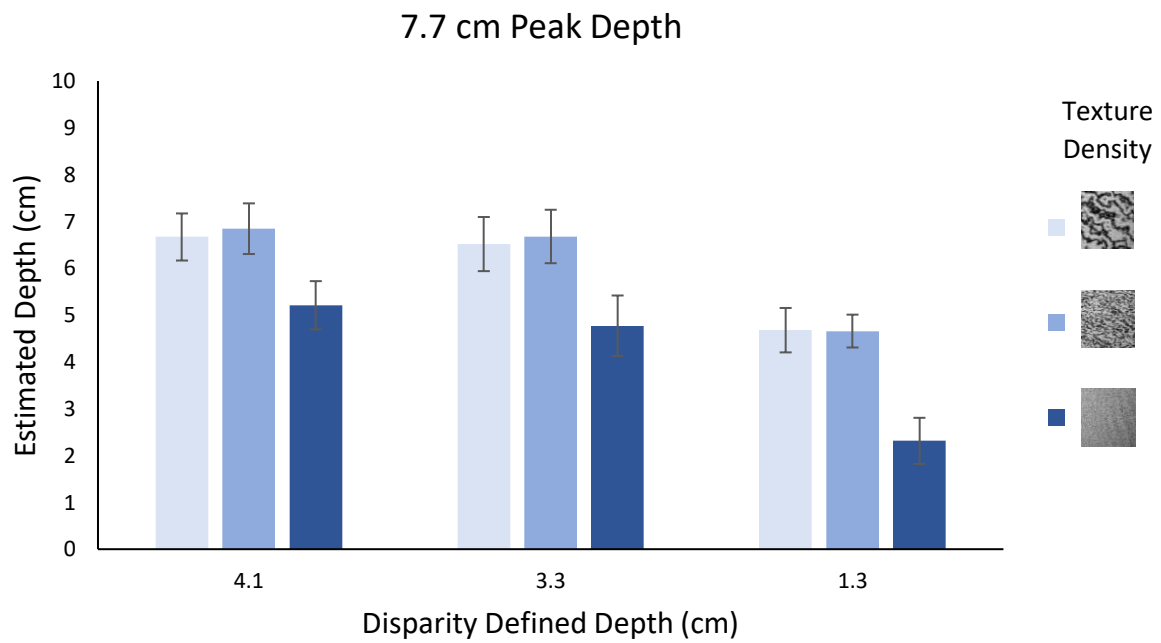
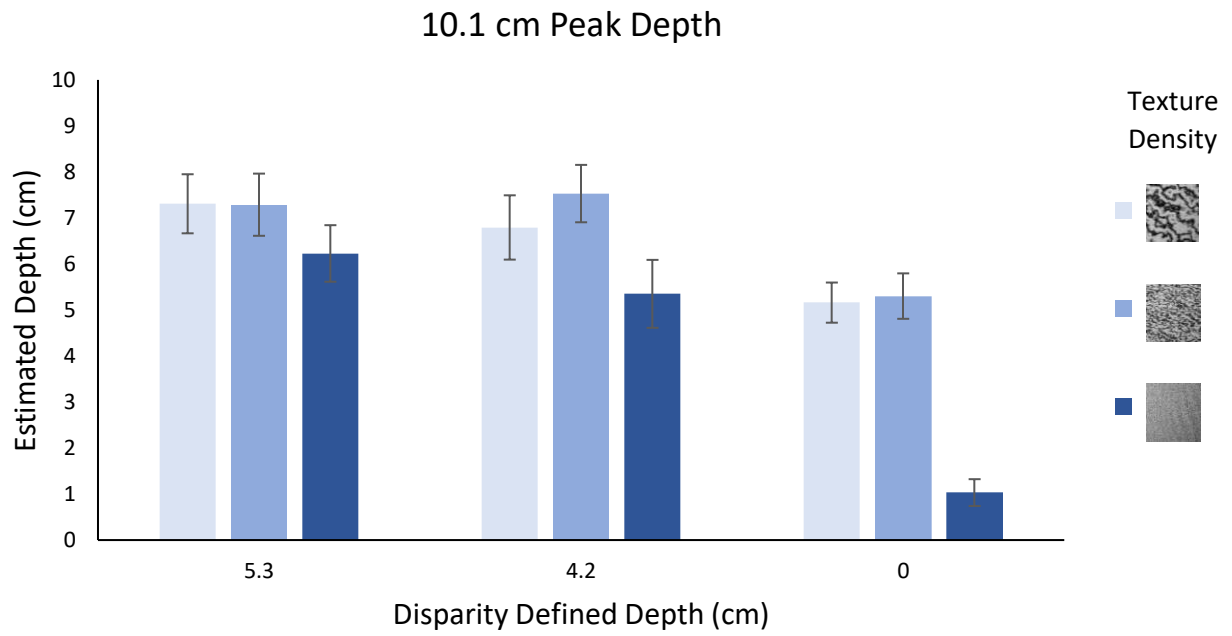
Figure 6. a) Three front surface width conditions. b) Three texture densities contained in the monocular region selected from Experiment 1. c) Examples of the 3 cm front width containing binocular disparity along the sides of the object (top) and the 6.2 cm front width containing no binocular disparity along the sides of the object (bottom). Both are arranged for cross fusion.

Procedure

The procedure was the same as that used in Experiment 1; observers were asked to match the width the perceived depth of the object using an adjustable horizontal line positioned below the stimulus. The length of the line was adjusted to indicate the amount of perceived depth from the front surface of the object to its furthest edge. Stimuli remained visible until a response was made. Peak depths of 10.1 cm were repeated ten times for each texture density while peak depths of 5.5 cm and 7.7 cm trials were included to avoid response stereotyping, these were seen five times. A total of 27 conditions were randomly interleaved (three front face widths x three texture densities x three peak depths) resulting in 90 trials for 10.1 cm, 45 trials for 7.7 cm, and 45 trials for 5.5 cm.

Results and Discussion

Figure 7 shows observers' depth estimates as a function of the size of the binocular region in the object. The results show that as binocular disparity decreases, depth estimates decrease but are consistently greater than the amount predicted from the disparity alone. As anticipated, the impact of monocular texture is most apparent when there is little to no binocular disparity (see the rightmost set of bar graphs in for all peak depths in Figure 7). Overall, depth was underestimated in the 10.1 cm peak depth condition and estimates were generally more accurate in the 7.7 cm and 5.5 cm peak depth conditions.



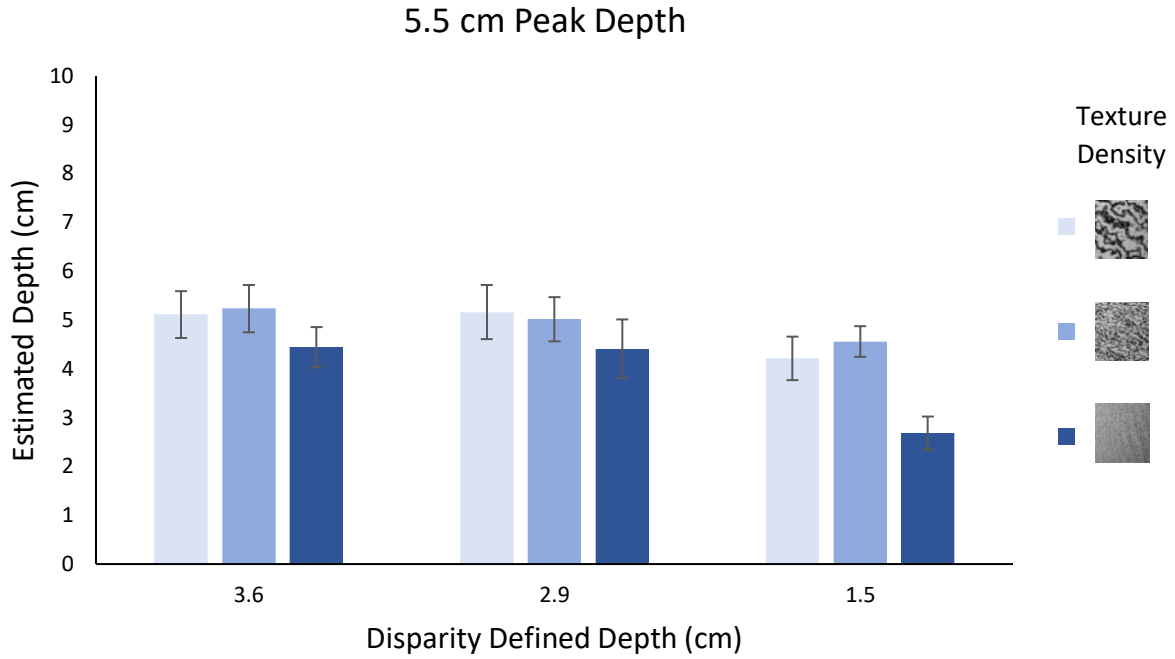


Figure 7: Estimated depth is plotted as a function of the disparity defined depth for each texture density in the monocular region. Each plot represents a different peak depth condition. Error bars show standard error of the mean.

Source	d.f.	F
Binocular Disparity	2, 14	40.409***
Texture Density	2, 14	14.537***
Peak Depth	2, 14	21.500***
Binocular Disparity * Peak Depth	4, 28	17.362***
Binocular Disparity * Texture Density	4, 28	10.586***
Peak Depth * Texture Density	4, 28	14.569***
Binocular Disparity * Peak Depth * Texture Density	8, 56	6.325***

* indicates $p < 0.05$. ** indicates $p < 0.01$. *** indicates $p < 0.001$

A three-way repeated measures ANOVA was used to analyze the results. There was a statistically significant three-way interaction between binocular disparity, peak depth, and texture density, $F(8, 56) = 6.325$, $p < 0.05$; $\eta_G^2 = 0.026$. A simple two-way interaction analysis revealed significant interactions in the binocular disparity conditions. For the 3 cm front face width containing the most binocular disparity, there was a significant effect of texture density ($F(2,14) = 20.7$, $p < 0.01$; $\eta_G^2 = 0.119$) and peak depth ($F(1.15,8.08) = 20.0$, $p < 0.01$; $\eta_G^2 = 0.253$), but no significant interaction between texture density and peak depth was observed. For the 4.5 cm front face width that contained intermediate amounts of binocular disparity, there was a significant effect of texture density ($F(1.10,7.71) = 10.3$, $p < 0.05$; $\eta_G^2 = 0.203$), peak depth ($F(1.17,8.18) = 21.3$, $p < 0.01$; $\eta_G^2 = 0.212$), and a significant interaction between texture density and peak depth ($F(4,28) = 9.54$, $p < 0.01$; $\eta_G^2 = 0.045$). For the 6.2 cm front face width containing the least amount of binocular disparity, there was a significant effect of texture density ($F(1.12,7.86) = 24.0$, $p < 0.01$; $\eta_G^2 = 0.457$) and a significant interaction between texture density and peak depth ($F(4,28) = 12.4$, $p < 0.01$; $\eta_G^2 = 0.112$), but no significant effect of peak depth was observed.

Simple main effects were run and the comparisons revealed that estimated depth differed significantly across texture densities for 6.2 cm front face width (containing the smallest amounts of binocular disparity) when compared with the 3 cm front face width ($p=0.016$, $p=0.039$, and $p=0.0001$, with increasing texture density) and the 4.5 cm truncated width ($p=0.024$, $p=0.005$, and $p=0.0002$, with increasing texture density).

It was predicted that by reducing the amount of binocular disparity in the object, familiar 3D shape information would be degraded, and depth estimates would decrease. Comparison of the condition with the most binocular disparity (3 cm face width) with the most extreme case (6.2 cm) containing little to no binocular disparity, shows a substantial decrease in perceived depth. As seen in Experiment 1, it appears that 3D shape information is interpolated from the disparity and used to estimate depth, so when this information is reduced there is a corresponding reduction in depth estimates. However, it is important to note that for most conditions, depth percepts are not eliminated, just reduced.

It is also notable that when disparity is present along the sides of the object, it appears that the monocular regions are used to aid depth estimates. If depth were perceived only from disparity in the 3 cm and 4.5 cm face width conditions, then more extreme underestimations would have been observed. This is particularly evident in the 5.5 cm and 7.7 cm peak depth conditions where estimates exceed that of the disparity in the object. For example, the 3 cm front face width condition contains 3.6 cm of binocular disparity for the 5.5 cm peak depth. In all three texture density conditions for this peak depth, average estimated depth exceeded the object's disparity of 3.6 cm (average estimates were 5.11 cm, 5.23 cm, and 4.44 cm from the lowest to highest texture density). Similarly, stimuli contained 4.1 cm of binocular disparity in the 3 cm front width for the 7.7 cm peak depth and estimates were greater than 4.1 cm across all texture densities (average estimates were 6.67 cm, 6.85 cm, and 5.21 cm from the lowest to highest texture density).

The impact of monocular texture density was greatest in the 10.1 cm peak depth condition when the stimulus contained no binocular disparity. Depth was notably underestimated in that condition when presented with the highest monocular texture density. Compared with the

other texture density conditions, the highest density had minimal monocular depth information (e.g. texture foreshortening and linear perspective). This demonstrates that when the observer is unable to use binocular disparity to extrapolate 3D shape, they use the texture information located in the monocular regions to help make depth judgements. Under these conditions, when the texture information is too dense to provide reliable perspective cues, perceived depth is at its lowest.

EXPERIMENT 3

In Experiment 2, observers significantly underestimated the depth of the objects when the texture density difference was very high, suggesting that they were sensitive to this manipulation. However, the binocular disparity information along the side regions of the object was also most limited in those conditions. This combined with the texture difference between the sides and front surfaces may have resulted in a loss of object coherence which in turn influenced perceived depth. Thus, in the current experiment, we investigated the combined effect of perceived object coherence and occlusion geometry. It has been shown that the violation of natural viewing geometry affects perceived depth in half-occlusions. That is, perceived depth does not increase as a function of monocular zone size when the sign of the occluded region's depth is reversed (Nakayama & Shimojo, 1990). It is not yet understood whether a violation of natural viewing geometry would affect how *self-occluded* objects are perceived. In the following study, we examined perceived object coherence between valid and invalid viewing conditions by manipulating the width of gaps between the monocular and binocular surfaces. We predicted that for both valid and invalid viewing conditions, observers would experience a decrease in object coherence as gap width increased. Further, if the visual system is sensitive to the viewing

geometry in interpreting half-occlusions, then we would expect that perceived object coherence would be degraded under invalid viewing conditions.

Observers

9 observers participated in Experiment 3 and all had previous experience in psychophysical experiments. Three were male and six were female. Ages ranged from 20 to 35 years with an average age of 28 years. Observers had self-reported normal or corrected-to-normal visual acuity and stereoacuity of at least 40 seconds of arc as assessed using the RANDOT™ stereoacuity test.

Stimuli

The stimuli were half-cylinders textured with a non-repeating dot pattern (51.6 cd/m²). The binocular region of the object was a front face that was the same size as the previous experiments, but with the addition of a 5.16 arcmin protrusion that was scaled and located in the middle of the front face. The purpose of the protrusion was to introduce variation in shape so that the stimulus would not readily appear as a flat plane occluding a separate distant plane. Gap widths (0, 6.18, 12.36, 18.54, 24.72, 30.90, and 37.08 arcmin) were introduced between the edges of the front face and the inner edges of the monocular textures on both sides.

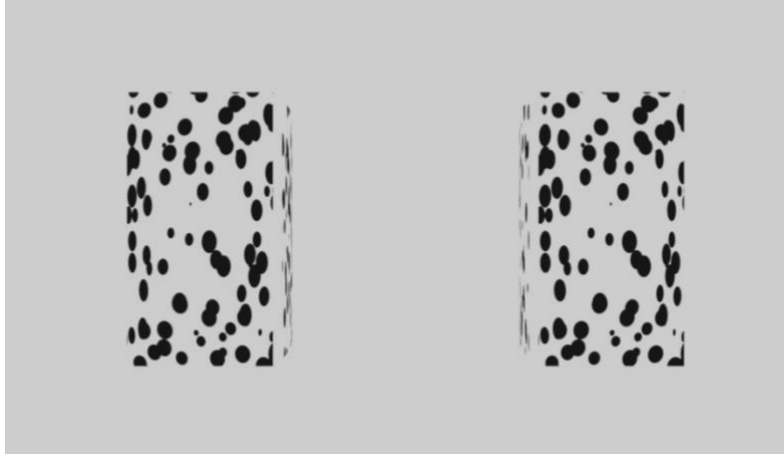


Figure 8: Example of self-occluded stimulus containing gaps in monocular regions arranged for cross fusion.

The stimuli were presented under ecologically valid and invalid viewing conditions (see Figure 9). In the ecologically valid condition, the size of the monocular regions remained constant regardless of the gap width. This meant that the outer edges of the monocular textures remained in the same position and as the gap width increased, the amount of monocular texture decreased.

In the ecologically invalid condition, the distance to the outer edges of the monocular regions increased as the gap width increased (see Figure 9). That is, the monocular textured areas were a constant size and were laterally shifted outward away from the central surface to create the gaps. When this occurred the monocular textured area of the object crossed outside the line of sight that was constraining the monocular region (see Figure 9). Under the assumption that the object shape is constant, the natural viewing geometry is violated. That is, although the side surfaces should have become visible to both eyes, the texture remained monocular. I will refer to these stimuli as ecologically ‘invalid’ but acknowledge that these stimuli are physically realisable but unlikely, which has been similarly described by Grove, Gillam, and Ono (2002).

Further, the texture defined depth information will be inconsistent with the 3D shape of the object.

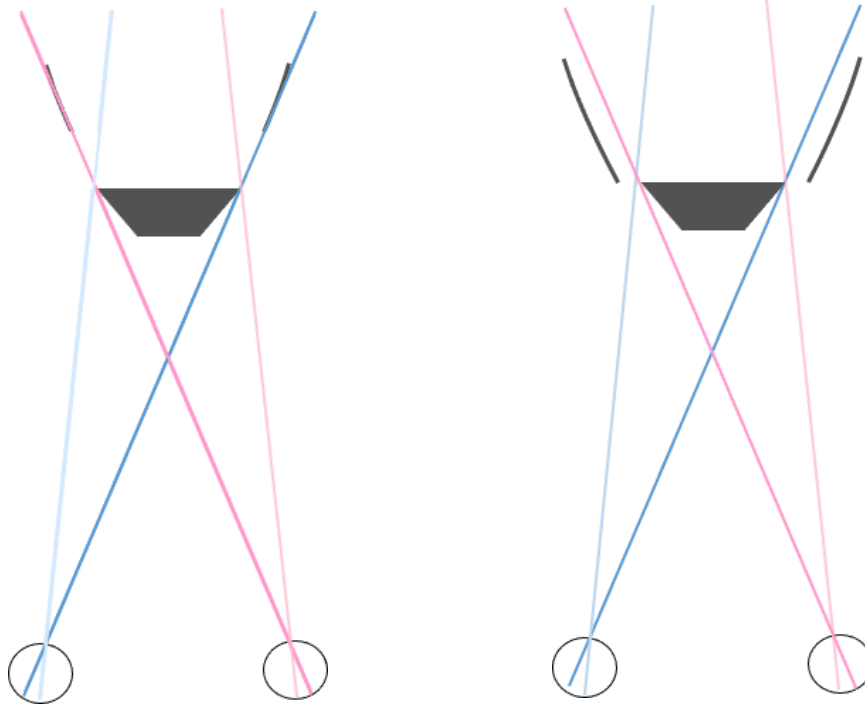


Figure 9: Top down view of ecologically valid condition (left) and ecologically invalid condition (right). Monocular side area is constrained by the line of sight in valid viewing and the monocular side area crosses the line of sight under invalid viewing.

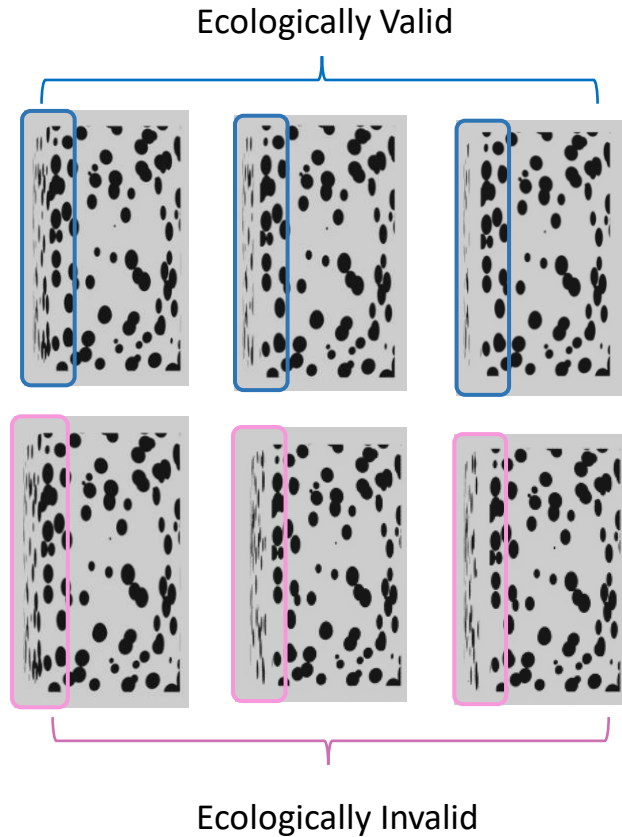


Figure 10: Top row depicts valid conditions in which the monocular region is fixed and increasing the gap width decreases the occluded texture. Bottom row depicts invalid conditions in which monocular regions increase as gap width increases.

Procedure

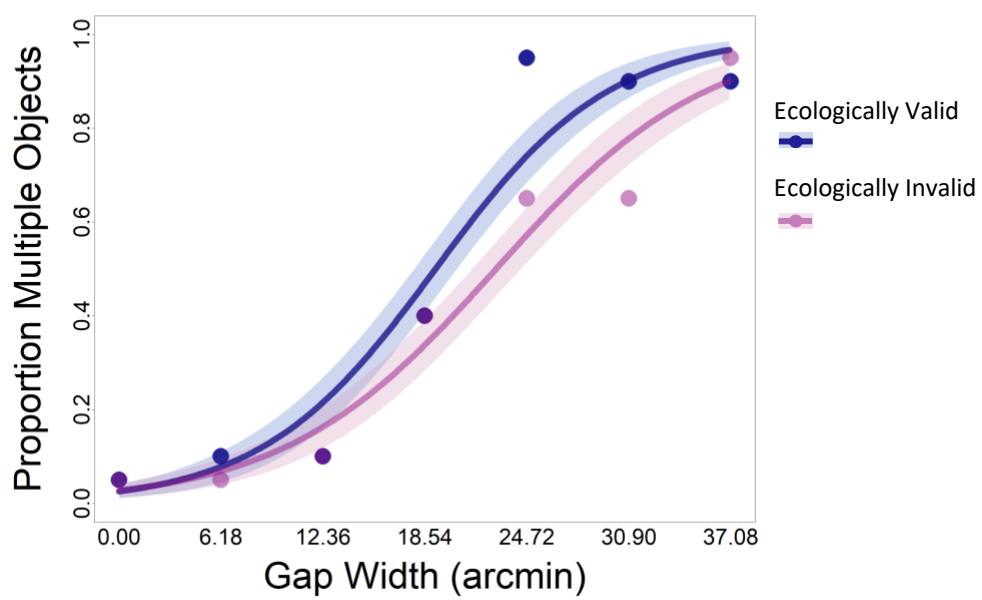
In this experiment, a two-alternative forced-choice method of constant stimuli was used. On each trial, observers were asked to indicate if they perceived a single object or multiple objects using two buttons on a gamepad. The stimuli were presented for 1 second followed by a response screen. All gap widths of 0, 6.18, 12.36, 18.54, 24.72, 30.90, and 37.08 arcmin were repeated 20 times for ecologically valid and invalid conditions. A total of 14 conditions (two

main viewing conditions x seven gap widths) were randomly interleaved resulting in a total of 280 trials.

Results and Discussion

Figure 10-a shows the proportion of trials that observers perceived the stimuli as consisting of multiple surfaces as a function of gap width. Observer's individual psychometric functions were fit using cumulative normal functions and PSEs were identified as the 50% point on the psychometric function. There is an increase in the proportion of trials that observers perceive multiple objects as gap width increased. A two-way repeated measures ANOVA was used to analyze the results. There was a significant main effect of gap width ($F(2.14,17.1) = 77.939, p > 0.01; \eta_G^2 = 0.811$), and viewing condition ($F(1,8) = 30.517, p > 0.01; \eta_G^2 = 0.044$). No significant interaction between gap width and viewing condition was observed ($F(6,48) = 1.819, p > 0.05; \eta_G^2 = 0.025$). A repeated measures *t*-test revealed a significant difference between the PSEs for the valid and invalid viewing conditions ($t = -5.0374, p = 0.0010$). No significant difference between the JNDs for the valid and invalid viewing conditions was observed ($t = -1.914, p = 0.0919$).

a)



b)

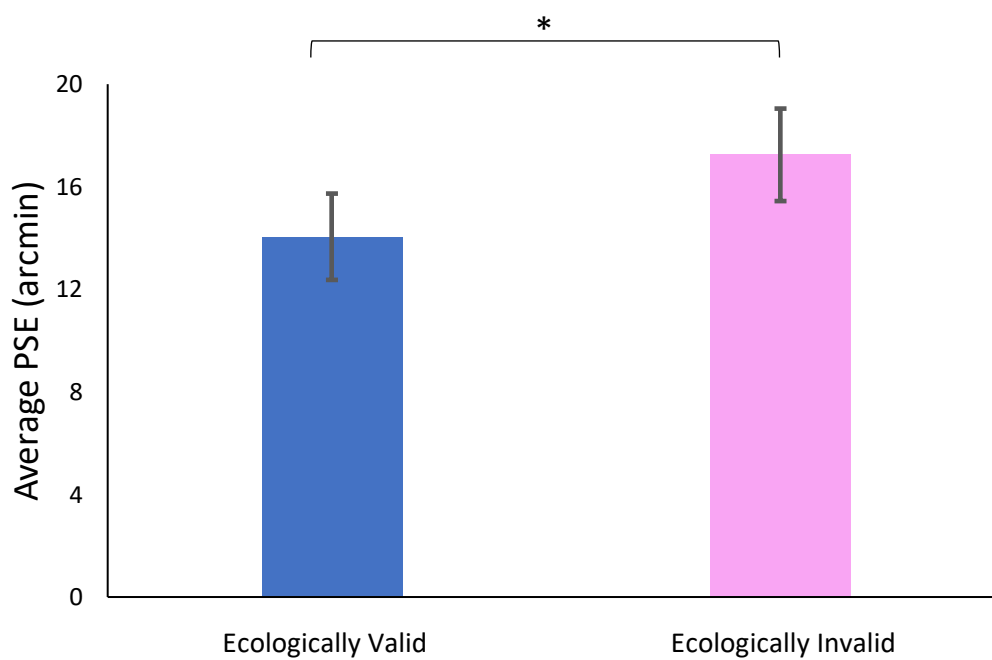


Figure 11: a) Psychometric function of representative observer fit with cumulative normal distribution function depicting the proportion of trials with multiple objects as a function of gap width. b) Average PSEs for ecologically valid and invalid viewing conditions. Error bars show standard error of the mean.

As predicted, larger gaps produce a percept of multiple objects for a greater proportion of trials. However, in the invalid viewing condition, when the texture extends outside the monocular occlusion zone, observers are more tolerant to the presence of gaps and more likely to continue to see a coherent object. This could be due to a type of perceptual suspension of disbelief, i.e. because the object is already ‘impossible’ the observers are more likely to accept the presence of breaks or gaps in the texture. Alternatively, in the valid viewing condition, as the gap is increased the texture within the monocular region is correspondingly reduced. Thus, there is less information to support the presence of a single object. In contrast, while the invalid conditions have the same gap widths, the size of the textured regions is constant so at the larger gap condition there is more texture, and the associated perspective information, available to support the percept of a common object. Considered with the lack of impact of texture properties on perceived depth in Experiment 1, these data suggest that the texture in the monocular regions is useful for the qualitative percept of object shape, but not as useful for the quantitative percept of depth. However, in Experiment 2, monocular texture properties showed greater impacts on quantitative depth when the object did not contain binocular disparity and monocular information was more heavily relied on as a consequence. The stimuli in Experiment 3 contain object shape information from binocular disparity and do not reflect Experiment 2’s conditions described above. Therefore, the nature of the texture within the binocular and monocular parts of the object does not seem to impact the amount of perceived depth, but it does impact the perceived

coherence of the object. This indicates that the size of the monocular region combined with object shape information drives depth estimates and texture information does not.

EXPERIMENT 4

In Experiment 3, we investigated the interaction between the validity of occlusion geometry and the perception of object coherence; the results showed that coherence was maintained across larger gaps in the invalid viewing conditions. This means that monocular sides were seen as forming part of a common object even though this solution was physically impossible. It is not clear from Experiment 3 whether the valid and invalid versions of the objects were seen to have the same 3D shape. Therefore, in Experiment 4 we assess perceived depth for these same stimuli. It has been shown that in traditional half-occlusion arrangements, the magnitude of perceived depth does not follow minimum depth constraint predictions when the stimuli are not ecologically valid (Nakayama & Shimojo, 1990). It should be noted that in those experiments the ‘invalid’ conditions were generated by exchanging the left and right eye’s views so that the monocular region was presented to the opposite eye. Arguably this is a more extreme manipulation than simply extending the monocular region outside the occlusion zone as we have done. However, we predict that if the visual system is sensitive to the ecological validity of self-occlusions then perceived depth should not increase with increasing monocular zone width in our stimuli.

Observers

Nine observers participated in Experiment 4. Five were male and four were female. Ages ranged from 20 to 38 years with an average age of 28.3 years. All had previous experience in

psychophysical experiments and six participated in Experiment 3. Observers had self-reported normal or corrected-to-normal visual acuity and stereoacuity of at least 40 seconds of arc as assessed using the RANDOT™ stereoacuity test.

Stimuli

The stimuli used in Experiment 4 were a subset of gap widths from Experiment 3 (see Figure 9) selected so that they covered a range of percepts from coherent to multiple objects based on the results of Experiment 3. Gap widths ranged from 6.18 to 30.90 arcmin.

As in Experiment 3, we manipulated the validity of the occlusions by varying the size of the monocular regions. In the valid conditions, the outer edges of the monocular regions remain in the same position regardless of gap width. As a result, the predicted depth from the minimum depth constraint, computed from the outer edge would also remain constant. In the invalid condition, the outer edges of the monocular portions shift outwards laterally as gap width increases. In this case, although not geometrically valid, if the line of site to the outermost edge of the monocular region is used to define the depth of the edge, we would predict increasing perceived depth as the gap width increases.

Procedure

The procedure used in Experiment 4 was the same as Experiment 1 and 2. On each trial, observers were asked to estimate the perceived depth of the object using an adjustable horizontal line positioned below the stimulus. The length of the line was adjusted to indicate the amount of perceived depth from the peak of the object to its furthest edge. Stimuli remained visible until the observer responded by pressing a button on the gamepad to move to the next trial. Gap widths of 6.18, 12.36, 18.54, 24.72, and 30.90 arcmin were each viewed ten times under ecologically valid

and invalid conditions. A total of ten conditions (two viewing conditions x five gap widths) were interleaved resulting in a total of 100 trials.

Results and Discussion

Figure 11 shows estimated depth as a function of gap width along with minimum depth constraint predictions for valid and invalid conditions. While the amount of depth varied across observers, the dependence on gap width did not. A two-way repeated measures ANOVA was used to analyze the results. No significant two-way interaction between the gap width and the viewing condition was observed, $F(4, 32) = 1.100$, $p > 0.05$; $\eta_G^2 = 0.001$ and no significant main effects of gap width ($F(4, 32) = 0.569$, $p > 0.05$; $\eta_G^2 = 0.002$) or viewing condition were observed ($F(1, 8) = 3.795$, $p > 0.05$; $\eta_G^2 = 0.009$). The lack of effect of gap width and viewing condition is likely driven by the constant size of the monocular region in the valid condition because the minimum depth constraint predicts a slope of zero and depth estimates should therefore be constant. It appears that the minimum depth constraint is not used to estimate depth in the invalid case, so no significant impact of viewing condition was observed. Paired *t*-tests (Bonferroni corrected) were used to compare the mean slope of estimated depth and the minimum depth constraint predictions. The analysis revealed that there was no difference between the valid condition and the predicted slope ($t(8) = -0.7581$, $p = 0.4701$). The average slope in the invalid condition also did not differ significantly from the prediction ($t(8) = 1.7881$, $p = 0.1116$), but it also did not differ significantly from the mean slope of the depth estimates in the valid condition ($t(8) = -1.8799$, $p = 0.0969$). Overall it appears that there was no effect of gap width on perceived depth in either configuration, this result is consistent with the minimum depth constraint predictions for the valid occlusion condition, but not for the invalid condition (see Figure 13).

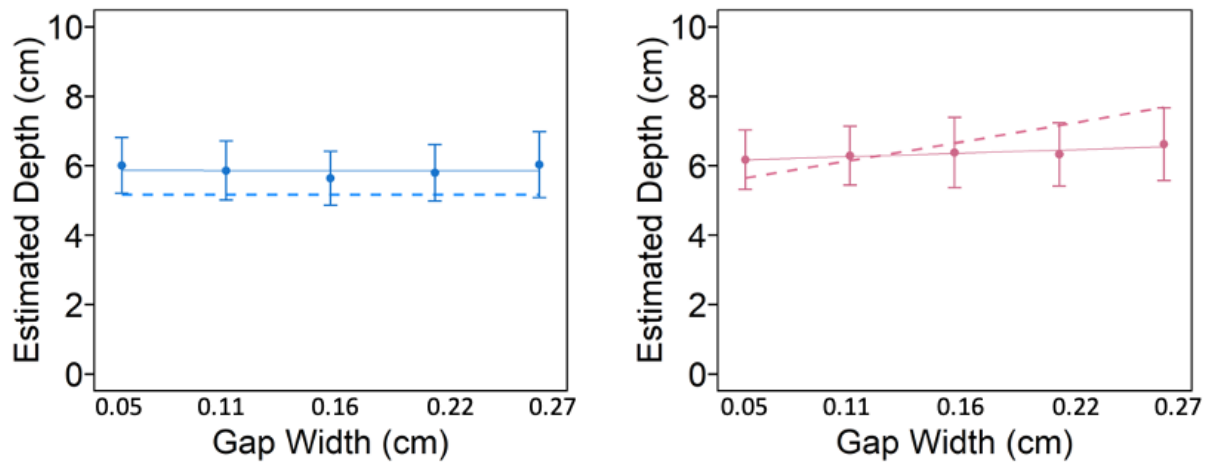


Figure 12: Estimated depth as a function of the gap width in the monocular region for valid (left) and invalid (right) conditions. Error bars represent standard error of the mean. Dashed lines show predictions from the minimum depth constraint – solid lines show average depth estimates.

Quantitative depth percepts from self-occlusions do not consistently follow model predictions based on binocular viewing geometry. When the occluded region is consistent with natural viewing geometry, we do see the expected zero slope but there is a consistent over-estimation of depth across all gap sizes. This is not likely due to the estimation method because unlike the overestimates in the valid condition, peak depth was generally underestimated in Experiment 1 and 2. It should be noted though that the minimum depth constraint does not limit the maximum possible amount of perceived depth when the depth is in the uncrossed direction, so such overestimates are not inconsistent with the geometric prediction (Tsirlin, Wilcox, & Allison, 2010).

The same pattern of estimates was seen in the valid and invalid viewing conditions despite the fact that the minimum depth constraint predicted a slope of 0.41 in the latter. It appears that when monocular region sizes exceed the constraints of natural viewing geometry

depth estimates do not increase as a function of gap width. Instead, depth estimates default to the depth perceived of the largest geometrically valid monocular region size.

Taken together the results of Experiments 3 and 4 show that there is substantial tolerance to violations of 3D viewing geometry when judging whether parts belong to the same object. It appears that the 3D shape of these objects, whether they are geometrically valid or not, is relatively stable. While the peak depth is overestimated relative to that predicted from the minimum depth constraint, it is constant across our gap manipulation, suggesting that the outer edges are used to define this value under valid viewing conditions. Under invalid conditions, depth does not increase as gap widths increase suggesting that invalid viewing geometry does not follow minimum depth constraint predictions despite the percept of a coherent object.

EXPERIMENT 5

The results of Experiment 4 demonstrated that depth estimates do not consistently follow minimum depth constraint predictions and that the amount of perceived depth is not tolerant to invalid viewing geometry as estimated depth does not increase with monocular region size when natural viewing is violated. As outlined in the Introduction (Figure 2), the minimum depth constraint is calculated from the 3D viewing geometry which includes the viewing distance, interocular separation, and the width of the occluding surface. The line of sight to the edge of the occluding surface defines the minimum depth of the outer edge of the monocular region. It is not clear from the results of Experiment 4 whether observers were relying on this minimum value to define the depth of the surface because in the valid condition it did not change. In Experiment 5, we used densely textured surfaces with no blank regions (gaps) and varied the monocular region

width for valid and invalid configurations to determine if the minimum depth constraint is used to define perceived depth under valid self-occlusion conditions.

Observers

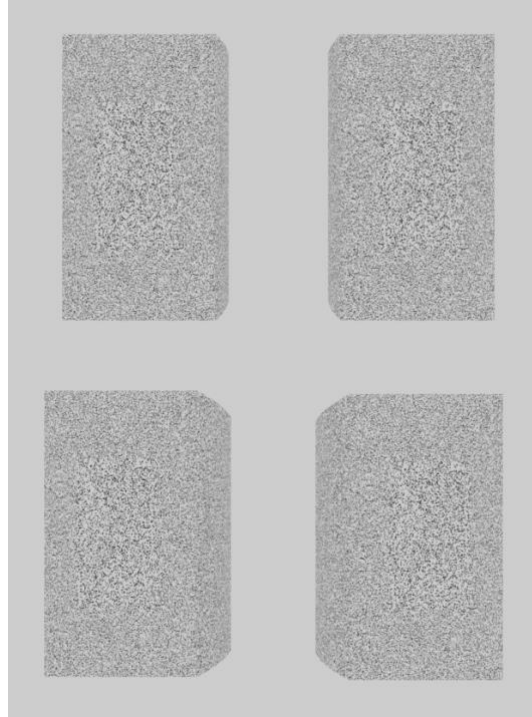
Eight observers participated in Experiment 5. Two were male and six were female. Ages ranged from 21 to 35 years with an average age of 28.9 years. Two were inexperienced with psychophysical experiments. Of the remaining six experienced observers, five participated in Experiment 4. Observers had self-reported normal or corrected-to-normal visual acuity and stereoacuity of at least 40 seconds of arc as assessed using the RANDOT™ stereoacuity test.

Stimuli

The stimuli were comprised of the same general truncated half-cylinder shape and structure used in Experiment 4, but two main modifications were introduced. First, the texture density of the random dot pattern was increased (55.2 cd/m^2) for greater texture consistency throughout binocular and monocular regions of the object. The second modification was the elimination of the gaps in the monocular regions (see Figure 12).

The truncated front face was positioned at a crossed disparity of 0.09 degrees. Ecologically valid monocular region sizes were 0, 0.17, 0.34, 0.51, 0.69, and 0.86 degrees in angular size. Ecologically invalid monocular region sizes were 1.03, 1.2, 1.4, and 1.55 degrees in angular size. Preliminary trials showed that invalid region sizes that exceeded 1.55 degrees in angular size produced binocular rivalry when fused. Ecologically invalid conditions were rendered using the same method as Experiment 4 where stereopairs only contain monocular sides.

a)



b)

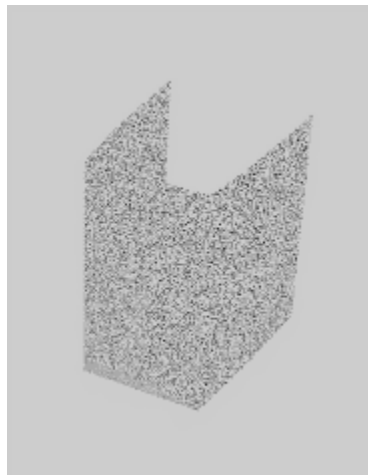


Figure 13: Examples depicting (a) ecologically valid (top) and invalid (bottom) self-occluded stimuli arranged for cross fusion and (b) side-view of stimulus.

Procedure

Previous studies examining depth from half-occlusions have used adjustable disparity probes and found that perceived depth closely followed minimum depth constraint predictions

(Nakayama & Shimojo, 1990; Tsirlin, Wilcox, & Allison, 2010). We used a similar method here to make this experiment more comparable to these previous studies. The disparity probe was a black dot that was 0.25 degrees in diameter. Stimuli were presented for 1 second followed by presentation of two dots: the probe dot and reference dot of the same size. The reference dot had the same disparity as the peak of the stimulus, which was 0.09 degrees, and was positioned in the middle of the display above the adjustable probe. Observers adjusted the disparity of the probe relative to the reference dot using a gamepad to indicate the perceived depth between the peak and the furthest edge. A total of 10 conditions (6 ecologically valid monocular region sizes and 4 invalid region sizes) were interleaved and repeated 10 times apiece resulting in a total of 100 trials (60 trials for valid and 40 trials for invalid).

Results and Discussion

Figure 13 shows estimated depth as a function of the monocular region size along with minimum depth constraint predictions for valid and invalid conditions. There is a change in perceived depth with monocular region size in the valid conditions, but estimates were variable and did not follow minimum depth constraint predictions closely. Similar to Experiment 5, in the invalid conditions, there is no change in perceived depth with increasing monocular region size.

A linear mixed model was used to analyze the results because there was an unequal number of levels in the valid and invalid conditions. Six of the monocular region sizes were valid and four were invalid, which was due to the binocular rivalry when the monocular region became too large as described above. Monocular region size and viewing condition were treated as fixed effects and the observer was considered a random factor. There was a significant effect of viewing condition, $F(1,8) = 41.4505$, $p < 0.05$, but no significant effect of monocular region size was observed, $F(1,8) = 3.9249$, $p > 0.05$. Paired *t*-tests with Bonferroni correction comparing

mean slopes between estimated depth and predicted depth in the valid condition revealed that the average slope of depth estimates differed significantly from the predicted depth slope ($t(7)=-5.2141$, $p=0.0012$). The mean slope of estimated depth in the invalid condition also differed significantly from the predicted depth slope ($t(7)=-7.0253$, $p=0.0002$). Additionally, the mean slope of the valid condition differed significantly from a slope of zero ($t(7)=3.2521$, $p=0.0140$) while the mean slope of the invalid condition was not significantly different from zero ($t(7)=0.5618$, $p=0.5918$).

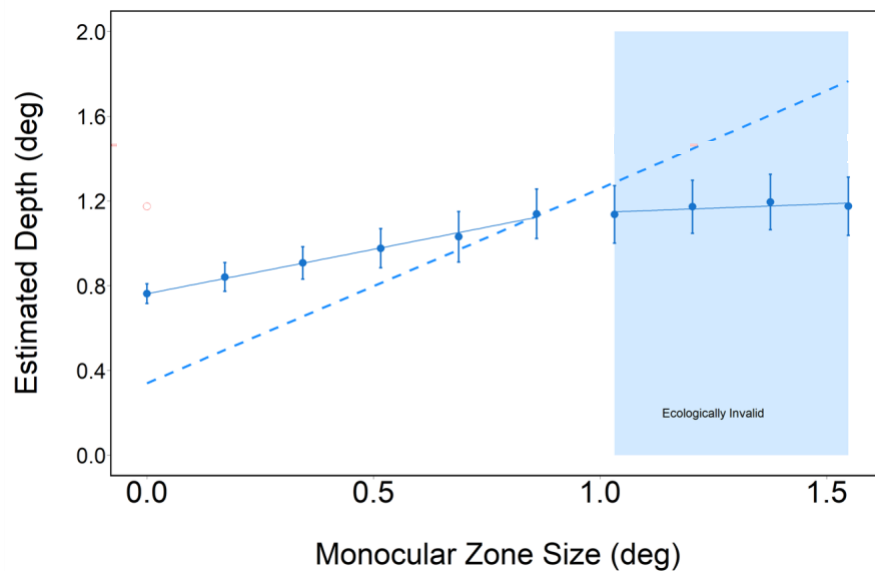


Figure 14: Mean estimated depth as a function of monocular zone size depicting ecologically valid and invalid (shaded in blue) conditions. Dotted blue line represents the minimum depth constraint prediction. Error bars represent the standard error of the mean.

When stimuli are ecologically valid, perceived depth increases as a function of monocular region size but estimated depth does not follow minimum depth constraint predictions. The slope of estimated depth in the valid condition is shallower and depth is

overestimated at most monocular region sizes. While in the invalid condition it appears that the estimates plateau when the monocular regions extend into the zone where they should be visible to the other eye, the difference between the slopes was not statistically significant. This is likely due to a high level of between observer variability. These results are similar to Experiment 4 in that depth is generally overestimated in the valid condition. However, overestimates become less extreme as the monocular region sizes increase and approach ecologically invalid conditions. This suggests that the monocular region sizes are not the only stimulus property used to perceive depth. Object shape from disparity is likely aiding depth estimates in conjunction with monocular region sizes as described in Experiment 2.

GENERAL DISCUSSION

The series of experiments examined here evaluated the impact of monocular self-occlusions on depth perception. This research was motivated by previous work that thoroughly investigated depth from two object half-occlusion arrangements. There has been less empirical attention given to *self-occlusions* and whether depth from these is perceived in the same manner as half-occlusions.

In Experiment 1, the impact of texture differences between the binocular and monocular regions of self-occluded half-cylinders was examined. The texture manipulations were motivated by previous work regarding two-object half occlusion arrangements where monocular region textures were consistent or inconsistent with the occluding surface (Gillam & Borsting, 1988; Grove & Ono, 1999; Grove, Gillam, & Ono 2002). It was predicted that if monocular texture differences in self-occlusions impact perceived depth the same way as half-occlusions, then

depth should be underestimated when the monocular texture is inconsistent with that of the binocular region's texture (Grove, Gillam, & Ono, 2002). Surprisingly, depth estimates did not vary substantially across monocular texture manipulations even in the presence of large self-occluded monocular regions. There was concern that observers could potentially extrapolate the volumetric shape of the self-occluded half-cylinders from the surface peak and use the shape information to make their estimates. It has been shown that volumetric shape perception of curved objects such as spheres and cylinders occurs even when occluding contours are present from self-occlusions (Watanabe & Idesawa, 2001).

In Experiment 2, the role of the familiarity of the 3D shape was evaluated by modifying the amount of curvature at the surface peak to make it an uninformative cue for object shape and to reduce the amount of disparity to extrapolate shape from. The surface peak was truncated at three different locations along the front face of the object, which produced decreasing binocular disparity and increasing monocular region sizes as the front surface width increased. Each front surface width and texture density were tested at three peak depths. At the largest front surface width there was little to no binocular disparity information as the sides of the object were completely monocular. It was predicted that the reduction of familiar object shape information would result in depth underestimates as the monocular texture density increased. The results showed that as the amount of binocular disparity decreased in the object observers perceived less depth, particularly at the highest texture density showing patterns more consistent with studies regarding two object arrangements (Grove, Gillam, & Ono, 2002). Underestimations were notably lower when stimuli contained the highest texture density with no binocular disparity as the texture was too fine and did not show reliable perspective information. This suggests that depth is perceived from monocular texture information when disparity is absent and if the

monocular texture is unreliable, depth estimates are degraded. In addition to this, the results for the largest front surface width at the highest texture density showed a trend where depth estimates decreased as the peak depth increased. While the results may seem surprising, there were small amounts of binocular disparity along the side regions of the object at shallower peak depths and no binocular disparity at the largest peak depth. The presence of binocular disparity aids in perceiving object shape and depth (Rogers & Collett, 1989).

The results of Experiment 2 showed significant depth underestimates when the stimuli contained the smallest amount of binocular disparity along the side regions at the highest texture density. If the stimulus contained side regions that were completely monocular with texture properties inconsistent with the binocular front face, the stimulus could have been perceived as a more distant surface containing a different texture being occluded by a nearer surface. In Experiment 3, we assessed the possibility that observers did not perceive this stimulus configuration as a coherent object by introducing varying gap widths in the monocular regions under geometrically valid and invalid viewing conditions. Observers indicated whether they perceived a single coherent object or multiple objects depending on the size of the gap width in the monocular region. The results showed that larger monocular gap widths produced a percept of multiple objects for a greater proportion of trials, but there was a surprising difference between ecologically valid and invalid viewing conditions where a greater gap width was needed for ecologically invalid stimuli to be perceived as multiple objects. This showed a surprising tolerance to invalid viewing geometry in self-occlusions for qualitative percepts. The invalid conditions used in this experiment (see Figure 9) differ from that used by Nakayama and Shimojo (1990) in that the size of the monocular region made the object invalid instead of reversing the sign of the occluded region's depth. It is also important to note that the invalidity of

the stimuli used by Nakayama and Shimojo (1990) has been shown to be geometrically possible as there are rare cases where such occlusion arrangements can exist, so one could argue that those stimuli are ecologically valid (Assee & Qian, 2007; Wilcox & Harris, 2009).

The tolerance to invalid self-occlusion arrangements described in Experiment 3 was observed for qualitative depth percepts, but it is unclear whether this the amount of perceived depth will increase with monocular region size when natural viewing geometry is violated. In Experiment 4, quantitative depth from self-occlusions was investigated under valid and invalid viewing conditions using a subset of stimuli from Experiment 3. Application of the minimum depth constraint to these stimuli predicted that depth estimates would remain constant in the valid condition as the monocular region size remained constant and depth estimates would increase with a slope of 0.41 in the invalid condition. The results revealed that depth estimates do not consistently follow minimum depth constraint predictions. Although the slope was consistent with the predicted valid slope, depth was overestimated in the valid condition. The estimates in the invalid condition displayed similar trends and did not follow model predictions. Previous evidence showing that quantitative depth is sometimes inconsistent with minimum depth constraint predictions have typically referred to binocular matching of textured edges between the occluder and edge of the monocular region (Gillam et al., 2003; Liu et al., 1997; Tsirlin, Wilcox, & Allison, 2010; Tsirlin, Wilcox, & Allison, 2011). However, the stimuli used in Experiment 4 do not contain strongly defined textured edges that could be matched to the occluding portion of the object. It is also possible that in the invalid test conditions observers perceived distortions in the 3D shape of the stimulus to compensate for the violation of the viewing geometry. However, this seems unlikely given the consistency of the depth estimates in those conditions.

Minimum depth constraint predictions in Experiment 4 for valid occlusion conditions showed no increase in predicted depth across all gap widths because the size of the monocular regions in valid conditions was always the same regardless of the gap width. Therefore, it was difficult to thoroughly evaluate whether an increase in monocular region sizes would also cause increases in perceived depth as predicted by the minimum depth constraint. In Experiment 5, gap widths were replaced with texture information that was consistent throughout the whole object and monocular region sizes increased linearly for both ecologically valid and invalid viewing conditions. For geometrically valid stimuli we predicted that perceived depth would increase as a function of the monocular region size. While perceived depth did increase with monocular region size, the slope and y intercept values were inconsistent with the predicted values. Depth was overestimated in all valid conditions apart from the largest monocular region size. In the invalid viewing conditions, depth estimates again did not increase as a function of the monocular region size. Estimated depth was consistent with that of the largest valid monocular region size. As outlined above, these invalid stimuli may have been interpreted as valid objects with a different shape. Again, the consistency of the depth estimates argues against this.

Self-Occlusions and Half-Occlusions

As outlined above, the experiments presented here suggest that depth perception in self-occlusions is processed differently from half-occlusions. At first this may be surprising given that previous investigations of depth from two-object half-occlusions show that depth estimates closely follow minimum depth constraint predictions, but there are important features specific to traditional half-occlusion arrangements that distinguish them from self-occlusions (Nakayama & Shimojo, 1990; Tsirlin, Wilcox, & Allison, 2010). The stimuli used in previous studies were composed of near and far fronto-parallel planes that did not contain volumetric shape

information from disparity. Monocular textures in these fronto-parallel surfaces also contained no texture gradient information (see Figure 14-b). Therefore, the size of the monocular region that is defined by the outer edge of the occluded surface is the only information available to support depth judgements.

However, volumetric shape from binocular disparity is often present in self-occlusions in the foreground portion of the object. When judging the amount of depth within a single object, there is opportunity for the amount of binocular disparity to impact the reliance on the minimum depth constraint. That is, when there is sufficient binocular disparity present in the foreground portion of the object a clearer percept of object shape is obtained, and the monocular regions appear as a continuation of the perceived shape. It is possible that depth overestimates in Experiment 5 were due to this 3D shape information as the binocular portion of the stimulus resembled the 3D structure of a truncated pyramid. This 3D structure could have been used to estimate the amount of depth instead of only using the size of the monocular regions. In addition, it is also important to consider the volumetric shape of the monocularly occluded regions. For example, if a self-occluded object is curved, the texture gradient in the monocular regions will correctly reflect the shape of the object by scaling with the curvature and distance of the surface (see Figure 14-a). It has been shown that perceived depth increases when viewing curved stereoscopic surfaces that contain valid texture cues versus a curved surface with texture cues consistent with a fronto-parallel plane (Johnston, Cumming, & Parker, 1993). Johnston, Cumming, and Parker (1993) used stimuli that had no monocularly occluded regions and still reported an effect of texture on perceived depth. Thus it might be reasonable to expect that texture cues would play an even more important role in perceiving depth from self-occlusions because only texture cues are available in the monocular region. In Experiment 2, 3D object

shape was reduced and there was some evidence that depth estimates were greater when the monocular texture information was consistent with texture of the front surface. Similarly when texture manipulations have been introduced in half-occlusions, underestimates occurred when the monocular texture was inconsistent with that of either the occluding surface or the background (Grove, Gillam, & Ono, 2002).

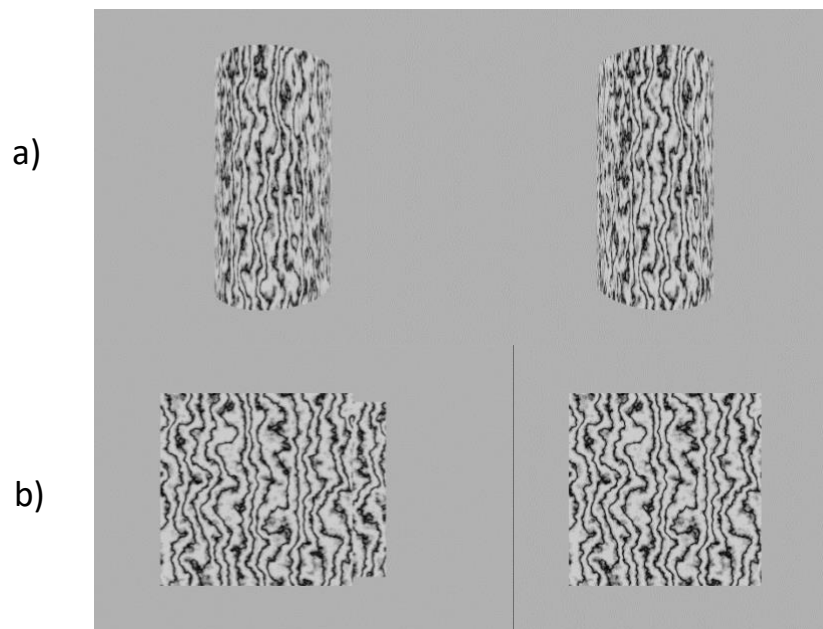


Figure 15: Examples of (a) curved self-occluded stimulus and (b) half-occluded stimulus arranged for cross fusion.

Another potential consideration for self-occlusions are the sizes of the monocular regions. Depending on how steep the peak of the surface is and the depth of the object, the stimuli in the experiments described here contained large monocular regions relative to the sizes of half-occlusions reported in previous literature. For example, in Tsirlin et al's (2010) experiments the largest monocular region size was approximately 0.22 degrees in angular size. Four of the five monocular self-occlusions in the valid condition were wider than this (0.34 to 0.86 degrees in angular size). While it is possible that a stronger dependence on the minimum depth constraint

may be seen for smaller occlusion regions, this seems unlikely given the large overestimates and shallow slopes seen in Figure 13 over a similar range of sizes (0.17 to 0.34 degrees in angular size).

Self-Occlusions in Natural Scenes

Monocular regions that arise from occlusions commonly occur in real-world viewing environments and natural scenes. It has been shown that when a scene is populated with multiple small-scale objects, more monocular regions occur from half-occlusions than self-occlusions (Langer & Mannam, 2012; Başgöze, White, Burge, & Cooper, 2020). Consistent with this, monocular regions in natural scenes were found to contain features such as texture and shading that were more visually consistent with background elements than to neighbouring foreground elements (Başgöze et al., 2020). This demonstrates that when monocular regions arise from occlusion, the foreground elements typically possess different texture and shading properties from the background's occluded regions as would be expected from two-object half-occlusions.

The lower prevalence of self-occlusions in natural scenes and real-world environments could explain why perceived depth differs for these stimuli compared to separate object half-occlusions. That is, the visual system may simply rely on the more commonly experienced 3D shape information to interpret the depth of such self-occluded 3D objects. This proposal is consistent with the results seen in Experiment 1 where we found that disparity from the curvature at the peak of the object provided reliable shape information and there was no effect of the monocular texture.

Future Work

To further investigate the role monocular texture in self-occlusions, linear perspective cues will be specifically examined in the monocular regions. Researchers have shown that when both stereoscopic and perspective information is present, there is little impact of the monocular texture on the amount of perceived depth (Dobias & Papathomas, 2014). However, the perspective information and texture foreshortening in stimuli may have greater impact in a self-occlusion as it can be completely monocular depending on the object shape. For example, strong texture gradient information is present along the curved sides of an object in both the binocular and monocular regions (see Figure 14-a). Monocular texture manipulations greatly impacted perceived depth in Experiment 2 when the stimuli did not contain binocular disparity information, so it is possible that reversing correctly mapped 2D perspective in the texture will also affect perceived depth. Correctly rendered and reversed perspective will be presented in the monocular regions and their impact on perceived depth will be compared in additional experiments. If perceived depth in self-occlusions is reliant on depth from 2D perspective, then unreliable texture such as reversed perspective should degrade depth estimates when compared to correctly rendered perspective.

In another line of experimentation, we will evaluate the role of 2D texture cues in the interpretation of half-occlusions by using images of natural objects and scenes. Such natural objects can contain varying object materials that produce monocular shading cues such as specular highlights and shadows. This work will build on existing evidence showing shorter reaction times for correct versus incorrect depth ordering of self-occlusions in natural scenes (Wilcox & Lakra, 2007). We will assess the impact of such occlusions in natural scenes on the amount of perceived depth in these objects.

Conclusion

In this thesis I have demonstrated that monocular *self-occlusions* impact depth percepts differently than two-object half-occlusions. Unlike half-occlusions, monocular texture differences do not impact depth estimates when the object contains binocular shape information. Additionally, there is an unexpected tolerance to invalid viewing geometry involving qualitative percepts from self-occlusions, but this tolerance depends critically on the context in which they are viewed because quantitative depth percepts do not increase as a function of monocular region size when the stimuli violate 3D viewing geometry. Stereoscopic shape information plays an important role in the degree to which monocular texture properties are used to estimate perceived depth. Monocular texture cues that reflect shape information from self-occluded objects are only used when binocular disparity is absent and 3D shape is ambiguous. Further, unlike half-occlusions, depth estimates from self-occlusions are not accurately predicted by the minimum depth constraint. These differences are important because they emphasize that two-object arrangement half-occlusions and self-occlusions are separate phenomena that cannot both be explained by the same set of constraints.

REFERENCES

- Assee, A., & Qian, N. (2007). Solving da Vinci stereopsis with depth edge-selective V2 cells. *Vision Research*, 47, 2585-2602.
- Badcock, D. R., & Schor, C., M. (1985). Depth-increment detection function for individual spatial channels. *Journal of the Optical Society of America, A*, 2(7), 1211-1216.
- Başgöze, Z., White, D.N., Burge, J., & Cooper, E.A. (2020). Natural statistics of depth edges modulate perceptual stability. *Journal of Vision*, 20(8), 1-21.
- Blakemore, C. (1970). The range and scope of binocular depth discrimination in man. *The Journal of Physiology*, 211(3), 599.
- Brainard, D. H. (1997) The Psychophysics Toolbox. *Spatial Vision*, 10, 433-436.
- Christophers, R., & Rogers, B. (1994). *The effect of viewing distance on the perception of complex random-dot stereograms*. Paper presented at the Investigative ophthalmology & visual science.
- Coutant, B. E., & Westheimer, G. (1993). Population distribution of stereoscopic ability. *Ophthalmic and Physiological Optics*, 13(1), 3-7.
- Da Vinci, L. (1877). *A treatise on painting (English translation)*. George Bell and Sons. 1.1.
- Dobias, J. & Papathomas, T.V. (2014). The role of linear perspective cues on perceived depth magnitude of the surfaces they are painted on: Proper-, reverse-, and flat-perspective paintings. *Perception*, 43(9), 989-1000.
- Forte, J., Peirce, J. W., & Lennie, P. (2002). Binocular integration of partially occluded surfaces. *Vision Research*, 42(10), 1225–1235.
- Gillam, B. & Borsting, E. (1988). The role of monocular regions in stereoscopic displays. *Perception*, 17(5), 603-608.
- 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.

- Grove, P.M. & Ono, H. (1999). Ecologically invalid monocular texture leads to longer perceptual latencies in random-dot stereograms. *Perception*, 28(5), 627-639.
- Grove, P.M., Gillam, B.J., & 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*, 49(22), 2666-2685.
- Hartle, B., & Wilcox, L. M. (2016). Depth magnitude from stereopsis: Assessment techniques and the role of experience. *Vision Research*, 125, 64-75.
- Howard, I. P., & Rogers, B. J. (1995). *Binocular vision and stereopsis*. New York, NY: Oxford University Press.
- Howard, I. P., & Rogers, B. J. (2012). *Perceiving in Depth. Volume 2: Stereoscopic Vision*. New York, NY: Oxford University Press Inc.
- Johnston, E.B., Cumming, B.G., & Parker, A.J. (1993). Integration of depth modules: stereopsis and texture. *Vision Research*, 33(5/6), 813-826
- Langer, M.S. & Mannan F. (2012). Visibility in three-dimensional cluttered scenes. *Optical Society of America*, 29(9), 1794-1807.
- Lawson, R.B. & Gulick, W.L. (1967). Stereopsis and anomalous contour. *Vision Research*, 7, 271-297.
- Liu, L., Stevenson, S. B., and Schor, C. M. (1997). Binocular matching of dissimilar features in phantom stereopsis. *Vision Research*, 37(5):633-644.
- Nakayama, K. & Shimojo, S. (1990). Da vinci stereopsis: depth and subjective occluding contours from unpaired image points. *Vision Research*, 30(11), 1811-1825.
- Ogle, K. N. (1952). On the limits of stereoscopic vision. *Journal of Experimental Psychology*, 44(4), 253.
- Ogle K. N, & Weil M. P. (1958). Stereoscopic vision and the duration of the stimulus. *A.M.A. Archives of Ophthalmology*, 59, 4–17.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming

- numbers into movies. *Spatial Vision*, 10(4), 437-442.
- Rogers, B.J., & Collett, T.S. (1989). The appearances of surfaces specified by motion parallax and binocular disparity. *The Quarterly Journal of Experimental Psychology*, 41(4), 697-717.
- Schor, C. M., & Wood, I. (1983). Disparity range for local stereopsis as a function of luminance spatial frequency. *Vision Research*, 23(12), 1649–1654.
- Schor, C., Wood, I., & Ogawa, J. (1984). Binocular sensory fusion is limited by spatial resolution. *Vision Research*, 24(7), 661-665.
- Siderov, J., & Harwerth, R., S. (1995). Stereopsis, spatial frequency, and retinal eccentricity. *Vision Research*, 35(16), 2329-2337.
- Tsirlin, I., Wilcox, L.M., & Allison, R.S. (2010). Monocular occlusions determine the perceived shape and depth of occluding surfaces. *Journal of Vision*, 10(6), 1-12.
- Tsirlin, I., Wilcox, L.M., & Allison, R.S. (2011). Disparity biasing in depth from monocular occlusions. *Vision Research*, 51(14), 1699-1711.
- Watanabe, O., & Idesawa, M. (2001). Effect of an occluding contour on the volume perception in binocular stereopsis. *Japanese Journal of Applied Physics*, 40, 958-960.
- Westheimer, G. (1979). The spatial sense of the eye. Proctor lecture. *Investigative Ophthalmology & Visual Science*, 18, 893-912.
- Westheimer, G., & McKee, S. P. (1979). What prior uniocular processing is necessary for stereopsis?. *Investigative Ophthalmology & Visual Science*, 18, 614-621.
- Wheatstone, C. (1838). Contributions to the physiology of vision - Part the first. On some remarkable, and hitherto unobserved, phenomena of binocular vision. *Philosophical Transactions of the Royal Society of London*, 128, 371-394.
- Wilcox, L.M. & Lakra, D.C. (2007). Depth from binocular half-occlusions in stereoscopic images of natural scenes. *Perception*, 36(6), 830-839.