# THE IMPACT OF 2-D AND 3-D GROUPING CUES ON DEPTH FROM BINOCULAR DISPARITY

LESLEY M. DEAS

A DISSERTATION SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN PSYCHOLOGY
YORK UNIVERSITY
TORONTO, ONTARIO

SEPTEMBER 2015

# ABSTRACT

Stereopsis is a powerful source of information about the relative depth of objects in the world. In isolation, humans can see depth from binocular disparity without any other depth cues. However, many different stimulus properties can dramatically influence the depth we perceive. For example, there is an abundance of research showing that the configuration of a stimulus can impact the percept of depth, in some cases diminishing the amount of depth experience. Much of the previous research has focused on discrimination thresholds; in one example, stereoacuity for a pair of vertical lines was shown to be markedly reduced when these lines were connected to form a rectangle apparently slanted in depth (eg: McKee, 1983). The contribution of Gestalt figural grouping to this phenomenon has not been studied.

This dissertation addresses the role that perceptual grouping plays in the recovery of suprathreshold depth from disparity. First, I measured the impact of perceptual closure on depth magnitude. Observers estimated the separation in depth of a pair of vertical lines as the amount of perceptual closure was varied. In a series of experiments, I characterized the 2-D and 3-D properties that contribute to 3-D closure and the estimates of apparent depth. Estimates of perceived depth were highly correlated to the strength of subjective closure. Furthermore, I highlighted the perceptual consequences (both costs and benefits) of a new disparity-based grouping cue that interacts with perceived closure, which I call ‘good stereoscopic continuation’. This cue was shown to promote detection in a visual search task but reduces depth percepts compared to isolated features.

Taken together, the results reported here show that specific 2-D *and* 3-D grouping constraints are required to promote recovery of a 3-D object. As a consequence, quantitative depth is reduced, but the object is rapidly detected in a visual search task. I propose that these phenomena are the result of object-based disparity smoothing operations that enhance object cohesion.

## ACKNOWLEDGEMENTS

I want to thank my Supervisor, Laurie Wilcox, for everything over the past four years. I have learned a great deal from you, and your support and kindness beyond the lab has meant the world to me. You are a wonderful colleague and friend. To any students reading this, you are very lucky to work with her.

I gained constructive and valuable insight from the many colleagues that I discussed this project with. In particular, I wish to thank Robert Allison, Richard Murray and James Elder for their thoughtful comments as this work progressed.

I want to thank my colleagues and lab-mates. Above all, Inna Tsirlin and Arthur Lugtigheid were my mentors and great friends. Their support continues to be invaluable. Hopefully we will all meet again soon, somewhere in the world. Bob Hou has been there from Day One, and I am grateful for all his help. To my current lab-mates, you are a smart bunch and I wish you the best of luck in the future.

This dissertation is dedicated to my family, to whom I owe everything. Without their tremendous support and inspiration, I never would have had the courage to get on that plane. Mum, you never have to explain ‘stereopsis’ again. Dad, my wee phone calls always make me smile. Susan, thanks for being my cheerleader. And to Mike, for always believing in me.

## TABLE OF CONTENTS

# LIST OF FIGURES

# CHAPTER 1

## 1.1 Objectives overview

It is well established that stimulus configuration can have a substantial impact on stereopsis. Many of these influences are believed to be due to relatively low-level processing constraints. One strong influence is the position and disparity of neighbouring features, for instance ‘disparity pooling’ (eg, Westheimer, 1986; Parker & Yang, 1989; Vreven, McKee & Verghese, 2002), ‘attraction and repulsion’ (Westheimer & Levi, 1987) or ‘disparity contrast’ effects (eg, Anstis, Howard & Rogers, 1978; Graham & Rogers, 1982; Westheimer, 1986; Kumar & Glaser, 1991). Another example of the effect of nearby features on perceived depth is the reported insensitivity of the stereoscopic system to smooth gradients of horizontal disparity (eg: Gillam, 1968; Wallach & Bacon, 1976; Rogers & Graham, 1983; Gillam, Flagg & Finlay, 1984; Stevens & Brookes, 1988; Mitchison & McKee, 1990; Mitchison & Westheimer, 1990). However, there is evidence that in addition to these low-level interactions, mid-level perceptual organization can also have a dramatic effect on depth perception. For example, stereoacuity for a pair of vertical lines can be dramatically reduced when connecting lines are added to form a closed object, and can become unmeasurable for some observers (Westheimer, 1979; McKee, 1983; Mitchison & Westheimer, 1984; Kumar & Glaser; 2002; Zalevski, Henning & Hill, 2007). It appears that when the vertical lines are perceived as part of a unified object, the stereoscopic system is unable to extract the disparity information with the same precision. This occurs in spite of the fact that the local disparity signal from the vertical contours has not changed in the two configurations. Such effects cannot be explained solely in terms of disparity interactions; rather the perceptual interpretation of stimulus arrangements appears to modulate perceived depth. This implicates higher-level processes that are involved in the recovery of global shape representation, which likely implement classic Gestalt grouping principles. While there have been some attempts to explore the impact of specific 2-D organizational cues on stereoacuity, or the minimum resolvable disparity difference, (Liu, Jacobs & Basri, 1999; Yin, Kellman & Shipley, 2000; Lu, Tjan & Liu, 2006) there has been no systematic effort to evaluate the impact of perceptual organization on suprathreshold depth perception within a single figure. To this end, I extend the threshold elevation experiments reported by McKee (1983) to suprathreshold magnitude estimates and explicitly manipulate Gestalt grouping cues to determine their role in the reduction of perceived depth in these configurations.

## 1.2 Background

### 1.2.1. Measuring Stereopsis

Stereoscopic vision refers to the three-dimensional percept obtained when two-dimensional retinal images are combined. These images are slightly offset on the left and right retinas because the eyes are horizontally separated. The small geometric difference between these images is known as retinal disparity, and serves as a binocular cue for the visual system to compute the depth of objects relative to fixation. Disparity has two types of signal: absolute and relative disparity. Absolute disparity is the difference in the two retinal locations corresponding to a single point and provides a depth estimate relative to where the observer is fixating (Marr, 1985). Relative disparity, on the other hand, is independent of fixation and is represented by the difference between the absolute disparities of two objects. It has been shown that, compared to absolute disparity, relative disparity is a stronger cue and provides more precise depth judgments (Westheimer, 1979).

Stereoacuity is a measure of the smallest binocular disparity that can reliably be discriminated. Under ideal conditions, the human eye is able to discriminate relative disparity with remarkable precision. Discrimination thresholds for a pair of vertical lines can be as low as 2 to 6 seconds of arc for practiced observers, but is highly variable in the general population (Ogle, 1953; Blakemore, 1970; Westheimer, 1979; Westheimer & McKee, 1979; McKee, 1983; Badcock & Schor, 1985; McKee, Levi, & Bowne, 1990; Kumar & Glaser, 1992; Andrews, Glennerster & Parker, 2001). Ideal conditions are simple stimuli of high contrast and sharp edges, which are viewed at an optimal size and separation. However, many variables dramatically influence stereoacuity including contrast, spatial frequency/size, modulation frequency, stimulus duration and configuration.

Though thresholds provide valuable information regarding the minimum detectable disparity signal, measuring the detectability of a visual stimulus does not often provide insight into the magnitudes of the evoked sensations. Given that most of our everyday visual processing involves objects and features that are easily detected, it is also important to understand how we experience suprathreshold sensory stimuli. Research has shown that differences in magnitudes of sensory experiences cannot reliably be predicted from threshold data (eg: Medjbeur & Tulunay-Keesey, 1986; Pattanaik et al. 1998). In depth perception, the relationship between stereoscopic thresholds and suprathreshold depth percepts is not well understood: although it is well established that thresholds for stereoscopic stimuli are elevated substantially under some conditions (eg, low luminance, reduced contrast, image blur, context and configuration), it is much less clear how these conditions affect the suprathreshold appearance of these stimuli. Research has suggested that not all conditions that increase the stereoscopic threshold result in reduced depth of targets with suprathreshold disparity (Patel, et al., 2009; Bedell, Gantz & Jackson, 2012). In an effort to better understand the impact of configuration on the perception of depth, this research focusses upon the suprathreshold perceived magnitude of depth from disparity.

The development of psychophysical sensory scaling techniques provided a means to assess suprathreshold properties of the human visual system and allowed the quantification of more everyday sensations. These techniques were first used by Richardson and Ross (1930), and were elaborated by S.S. Stevens (1953, 1955, 1975) based on the work of Fechner and Weber (for review see Marks & Gescheider, 2002). Suprathreshold depth from disparity has been assessed in a number of ways, all of which rely on some internal transformation of the perceived depth from disparity to a value that can be quantified. For example, in a matching task, observers estimate the magnitude of depth they perceive by giving a haptic response (a cross-modal task) or adjusting the magnitude of an on-screen reference, such as a virtual ruler (an intra-modal task). It is well known that the perception of suprathreshold depth based on stereopsis alone is typically not veridical (Foley, 1980; Bülthoff, Fahle, Wegmann, 1991; Johnston, Cumming & Parker, 1993; Johnston, Cumming & Landy, 1994; Volcic et al;. 2013). In general, there is a viewing distance where depth perception is veridical, averaging at 80cm or the length of the observer’s arm (Johnston, 1991; Volcic et al, 2013). At distances beyond arm’s reach, depth is underestimated but at shorter distances, depth is typically overestimated (Foley, 1980; Richards, 2009; Volcic et al, 2013).

### 1.2.2. Configuration and Stereopsis

Perceived three-dimensional shape is determined by the way individual stimulus components are arranged in depth. However, a number of studies have shown that stereoacuity can be degraded as parts are grouped to form objects (eg, Westheimer, 1979; McKee, 1983; Mitchison & Westheimer, 1984; Fahle & Westheimer, 1988; Kumar & Glaser, 1992; Liu, et al., 1999; Yin, et al., 2000; Vreven et al. 2002; Lu, et al., 2006; Zalevski, et al., 2007). These do not need to be complex shapes, as even simple shape components are prone to configural effects. As noted above, a compelling example of the impact of configuration on stereopsis was demonstrated by McKee (1983). As part of her assessment of the spatial requirements for fine stereoacuity, McKee (1983) measured discrimination thresholds for two parallel vertical lines (depicted in Figure 1.1A). Excellent stereoacuity was achieved for the isolated lines but was substantially reduced when the lines were connected by two horizontal lines to form a rectangle, apparently slanted in depth (see also Westheimer, 1979; Mitchison & Westheimer, 1984; Kumar & Glaser, 2002, Zalevski, et al. 2007). Figure 1.1A reproduces the data from McKee's (1983) study, and shows how thresholds for the same pair of vertical lines increased with various degrees of figural connectivity and/or closure. Thresholds were highest for a fully-closed figure but changing the position of the horizontal connecting lines to form a ladder configuration also degraded sensitivity. The negative effects on thresholds occurred even though the same disparity signal was always present in the vertical lines. Subsequent studies showed that thresholds were also disrupted when the connecting lines contained gaps (see Figure 1.1B) (Mitchison & Westheimer, 1984). Moreover, Mitchison & Westheimer (1984) reported that stereoacuity for two columns of dots were markedly reduced when identical columns flanked the target pair at equal lateral separations to form a slanted plane. These columns provided additional binocular disparity information which could have aided depth discrimination, yet discrimination was degraded. In another study, Fahle and Westheimer (1988) showed that a single horizontal contour of dots was also susceptible to configural influences. They measured stereoacuity between a pair of small dots and found that the addition of one dot between the target pair was sufficient to increase discrimination thresholds. Further, thresholds increased systematically as more dots were added to form a linear disparity gradient. Taken together, the studies described above support the conclusion that disparity sensitivity is reduced once the previously isolated elements are integrated to form a common object.

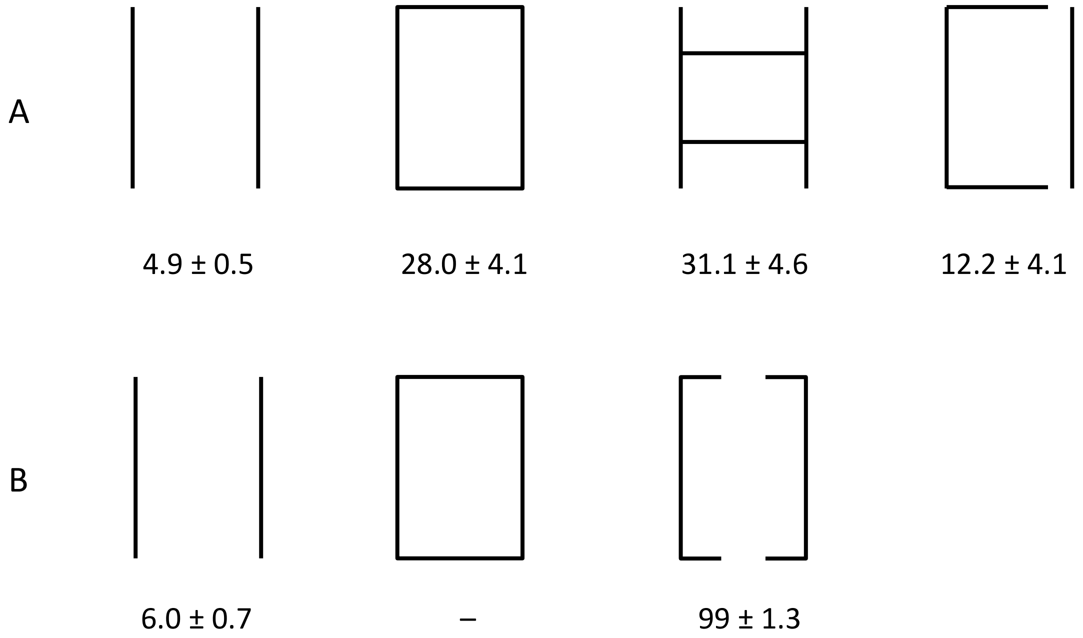

*Figure 1.1 Depictions of stimulus configurations used by (A) Mckee (1983) and (B) Mitchison and Westhimer (1984). The vertical lines in each condition were displaced in depth to measure stereoacuity and are identical in each experiment. Each panel reports the threshold (in seconds of arc) for the same observer in both experiments (note the ‘ – ‘ represents an unmeasurable threshold).*

This ‘degraded depth effect’ for unified figures has been explained in terms of disparity pooling or averaging (McKee, 1983; Fahle & Westheimer, 1988; see also Vreven, et al., 2002), saliency (Mitchison & Westheimer, 1984), cue conflict with perspective (Zalevski et al. 2007), re-definition of the fronto-parallel plane (Kumar & Glaser, 1992; Glennerster & McKee, 1999) and an insensitivity of the visual system to continuous (ie, slanted) disparity gradients (Gillam, et al., 1984; Stevens & Brookes, 1988; van Ee & Erkelens, 1996). I propose an alternative explanation: that the perceptual organization of elements to form a coherent figure or object has a direct impact on disparity processing. Such an explanation requires feedback from mid-level processes that are involved in the representation of shape or objects, and implicates well-known Gestalt grouping principles. I propose that this is an overlooked component of reduced depth effects.

Previous research has shown that top-down effects of 2-D organizational cues can impact the percept of depth. This has largely been related to the grouping of occluded contours, where researchers have carefully manipulated isolated grouping cues in amodal completion arrangements and assessed their effects on slant perception or disparity thresholds (Liu, Jacobs & Basri, 1999; Yin, Kellman & Shipley, 2000; Fanton, Gerbino & Kellman, 2004, 2005; Liu & Schor, 2005; Hou et al. 2006). For example, a number of studies have shown that the perceived slant difference between two rectangles is reduced when the fragments are grouped via collinearity behind an occluder (Yin, et al., 2000; Fanton, et al., 2004, 2005; Liu & Schor, 2005; Hou et al. 2006). Liu et al. (1999) demonstrated that, in addition to collinearity, the shape of a bounding contour (convex vs concave) is the critical determinant of whether a figure was perceived as coherent. In Yin et al.'s (2000) experiments they showed that integration of flanking surfaces behind an occluder reduced disparity sensitivity (d'). Integration was critically dependent on the similarity of visible surface features (e.g., colour and texture) as well as the presence of collinear edges. These studies provide convincing evidence that figural interpretation constrains depth thresholds and provide important evidence against a strictly hierarchical model of disparity processing, as low-level operations are clearly modulated by high-level contextual effects.

As early as 1930, the Gestalt psychologists acknowledged the importance of depth in perceptual organization. One of the first illustrations of the effects of 2-D figural grouping on stereoscopic depth magnitude was provided by Kopfermann (1930). She drew different components of closed line figures (e.g., fragments of triangles or rectangles) on glass plates and slotted the segments into a light-proof box at separations of 2cm. She found that the perceived relative depth of the figure's components critically depended on the perceived coherence of the figure. If the stimulus was seen as distinct unconnected units, the relative depth of the individual fragments was veridical. On the other hand, if the line patterns were formed as single object by the observer, the percept of depth was eliminated (for a summary of this work see Hartmann, 1935). The Gestaltists argued that the good Gestalt created via the 2-D grouping cues dominated the disparity signal provided by stereopsis (Hartmann, 1935; Koffka, 1935). It is possible that the same logic underlies the degraded depth effect demonstrated by McKee (1983) and others. In all the studies described above, the focus of the research was primarily on the impact of specific 2-D cues on threshold elevation or slant perception. There has been no systematic evaluation of potential grouping properties, nor have researchers considered the possibility of disparity-based grouping. While three-dimensional objects may be subject to the same grouping principles (Koffka, 1935), there may be additional organizational factors that influence perception. One aim of this research is to evaluate whether the rules of classic 2-D Gestalt organization extend to stereoscopic 3-D stimuli by isolating both 2-D and 3-D grouping components and evaluating their impact on depth perception.

### 1.2.3. Perceptual organization: Closure

Gestalt psychologists questioned “Why is it that we see things rather than the spaces between them?” (Koffka, 1935). They described the integration of parts into coherent structures as an outcome of figural grouping according ‘Laws of Perceptual Organization’ (Wertheimer, 1923, for condensation see Ellis, 1938; Kohler, 1926, 1930; Koffka, 1935). According to these principles, visual elements are integrated by virtue of certain properties that are present in the image, including proximity, similarity, good continuation, common fate, and closure.

The primary grouping cue of interest in this dissertation is the ‘Factor of Closure.’ In Gestalt terminology, closure assumes that all else being equal, elements that form a closed figure tend to be grouped together:

> “Ordinary lines, whether straight or curved, appear as lines not as areas… If a line forms a closed, or almost closed, figure, we see no longer merely a line on a homogeneous background, but a surface figure bounded by the line.” (Koffka, 1935, pp. 150).

Closure was assigned particular importance in the formation of objects as the closure of boundaries leads to the representation of an independent and stable shape or object; whereas single components are unstable structures (Wertheimer, 1923; Kohler, 1930; Koffka, 1935). As such, closure was described as a binary property, whereby closure is either present or absent in the shape. This qualitative definition of closure has since evolved to a more quantifiable attribute. In mathematical terms, perfect closure is easily defined, as contours are either connected or not (eg, Matthews & Howell, 2012). However, this is not the case with a perceptual description, where grouping can tolerate occlusions and gaps that are present in the natural world.

Closure has been described as an emergent perceptual feature that can be extracted pre-attentively as a simple property (Pomerantz, Sage & Stoever, 1977; Treisman & Patterson, 1984; Treisman & Gormincan, 1988; Donnelly, Humphreys, & Riddoch, 1991; Conci, Muller & Elliot, 2007; Pomerantz & Portillo, 2011). Support for this proposal is evident from studies that show significant pop-out effects for closed arrangements of fragments in cluttered displays (Kovacs & Julesz, 1993; Pettet, McKee & Grzywacz, 1998; Yen & Finkel, 1998; Braun, 1999; Mathes & Fahle, 2007). Alternatively, closure has been cast as a bridge between 1-D contour and 2-D shape formation (Gillam, 1975; Elder & Zucker, 1993, 1994). This interpretation of closure has some correspondence to mathematical notions of closure, but also has properties specific to a perceptual context. Several studies suggest that the role of perceptual closure in grouping of two-dimensional shape is modulated by the presence of specific configural components (eg: Elder & Zucker, 1993, 1994; Kimchi, 2000; Spehar, 2002; Barenholtz, et al., 2003; Hadad & Kimchi, 2008; Badcock, Haley & Dickinson, 2015). For example, in a series of visual search tasks, Elder and Zucker (1993) demonstrated more efficient search for closed contours than open versions of the same stimulus. Target detection time was optimal for connected components, and this pop-out effect was employed as a measure of closure for subsequent stimuli. They showed that perceptual closure can be modulated by altering the properties of the shape. For instance, results showed that reversing the contrast of the connecting lines eliminates perceptual closure, suggesting that perceptual closure operates only upon contour elements of a consistent contrast sign. Moreover, breaking the connection slightly increased reaction times, and search times increased monotonically as the size of the gap between inducing fragments increased (see also Gillam, 1975). On the basis of these findings, Elder and Zucker (1993, 1994) proposed that perceptual closure is best described as a continuum, rather than an all-or-none phenomenon.

More recent investigations of the perceptual consequences of closure have focused on the cues that drive the percept. These studies have shown that perceptual closure exhibits a strong dependence on the distribution of discontinuities within the stimulus. For example, using Elder and Zucker's (1993) stimulus, Spehar (2002) demonstrated that the impact of contrast reversals was dependent on the location of the contrast change: efficient search was maintained when contrast reversals occurred along the straight contour segments, and was only degraded when they occurred at the contour intersections (orientation discontinuities). Similarly, the impact of disrupting connectivity also exhibits a strong dependence on the distribution of the gaps. Depending on the location of the gap in a rectangle, collinearity will either be present or absent between elements. Hadad and Kimchi (2008) showed that this was a critical distinction to inducing closure: they showed that rapid grouping was only observed when collinearity between line end-points was present; whereas, gaps located at object corners did not lead to closure. The importance of collinearity and proximity was earlier demonstrated by Gillam (1975), who found that the degree to which two ambiguously rotating lines in depth were perceived as one unit was dependent on collinearity, and that grouping strength increased as fragments were positioned closer together (see also Gillam & Grant, 1984; Gillam, 1992). Taken together, the results of these studies show that closure is a global property, one that is a synthesis of many configural features, which can vary in strength.

A potentially important consideration that will be raised in the experiments outlined below, is the fact that previous studies of perceptual closure rely on face validity to operationally define a 'closed' figure. While this may be reasonable for some stimuli (eg, a rectangle), it may be difficult to maintain for variants of a stimulus, and the observed impact of a given manipulation on performance may not reflect its impact on perceived closure. To validate my assumption that the stimulus manipulations have the intended effect on perceived closure, I will assess perceived closure directly using rating scales. This approach validates the operational definitions and provides an objective means of relating the stimulus manipulation to the strength of perceived closure and the resultant impact on perceived depth.

### 1.3. Chapter Outline

In the following set of experiments, I assess the relationship between figural grouping and the perception of depth from suprathreshold disparity. Using a series of novel methods, I explore the critical factors underlying the effects of grouping and identify a new grouping cue that has previously not been discussed in the literature concerning perceptual organization.

In Chapter 2, I assessed the influence of figural closure on estimates of perceived depth magnitude from disparity. This experiment extended the threshold paradigm used by McKee (1983) to assess the impact of configuration on the suprathreshold appearance of stimuli. That is, is the perceived separation in depth of two isolated elements reduced when they are connected to form parts of a closed object? McKee's stimulus was modified to permit direct comparisons of depth magnitude estimates for lines when they are isolated, perceived as part of a larger figure, or parts of separate objects.

Having established the suprathreshold impact of perceptual grouping, in Chapters 3 and 4 I identified the stimulus properties that contribute to perceptual closure in depth. In a series of experiments, I systematically evaluated the impact of specific 2-D and 3-D contextual properties on depth magnitude estimates. In Chapter 3, the role of similarity was evaluated whilst maintaining element connectedness. In Chapter 4, I evaluated the importance of connectedness, proximity, good continuation, and other cues to form. In all experiments, a suprathreshold depth magnitude task was employed in conjunction with subjective measures of grouping. The results provide a framework for the rules that govern closure in depth.

In Chapter 5, I evaluated the role of disparity-based grouping by isolating the 3-D properties of a stimulus. Two methodologies were employed in the Chapter. In addition to measuring depth magnitude percepts for various configurations, I assessed visual search performance for these same stimuli.

# CHAPTER 2

## Suprathreshold Depth Magnitude Estimation Within- Versus Between-objects

Experiment 2.1 was published in the Journal of Vision: Deas, L. M., & Wilcox, L. M. (2014). Gestalt grouping via closure degrades suprathreshold depth percepts. *Journal of Vision, 14(9),* 14, 1-13. doi: 10.1167/14.9.14. Statement of contribution: Deas and Wilcox developed the study concept and design. Deas performed data collection and analysis. Both authors collaborated on data interpretation and preparation of the manuscript.

### 2.1. Introduction

The impact of closure on stereoscopic depth perception has never been directly studied. This is surprising because it is well known in the grouping literature that closure is an important cue to 2-D object formation and can have a marked impact on a variety of performance measures (eg: visual search, detection, spatial acuity). It would be reasonable to assume that there are constraints on 3-D shape formation – as documented with 2-D grouping – but we cannot simply assume that the rules of 2-D organization apply to 3-D displays, where complex disparity interactions might be at play. Thus, the aim of the experiments in Chapter 2 was to evaluate the perceptual consequences of figural grouping via closure on percepts of depth. To do so, I built upon the simple stimulus employed by McKee (1983) and others that assessed configural changes on stereoacuity. As discussed in the Section 1.2.2, these studies showed that depth discrimination thresholds for a pair of vertical lines can be markedly elevated when the lines are connected as a closed rectangle (Westheimer, 1979; McKee, 1983; Mitchison & Westheimer, 1984; Zalevski et al., 2007). The simplicity of this stimulus made it relatively straightforward to isolate and directly compare the effects of closure on perceived depth. The previous studies were restricted to measuring stereoacuity, but it is not known what impact (if any) the configural influences on stereoscopic thresholds had on suprathreshold perceived magnitude. In other words, given that thresholds are substantially elevated when isolated elements are connected, how does this same manipulation affect the appearance of suprathreshold versions of the stimuli? It might be expected that under conditions yielding high stereoacuity (small JND), a given suprathreshold disparity would produce greater magnitude estimation than when stereoacuity is low. However, studies have shown that there is not always a direct relation, that is, not all conditions that increase the stereoscopic threshold reduce the perceived depth of targets presented at suprathreshold disparities (Patel, et al., 2009; Bedell, et al., 2012).

## 2.2. General Methods for Experiments 2.1 and 2.2

**Observers** Eighteen observers participated in Experiment 2.1 and 2.2. Eleven were experienced stereoscopic observers. Seven were paid undergraduate students at York University with no prior experience with psychophysical tasks. All observers had normal or corrected-to-normal visual acuity and were able to discriminate disparity of at least 40 seconds of arc on the Randot$^{TM}$ stereoacuity test.

**Stimuli** The stimulus was composed of four identical vertical lines, positioned symmetrically about the mid-point of the display. Three conditions were created by manipulating whether various pairs of lines were connected with horizontal lines to form closed objects, as illustrated in Figure 2.1:

(A) Isolated Lines: Vertical lines were presented in isolation; observers judged the relative depth of the central pair.

(B) Closed Object: Two horizontal lines connected the end points of the central vertical pair. The target lines are the same as in (A), but now they form the vertical sides of a single closed rectangle.

(C) Segmented Objects: Each outer line pair was connected to create two closed rectangles. The central target lines formed the vertical sides of two discrete objects.

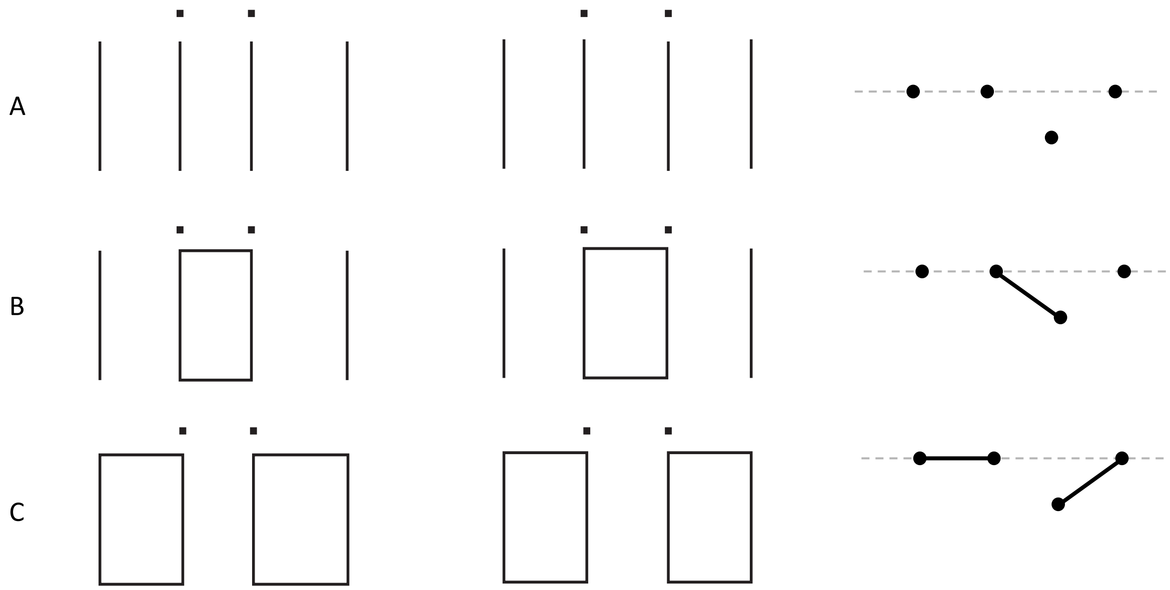

*Figure 2.1. Stereograms of the stimulus configurations used in Experiment 2.1 arranged for crossed fusion. Observers judged the relative depth of the two central vertical lines in each condition. Each row depicts one condition: (A) Isolated lines (B) Closed Object and (C) Segmented Objects. In each of the above stereopairs, the rightmost line of the central target pair has the same crossed disparity.*

The stimulus was white (59.1 cd/m$^2$) on a mid-grey background (15.6 cd/m$^2$). Each line measured 3.20° x 0.03° and was laterally separated from its neighbour by 2.10°. The connecting horizontal lines had the same width (0.03°) and luminance as the vertical lines. Each closed object subtended 3.20° x 2.16°. The monocular image of the stimulus was symmetrical about the mid-point of the display both horizontally and vertically. When connected (Figure 2.1 B, C), objects looked like slanted planar surfaces rotated about a vertical axis.

On each trial, one line of the central pair was presented at one of a range of crossed disparities relative to the other three lines (0°, 0.06°, 0.12°, 0.18°, 0.24° and 0.32°). Pilot testing ensured that all disparities were within Panum's fusional area. To create the binocular disparity each half-image was shifted in opposite directions by half the disparity. The experiment consisted of 18 conditions (6 disparities x 3 configurations), with each condition presented 10 times in random order (5 left line disparate and 5 right line disparate). All 180 trials were completed in a single session.

**Apparatus: Stimulus Presentation** Stimuli were generated using the Psychtoolbox package (Brainard, 1997; Pelli, 1997) for MATLAB on a Mac OS X computer. They were presented on a pair of LCD monitors (Dell U2412M) in a mirror stereoscope arrangement at a viewing distance of 57cm. The monitor resolution was 1920 x 1200 pixels with a refresh rate of 75Hz. At this resolution and viewing distance, each pixel subtended 1.60 arcmin of visual angle. The monitors were carefully calibrated and matched prior to testing and the gamma functions linearized. A chin rest stabilized head position during testing. The interocular distance for each observer was measured with a Richter digital pupil distance meter$^{TM}$.

**Apparatus and Procedure: Depth magnitude estimation** Depth estimates were made using a purpose-built touch sensitive sensor. A rectilinear SoftPot membrane potentiometer (SpectraSymbol) was mounted on thin aluminum bar. The sensor strip was 200mm long and 7mm wide with a resistance of 10 kOhm. The potentiometer allowed linear measurements across the 200mm length, with a resolution of approximately 0.2mm. Responses were read using an analog to digital converter and a 16-bit micro controller. A rod at one end of the sensor was used to position the thumb, and its distance from the start of the sensor strip was adjusted prior to testing for each observer (this took into account differences in thumb thickness). The recorded voltage was converted to millimetres using a MATLAB script.

At the beginning of each trial, a white fixation cross (1.5° x 1.5°) was presented at the center of a screen for 750ms. In each trial, observers were asked to indicate the amount of depth they perceived between the two central test lines. They did this by positioning their thumb against the adjustable rest at the base of the sensor, and pressing the side of the nail of their index finger at some point along the sensor to indicate depth magnitude. A small red LED positioned in front of the stereoscope mirrors, and 10.8° below the line of sight to the stimulus, illuminated when sufficient pressure was applied to the sensor strip. Observers were free to adjust their fingers until satisfied with their estimation. They then pressed the spacebar to record the response and move on to the next trial. Between trials, observers were asked to reposition their finger at the base of the sensor (the small LED was off when no pressure was applied to the strip). Prior to testing, observers completed a brief practice session of 30 trials (using the isolated line stimuli) to familiarize themselves with the depth estimation technique.

The sensor measurement technique was validated in a separate study (Hartle & Wilcox, 2015) in which observers estimated depth between a similar pair of vertical lines using three methods (in random order): estimates were compared using the touch-sensor described here, thumb and index finger separation measured manually using a digital caliper and a virtual ruler displayed on a computer screen with an adjustable cursor. While the ease of measurement (and preferred method) varied across observers, all observers consistently overestimated the depth for the range of disparities used in these experiments. Importantly, observer's magnitude estimates remained consistent across multiple trials and measurement methods.

**Theoretical depth from disparity** To simplify the comparison of on-screen theoretical depth to observers' estimated depth, the stimulus disparities were converted to theoretical depth in millimetres for each experiment. The conventional formula was used, which relates disparity to predicted depth at a known viewing distance (57cm): $Depth = (d * D^2 / IOD)$, where $d$ is the relative disparity, $D$ is the viewing distance and $IOD$ is the inter-ocular distance (see Howard & Rogers, 2012, pp. 457). The average interocular difference for the observers that participated in each experiment was applied. For Experiments 2.1 and 2.2, the theoretical depth between the two test lines corresponding to crossed disparities of 0°, 0.06°, 0.12°, 0.18°, 0.24° and 0.32° degrees were 0, 5.5, 11, 16.5, 22 and 27.5 mm respectively (with average IOD = 61.7mm).

## 2.3. Results and Discussion for Experiment 2.1

Figure 2.2 shows the mean estimated depth for each condition as a function of the predicted separation in depth. As the disparity between the target pair increased, estimated depth increased in all conditions. However, there was a clear difference in the amount of depth perceived in the connected versus unconnected conditions. That is, when horizontal lines connected the target pair to form a closed object, the estimated depth was consistently smaller than in the isolated lines condition. By comparison, depth percepts for the isolated lines and segmented objects conditions were very similar at all disparities and consistently larger than the closed object condition. This pattern of results was confirmed statistically. Because observers were told that some stimuli would have zero disparity, this response became stereotyped to the base of the sensor strip and had no associated variance. To avoid biasing the model fits, the zero-disparity estimates were excluded from analyses. A repeated measures ANOVA showed main effects of Configuration, $F(1, 20) = 24.48$, $p < 0.0001$; $\eta^2 = 0.59$, and Disparity $F(1, 26) = 84.44$, $p < 0.0001$; $\eta^2 = 0.83$. There was no significant Configuration x Disparity interaction, $F(3, 54) = 1.96$, $p = 0.056$; $\eta^2 = 0.10$. Simple effects analyses was used to compare the differences among configurations. These contrasts revealed that perceived depth for the isolated lines ($F(1,17) = 31.22$ , $p<0.0001$; $\eta^2 = 0.36$) and segmented objects ($F(1,17) = 23.34$ , $p<0.0001$; $\eta^2 = 0.58$) conditions were significantly higher than for the closed objects. There was no difference between the results of the isolated lines and segmented objects conditions ($F(1,17) = 0.34$, $p=0.57$; $\eta^2 = 0.02$).

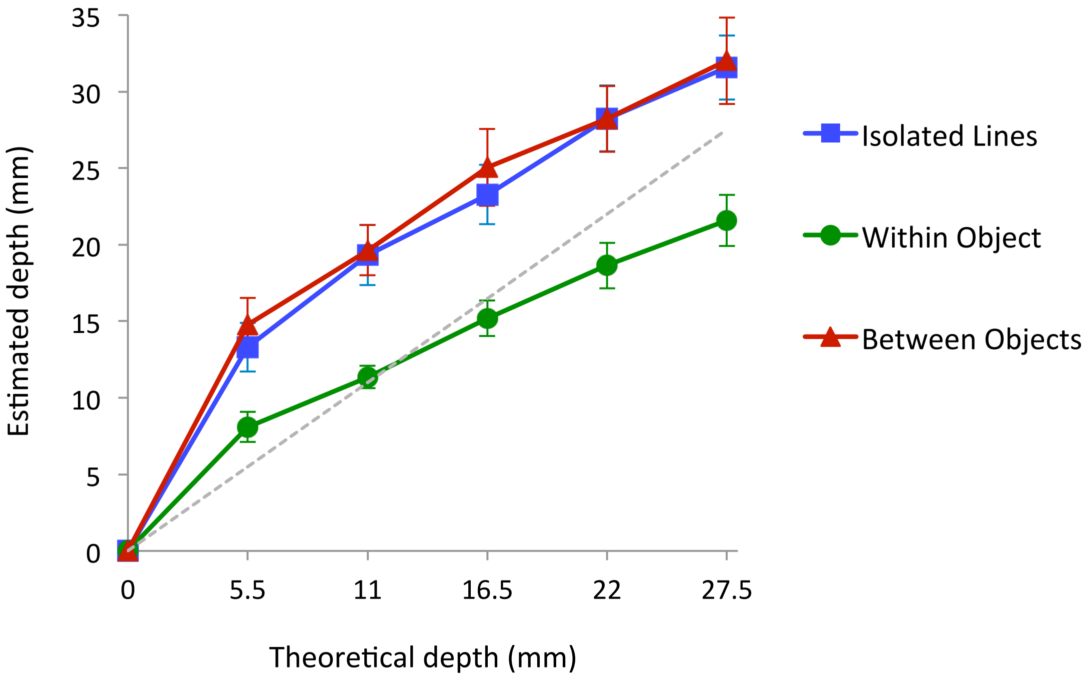

*Figure 2.2. Perceived depth as a function of depth offset for three stimulus configurations: isolated lines (blue squares), within object (green circles) and between objects (red triangles). The abscissa shows the theoretical depth, and the ordinate shows the amount of estimated depth (mm). The dashed line indicates a gain of one and error bars show ± one standard error of the mean.*

As shown in Figure 2.2, depth percepts were largest when the depth judgement involved stimuli perceived to be either objects in their own right (isolated lines) or parts of two separate objects. On the other hand, depth estimates for these lines were consistently smaller when they formed edges of a single closed object. It could be argued that depth estimates in the closed object condition are closest to veridical, and that the isolated lines data reflect *enhanced* depth percepts via segmentation. However, the manual estimation technique used here has been shown to yield a constant depth overestimation for all observers at the disparity range tested here (Hartle & Wilcox, 2015). In that study, experienced and naïve (no previous psychophysical experience) observers estimated the depth between a single pair of vertical lines and the average depth estimates were consistently ≈5mm above veridical (for depths up to 20.3mm). Previous studies have shown that the perception of depth based on stereopsis is overestimated at short distances (<80cm) (Foley, 1980; Johnston, 1991). The viewing distance here was 57cm and this may have contributed to the overestimation of depth for isolated lines. For the purposes of these experiments, the focus of comparison is the differences between conditions, rather than their absolute depth.

These results show that perceived depth from disparity is contingent on observers' figural interpretation of the stimulus. The amount of depth perceived between two connected vertical lines was consistently and significantly less than between those components in isolation. Thus, the contextual effects shown by McKee (1983) and others are not limited to threshold discrimination tasks, but also influence the perceived magnitude of suprathreshold disparity between two features. I argue that the degraded depth effect is a consequence of figural closure, and these results are consistent with that proposal. Alternatively, it is possible that when using line stimuli with well-defined boundaries, the competition between shape-from-stereo and shape-from-perspective interfered with depth perception. In Experiment 2.2, I assessed the potential role of linear-perceptive on the magnitude of depth reported in these stimuli.

## 2.4 Experiment 2.2

### 2.4.1. Introduction

An alternative explanation for the degraded depth effect shown in Experiment 2.1 is that the disparity signal conflicts with the monocular perspective cues in the lines and shapes. Monocular cues, such as linear perspective, can provide valuable cues to relative depth. In the stimuli used above, the height of the lines was not adjusted according to the changing disparity and therefore the disparity information conflicted with the linear perspective information. That is, horizontal disparity information signaled that one line is closer than the other, but relative size and foreshortening information suggested that both lines lie on the same depth plane**.** Cue conflict may have been more salient in the closed object where additional perspective-based cues were provided by the enclosed obtuse and acute angles created at the intersection points with the vertical lines, and the orientation of the horizontal connecting lines (see Figure 2.3).

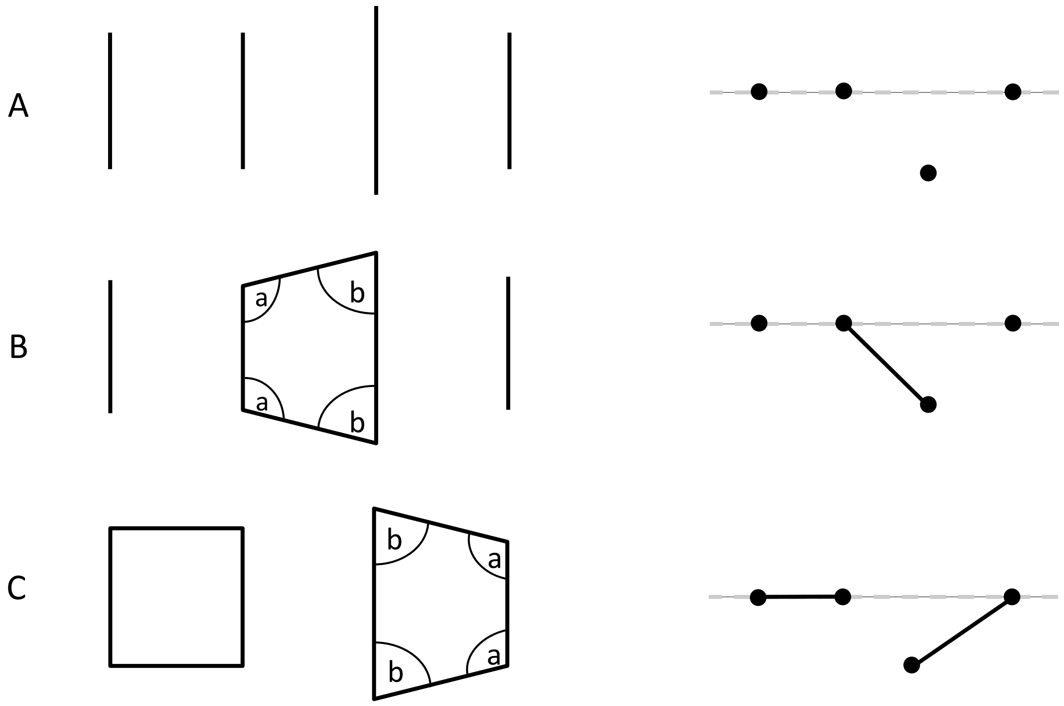

*Figure 2.3 Illustrations of the linear perspective information signalled by the isolated line (A), closed object (B) and segmented objects (C) stimuli. Monocular information about depth in each stimulus comes from the difference in the projected length of the lines (not drawn to scale). In (A) perspective is the only source of monocular information. In (B) and (C) additional linear perspective information is provided by the relative size of the disparate line and the projections from the connecting lines: the lines converge to different angles, where angle (a) is larger than angle (b).*

In Experiment 2.2, the impact of cue conflict on the degraded depth effect for closed figures was assessed. Depth magnitude for the same set of stimuli were compared with and without conflicting linear perspective information. Cue conflict was removed by manipulating the length of the disparate test line to be consistent with the distance signalled by changing disparity, so that the closer line subtended a larger visual angle. If cue conflict accounts for the attenuated depth percepts from closed objects shown in Experiment 2.1, then its removal should eliminate the disruptive effect and restore depth estimates to the level obtained using isolated lines.

### 2.4.2 Stimuli

The three conditions were the same as those described in Experiment 2.1: isolated lines, closed object and between two objects. At zero disparity, each vertical line measured 3.20° x 0.03° and was laterally separated from its neighbour by 2.10°. Perspective foreshortening was introduced by altering the test line to match the disparity for a given trial. This was achieved by measuring the interocular distance for each observer and calculating the perspective projection for each eye. The stimuli are depicted in Figure 2.3 (not to scale). For the closed objects, the horizontal disparity and perspective cues were consistent with the real rotation of a plane containing the stimuli about a fixed vertical axis in space, thus providing consistent information about relative depth.

### 2.4.3 Results and discussion

Figure 2.4 shows the results for Experiment 2.2, with depth estimates reported in Experiment 2.1 included for comparison. The amount of depth perceived within a closed object was reduced compared to isolated lines. Moreover, depth estimates for the between object condition were very similar to the isolated lines. These results were consistent with the pattern reported in Experiment 2.1, and showed that the degraded depth effect is maintained even when cue conflict is removed. A repeated measures ANOVA showed main effects of Configuration, $F(1, 20) = 29.66$, $p < 0.0001$; $\eta^2 = 0.71$, and Disparity $F(1, 26) = 97.05$, $p < 0.0001$; $\eta^2 = 0.77$. The Configuration x Disparity interaction was significant, $F(3, 54) = 1.22$, $p = 0.048$; $\eta^2 = 0.24$. Contrasts among configurations (using simple effects analyses) revealed that perceived depth for the isolated lines ($F(1,17) = 43.29$ , $p<0.0001$; $\eta^2 = 0.45$) and between objects ($F(1,17) = 37.43$ , $p<0.0001$; $\eta^2 = 0.61$) conditions were significantly higher than for the closed objects.

There was no difference between the results of the isolated lines and between objects conditions ($F(1,17) = 0.22$, $p=0.60$; $\eta^2 = 0.08$).

Now compare results the results with Experiment 2.1 (also shown in Figure 2.4). A repeated measures ANOVA (between-subjects) was conducted on each condition across the experiments. There was no significant difference for any comparison at the p=0.05 level.

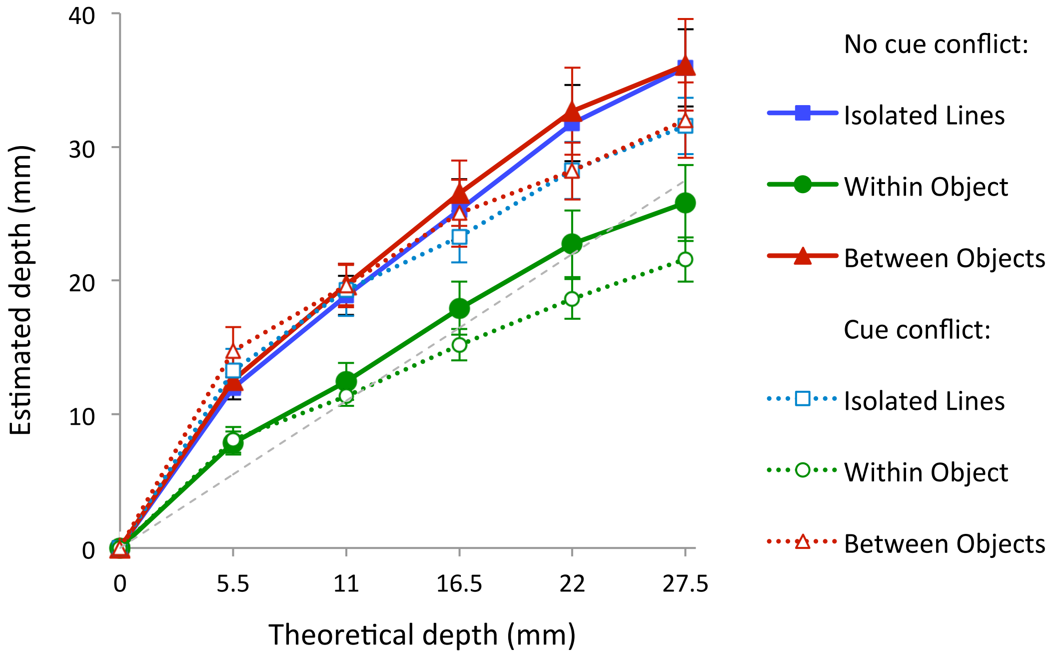

*Figure 2.4. Perceived depth as a function of depth offset for stimulus configurations with perspective-based cue conflict removed. Each solid line represents the results where perspective-based cue conflict was removed: isolated lines (blue squares), within object (green circles) and between objects (red triangles). Results for Experiment 2.1 (with cue conflict) are included for comparison, represented by dashed lines. The abscissa shows the theoretical depth, and the ordinate represent the amount of estimated depth. The dashed line grey indicates a gain of one and error bars show ± one SEM.*

Now compare the results with and without cue conflict. Although there was a trend towards increased depth at larger disparities, this was not significant for any condition ($p > 0.05$, pairwise comparisons). It is notable that the between observer variance in the estimates increased for all conditions, compared to the estimates obtained in Experiment 2.1. While variance tends to increase with magnitude (Weber's law), this increase may also be related to individual differences in dependence on specific depth cues. For instance, when multiple cues are present, some observers rely primarily on perspective information, while others are able to solely use binocular disparity (eg: Allison & Howard, 2000; Sato & Howard, 2001; van Ee, van Dam & Erkelens, 2002; Zalevski et al. 2007), a difference that may be a function of experience with stereoscopic tasks (Hartle & Wilcox, 2015). When perspective information is varied, observers that exclusively use disparity information show little or no change in depth estimates; whereas observers that are most influenced by perspective cues show large changes in their estimates (Sato & Howard, 2001; Hartle & Wilcox, 2015). Thus, the removal of conflicting perspective information may have changed the estimates of some observers, but not others. Nevertheless, for the stimuli used here, the reduction in depth for closed objects was robust against changes in linear perspective. Thus, perspective conflict may contribute to, but is not the sole explanation for, the reduced depth percepts for closed objects.

## 2.5 General Discussion for Experiments 2.1 and 2.2

The aim of Chapter 2 was to directly assess the impact of figural closure on suprathreshold estimates of depth from disparity. The stimuli were designed to compare depth percepts for the same pair of vertical lines, but with different figural interpretations. In the context of Gestalt organization, the addition of the horizontal line segments to the isolated line stimulus changed the interpretation to a single object by virtue of figural closure. The results showed that amount of depth estimated across the edges of a closed object was consistently smaller than the depth between the edges in isolation or between parts of two separate objects. This research extends work at the threshold-level (described in the Introduction), which has shown an increase in depth discrimination thresholds when a closed object is created or implied (Westheimer, 1979; McKee, 1983; Mitchison & Westheimer, 1984; Zalevski et al., 2007). In a preliminary study, I replicated this finding using the modified version of McKee's stimulus employed here (Deas & Wilcox, 2012). It is not obvious if, or how, the observed loss of precision at threshold relates to the decreased depth magnitude percepts for suprathreshold depth differences shown here. In both cases 'performance' is degraded (thresholds increase and perceived depth decreases), but the underlying source of the disruption may be quite different. Similar differences between threshold and suprathreshold stimulus dependencies have been shown for other aspects of visual processing (eg. contrast sensitivity as a function of spatial frequency); however, the exact nature of the relationship between threshold and suprathreshold effects of perceptual grouping is unclear.

The impact of the configural manipulations on perceived depth cannot be accounted for by effects of depth contrast (Werner, 1937; Graham & Rogers, 1982; van Ee, Banks & Backus, 1999; Sato & Howard 2001). Depth contrast describes the illusory inclination induced between two planar surfaces: this phenomenon illustrates the fact that the apparent depth of a stimulus is affected by not only the disparity within a stimulus but also by neigbouring disparities. For example, when a stimulus in a frontal plane is surrounded by an inclined surface (as in Figure 2.1C), it appears inclined in the opposite direction. If such an effect was induced in these configurations, then we might have seen enhanced depth estimates between the two objects, as the edge of the frontal surface would appear to recede in depth. However, the similarity in depth estimates between two objects and the isolated lines suggests that this was not the case. It is possible, however, that attentional factors may contribute to the results for the between-object conditions. In this case, observers were instructed to judge the depth of the centre lines only, therefore they may have disregarded the depth profile of the configuration on either side of fixation. Additional experiments would be needed to assess this possibility.

In addition, the difference in magnitude estimates across the within- and between- object conditions show that the depth reduction is not due to local effects that are presumed to occur at early visual processing stages. In these cases, the disparity information in the target lines was identical, and the same local features were present in the form of L-junctions at the top and bottom of both target lines (but at different orientations). Yet, the amount of depth perceived was only reduced when estimating within an object. Moreover, the results for the between-object condition were virtually identical to those obtained in the isolated line configuration. These figural manipulations allowed the exclusion of explanations based on disparity interactions from neighboring components, such as averaging or inhibition, or cue conflict with perspective (confirmed in Experiment 2.2). Instead, I argue that the reduction of depth percepts shown here is a mid-level phenomenon based on figural grouping cues. In this case, I propose that depth depends on perceived closure. Connecting the vertical line pair with horizontal line segments transformed the interpretation of the display from individual line fragments to a single 'whole' object via perceived closure. As a consequence, perceived depth from disparity was reduced. However, as discussed in Chapter 1, perceptual closure is a complex grouping factor, one that has been defined as both an all-or-none emergent property and a perceptual continuum: the results of Experiment 2.1 are consistent with both of these interpretations. To clarify the relationship between the degraded depth effect and perceived closure, the next experiments evaluated the contribution of specific properties to the amount of depth perceived in a connected object.

# CHAPTER 3

## Evaluating the Role of Perceived Closure to Degraded Depth Percepts

Experiment 3.1 was published in Journal of Vision in 2014: Deas, L. M., & Wilcox, L. M. (2014). Gestalt grouping via closure degrades suprathreshold depth percepts. *Journal of Vision, 14(9),* 14, 1-13. doi: 10.1167/14.9.14.

Statement of contribution: L.M. Deas developed the study concept and design, and performed data collection and analysis. Both authors collaborated on data interpretation and preparation of the manuscript.

### 3.1 Introduction

As outlined in Chapter 1, previous researchers have described closure as a grouping cue in its own right, one that is dependent on multiple cues. Some researchers suggest that closure is an all-or-none property, where closure emerges in a configuration by the addition of contextual information (Pomerantz et al., 1977; Treisman & Patterson, 1984; Treisman & Gormincan, 1988; Donnelly, et al., 1991; Pomerantz & Portillo, 2011). Other researchers claim that closure is graded in nature and forms a perceptual continuum (eg: Elder & Zucker, 1993, 1994; Kimchi, 2000; Spehar, 2002; Barenholtz, Cohen, Feldman, & Singh, 2003; Hadad & Kimchi, 2008; Bertamini & Wagemans, 2013). The results of Experiment 2.1 were consistent with both interpretations.

One way to assess the impact of perceptual grouping on target stimuli is to systematically manipulate alternative grouping cues and evaluate the impact on performance. This approach has been used by several researchers to assess the impact of grouping on visual detection and spatial acuity (Malania, Westheimer & Herzog, 2007; Sayim, Westheimer & Herzog, 2008, 2010; Saarela, Sayim, Westheimer & Herzog, 2009; Pomerantz & Portillo, 2011; Manassi, Sayim & Herzog, 2012; Herzog & Manassi, 2015). For example, in their investigations of the influence of Gestalt cues on crowding, Herzog and colleagues demonstrated that Vernier acuity is highly dependent on the ‘amount of grouping’ in flanking lines introduced as crowding features (Malania, et al., 2007; Sayim, et al., 2008, 2010; Saarela, et al., 2009; Manassi, et al., 2012; Herzog & Manassi, 2015). In one example, when a Vernier target (consisting of two equal-length vertical lines) was embedded in an array of flankers that had the same height and same colour as the target, alignment thresholds were markedly poorer than when the Vernier target was presented on its own. However, thresholds improved when the contrast of the flanking lines was reversed, as if grouping by similarity segmented, or 'ungrouped,' the flanking lines from the Vernier configuration. In Experiment 3.1, I used a similar approach to determine the extent to which the reduction in perceived depth for closed object is modulated by perceptual closure. I adapted the configuration used in Experiment 2.1 by reversing the contrast of the horizontal connecting lines to evaluate the influence of contrast similarity on perceived depth magnitude for the closed object stimulus. This configuration was presented in isolation or as part of a larger configuration which contained alternate grouping solutions (see Figure 3.1). If contrast similarity is necessary for grouping, then reversing the contrast of the connecting line should eliminate the reduction in perceived depth. In addition to estimating perceived depth magnitude, subjective ratings of closure were obtained in order to formalize the relationship between perception and the stimulus manipulations. This is an important consideration, as an observed impact of a given manipulation on performance may not reflect its impact on perceived closure. As described in Section 1.2.3, previous studies of perceptual closure depended on face validity to operationally define a 'closed' figure. Combining depth magnitude estimates with subjective ratings provides an objective means of relating the stimulus manipulation to the strength of perceived closure and the resultant impact on perceived depth.

## 3.2 Methods

**Observers** Nine participants from Experiment 2.1 took part in Experiment 3.1. Five of these were undergraduate students who, prior to the first experiment, had no previous experience with psychophysical tasks. The remaining four participants were experienced observers.

**Apparatus and Procedure** The apparatus and depth magnitude assessment was the same as that described in Section 2.2. All configurations were randomly interleaved and presented ten times per disparity.

*Subjective ratings* To verify the effect of the grouping cue manipulations on the interpretation of the stimuli, the nine participants were asked to evaluate the extent to which the central vertical test pair appeared as part of a distinct object with 0 = not an object to 10 = a strong sense of an object. The stimuli were displayed on the stereoscope at the largest test disparity and ratings were obtained for monocular (left-eye view) and binocular (stereoscopic) viewing.

**Stimuli** In addition to the original isolated line and closed object conditions (Figure 3.1 A, B), five new configurations were tested (Figure 3.1 C–G). The specific hypotheses are outlined below. In Figure 3.1C, the contrast polarity the connecting lines was reversed (black = 3.1 cd/m$^2$). This removed the uniform contrast properties of the elements forming the closed object but maintained connectedness. In the stimuli depicted in Figure 3.1 D-G, the closed object was flanked by equidistant horizontal lines, eight above and eight below the horizontal connecting lines. These flankers were added to provide alternative grouping solutions for the horizontal connectors thus placing flanker/connector proximity, collinearity and similarity in competition with closure. All flankers had the same dimensions as the connecting horizontal lines (2.1° x 0.03°) and were separated from their nearest neighbours by 0.23°. The disparity gradient of the flankers matched that of the connecting lines on a given trial. There were four variants of the flanker stimuli. In the two uniform flanker configurations, all horizontal lines either had the same contrast as the vertical lines (Figure 3.1D) or had reversed contrast (Figure 3.1E). In two additional configurations, the contrast of the flanking lines was alternated, starting from either the same- (Figure 3.1F) or reversed- polarity connecting line (Figure 3.1G). These conditions served as a control for the effects of adding more disparity information above and below the central figure.

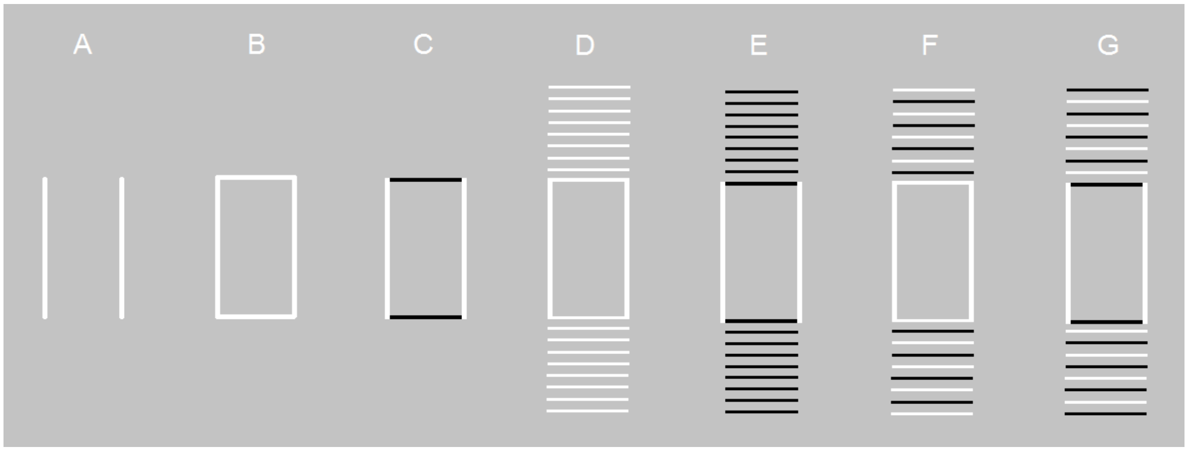

*Figure 3.1. Stimuli used in Experiment 3.1. The outer vertical lines are not included in the figure due to space considerations, but were present during testing (as illustrated in Figure 2.1). Stimuli (A) and (B) are the isolated line and closed object stimuli used in Experiment 2.1 which were re-tested for comparison. In configurations C-G, the grouping cues were varied, either in isolation or in combination, to influence perceived closure: (C) Reversed-contrast connectedness. (D) Uniform connectedness, proximity, similarity and similarity of orientation (E) Reversed-contrast connectedness, proximity, similarity and similarity of orientation. (F) Uniform connectedness, proximity and similarity of orientation. (G) Reversed-contrast connectedness, proximity and similarity of orientation.*

### 3.3 Results & Discussion

Perceived depth averaged across observers for each condition is shown in Figure 3.2. Individual figures depict the results for each stimulus type. Figure 3.2A shows the depth estimates for the original isolated lines and closed object conditions. A repeated measures ANOVA confirmed that the estimated depth between the vertical target lines in the closed object (Figure 3.2A) was significantly less than in the isolated vertical lines, $F(1, 8) = 8.68$, $p = 0.019$; $\eta^2 = 0.52$. Zero-disparity estimates are excluded from analyses, as explained in Section 2.3. These results are plotted in Figures 3.2 B-D to aid comparison.

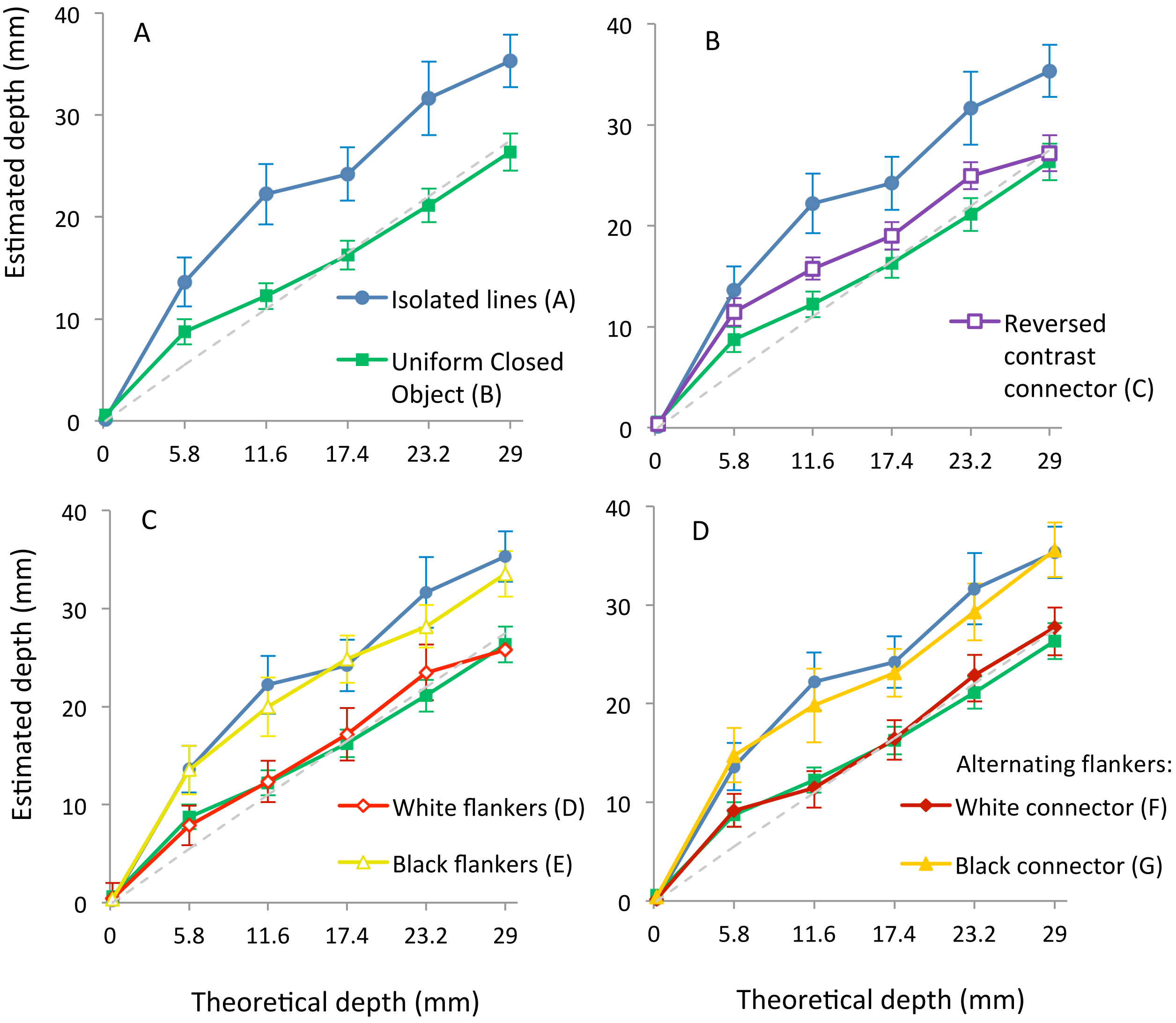

*Figure 3.2. Perceived depth as a function of depth offset (in mm) for Experiment 3.1 (n=9). In each figure the abscissa shows the theoretical depth, and the ordinate shows the estimated depth. See Figure 3.1 for illustration of stimuli. (A) Replication of Experiment 2.1 for isolated lines (blue circles) and closed objects (green squares). For comparison, these data are re-plotted in subsequent figures. (B) The polarity of the horizontal connecting line was reversed (purple open squares). (C) Results for uniform flankers with the same contrast polarity as the connecting line, either all white (red diamonds) or black (yellow triangles). (D) Alternating contrast flanking lines extending from either white (red diamonds) or black (orange triangles) connectors. The dashed line indicates a gain of one and error bars show ± one SEM.*

***Reversed contrast connectors*** Figure 3.2B shows the depth estimates obtained when the contrast of the connecting lines was reversed. While this change to the stimulus slightly increased the magnitude of depth percepts relative to the closed object condition, the amount of depth remained less than that seen for isolated lines. A repeated measures ANOVA confirmed that there was a significant effect of Configuration, $F(1, 9) = 8.43$, $p = 0.016$; $\eta^2 = 0.51$ and Disparity, $F(2, 18) = 104.28$, $p < 0.0001$; $\eta^2 = 0.93$. However, the Configuration x Disparity interaction was not significant, $F(4, 36) = 2.56$, $p = 0.077$; $\eta^2 = 0.24$. Simple effects analyses was used to compare the differences among configurations. Contrasts revealed a significant difference between the isolated lines and reversed-contrast connector conditions, ($F(1, 8) = 20.14$ , $p = 0.006$). In addition, a significant difference was found between the conditions with reversed and uniform contrast connecting lines $F(1, 8) = 8.48$ , $p = 0.035$). Subjective ratings (Figure 3.3) for the reverse contrast lines configuration were in the mid-range of the scale (rating = 5.8), confirming that perceived closure was not as strong for this stimulus as found in the original (uniform contrast) closed object condition (rating = 10). The incomplete disruption of perceptual grouping reflected in these ratings is consistent with the fact that perceived depth lies between the depth estimates for grouped and ungrouped configurations.

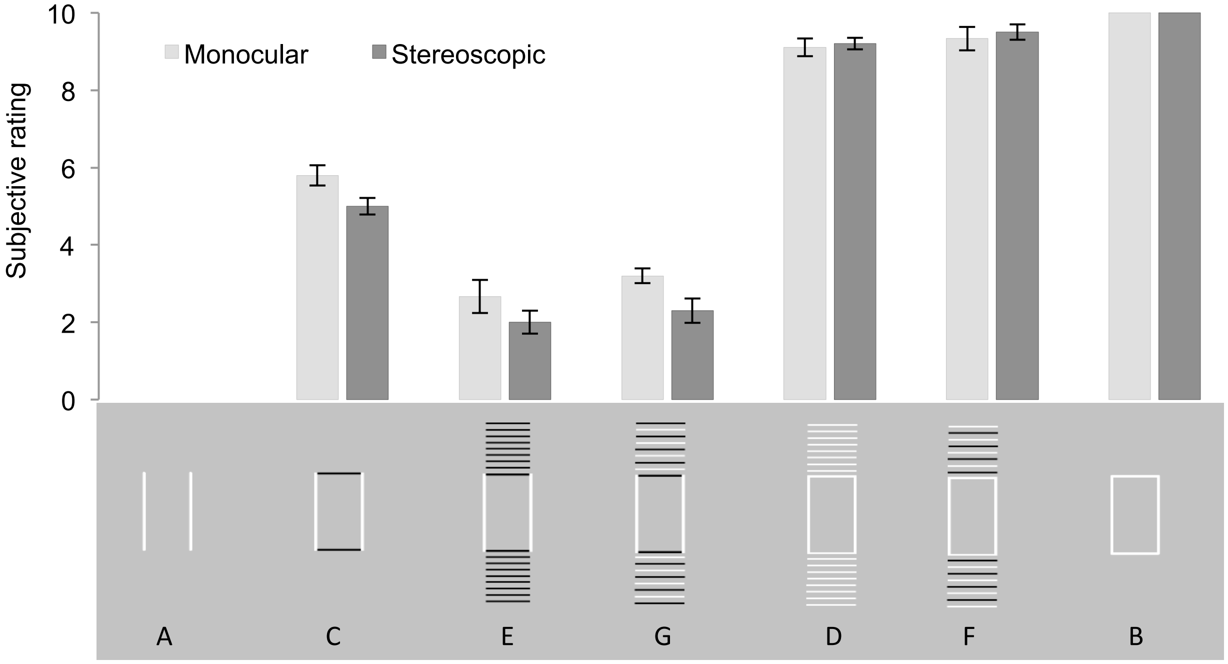

*Figure 3.3. Average subjective ratings for the seven stimulus configurations used in Experiment 3.1 (n=9). Ratings range from 0 (not an object) to 10 (a strong sense of a closed object). The stimuli are depicted below each rating (see Figure 3.1). Error bars represent ± one standard error of the mean.*

***Alternative grouping solutions*** Flanking lines (Figure 3.1 D-G) were added above and below the horizontal connectors to provide an alternative grouping solution for the connecting lines to establish if this would enhance the perceptual segmentation of the horizontal lines from the vertical target pair. Four versions of this configuration were tested in which the horizontal lines had the same contrast polarity as the vertical lines (Figure 3.1D), reversed contrast polarity (Figure 3.1E), or alternating contrast polarity (Figure 3.1F, G).

*Uniform flankers* First consider the results obtained when all flanking lines had the same contrast polarity as the connecting lines, and was either the same as the vertical lines (Figure 3.1D) or had reversed polarity (Figure 3.1E). In these stimuli, proximity, similarity of colour and similarity of orientation cues competed with closure of the central object. As shown in Figure 3.2C, it is clear that the amount of perceived depth depended critically on the relative contrast polarity of the horizontal lines. When the horizontal lines had the same contrast as the vertical lines, estimated depth was very similar to the uniform closed object at all test disparities. On the other hand, when the contrast of the horizontal lines was reversed, the amount of perceived depth increased to the levels reported for the isolated lines. Statistical analyses confirmed these observations. There were significant main effects of both Configuration, $F(3, 24) = 4.77$, $p = 0.01$; $\eta^2 = 0.37$ and Disparity, $F(1, 15) = 99.74$, $p < 0.0001$; $\eta^2 = 0.92$ and no Configuration x Disparity interaction, $F(12, 96) = 1.31$, $p = 0.23$; $\eta^2 = 0.14$. Contrasts (using simple effects analyses) revealed that estimates for the same-contrast configuration were significantly different from the isolated lines ($F(1, 8) = 5.97$ , $p = 0.04$; $\eta^2 = 0.43$), but were not different from the uniform closed object ($F(1, 8) = 0.02$, $p = 0.89$; $\eta^2 = 0.003$). In contrast, estimates in the reversed contrast condition were not significantly different from the isolated lines ($F(1, 8) = 0.39$ , $p = 0.55$; $\eta^2 = 0.05$), but were different from the uniform closed object ($F(1, 8) = 4.95$, $p = 0.05$; $\eta^2 = 0.43$). Additional analyses showed that estimates in the reversed contrast condition were significantly different from estimates reported in the reversed-contrast object without flankers (F(1,8) = 8. 42, $p$ = 0.037; $\eta^2$ = 0.33).

These results show that the amount of perceived depth was modulated by the relative contrast polarity of the connecting line in the presence of an alternative grouping solution. That this was related to perceived closure of the central figure was confirmed by the ratings data for these configurations. When the connector and flankers were white (same polarity), the closed object ratings were high (9.1), and perceived depth was similar to the level obtained when observers viewed the closed object configuration. However, when the both the connectors and the flankers were black (reversed polarity), closure ratings dropped to 2.7 and perceived depth increased to that obtained when viewing isolated lines.

*Alternating flankers* In configurations F and G in Figure 3.1, contrast similarity in the flanking units was controlled for to determine if proximity and similarity of orientation was sufficient to break the perceived closure. In both of these configurations, alternative grouping solutions for the horizontal connectors were available via proximity and collinearity. The results in Figure 3.2D are very similar to those shown in Figure 3.2C: in the presence of the alternative grouping solution (flankers), the contrast polarity of the connecting horizontal lines, relative to the vertical lines, determined the amount of perceived depth. As in Figure 3.2C, when the connectors were black (reverse polarity) depth percepts were virtually identical to those recorded for the isolated lines. When the connectors were white (same polarity) depth percepts followed those obtained in the uniform closed object condition. Again, there were main effects of both Configuration, $F(2, 16) = 5.52$, $p = 0.015$; $\eta^2 = 0.41$ and Disparity, $F(4, 32) = 121.51$, $p < 0.0001$; $\eta^2 = 0.94$, and no Configuration x Disparity interaction, $F(12, 96) = 1.26$, $p = 0.26$; $\eta^2 = 0.14$. Contrasts with these alternating flanker conditions showed the same pattern of results as the uniform flankers described above. When the connector had the same contrast polarity as the vertical lines, depth estimates were significantly different from those obtained in the original isolated lines condition ($F(1, 8) = 6.95$ , $p = 0.03$; $\eta^2 = 0.7$) but not different from the uniform closed object ($F(1, 8) = 0.06$ , $p = 0.81$; $\eta^2 = 0.10$). In comparison, when the contrast polarity of the connector was reversed, estimates were not statistically distinguishable from the isolated lines ($F(1, 8) = 8.68$ , $p = 0.02$; $\eta^2 = 0.52$)

but were different from the uniform closed object ($F(1, 8) = 5.81$, $p = 0.04$; $\eta^2 = 0.42$) and reversed-contrast object ($F(1, 8) = 11.62$, p=0.021 $\eta^2 = 0.29$)

As was the case in the matched flankers conditions, the results show that the depth percepts reported for the alternating polarity flanker conditions corresponded well with the participants' closure ratings. Again, when the horizontal connectors were white (Figure 3.1F) the depth percepts were low, and the closure ratings were high (9.3). In contrast, when the horizontal connector was black (Figure 3.1G), the depth percepts increased and perceive closure ratings dropped to 3.2. From this, it is obvious that when combined with an alternative grouping solution, the similarity of the horizontal connecting lines to the vertical lines was the critical determinant of perceived closure and depth magnitude percepts. The systematic increase in depth reported for the reversed contrast object and then

## 3.4 General Discussion for Experiment 3.1

The experiments presented in Chapter 3 demonstrate that the disruptive effect of connecting the central vertical line pair critically depends on the degree to which the target lines were perceived as part of a closed object. Without altering the configuration or spatial properties of the vertical and horizontal lines, the perceived separation in depth between the vertical lines was modulated simply by changing the degree to which they were perceived as a part of a closed object. In this instance, figural closure was manipulated by reversing the contrast of the connecting line and providing alternative classical grouping cues in the form of proximity and similarity (colour and orientation). In addition to measuring depth magnitude, each manipulation was validated by asking observers to rate the degree of closure of the central target pair. By using these methods in combination, the results showed that any manipulation that led to a decrease in perceptual closure caused an associated increase in perceived depth. For instance, closure was strongest when the spatial properties of the connecting lines were most similar to the vertical lines (Figure 3.2A). But by reversing the contrast polarity of the horizontal lines relative to the vertical test line, it was possible to reduce perceived closure and obtain slight enhancements in depth percepts (Figure 3.2B). Thus, connectedness itself is not sufficient to induce a strong percept of closure. But estimates were still degraded compared to the isolated lines. It is possible that, with no other grouping options available, the connector had no choice but to group with the vertical lines. Indeed, when alternative grouping options were available (via proximity and similarity), the horizontal connecting line was assigned as a strong or weak closure cue based respectively on uniform versus reversed contrast properties. For these flanker conditions, the reduced depth effect was eliminated only when the contrast of the connecting line was reversed relative to the vertical lines. The systematic increase in depth reported for the reversed contrast object, and then adding flankers, highlights the graded nature of closure.

Importantly, the changes in depth occurred even though the horizontal elements were still physically connected to the vertical line segments. These results suggest that contrast similarity is a strong cue to grouping that can induce perceptual closure in the presence of connectivity: closure was strongest when contrast properties of the connecting lines matched the vertical lines. The combination of similarity and element connectedness echo Palmer and Rock's (1994) 'uniform connectedness' principle, but I do not propose that these attributes form entry-level units, nor do the data speak to this possibility. Rather, the current experiment has provided the first evidence that the degraded depth effect is tied to figural closure and suggests that perceptual closure is graded in nature. In the following experiments, I further explored the features that drive the strength of grouping, and as a consequence, the amount of depth perceived within an object.

# CHAPTER 4

## 2-D and 3-D Determinants of Perceptual Closure

### 4.1 Introduction

In Chapter 3, I showed that the amount of perceived depth in a rectangular object is modulated by perceptual closure. In a perceptual description of closure, closure is a non-local property that involves the combination of multiple grouping components. For example, the perception of the simple rectangle used in the preceding experiments is a global Gestalt, but is likely to be a synthesis of classic Gestalt cues and two-dimensional shape properties. Several studies suggested that 2-D perceptual closure depends on the presence of specific cues (Kovacs & Julesz, 1993; Elder & Zucker, 1993, 1998; Kimchi, 2000; Spehar, 2002; Conci, et al., 2007; Mathes & Fahle, 2006; Hadad & Kimchi, 2008). For example, Kimchi (2000) showed that closure of two vertical contours is dependent on the presence of corners formed at line intersections (see also Hadad & Kimchi, 2008). Moreover, Conci et al. (2007) specified that corner (or L-junction) components must be collinear in order to group as a closed object. However, it is not known how these findings might apply to stereoscopic stimuli, as this research was limited to the study of 2-D stimuli. To precisely characterize the role of closure in depth, it is important to evaluate the contribution of individual properties to the reduced depth phenomenon observed in these experiments.

In the following experiments, I used the isolated line and connected object stimuli (described in Section 2.2) as a baseline comparison. I systematically varied the stimulus properties in the connected object that may have contributed to perceived closure and measured their impact on depth magnitude estimates (verified by subjective ratings). The rationale was that any decrease in perceived closure should lead to a corresponding increase in perceived depth. Therefore, when closure occurs, depth estimates should remain at the reduced levels seen for the connected rectangles. However, when closure is eliminated, depth estimates would be the same as those obtained for the isolated lines condition.

## 4.2 General Methods for Experiments 4.1 – 4.4

**Observers** Eleven observers participated in Experiments 4.1 – 4.4. Two observers had participated in Experiment 3.1, but were not experienced stereoscopic observers. Nine observers had no previous experience with psychophysical tasks.

**Stimuli** In all configurations presented in this series of experiments, four identical vertical lines were positioned symmetrically about the midpoint of the display. The baseline comparison consisted of the set of four vertical lines (central test pair in isolation) compared to a ‘closed connected object’ version in which the central pair was fully connected by horizontal lines, as described in Section 2.2 for Experiment 2.1 (Figure 2.1A and 2.1B). The stimulus was white (59.1 cd/m2) on a mid-gray background (15.6 cd/m2). Each vertical line measured 3.2° x 0.03° and was laterally separated from its neighbour by 2.13°. The horizontal lines used between the vertical test pair all had the same width as the vertical lines (0.03°) but the length varied depending on the experimental manipulation. In each experiment, the configuration of the central pair was modified as outlined in the individual sections below.

**Apparatus and Procedure** The experiments were conducted using the mirror stereoscope described in Section 2.2. The viewing distance was set to 64cm and the monitor resolution was 1920 x 1200 pixels. At this resolution and viewing distance, each pixel subtended 1.45 min of visual angle. As in Experiments 2.1 - 3.1, observers were asked to assess the amount of depth between the central line pair. Measurements were made using a pressure-sensitive sensor as described in Section 2.2. One line of the target pair had a crossed disparity of 0°, 0.048°, 0.096°, 0.144°, 0.192° or 0.24°. In all experiments, the conditions were randomly interleaved and presented ten times per disparity. Trials were completed in two blocks in a single test session. Subjective ratings of closure for all configurations were obtained from the nine observers that participated in the depth magnitude experiment (see Section 2.2). The ratings stimuli were presented on the stereoscope at the largest test disparity and ratings were obtained monocularly (left-eye only) and stereoscopically.

## 4.3 Experiment 4.1

### 4.3.1 Introduction

Connectedness has been proposed as an important property of perceptual organization (Rock & Palmer, 1990), but is not necessary for the perception of closure (eg: Koffka, 1935; Gillam, 1975; Kovacs & Julesz, 1993; Elder & Zucker, 1993; Yen & Finkel, 1998; Pettet, et al., 1998; Kimchi, 2000; Tversky, Geisler & Perry, 2004; Hadad & Kimchi, 2006, 2008; Gerhardstein et al., 2012). Studies have demonstrated that perceptual closure can tolerate breaks in connectivity, however it may be dependent on alternative cues. Some studies specify that the location of the discontinuity is the critical factor in determining whether fragments are grouped (eg: Kimchi, 2000; Hadad & Kimchi, 2008; although see Elder & Zucker, 1993). Additionally, some researchers have suggested that the proximity between the closure-inducing image fragments influences the strength of closure (Gillam, 1975; Elder & Zucker, 1993, 1994; see also Conci, et al., 2007). In this experiment, the impact of connectedness and proximity on perceived closure was assessed by measuring depth percepts for stimuli with discontinuous horizontal lines. Stimuli were designed to evaluate the importance of the location and size of discontinuities (gaps) in mediating perceptual grouping.

### 4.3.2 Stimuli

The stimuli are shown in Figure 4.1. The baseline condition was re-tested to aid comparison. Two additional sets of stimulus configurations were tested in which connectedness was removed by introducing gaps along the horizontal lines at different locations:

> (A) Centred-gap (Figure 4.1 A, C): A portion of the connecting line was removed from the centre of both horizontal lines. This stimulus appeared as inward facing square 'brackets.'
>
> (B) Split–gaps (Figure 4.1 B, D): Equal portions of the horizontal lines were removed at the Ends of the horizontal lines. In this case, the horizontal lines did not physically connect with the vertical test lines. This stimulus will be referred to as a 'bridge' configuration.

Three gap sizes were assessed for each set of stimuli (0.5°, 1° and 1.5°) to evaluate the role of proximity. These gap sizes corresponded to 25%, 50% and 75% of the length of the original connecting horizontal line. Thus, for a given gap size, the total length of the horizontal lines was the same in corresponding configurations. In all stimuli, the horizontal lines traversed a linear path through depth between the disparate vertical test lines (which remained unchanged).

The stimuli described above were symmetric about both the horizontal and vertical axes. Mirror symmetry has been proposed as Gestalt organizing principle, and may influence object formation. In addition, the symmetric blank spaces through the object may be perceived as an occluder (see Gillam, Grove & Layden, 2010) and the lines may interpolate across the gap. To evaluate the effect of symmetry two additional stimuli were included (Figure 4.1 A, B, last row); in these figures the gaps (50% gap size) were offset to the left or right of the stimulus mid-line.

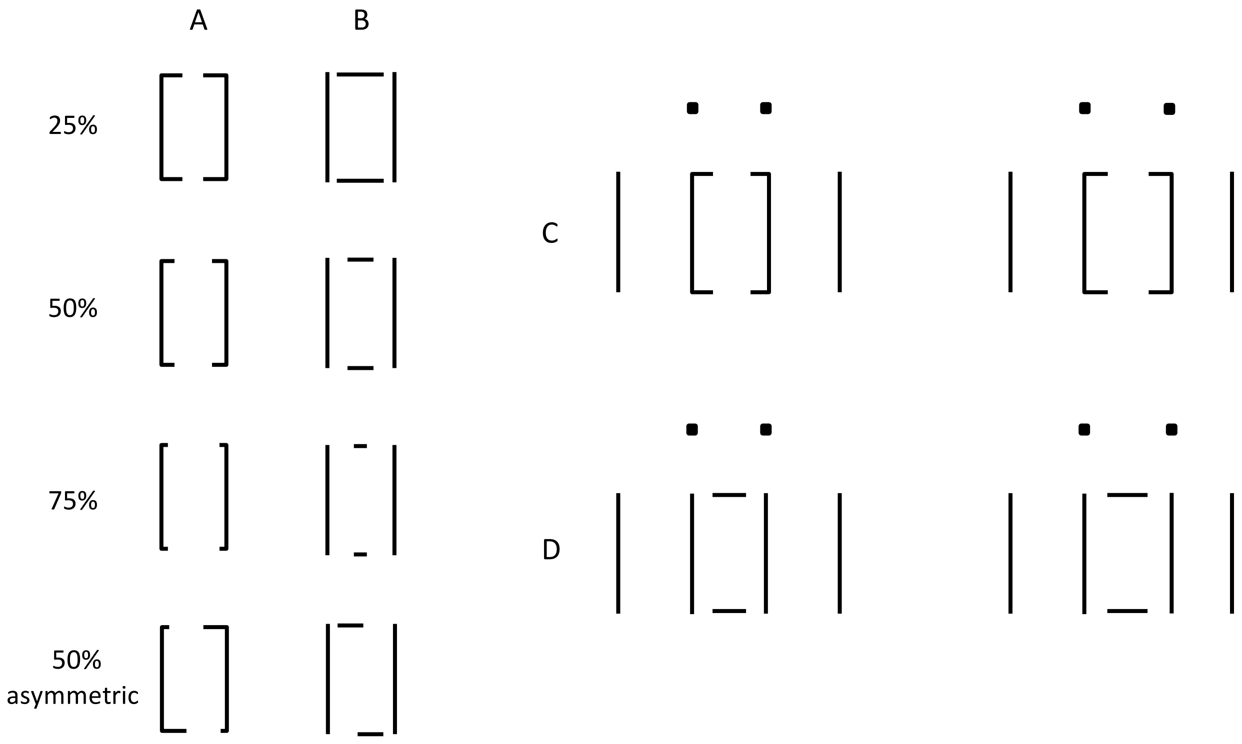

*Figure 4.1. Illustration of the stimulus configurations used in Experiment 4.1. Columns (A) and (B) show the two sets of stimulus configurations: (A) 'Bracket' configurations where the gap is positioned in the centre of the horizontal lines, and (B) 'Bridge' conditions where the gaps are distributed equally at the ends of the horizontal lines. Each row represents a different gap size, which is expressed as the fraction*

*of the horizontal contour removed: 25%, 50%, 75% and 50% asymmetric, corresponding to 0.5°, 1°, 1.5° and 1° asymmetric respectively. Figures (C) and (D) are stereograms arranged for cross-fusion, depicting one version of each stimulus (50% gap). In each stereopair, the right line of the central target pair has the same crossed disparity. In all trials, observers judged the relative depth of the central vertical lines.*

### 4.3.3 Results and Discussion

The depth estimates obtained for these stimuli are shown in Figure 4.2. For simplicity, the results for each configuration (brackets and bridges) are plotted separately. The depth estimates for the baseline condition (isolated lines and connected object) were included in both figures. Again, a repeated measures ANOVA confirmed the baseline pattern where the estimated depth within the closed object was much less than in the isolated vertical lines ($F(1, 10) = 36.73$, $p < 0.0001$; $\eta^2 = 0.78$).

As shown in Figure 4.2, there was a pronounced difference in the amount of depth perceived between the different types of configurations. When the gap was positioned in the centre ('brackets'), the amount of reported depth was very similar to the estimates obtained for the connected object (Figure 4.2A). There was no effect of gap size, showing that proximity and line length were not determinants of grouping for these stimuli. Moreover, the estimates for the asymmetric stimuli were also reduced, suggesting that symmetry (or matching gap location) did not contribute to figural closure. The reduced depth estimates are complimented by high ratings of figural closure for all gaps sizes (range = 7.1 – 9.1). Statistical analysis (repeated measures ANOVA) compared the connected object to all bracket configurations. This revealed no effect of Configuration ($F(1, 15) = 1.21$, $p = 0.31$) and no Configuration x Disparity interaction ($F(3,27) = 2.12$, $p = 0.125$). These results suggest that breaking connectivity in the centre of the object does not diminish perceptual closure.

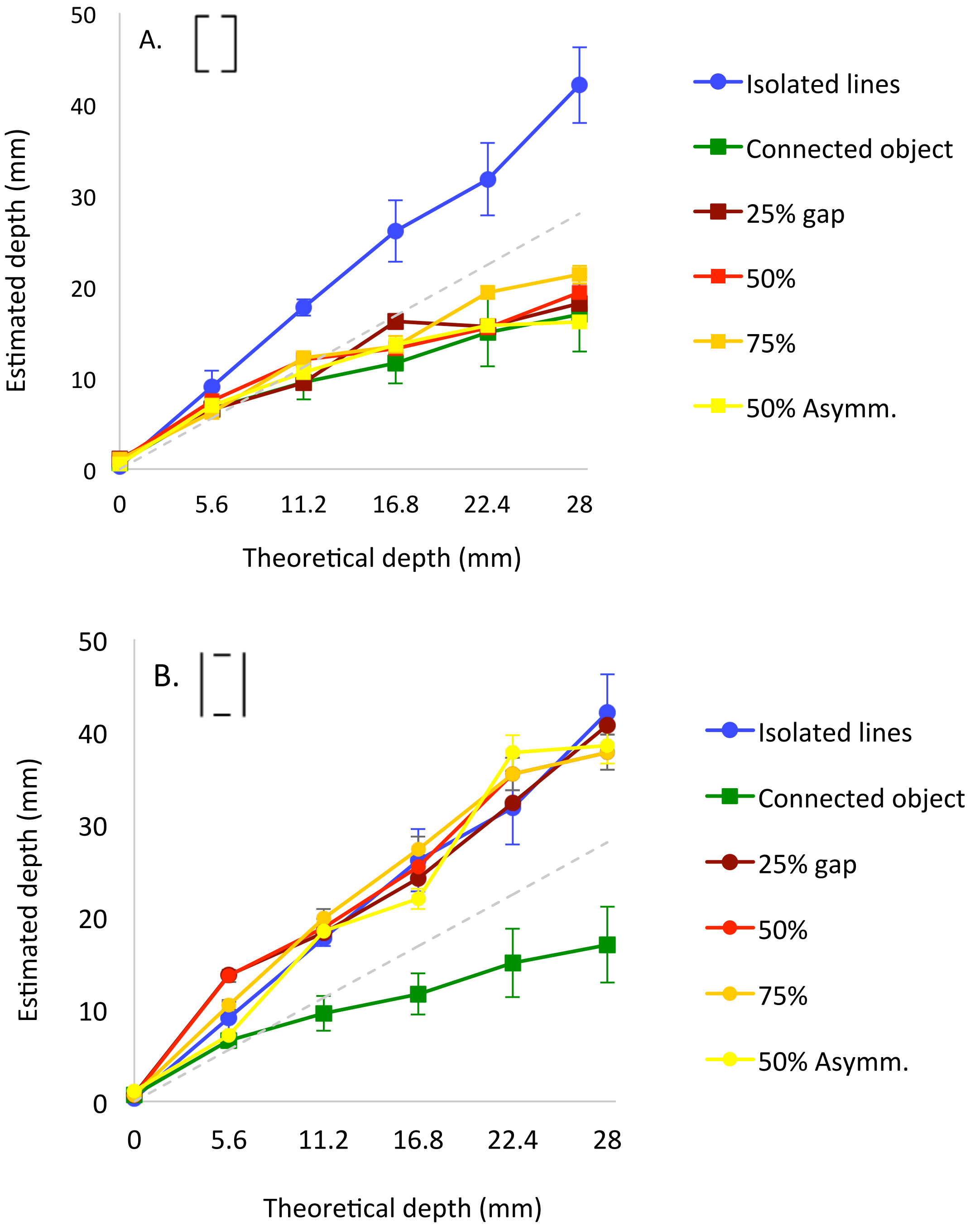

*Figure 4.2. Perceived depth as a function of depth offset (in mm) for Experiment 4.1 (n=11). The baseline comparison of isolated lines (blue) and connected object (green) is included in both figures. (A) Estimated depth reported for ‘bracket’ configurations possessing a centred gap. Each line represents estimates for one gap size, listed as the percentage of the stimulus width: 25%, 50%, 75% and 50% asymmetric (for*

*stimulus description see Section 4.3.2). (B) Depth estimates for 'bridge' stimuli, where the gaps (same sizes) were distributed at the corners of the rectangle. In each figure the theoretical depth (see Section 2.2) is shown on the abscissa, and the ordinate shows the estimated depth. The dashed line indicates a gain of one and error bars show ± one standard error of the mean.*

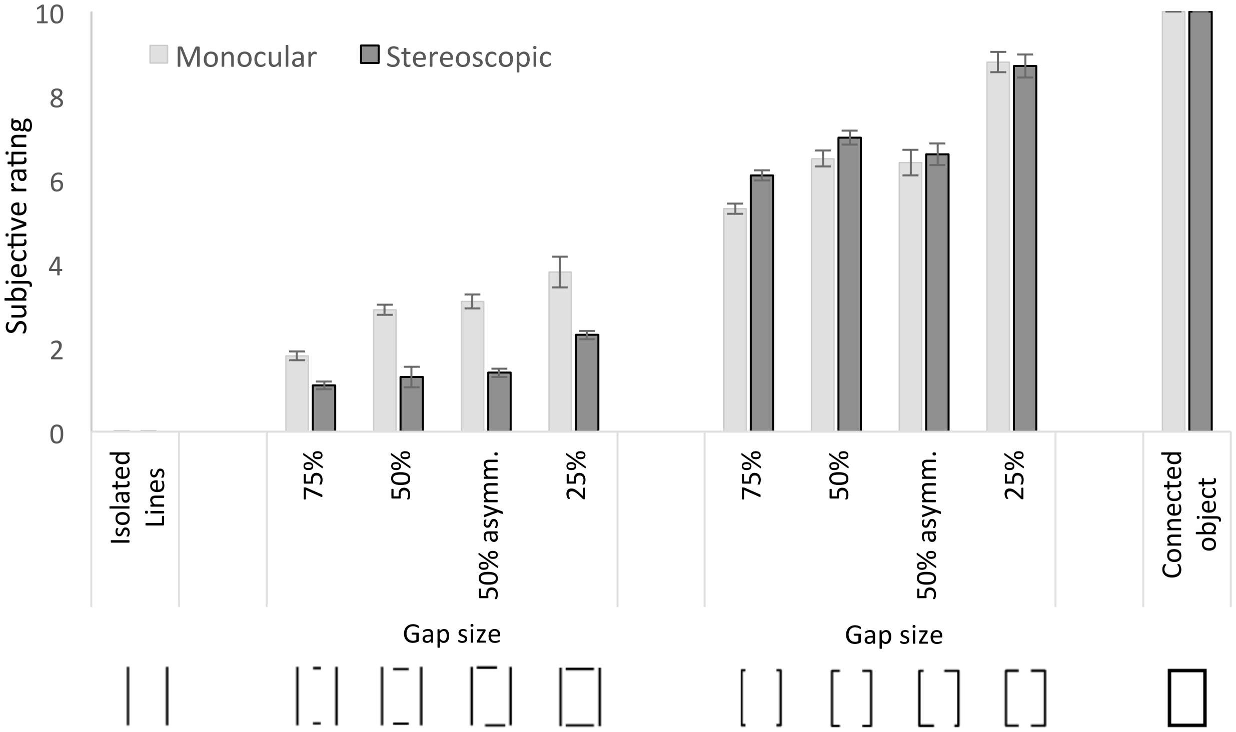

*Figure 4.3. Subjective ratings for the ten stimulus configurations used in Experiment 4.1 (n=11). Perceptual closure was rated on a scale from 0 (not an object) to 10 (a strong sense of a closed object). Ratings were obtained for monocular viewing and binocular (stereoscopic) viewing at the largest test disparity. The error bars represent ± one standard error of the mean.*

In the second type of configuration, the gaps were positioned at the ends of the connecting lines. As shown in Figure 4.2B, depth estimates for these stimuli were increased to levels reported using isolated lines. Again, neither proximity nor symmetry had an effect on the amount of reported depth for these stimuli. Statistical analyses compared the estimates reported for the isolated lines condition to all four bridge configurations. This analysis showed that there was no effect of Configuration (*F*(2, 14) = 0.87, *p* = 0.406) and no Configuration x Disparity interaction (*F*(4, 40) = 1.47, *p* = 0.227). The closure ratings for these configurations were much lower than those obtained for both the connected object and bracket stimuli (range = 1.5 – 3.6). Notably, the ratings for the stereoscopic stimuli were lower than the monocular 2-D versions. This suggests that the addition of disparity to the bridge stimuli enhanced the perceived segmentation between the vertical lines**,** akin to a perceptual 'repulsion' effect (see Westheimer & Levi, 1987). However, depth estimates for the bridge stimuli were never higher than the estimates reported for the isolated lines.

The results of Experiment 4.1 show that the observed reduction in depth for closed objects does not require that the components be spatially connected. Rather, the position of the horizontal lines relative to the vertical targets is critical. That is, for the stimuli used here, closure was only disrupted when gaps were located at the point of intersection between the vertical and horizontal lines. Strong closure percepts were maintained so-long-as the intersections were intact. The presence of horizontal contour information at the end-points of the vertical lines created L-junction corners, and these components appear to play an important role in mediating perceived closure, and consequently relative depth, in these stimuli.

The importance of L-junctions has been previously studied in 2-D stimuli and implied in 3-D stimuli. Researchers have directly shown that L-junctions are important to 2-D figural closure. For instance, Hadad and Kimchi (2008) showed that rapid grouping of vertical line stimuli was only observed when L-junctions were present in the stimulus; when gaps were located at the object corners, figural grouping was eliminated. Saarinen & Levi (1999) showed that contour closure enhances the accuracy of shape recognition: when the global shape was composed of four inward-facing corners, a shape was rapidly identified, but recognition was degraded when the same components were oriented outward. Similarly, detection of a closed object in visual search displays made up of L-junctions is optimal when the target is composed of inward-facing collinear L-junctions and the distractors are open configurations (Donnelly, et al., 1991; Conci et al., 2007). In stereoscopic displays, McKee (1983) showed that introducing a gap at one side of the rectangle slightly improved thresholds (see Figure 1.1A). This configuration would be consistent with a decrease in perceptual closure as it only possessed 'closure-inducing' L-junctions at one side of the figure. Similarly, using the McKee (1983) stimulus, Mitchison and Westheimer (1984) showed that breaking connectivity in the centre of the rectangle – maintaining L-junctions at both sides of the figure – resulted in similar thresholds to a fully connected object (see Figure 1.1B). Thus, the presence of L-junctions appears to contribute to the reduction in perceived depth between the vertical line pair.

The results of this experiment suggest that L-junctions are critical features to perceptual closure. However, it is possible that any type of junction can disrupt depth from disparity. McKee (1983) showed that re-positioning the connecting lines between the two vertical lines increased stereoacuity for the vertical pair to the same degree as the connected rectangle (see Figure 1.1A). That 'ladder' configuration contained T-junctions, therefore figural closure may simply require that the horizontal and vertical lines intersect. In the next experiment, I assessed the importance of the nature of the junction formed by the horizontal connectors to the perception of closure.

## 4.4 Experiment 4.2

### 4.4.1 Introduction

In the configurations employed in the preceding experiment, it appeared that local L-junction information (formed at the end-points of the target lines) was necessary for the perception of closure. Alternatively, any type of junction between the horizontal and vertical lines may be sufficient to create a closed object. Classic Gestalt theory assumes that boundary and region (interior) information contribute equally to grouping (Koffka, 1935). However, there is psychophysical evidence that region-based connections provide a very weak cue to perceptual closure. For example, in Elder and Zucker's search experiments, contours with connections at the centre of the line pair – including luminance-defined lines and random-dot patterns – resulted in longer search times and degraded object recognition compared to stimuli with boundary-defined support (Elder & Zucker, 1993, 1994, 1998).

These observations may be tied to structural rules for object formation, where distinctions are made between the role of L- and T- junctions to boundary assignment. An L-junction typically forms a right-angle corner of a shape, whereas, T-junctions tend to signify occlusion and depth order (for images that lack stereopsis). As a result, it has been shown that T-junctions tend not to be perceived as integrated units; rather the two components are interpreted as belonging to distinct objects which happen to overlap (eg, Clowes, 1971; Huffman, 1971; Lowe, 1987; Nakayama, et al., 1989; Kellman & Shipley, 1991; Tse & Albert, 1998; Rubin, 2001; Kellman, 2001; Gillam, et al., 2010).

In Experiment 4.2, the role of local corner information was assessed using stimuli with either four L-junctions or four T-junctions. If L-junctions were critical to perceptual closure, then their removal would disrupt grouping and enhance depth estimates. On the other hand, if the presence of any type of junction or discontinuity at the vertical lines promotes grouping, then configurations with region-based T-junctions whould also result in reduced depth percepts.

### 4.4.2 Stimuli

Three configurations were tested and are shown in Figure 4.4 D-F. These stimuli were adapted from the '50% bracket' stimulus (1° gap) that was used in Experiment 4.1 (see Figure 4.1A). As shown above, this stimulus reduced depth percepts to the same degree as the connected object configuration (Figure 4.2A), and provided a baseline comparison here. The three T-junction stimuli were differentiated by the distance of the horizontal line fragments from the end-point of the vertical lines:

(D) 10% offset: The four horizontal lines were positioned 0.32° from the ends of the vertical line. This offset corresponded to 10% of the stimulus height. The vertical separation between the horizontal lines was 2.56° (80% of stimulus size).

(E) 40% offset: Each horizontal line intersected the vertical lines at a distance of 1.28°from the end-points. The horizontal lines were vertically separated by 0.64°.

(F) No collinearity: On each vertical line, one horizontal line fragment was positioned at 10% (0.32°) and the other line at 40% (1.28°). This removed collinearity from the configuration.

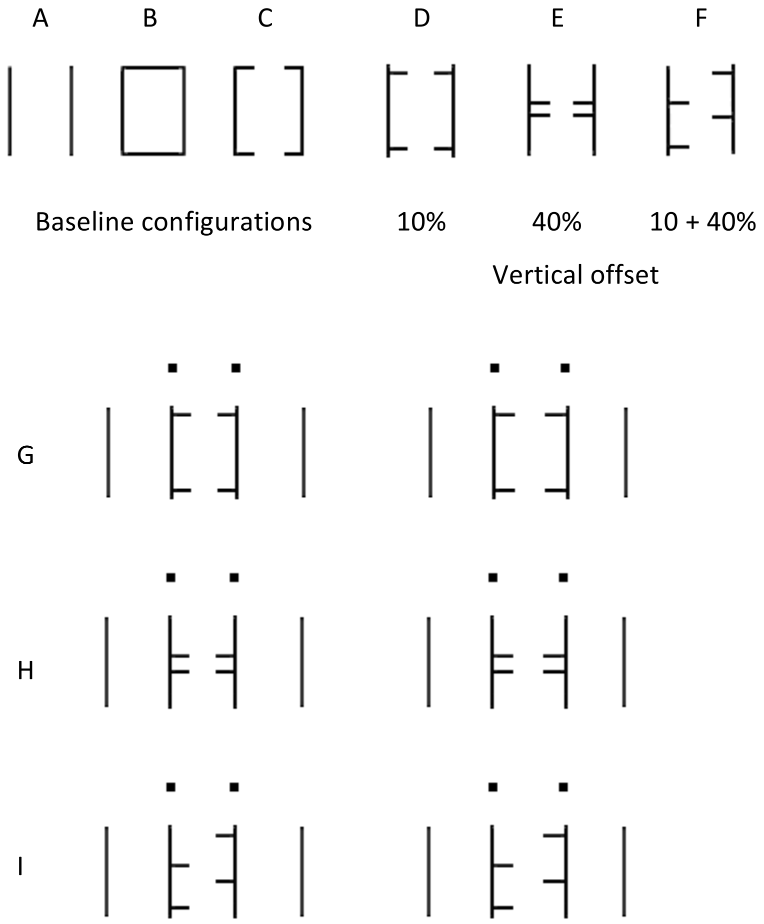

*Figure 4.4. Illustrations of the stimuli used in Experiment 4.2. Stimuli A – C are the isolated line, closed object and bracket stimuli that were re-tested to aid comparison. In configurations D – F the positions of the horizontal lines were varied to remove corners. Vertical offsets are expressed as the fraction of the stimulus height. (D) 10% (0.32°) from vertical line end-points; (E) 40% (1.28°) from end-points; (F) Collinearity is removed by positioning the lines on each fragment one each offset. Figures G - I are stereograms versions of stimuli D - F arranged for cross-fusion. In each of the stereopairs, the rightmost line of the central target pair has the same crossed disparity.*

### 4.4.3 Results + Discussion

Figure 4.5 shows the depth estimates obtained when the positions of the horizontal lines were varied to remove L-junctions. Again, the estimated depth between the isolated lines was significantly higher than observed in the connected object configuration ($F(1, 10) = 37.42$, $p < 0.0001$; $\eta^2 = 0.71$). Depth estimates for the bracket version (possessing L-junctions) were also reduced and were not significantly different from the connected object ($F(1, 10) = 0.66$, $p = 0.435$; $\eta^2 = 0.62$). This pattern of results was consistent with the report in Experiment 4.1 (Figure 4.2A).

First consider the results where the T-junctions were collinear (Figure 4.4 D & E). This change to the stimulus resulted in higher depth magnitude percepts relative to the connected object condition (and L-junction condition). However, the amount of depth remained less than estimates for the isolated lines. Depth estimates were very similar for both versions of the stimulus (10% and 40% vertical offsets), which showed that the distance of the intersection from the end-points was not a factor in depth estimation. The subjective ratings (Figure 4.6) were in the mid-range of the scale for both collinear configurations (10% = 4.7; 40% = 4.5 for stereoscopic presentation). These ratings confirmed that disrupting L-junctions weakened, but did not eliminate, perceptual closure: the intermediate closure ratings are consistent with the fact that perceived depth did not return to the level of the isolated lines. However, when collinearity between the horizontal lines was also disrupted (Figure 4.4F), the results showed that figural grouping was eliminated. In this case, the amount of perceived depth was increased to match the estimates reported for the isolated lines ($F(1, 10) = 0.29$, $p=0.603$) and the closure ratings were dramatically reduced (0.83).

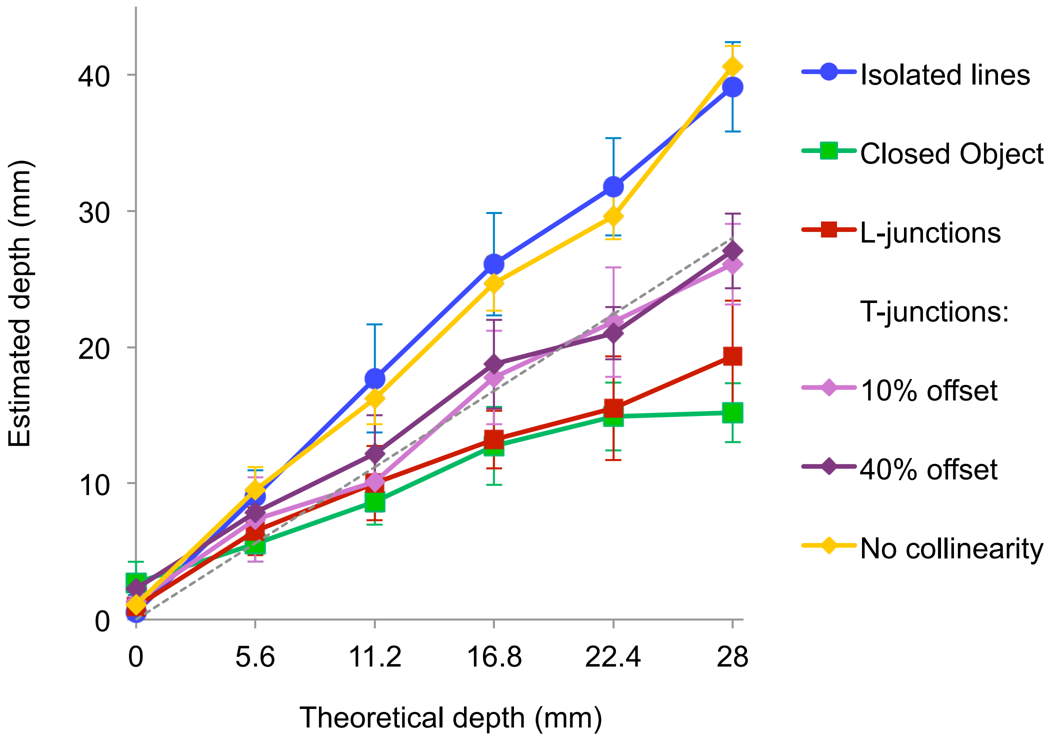

*Figure 4.5. Perceived depth as a function of depth offset for six stimulus configurations. Isolated lines (blue), closed object (green) and L-junction bracket (red) form the baseline condition. L-junction corners were removed by repositioning the horizontal line segments along the vertical lines to create T-junctions (see Section 4.4.2). Light purple lines represent the estimates for the 10% vertical offset. Dark purple depicts estimates for the 40% vertical offset. The yellow points represent estimates where collinearity was removed. The abscissa shows the theoretical depth, and the ordinate shows the estimated. The dashed line indicates a gain of one and error bars show ± one SEM.*

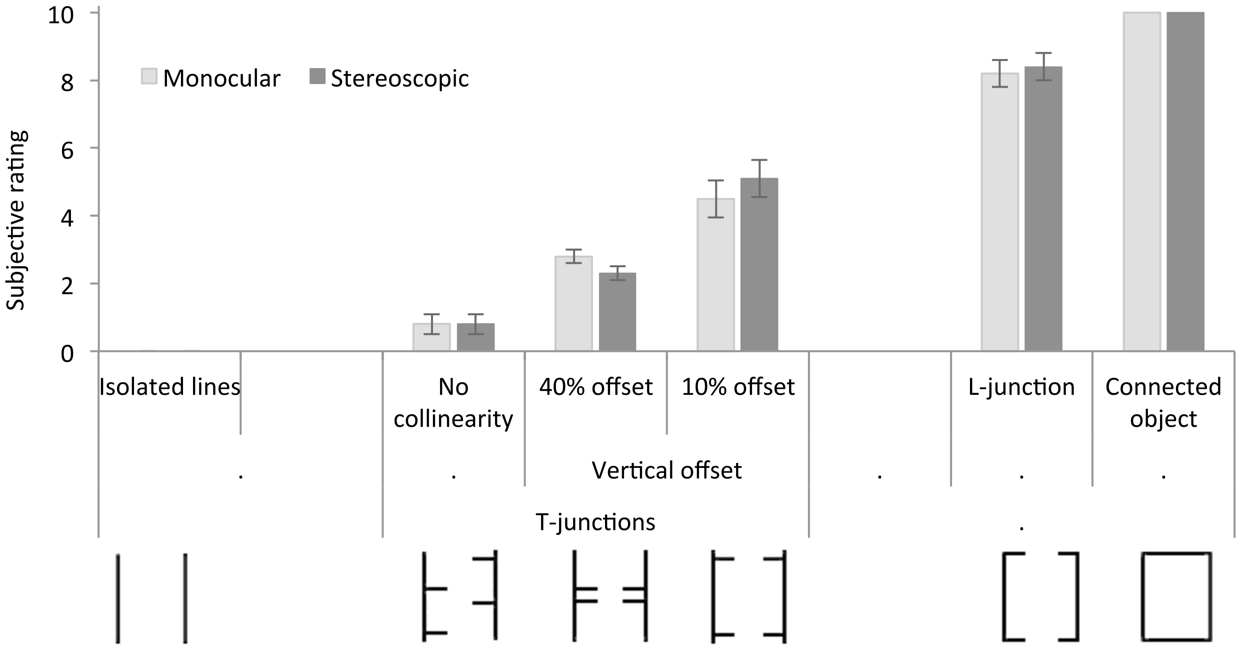

*Figure 4.6. Subjective ratings for the six stimulus configurations used in Experiment 4.2 (n=11). Ratings were on a 0 – 10 scale, where low figure ratings indicate weak perceived closure and high ratings indicate a strong sense of closure. Results are shown for ratings obtained for monocular viewing and stereoscopic binocular viewing at the largest test disparity. All error bars represent ± one SEM.*

Statistical analysis of all configurations revealed that the main effects of Configuration, F(4, 19) = 22.58, $p < 0.0001$, $\eta^2 = 0.69$ and Disparity, F(4, 28) = 266.82, $p < 0.0001$; $\eta^2 = 0.96$, and their interaction F(2, 21) = 12.23, $p < 0.0001$; $\eta^2 = 0.55$ were all significant. Depth estimates for all T-junction configurations were significantly different from the estimates obtained for the original closed object at the p = 0.05 level (10% offset: F(1,10) = 14.06, p =0.04; 40% offset: F(1,10) = 14.22, p =0.043; no collinearity: F(1,10) = 11.15, p <0.0001). Furthermore, estimates for both collinear conditions (Figure 4.4 D, E) were significantly lower than the estimates reported for the isolated lines (10% offset: F(1,10) = 12.87, p =0.004; 40% offset: F(1,10) = 13.29, p = 0.003), however, depth estimates for the non-collinear version were statistically similar to the estimates for the isolated lines (F(1, 10) = 19.76, p = 0.087).

The results of Experiment 4.2 showed that changing the position of the horizontal connectors relative to the vertical lines by even 10% of the stimulus height significantly reduced perceptual closure. However, perceptual closure was not completely eliminated when the T-junctions were collinear. Only when collinearity was disrupted (and L-junctions removed) was perceptual closure eliminated, suggesting that good continuation between L-junctions may be required to group image fragments. The increase in perceived depth in this non-collinear condition cannot be accounted for by the disruption of mirror symmetry, as the results of Experiment 4.1 showed that feature did not influence perceived closure for these line stimuli. Moreover, the increase in the separation between the horizontal fragments was unlikely to be a factor as I have also shown that proximity (up to 75% of lateral separation) does not differentiate depth estimates (Experiment 4.1, Figure 4.2).

## 4.5 Experiment 4.3

### 4.5.1 Introduction

The ability to connect visible contours across spatial gaps or occluders plays a crucial role in object perception. Collinearity (or good continuation) between contours has been shown to be an important determinant of grouping, in both 2-D (eg: Nakayama, et al., 1989; Hess, Field & Hayes, 1993; van Lier, 1999; Rubin, 2001; Elder & Goldberg, 2002; Shipley & Kellman 2003; Gillam et al., 2010) and stereoscopic configurations (Liu, et al., 1999; Yin, et al., 2000; Kellman, Garrigan & Shipley, 2005; Lu, et al., 2006). In the preceding experiment, the results showed that removing collinearity from T-junctions eliminated closure. Does the same result hold for disrupting the collinearity of L-junctions? In Experiment 4.3, I evaluated the importance of 2-D collinearity between the horizontal lines in promoting the percept of figural closure.

Previous research has shown that stereoscopic co-planar edges or contours must first align in the monocular image, otherwise grouping does not occur (Liu, et al., 1999; Yin, et al., 2000; Kellman, Garrigan & Shipley, 2005; Lu, Tjan & Liu, 2006). There is some evidence that perceptual grouping mechanisms can tolerate a small degree of misalignment: contour interpolation has been shown to break down when misalignment exceeds about 10 to 15 minutes of visual angle (Kellman, et al. 2005).

However, this threshold is likely affected by display and stimulus factors, such as viewing distance, stimulus size and the shape of the contours (eg: straight or curved).

### 4.5.2 Stimuli

Figure 4.7 shows the set of stimulus configurations that were designed to disrupt collinearity between the L-junction components. To achieve this, each central test line was shifted vertically in equal but opposite directions (counterbalanced across trials). Two vertical offsets were employed: 0.16° and 0.32° in each direction, for a total offset of 0.32° and 0.64°. These offsets correspond to 10% and 20% of the stimulus height respectively. The outer lines were not shifted and remained fixed for all conditions. As in the preceding experiments, the baseline condition (isolated lines and connected object) was re-tested. For the connected object, offsetting the vertical lines slightly increased the length of the connecting lines (2.1° at zero offset) and changed the angles at the corners from four right-angles to two obtuse and two acute angles. For this experiment, the line lengths increased to 2.12° and 2.18° for 0.32° and 0.64° offsets respectively and the corresponding 2-D acute angles were 81.87° and 74.05° (at zero disparity).

In addition to the baseline condition, two configurations were designed to evaluate the role of collinearity. The original 'L-junction bracket' stimulus was once again employed (Experiment 4.1 – 4.2) and collinearity was manipulated by altering the 2-D orientation of the horizontal lines, as shown in Figure 4.7:

(C) Collinear: Each horizontal line was oriented toward the end-point of the opposite line. The extension of the monocular image was a straight line.

(D) No collinearity: Each horizontal line was orientated perpendicular to the vertical line, forming a 90° L- junction component. For vertical offsets of 0.32° and 0.64°, the separation between the horizontal line end-points was 1.04° and 1.17° respectively (at zero disparity). The extension of the end-point of the left horizontal line to the right line would be doubly inflected.

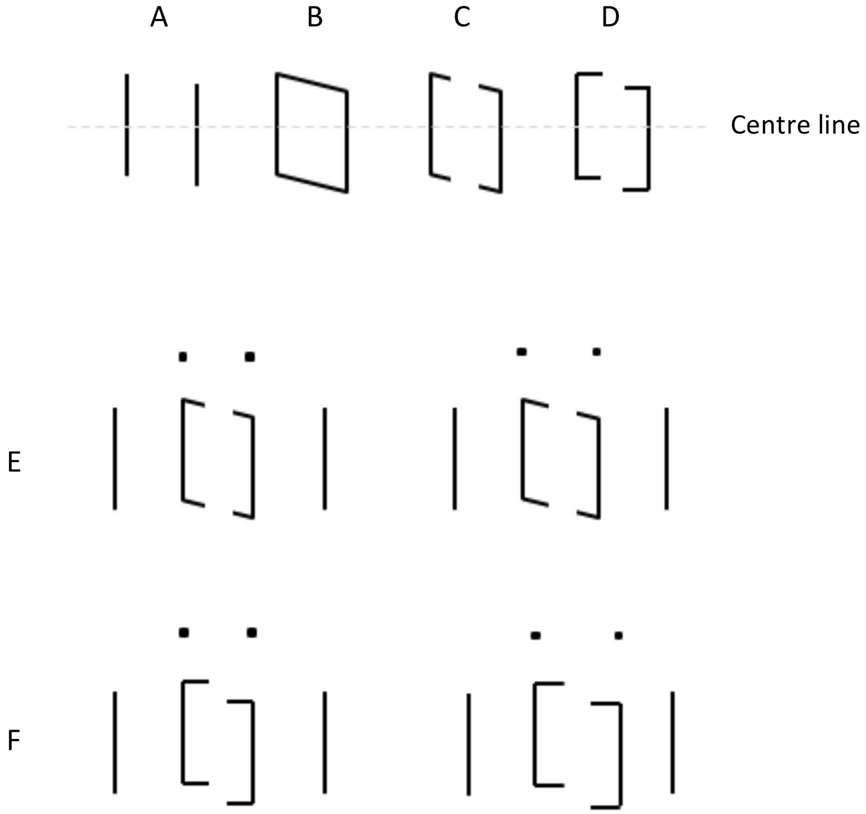

*Figure 4.7 Depictions of stimuli used in Experiment 4.3. In this experiment, the central vertical test lines were offset on the Y-axis by an equal amount in opposite directions. Two offsets were tested: 0.32° and 0.64° corresponding to 10% and 20% of the stimulus height. Isolated lines (A) and connected object (B) were modified and re-tested at each vertical offset. (C) L-junction stimulus possessing collinearity (D) L-junction stimulus with collinearity removed. Figures (E) and (F) are stereograms of stimuli (C) and (D) arranged for cross-fusion, with the right-most line of the central target pair containing the same crossed disparity. All stimuli were slanted on a continuous plane about the vertical axis.*

### 4.5.3 Results and Discussion

The configurations tested here were designed to assess the impact of the collinearity of L-junctions by offsetting image fragments on the vertical axis. Figure 4.8 plots the results of Experiment 4.3. Each sub-plot shows the results for each vertical offset (10% and 20%), and the associated baseline condition for each test offset. Statistically, the reduction in depth due to grouping remained at both vertical offsets: estimated depth for the connected object was much lower than the isolated vertical lines condition (10% offset: $F(1, 10) = 112.48$, $p < 0.0001$; $\eta^2 = 0.91$ ; 20% offset: $F(1, 10) = 134.88$, $p < 0.0001$; $\eta^2 = 0.93$).

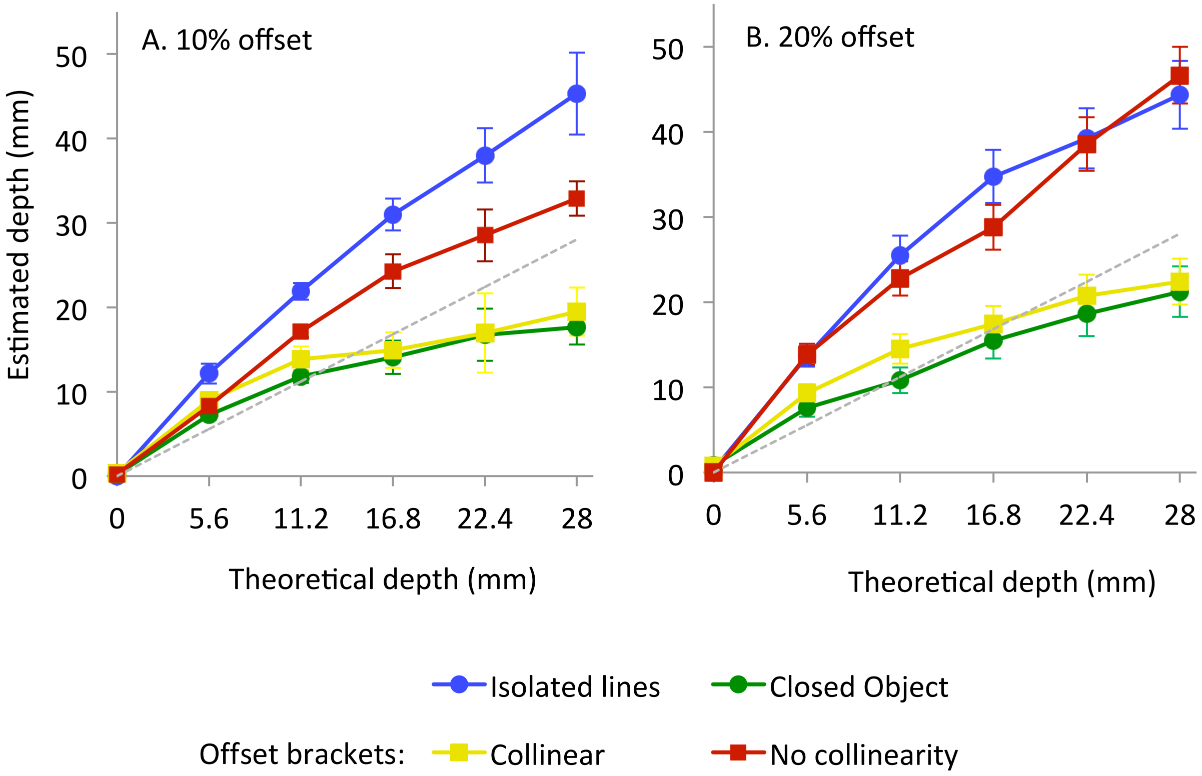

*Figure 4.8. Results for Experiment 4.3 where the impact of collinearity was assessed. In each figure, estimated depth is plotted as a function of depth offset (in mm). Each sub-plot shows the results for one vertical offset and includes the baseline comparison of isolated lines (blue) and connected object (green). (A) Vertical lines arranged with 10% vertical offset. (B) 20% vertical offset between image fragments. The theoretical depth is shown on the abscissa, and the ordinate shows the estimated depth. The dashed line indicates a gain of one and error bars show ± one standard error of the mean.*

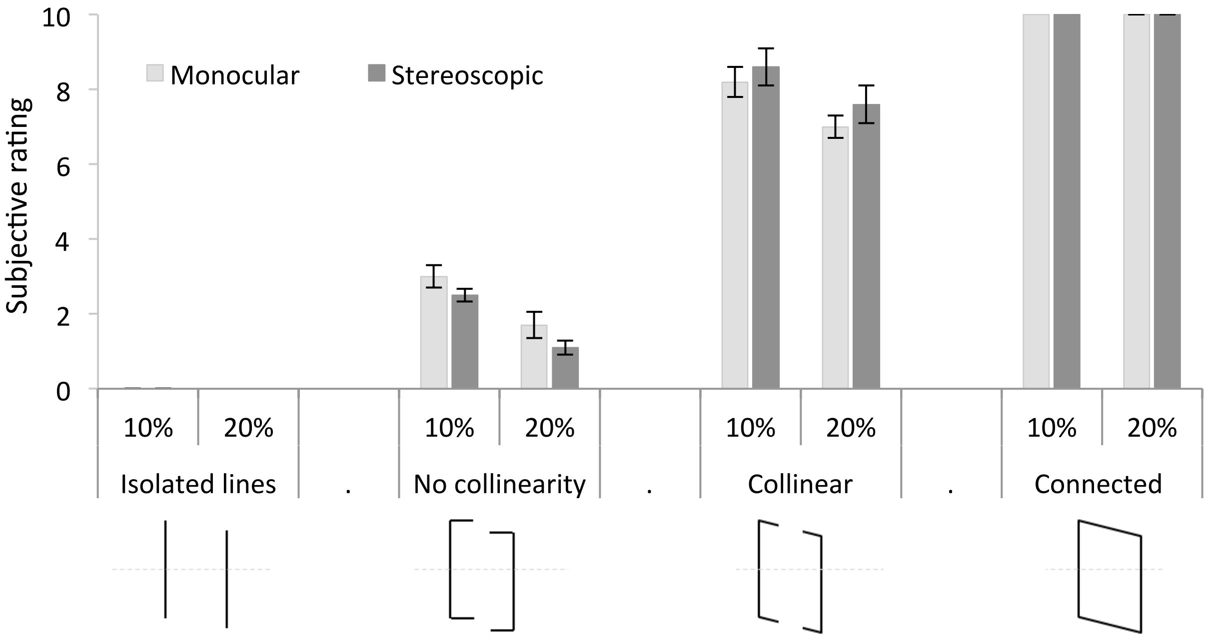

*Figure 4.9. Subjective ratings (0 – 10) for eight stimulus configurations (described in Figure 4.7). Ratings for 11 observers were obtained for a stimulus viewed monocularly and stereoscopically at the largest test disparity. Error bars represent ± one standard error of the mean.*

Results for the 10% vertical offset conditions are shown in Figure 4.8A. When the L-junction components are collinear, the amount of estimated depth was consistent with that obtained for the connected object. Strong percepts of closure were reported for this stimulus (8.0), confirming that collinearity between image fragments leads to figural closure. This result was first shown in Experiment 4.1 with no vertical displacement (Figure 4.2A). Now consider the configuration where collinearity was disrupted (illustrated in Figure 4.7D). This manipulation enhanced depth estimates to an intermediate level between the depth reported for the baseline condition, a change that was reflected by a decrease in the closure rating (2.7 for the stereoscopic version). These results suggested that a small deviation from collinearity can cause an associated reduction in grouping strength. That the increase in depth from misalignment was a graded effect was confirmed by the results obtained when the image fragments are offset by 20% (Figure 4.8B). Here, perceptual closure for the non-collinear condition was eliminated and depth estimates were restored to the level of the isolated lines.

Statistical analysis was run separately on each set of vertically offset stimuli. For the 10% vertical offset, the main effects of Configuration ($F(2, 15) = 65.92$, $p < 0.0001$, $\eta^2 = 0.73$) and Disparity ($F(1, 14) = 117.4$, $p < 0.0001$, $\eta^2 = 0.92$) were both significant, as well as the interaction ($F(3, 26) = 17.17$, $p < 0.0001$, $\eta^2 = 0.63$). The depth estimates reported for the connected object condition were not different from the estimates obtained for the collinear version ($F(1, 10) = 2.48$ , $p = 0.146$, $\eta^2 = 0.12$) but were significantly different from the non-collinear condition ($F(1, 10) = 32.43$, $p < 0.0001$, $\eta^2 = 0.76$). Moreover, results for the non-collinear condition were significantly different from those obtained in the isolated lines condition ($F(1, 10) = 202.58$ , $p < 0.0001$, $\eta^2 = 0.95$). The same analysis was applied to the 20% offset. Significant main effects were shown for Configuration ($F(2, 18) = 68.98$, $p < 0.0001$, $\eta^2 = 0.80$) and Disparity ($F(1, 14) = 52.34$, $p < 0.0001$, $\eta^2 = 0.83$), and their interaction ($F(3, 36) = 17.48$, $p < 0.0001$, $\eta^2 = 0.66$). Again, there was no difference between estimates reported for the connected object and the collinear configuration ($F(1, 10) = 3.72$ , $p = 0.077$, $\eta^2 = 0.25$) but there was a significant difference between estimates obtained when collinearity was removed ($F(1, 10) = 58.62$ , $p < 0.0001$, $\eta^2 = 0.81$). Estimates for the non-collinear version were similar to the isolated lines ($F(1, 10) = 1.79$ , $p = 0.21$, $\eta^2 = 0.14$).

Taken together, these results provided further evidence that collinear L-junctions are critical to the formation of a closed object between vertical lines. The results showed that a small disruption in collinearity (eg, 10% of the stimulus size) was sufficient to significantly reduce perceptual closure and enhance perceived depth. In this stimulus configuration, L-junctions were present but the horizontal lines were oriented away from a linear path of 2-D good continuation. However, the failure to completely restore depth estimates with this manipulation suggested that some degree of perceptual closure was maintained. Only with a larger offset, corresponding to 20% of the stimulus height, was perceptual closure eliminated and depth estimates restored.

It is well known that detecting collinearity (or good continuation) between fragmented contours is important to recover object boundaries. A gap or occluder between contours introduces uncertainty in determining which fragments should be grouped, especially in cluttered naturalistic scenes. Although the stimuli used in Experiment 4.3 do not present such an ambiguity, it is possible that collinearity is fundamental to figural grouping. Indeed, in models of interpolation (eg: Kellman & Shipley, 1991), fragmented contours will integrate only if when extended they form a smooth monotonic curve (although see Guttman, Sekuler & Kellman, 2003). According to Kellman and colleagues, the offset horizontal lines employed in this experiment would not interpolate because a double-inflection would be required to connect the end-points. Thus, these image fragments are interpreted as separate objects and do not cohere. On the other hand, the collinear lines would interpolate because the extension forms a smooth contour. This description of interpolation based on good continuation could account for the results presented here.

## 4.6 Experiment 4.4

### 4.6.1 Introduction

In the experiments presented to this point, I have established that closure-based Gestalt grouping is an important determinant of the perception of depth in simple figures. The addition of the horizontal line segments to the isolated line stimulus degrades depth percepts because the contours are transformed into objects via closure. This phenomenon appears to be critically dependent on the presence of collinear L-junctions. However, these grouping cues are two-dimensional properties. While it is clear that 2-D closure plays a key role in reducing perceived depth from disparity in these stimuli – and those of Westheimer (1979) and McKee (1983) – the impact of the disparity profile has not been considered. In the preceding experiments, the perceptual grouping manipulation necessarily introduced a within-object disparity gradient across all horizontal lines, specifically a smooth disparity profile. Here, I propose that the disparity profile itself contributes to the reduction in perceived depth via a form of good continuation in 3-D. Thus, the aim of Experiment 4.4 was to assess the hypothesis that 2-D good continuation has a disparity-based counterpart that also contributes to the reduction in perceived depth in closed figures. I call this disparity-based grouping principle ‘good stereoscopic continuation.’ If perceptual closure and the degraded depth effect are related to 3-D continuity, then introducing a depth step between image fragments should alleviate the depth reduction and increase estimates.

### 4.6.2 Stimuli

A square-bracket stimulus was created, as described in Section 4.3.2. The gap between the horizontal lines was 1°, or 50% of the distance between the vertical lines. The depth profile of the image fragments was manipulated to create two versions of the L-junction configuration, as shown in Figure 4.10:

> (A) Good stereoscopic continuation: The horizontal lines varied smoothly through depth according to a linear extrapolation between the vertical lines. This was the same as the stimulus used in Experiments 4.1 – 4.2.
>
> (B) Depth step: The horizontal lines were aligned on the same depth plane as the vertical line for a given test disparity. This gave the appearance of two fronto-parallel brackets separated by a depth step. This stimulus did not possess good stereoscopic continuation.

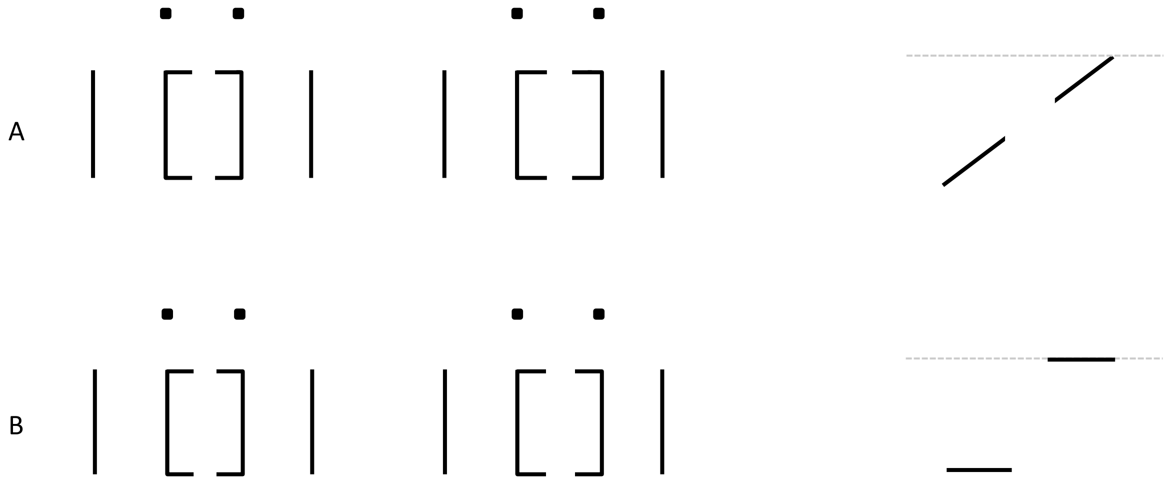

*Figure 4.10. Stereograms of stimuli employed in Experiment 4.4 to assess the impact of 3-D collinearity, termed 'good stereoscopic continuation'. Stimuli are arranged for cross-fusion and the right-most line of the central target pair has the same crossed disparity. The images on the right show the view-from-above of the disparity profile. (A) 3-D collinearity: Horizontal lines are slanted through depth on a*

*continuous disparity gradient. (B) Depth step: Image fragments are fronto-parallel and are positioned on two different depth planes to disrupt 3-D collinearity. Note that the baseline condition was also re-tested but is not included here.*

### 4.6.3 Results and Discussion

Figure 4.11 shows the results for Experiment 4.4. Once again, the disruptive effects of grouping on perceived depth was confirmed here: the amount of depth estimated in the isolated line condition was significantly different from that estimated in the connected object condition ($F(1, 10) = 38.13$, $p<0.0001$; $\eta^2 = 0.63$). The amount of depth perceived in each bracket condition critically depended on the disparity profile of the stimulus. For the stimulus possessing good stereoscopic continuation, the amount of estimated depth was very similar to the original connected object ($F(1, 10) = 0.83$ , $p = 0.41$, $\eta^2 = 0.81$). This result is consistent with the reduced depth effect observed in Experiment 4.1 for the same stimulus (see Figure 4.2). On the other hand, when good stereoscopic continuation was removed by creating a depth step between image fragments, the amount of perceived depth was increased to the levels reported for the isolated lines. In this instance, estimates were significantly different to the connected object ($F(1, 10) = 83.54$ , $p <0.0001$, $\eta^2 = 0.74$) and were similar to the isolated lines ($F(1, 10) = 0.93$ , $p = 0.55$, $\eta^2 = 0.67$). These results show that the degraded depth effect is determined, in part, by the orientation and relative position in depth of the image fragments. Perceived depth is low when the horizontal lines are slanted in depth, but high when the depth profile is abrupt.

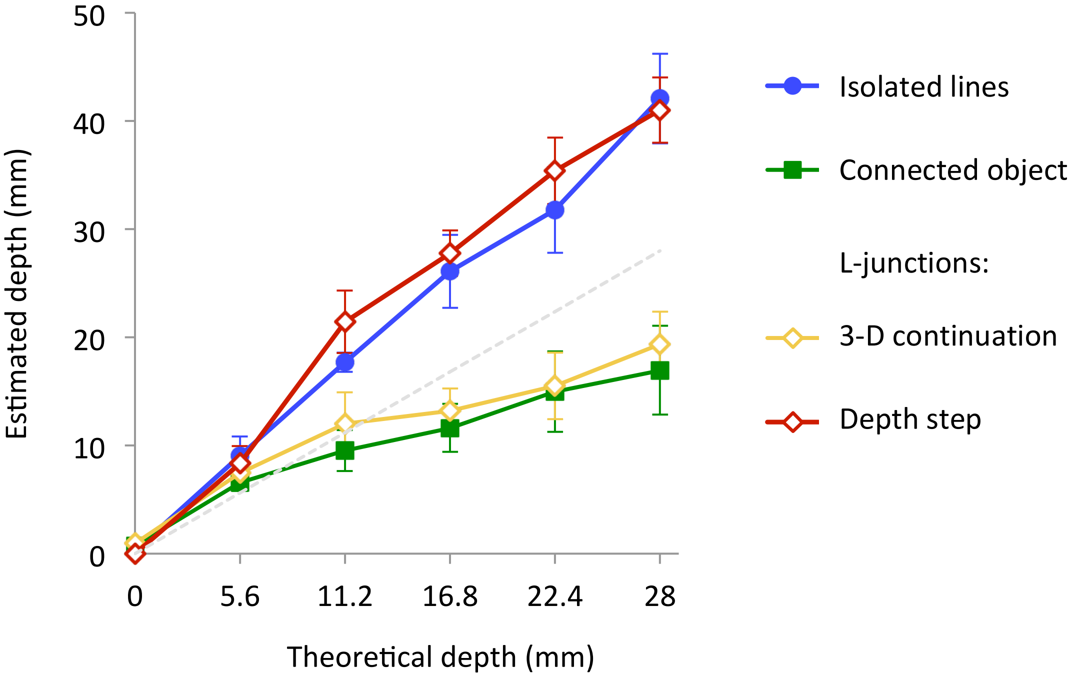

*Figure 4.11. Perceived depth as a function of depth offset for four stimulus configurations. The isolated lines (blue) and connected object (green) stimuli form the baseline condition. In two new configurations, the disparity profile of a square-bracket stimulus was varied (see Section 4.6.2 for stimulus description). The yellow points represents estimates for stimuli possessing 3-D continuation, and the red points depict estimates for conditions with a depth step. The abscissa shows the theoretical depth, and the ordinate represents the estimated depth. The dashed line indicates a gain of one and error bars show ± one SEM.*

It has been proposed that the visual system is insensitive to continuous changes in horizontal disparity and may be most sensitive to regions in which there is disparity discontinuity, such as a depth step, a gradient discontinuity, or a curvature of depth. Depth discontinuities may produce stronger signals than smooth disparity gradients because discontinuities are more informative (Gillam et al., 1984; Gillam, Chambers & Russo, 1988; Brookes & Stevens, 1989; Rogers & Cagenello, 1989), just as luminance discontinuities are more informative than luminance gradients. This may well be true, but cannot explain why depth percepts for some configurations with explicit representation of a continuous disparity gradient were not disrupted (eg: the bridge stimuli in Experiment 4.2). Rather, the results are consistent with the hypothesis that the degradation of depth for planar configurations is tied to perceptual closure via 2-D and 3-D grouping. Support for this interpretation can be seen in the subjective ratings for these configurations (shown in Figure 4.12 for both monocular and binocular viewing). When viewed monocularly, both versions of the square-bracket stimuli subjectively appeared as a closed object (good 3-D continuation: 8.6; two depth planes: 8.7). But when viewed stereoscopically, the ratings for the depth-step configuration were drastically reduced to 2.1. In contrast, ratings for the stimulus possessing good 3-D continuation remained high (8.5). Thus, disrupting good stereoscopic continuation appears to eliminate perceptual closure in depth. These observations show that object representation is governed by more than the 2-D cues present in the display – the organization of relative depth within the scene determines the perceptual outcome.

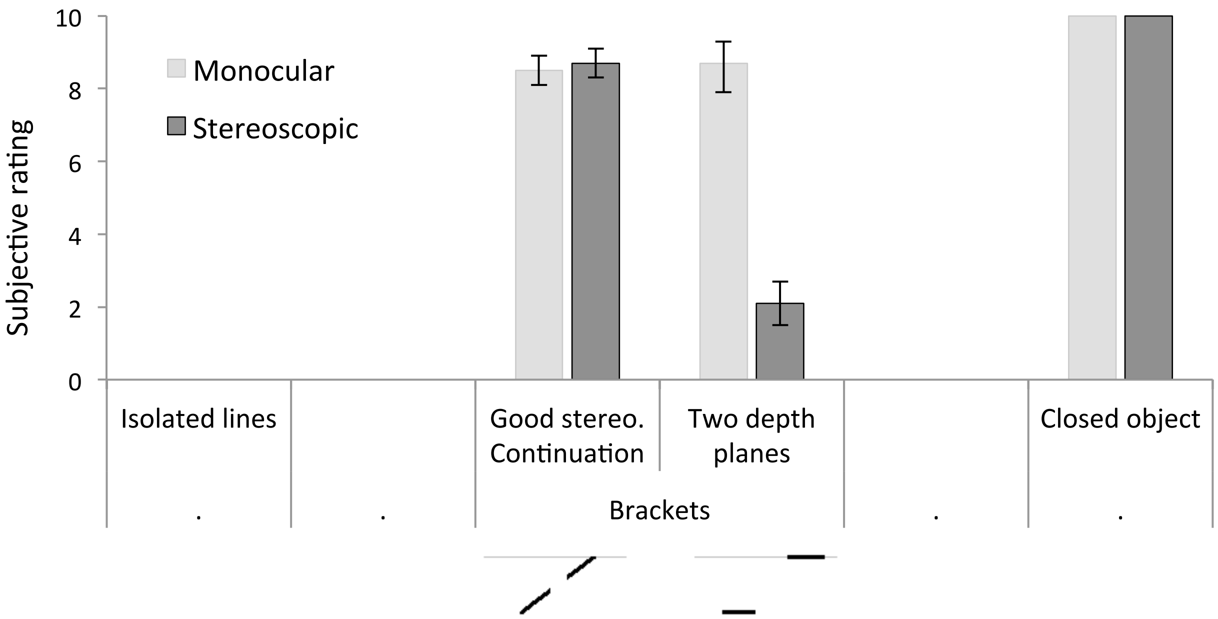

*Figure 4.12  Subjective ratings obtained for configurations with and without good stereoscopic continuation (in addition to baseline pair). Each condition was presented in a stereoscope at the largest test disparity and observers (n=11) rated the apparent closure (0 – 10) for monocular and stereoscopic viewing. Error bars represent ± one standard error of the mean.*

These results suggest that good stereoscopic continuation is a contributing factor in the degraded depth effect for a closed object presented in depth. Despite the constant 2-D shape, the amount of depth perceived within the stimulus was dependent on the disparity information in the horizontal line fragments. By introducing a depth discontinuity within the object – ie: a break in good stereoscopic continuation – the level of perceptual closure was dramatically reduced and the degraded depth effect was eliminated. The characteristics and implications of this disparity-based grouping cue are elaborated below.

## 4.7 General Discussion of Experiments 4.1 – 4.4

The preceding experiments were designed to systematically reveal the isolated contribution of 2-D and 3-D determinants of perceptual closure and their interaction with other configural properties. This work presents the first effort to simultaneously manipulate specific grouping components and validate depth percepts with subjective ratings. The results provide further support that the perceived separation in depth between a pair of vertical lines can be modulated by changing the degree to which they are perceived as a part of a closed object. Taken together, these results offer the first characterization of closure-based grouping through depth and highlight the role of perceptual closure and its interaction with other Gestalt grouping cues in determining the appearance of stereoscopic three-dimensional structure.

The reduced depth percepts reported in Chapters 2 and 3 were used as an indicator of grouping for the studies presented in this chapter. Depth magnitude was greatest in the isolated lines conditions and as the closed object interpretation was strengthened, the apparent depth decreased. Concurrently, observers were asked to quantify the degree of closure they perceived in each configuration. In all cases, there was a strong relationship between perceived closure and perceived depth. This correlation is illustrated as a scatterplot in Figure 4.13, where each point represents the average depth estimate reported at the largest test disparity as a function of subjective closure ratings for all stimuli assessed in Chapters 3 and 4. Linear regression analysis revealed that subjective ratings significantly predicted estimated depth ($\beta = -2.5$, $p<.0001$) and the two predictors explained 90% of the variance ($R^2=0.90$, *F*(1,30)=277.65, *p*<.0001). This correlation shows that the subjective ratings are a good predictor of the amount of depth that is perceived in these stimuli, and should be an important consideration when studying configuration and perception in depth. The graded changes in perceived depth and perceived closure are consistent with proposal that closure represents a perceptual continuum.

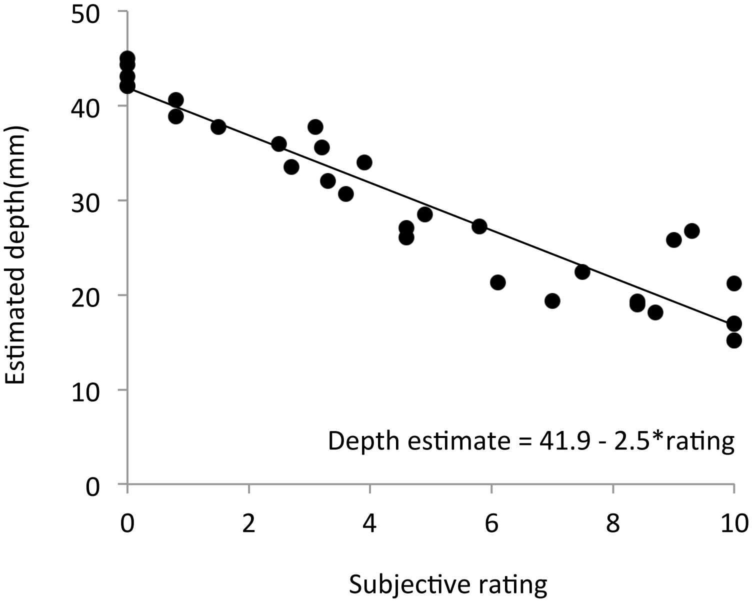

*Figure 4.13 Estimated depth (in mm) as a function of subjective figure ratings. Each point represents the average depth estimate obtained at the largest test disparity for each configuration employed in Chapters 3 and 4. The solid line shows the linear regression (see text for analysis).*

*Determinants of closure* In the preceding experiments, I separately manipulated several 2-D and 3-D stimulus properties to influence the degree to which the central line pair was grouped to form a closed figure. In doing so, I identified the cues that were necessary for perceptual closure. For the rectangular object assessed in these experiments, the inward orientation of the corner junctions (Experiment 2.1), the relative contrast of test lines to connectors (Experiment 3.1), the specification of L-junction continuations (Experiment 4.1 - 4.3), and the smooth transition through depth of the elements

(Experiment 4.4) all influenced the strength of grouping. While each of these components were important, none on their own were sufficient to generate a maximum level of closure. Instead, the interaction and global interpretation of these local components determined the representation of shape. Previous work has addressed the role of individual grouping properties in 2-D object formation: L-junctions (eg: Donnelly, et al., 1991; Saarinen & Levi, 1999; Conci, et al., 2007), contrast similarity (Elder & Zucker, 1993; Spehar, 2002) and collinearity (eg: Gillam, 1975, Shipley & Kellman, 1991; Kimchi, 2000; Hadad & Kimchi, 2008; Gillam, et al., 2010) have all been individually shown to be important to closure-based grouping in 2-D. But this was the first study to systematically assess the role of these properties in the representation of a stereoscopic 3-D object.

The presence of four collinear L-junctions, in particular, was shown to be important in promoting closure. The organization of such components has been shown to benefit performance in a number of measures using 2-D stimuli, including visual search (Donnelly, et al., 1991; Conci et al., 2007), modal completion (eg: Kanisza, 1976; Kellman, Yin & Shipley, 1998; Rubin, 2001; Anderson, Sing & Fleming, 2002) and object recognition (Elder & Zucker, 1998; Saarinen & Levi, 1999). According to a number of studies regarding contour completion, L-junctions are required to initiate contour interpolation and their removal eliminates or markedly reduces interpolation (eg: Shipley & Kellman, 1990; Albert & Hoffman, 2000). In other words, if a closed contour can be recovered through interpolation, then a closed object will be perceived and interpreted. Such an interpolation process may follow the constraint that two contours will interpolate if, when extended, they form a smooth monotonic uninflected contour (Kellman & Shipley, 1991; Field et al. 1993; Singh & Hoffman, 1998). However, segmented components might result if there is not enough visible contour information for successful interpolation or if contour fragments violate the constraints of interpolation. Thus, when I disrupted collinearity in Chapter 4 – in both 2-D (Experiments 4.2 – 4.3) and 3-D configurations (Experiment 4.4) – this criterion would have been violated since any continuation of horizontal lines on the left into lines on the right would be doubly inflected. Hence, the L-junction components do not group in the case.

The importance of collinearity through depth led me to identify a new grouping principle, which is based on the nature of the disparity gradient within an object. This disparity-based grouping cue – 'good stereoscopic continuation' – specifies the continuous gradient that is present within the intermediate stimulus features. Critically, this disparity-based grouping cue is contingent upon grouping in the monocular image (via 2-D cues). This was obvious given that reduced depth percepts from grouping did not occur for isolated lines, which had the same relative disparity in all conditions (for a given test disparity). Moreover, I have shown several examples of configurations that possess horizontal lines that traverse a linear path thorough depth and yet depth estimates were not disrupted. In these instances, the 2-D grouping constraints of collinear L-junctions were not satisfied (eg, the bridge stimuli in Figure 4.2B). Thus, 3-D continuation cue requires explicit 2-D input specified by either a connected contour or interpolation between horizontal fragments.

Taken together, these results show that perceptual closure requires that 2-D *and* 3-D constraints be satisfied. In cases where the image fragments meet both 2-D and 3-D constraints, the relative apparent separation in depth is dramatically reduced, compared to cases in which the configuration violates the constraints. If either the 2-D or 3-D properties are absent, then grouping does not occur and estimates are similar to relative disparity estimation of the isolated lines. Based on the specificity of the results, it is tempting to posit a hierarchy of processing for recovery of depth from objects. However, careful and considered efforts have been made to isolate and order the stages of depth perception with no definitive conclusion (Rock & Brosgole, 1964; Marr, 1982; Palmer & Rock, 1994; Palmer, Brooks & Nelson, 2003). Rather, it is more likely that depth processing is a fluid interaction, one that is not strictly hierarchical but involves feedback and feed forward interactions between processing levels (Lee & Mumford, 2003).

The interaction of 2-D and 3-D cues on stereopsis has been shown in previous work, including reports that stereoscopic thresholds are elevated when isolated components form an object with amodal completion (Liu, et al., 1999; Yin, et al., 2000; Hou, et al., 2006). For example, the perceived slant difference across two interpolated slanted rectangles can be reduced when the fragments are collinear (Yin, et al., 2000; Fantoni, et al., 2004, 2008; Kellman, et al., 2005; Liu & Schor, 2005; Hou et al. 2006).

Yin et al. (2000) showed that, in addition to collinearity, surface similarity (a 2-D property) is also an important attribute for grouping. Notably, contour interpolation is disrupted – and depth is restored – if one contour breaks away from the continuous plane through depth (via misalignment, opposite tilt/twist or opposite inclination). Even small amounts of misalignment substantially reduce completion effects (Kellman, et al., 2005). These reports are consistent with the results at suprathreshold level presented here, in so-far-as the perceived depth for a 2-D grouped object is dependent on the alignment in depth. However, the impact of amodal completion on depth percepts has been shown to be directly tied to the interaction of the inducing contours with a physical occluder, ie: the geometry of the interpolation is influenced by the properties of another intersecting surface. For the same pair of planar fragments, Fanton et al (2008) showed that the slant degradation in the presence of an occluder is lost when the occluder is removed. The stimuli used in Chapters 2 – 4 did not possess a physical occluder. Therefore, reduction in depth from grouping is an object-based phenomenon that occurs in isolation, without the interaction of other surfaces.

It is evident from the experiments reported in this chapter that both 2-D and 3-D attributes influence perceptual grouping. More specifically, I have shown that 2-D good continuation of L-junctions plays a critical role in closure-based grouping and, concurrently, discovered that this cue has a disparity-based grouping correlate (‘good stereoscopic continuation’) that may explain the degraded depth effect reported in these experiments. In the next chapter, I examine the perceptual consequences of good stereoscopic continuation in more detail.

# CHAPTER 5

## The Impact of Good Stereoscopic Continuation

The experiments in this Chapter were published in Journal of Vision in 2015: Deas, L. M., & Wilcox, L. M. (2015). *Perceptual grouping via binocular disparity: The impact of stereoscopic good continuation.* Journal of Vision, 15 (11), 1-13. doi:10.1167/15.11.11

Statement of contribution: L.M. Deas and L.M. Wilcox developed the study concept and design. Data collection and analysis was performed by L.M. Deas. Both authors collaborated on data interpretation and preparation of the manuscript.

### 5.1 Introduction

In Chapter 4, I proposed that good stereoscopic continuation was necessary for figural grouping in depth. When this disparity-based grouping information was absent, the disruption of perceived depth magnitude was eliminated. In previous experiments, I evaluated the impact of perceived closure on within-object depth percepts by manipulating both 2-D *and* 3-D grouping cues. In this chapter, I focused on the separate impact of good stereoscopic continuation in modulating grouping strength. To achieve this, I used a stimulus similar to that of Fahle and Westheimer (1988), who showed that stereoacuity for a pair of dots was elevated by as much as 100% when a third dot was added between them. It is possible that the threshold elevation seen in their experiments may be caused by disparity-based grouping. If so, then I would predict that disrupting the good 3-D continuation across the local dot elements would restore depth percepts. In addition, based on my previous hypothesis that the benefit of 3-D grouping is enhanced object cohesion, then breaking 3-D continuation should degrade subjective object ratings.

The benefits of a systematic continuous disparity cue have been shown in visual detection experiments. These experiments showed that in a volume of random elements, a smooth path of dots through depth was easier to detect than a jittered path (Uttal, 1983; Hess & Field, 1995; Hess, Hayes & Kingdom, 1997). However, these experiments did not explicitly manipulate good continuation while controlling for the depth of the end points; nor did they measure suprathreshold depth percepts. In Chapter 5, the 3-D cue of good stereoscopic continuation was assessed to determine its role in object perception. In addition to measuring depth magnitude for dot-paths, I conducted a visual search task to evaluate if disparity-based continuation had a positive impact on search time.

## 5.2 Experiment 5.1.1

### 5.2.1 Introduction

In this experiment, depth magnitude percepts were measured for a pair of target dots presented in isolation or as part of a multi-element linear path varying continuously through depth. Given the pattern of results seen from the experiments on figural closure (Chapters 2 – 4), I expected that two dots would be perceived as distinct objects, and that the addition of intermediate dots would lead to the perception of a single contour. If this is the case, then I predicted that perceived depth would be reduced as grouping strength increased. This pattern was observed by Fahle and Westheimer (1988), who showed that the discrimination threshold between a pair of small dots increased by the addition of one dot between the target pair. Moreover, thresholds continued to increase as more dots were added to the contour.

### 5.2.2 Methods

**Observers** Eleven observers participated in Experiment 5.1.1 and 5.1.2. Four were experienced stereoscopic observers with excellent stereoacuity (≤40arcsecs) and had participated in Experiments 2.1 and 3.1. Seven were undergraduate students with no prior experience with psychophysical tasks. These seven students were recruited based on a pre-screening assessment with the criteria of achieving at least 40 seconds of arc on the Randot$^{TM}$ stereoacuity test. All observers had normal or corrected-to-normal visual acuity.

**Apparatus** Stimuli were generated using the Psychtoolbox package (Brainard, 1997; Pelli, 1997) for MATLAB. Stereopairs were presented on a stereoscope comprising two LCD monitors (Dell U2412M) positioned at a viewing distance of 64cm. The monitor resolution was 1920 x 1200 pixels with a refresh rate of 75Hz. At this resolution and viewing distance, each pixel subtended 1.45 min of visual angle. The monitors were calibrated and matched prior to testing and the gamma functions linearized. Depth estimates were made using a purpose-built touch sensitive sensor (specifications are described in Chapter 2, Section 2.2). Each observer's inter-ocular separation was measured with a Richter digital pupil distance meter$^{TM}$.

**Stimuli** The stimulus comprised small (0.15° diameter) white (59.1 cd/m$^2$) dots arranged in a horizontal row on a mid-grey background (15.6 cd/m$^2$). At this size, each element's internal disparity gradient was not perceived. The path length was fixed at 2.6°, and the number of elements was varied from two (one at each end) to seven (see Figure 5.1). The baseline test stimulus was two dots, positioned symmetrically about the mid-point of the display and laterally separated by 2.6°. To create the 3, 5 and 7- element conditions, intermediate dots were added between the outer dots (maintaining equal element spacing in a given configuration). The element lateral separation for the 2, 3, 5 and 7 dot configurations were 2.6°, 1.3°, 0.65°, 0.43° respectively. On each trial, the outer test dots were presented at one of a range of disparities (0°, 0.04°, 0.08°, 0.12°, 0.16° and 0.2°). One dot had a crossed disparity and the other end had an equal amount of uncrossed disparity, such that the path was symmetric in depth about the zero disparity plane. The total disparity between the outer dots was 0°, 0.08°, 0.16°, 0.24°, 0.32° and 0.4° respectively. In the conditions with intermediate elements, these dots were positioned at a disparity determined by a linear extrapolation between the disparities of the two outer elements. The maximum disparity gradient (defined as $G = Db/Sb$, where $Db$ is the binocular disparity difference between the endpoints and $Sb$ is the binocular dot separation of the endpoints, averaged over both eyes) in the dot arrays was 0.17. At this gradient the stimuli are well within the fusable range, and much lower than the limit of 1 proposed by Burt and Julesz (1980). Note that for each disparity tested, the disparity gradient between the two end elements remained the same regardless of the number of intervening elements. Pilot testing prior to the main experiments verified that all disparities were within Panum's fusional area.

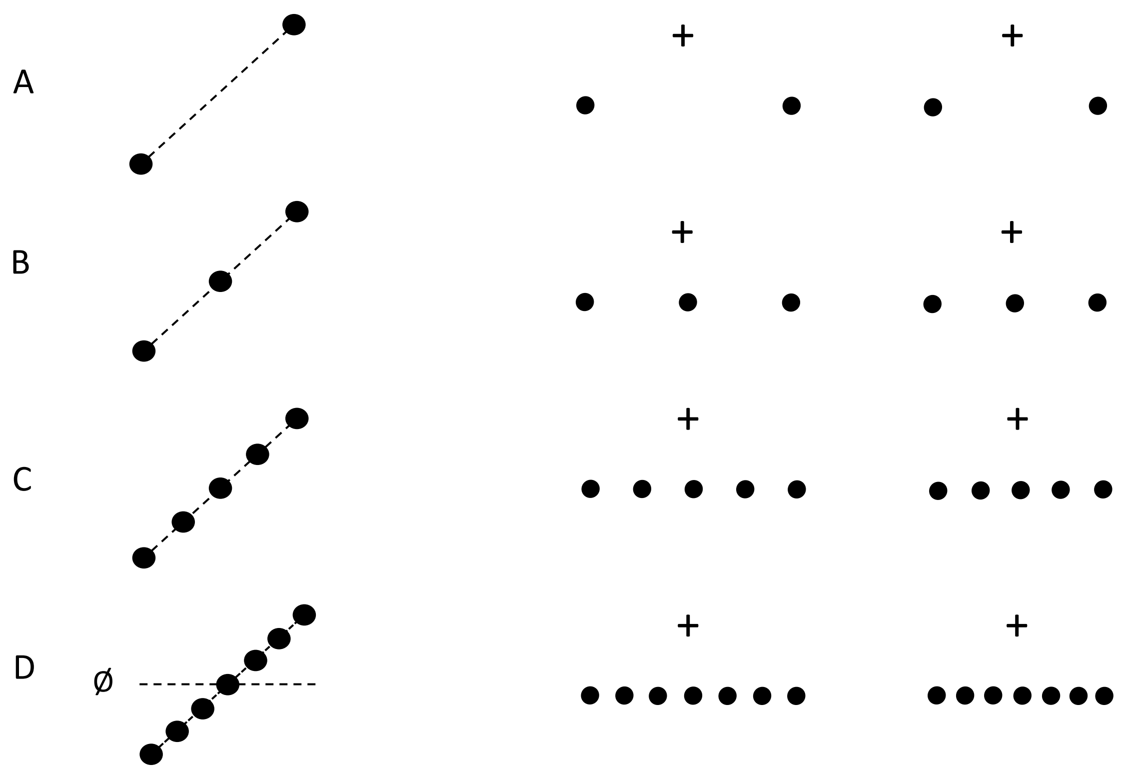

*Figure 5.1. Illustration (not to scale) of the stimulus configurations used in Experiment 5.1.1. Observers judged the amount of depth between the two outer dots in each condition. Each row depicts one condition, with systematic increases in the number of elements forming the contour. The first column shows the patterns as a view-from-above and the second column depicts stereograms of the stimulus configurations arranged for crossed fusion. The horizontal distance between the outer elements in each configuration was fixed and, for a given test disparity, the disparity gradient was constant for all path densities.*

**Procedure** Each trial started with a fixation cross (2° x 2°) that was presented for 750ms. On each trial, observers were asked to indicate the amount of depth they perceived between the two outer test dots (see Section 2.2 for description of magnitude estimation). Viewing time was unrestricted. Perceived depth was measured for four conditions: two (baseline), three, five and seven dots. The experiment consisted of 24 conditions (6 disparities x 4 configurations), with each condition presented 10 times in random order (5 times with left dot having crossed disparity and 5 times with right dot possessing crossed disparity). The 240 trials were completed in two blocks in a single session and the experiment took place in a darkened room. Prior to testing, observers completed a brief practice session of 30 trials to familiarize themselves with the depth estimation technique.

As in previous experiments, the stimulus disparities were converted to theoretical depth in millimetres for each experiment to allow comparison of on-screen angular disparity to observers' depth estimates. The average interocular difference (IOD) was used in each experiment (see Howard & Rogers, 2012, pp. 457). For Experiments 5.1.1 – 5.1.2, the theoretical depth between the two test lines corresponding to disparities of 0°, 0.08°, 0.16°, 0.24°, 0.32° and 0.4° degrees were 0, 9.5, 19.0, 28.5, 38.0 and 47.5 mm respectively (with average IOD = 60.1 mm for eleven observers).

In a follow-up session, subjective ratings of object cohesion were obtained for each stimulus in order to relate the stimulus manipulations to perceived depth. The eleven observers who had participated in the depth magnitude component were asked to rate each stimuli in terms of the 'extent to which the outer dots were perceived as part of one object'. Stimuli were displayed in the stereoscope at the largest test disparity (0.4°) and ratings were measured for each stimulus viewed monocularly (left-eye image) and binocularly.

### 5.2.3 Results and Discussion

Figure 5.2 shows the mean estimated depth for each configuration as a function of the theoretical separation in depth. As the disparity separating the target pair increased, estimated depth increased linearly for all configurations. However, the amount of depth reported depended on the number of dots in the configuration. Maximum depth was perceived in the two-dot condition, and in this case, closely followed theoretical predictions except at the largest disparity. In comparison, perceived depth was consistently and systematically reduced as more dots were added to the path. Note that the addition of one dot (3-element configuration) was sufficient to significantly reduce perceived depth of the outer elements and the disruptive effect increased with increasing dot number. Statistical analyses confirmed these observations. Because observers were told that some stimuli would have zero disparity, this response became stereotyped to the base of the sensor strip and had no associated variance. To avoid biasing the model fits, the zero-disparity estimates were excluded from analyses. A repeated measures ANOVA showed main effects of Configuration, $F(1, 16) = 32.59$, $p < 0.001$; $\eta^2 = 0.77$, and Disparity, $F(1, 14) = 83.35$, $p < 0.001$; $\eta^2 = 0.89$. The interaction between Configuration and Disparity was also significant, $F(4, 41) = 4.44$, $p = 0.004$; $\eta^2 = 0.31$. Simple effects analyses were used to compare the differences between configurations. These contrasts revealed that perceived depth for all conditions was significantly different at the $p = 0.01$ level.

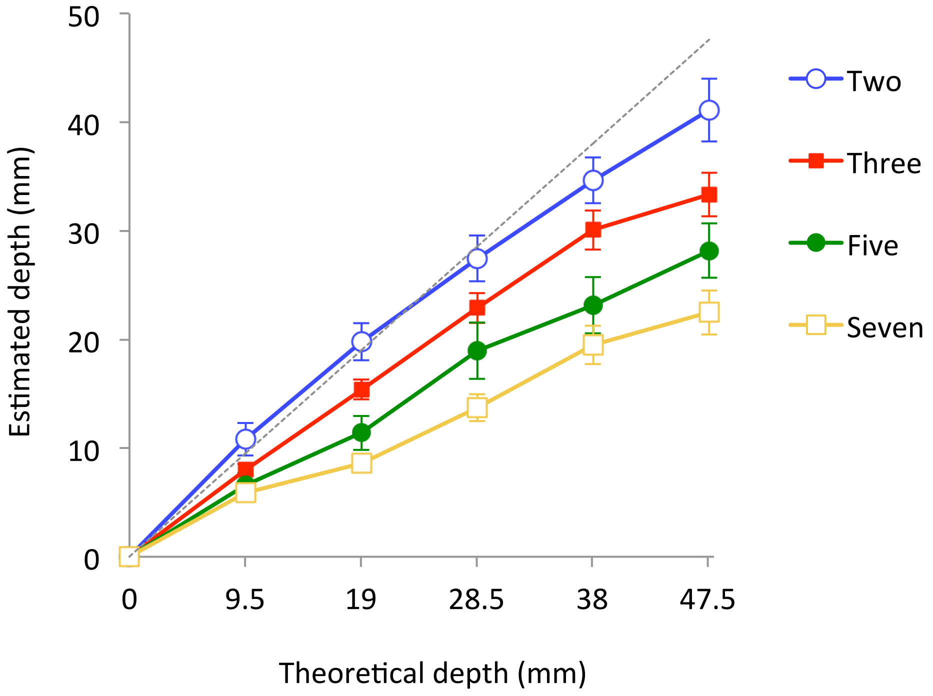

*Figure 5.2. Perceived depth as a function of the depth difference between the outer dot pair for four dot densities with smooth disparity gradients. Paths were composed of two (open blue circles), three (closed red squares), five (closed green circles) and seven (open yellow squares) dots. The abscissa shows the theoretical depth, and the ordinate shows the estimated depth. The dashed line indicates a gain of one and error bars represent show ± one standard error of the mean.*

The disparity gradient was necessarily affected by altering the relative disparity of the two end dots. However, for a given test disparity, the disparity gradient did not change with the addition of intermediate dots, as both the binocular distance (*Sb)* and the disparity difference (*Db)* remained unchanged. Despite the constant gradient, observers performed veridically in the two-dot condition but disparity was systematically misperceived with the addition of intermediate dots. Therefore, explanations based on loss of fusion due to violation of the disparity gradient limit could be ruled out. Rather, these results were consistent with the hypothesis that higher-level grouping mechanism underpin this phenomenon. In this instance, the addition of intermediate dots to the 2-dot condition changed the interpretation from distinct components to a single path via good continuation. This was reflected in the subjective ratings (depicted in Figure 5.3). Ratings for the 2-dot stimulus were consistently at zero (SD=0), suggesting that there was no sense of an object in this configuration. Note that this rating was consistent with those obtained for the isolated lines condition employed in Chapters 2 - 4. As intermediate dots were added to the stimulus, the object ratings systematically increased and were highest for the 7-dot contour (rating = 10, SD = 0 for stereoscopic viewing). These ratings showed that the strength of grouping in the dot-path was dependent on the amount of spatial support and, consequently, modulated the amount of depth perceived within the object. Taken together, I have shown that figural grouping via good continuation in depth negatively impacts the amount of depth perceived in a dot-path. I propose that this is directly linked to the path of good stereoscopic continuation – that disparity-grouping degrades depth. It is possible, however, that the addition of the intermediate elements alone interfered with depth estimates, and that this reduction in perceived depth was unrelated to the disparity relationships per se. In part two of this experiment, I evaluated this possibility.

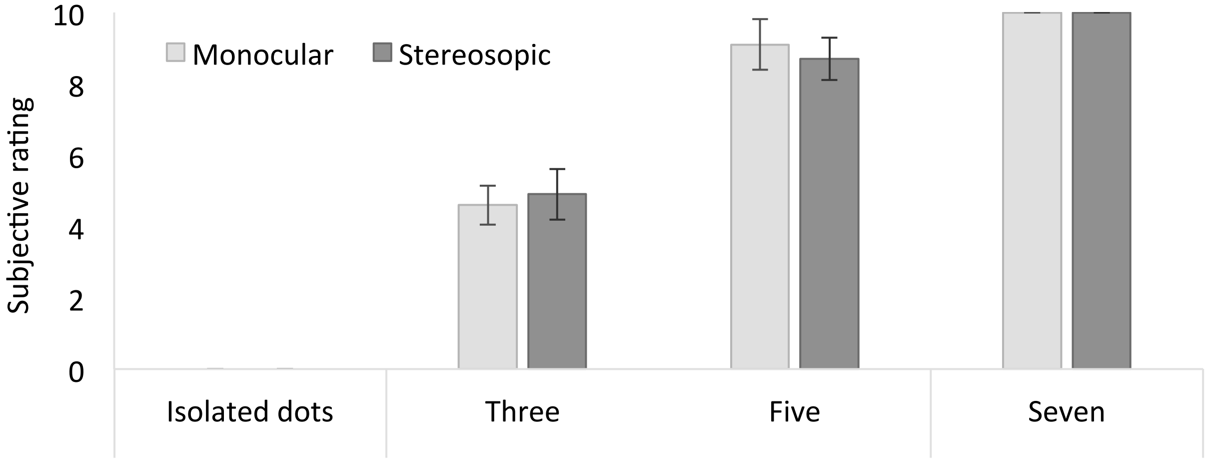

*Figure 5.3. Subjective ratings for the stimulus configurations presented in Experiment 5.1.1. Eleven observers rated stimuli presented in a stereoscope at the largest test disparity (0.4°). Each condition was rated for monocular and binocular viewing. Error bars represent ± one standard error of the mean.*

## 5.3 Experiment 5.1.2

### 5.3.1 Introduction

As outlined in Experiment 5.1.1, there was a significant reduction in perceived depth from disparity between two dots when intervening elements were added on a linear path through depth. This may be related to the operation of perceptual grouping via good stereoscopic continuation; alternatively, the reduction may simply be related to the 2-D cues present by virtue of more dots. In Experiment 5.1.2, I evaluated the importance of good stereoscopic continuation using a 5-dot configuration and varied the smoothness of the disparity profile. The benefit of using dots as the stimulus was that the magnitude of the jitter could be manipulated easier than with horizontal lines, therefore allowing quantification of different levels of jitter. By controlling the 2-D cues present in the dot path, I could be confident that an observed effect of good stereoscopic continuation was related to the disparity-relationships, perhaps disparity-based grouping. If good stereoscopic continuation was critical to the degraded depth effect in the preceding experiment, then jittering the disparity of the central elements should eliminate the effect and restore depth magnitude estimates to levels reported for isolated elements. However, if the reduction in perceived depth seen in Experiment 5.1.1 was due to the presence of additional elements, the pattern of results should be similar for all conditions.

### 5.3.2 Stimuli

Depth magnitude was measured for a 5-dot array with continuous and discontinuous disparity gradients, depicted in Figure 5.4. Two discontinuous disparity profiles were assessed, defined by different levels of jitter. The continuous disparity contour was generated as described in Experiment 5.1.1: intermediate dots were positioned in depth by dividing the disparity between the end points in equal parts. In the discontinuous conditions, the binocular position of each intermediate dot was randomly jittered in depth (in the Z-axis only) relative to the continuous plane. For ease of explication, disparity jitter is described relative to a linear interpolation in depth between the disparities of the end dots. The jitter step limit was 25% of the total disparity in each direction (crossed, uncrossed) which corresponded to half of the total disparity range in a given run (test disparities: 0.08°, 0.16°, 0.24°, 0.32° and 0.4°). Figure 5.4 illustrates the three conditions, the continuous disparity profile and example profiles for the discontinuous conditions (low and high). To generate low disparity jitter, each dot was repositioned by one ‘step’ on either side of the original position on the continuous plane. The position of each intermediate element was selected pseudo-randomly from disparities ranging from ± 25% of the test disparity (Figure 5.4B). To create high jitter, each dot was re-positioned by more than one step, but was constrained to not extend beyond the depth of the end dots (Figure 5.4C). In this instance, the disparity sign of the elements neighbouring the outer test dots would be reversed. In both jitter conditions, element position was determined according to these rules, with two additional constraints: no element was positioned along the original (linear) path and no dot extended beyond the disparity of the end dots. The disparity gradient limit was never violated within the path. Note that in all conditions, the path dot density was constant (ie: five identical dots aligned horizontally); the differences in the appearance of the contours arose from disparity alone.

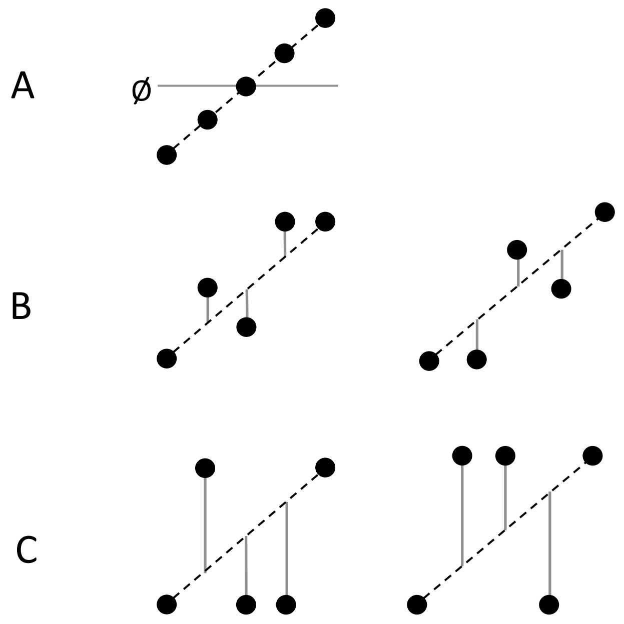

*Figure 5.4. A bird's eye view of the test conditions in Experiment 5.1.2. Observers estimated the depth between the outer test dots in horizontal contours of 5 dots. The dashed line indicates a linear path in depth between the two end dots. Three conditions were assessed, defined by different depth profiles. (A) Continuous disparity change (like that used in Experiment 5.1.1). (B) and (C) depict possible versions of jittered conditions, where dots were repositioned in depth according to the constraints outlined in the text. Solid grey lines represent the maximum displacement in depth (disparity jitter). (A) No jitter (B) 'Low' jitter, (C) 'High' jitter.*

**Observers and Procedure** Eleven observers participated in Experiment 5.1.2, all of whom had previously completed Experiment 5.1.1. Observers were asked to assess the amount of depth between the two outer dots as described in Experiment 5.1.1, for the same set of disparities (equally offset 0°, 0.08°, 0.16°, 0.24°, 0.32° and 0.4°). The three conditions were randomly interleaved for a total of 180 trials (3 conditions x 6 disparities x 10 trials per condition) and were completed in one test session.

### 5.3.3 Results and Discussion

Perceived depth (in mm) for the conditions in Experiment 5.1.2 is shown in Figure 5.5. First, the data obtained in the zero-jitter 5-dot condition and two isolated dots condition replicated the loss of perceived depth seen in Experiment 5.1.1 (Figure 5.2). These data are compared with the jittered disparity profile results also depicted in Figure 5.5 and showed that the amount of depth perceived within a simple dot path was contingent on its disparity profile. In all conditions, the path density was constant but the differences in disparity relationships modulated the amount of depth perceived in these stimuli. First consider the low jitter condition (depicted in Figure 5.4B), when dots straddled the line of continuous disparity by one step (less than 25% of the total disparity for a given trial). This change to the stimulus increased the magnitude of depth percepts relative to the continuous disparity condition; even this small amount of disparity jitter was sufficient to 'break' the disruptive effect. Displacing the interior elements by even larger increments (high jitter) resulted in a marked *overestimation* of the depth of the end dots, compared to two-dot condition. Statistical analyses revealed significant main effects of both Configuration, $F(2, 20) = 28.2$, $p < 0.0001$; $\eta^2 = 0.74$ and Disparity, $F(1, 15) = 50.87$, $p < 0.0001$; $\eta^2 = 0.84$, and a significant Configuration x Disparity interaction, $F(8, 80) = 6.51$, $p < 0.0001$; $\eta^2 = 0.39$. Contrasts (using simple effects analyses) revealed that estimates for the continuous-disparity configuration were significantly different from estimates reported for both the low jitter configuration ($F(1,10) = 18.50$, $p=0.015$) and the high jitter condition ($F(1, 10) = 60.12$, $p < 0.0001$).

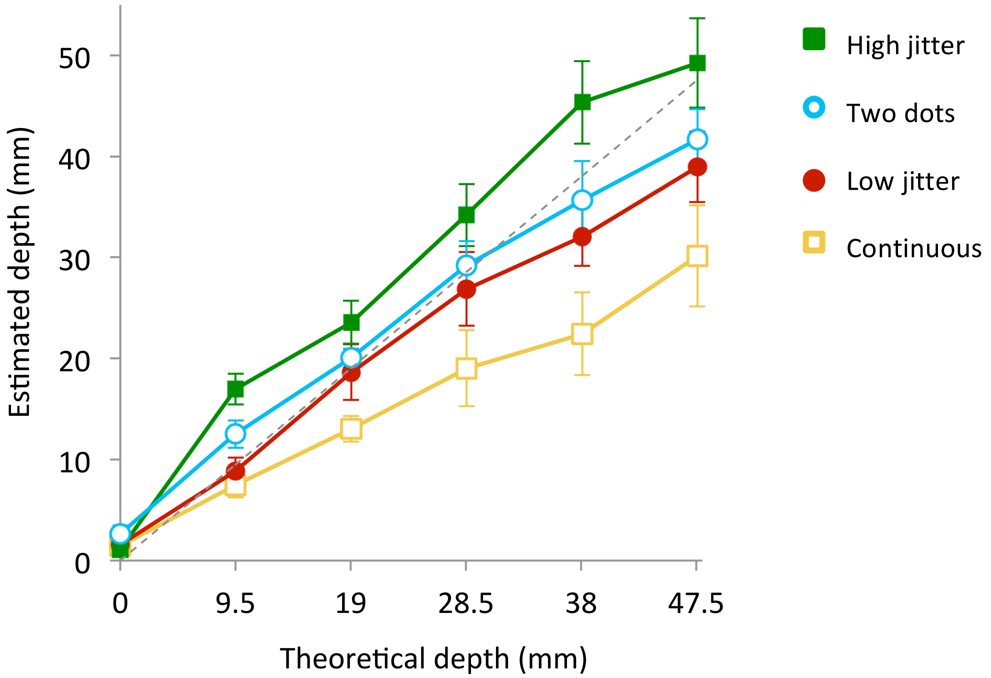

*Figure 5.5. Perceived depth for a 5-dot contour as a function of depth offset for three disparity profiles: continuous (orange open squares), discontinuous low jitter (red closed circles) and discontinuous high jitter (green closed squares). Perceived depth for two isolated dots (blue open circles) are included for comparison. The abscissa shows the theoretical depth, and the ordinate shows the estimated depth. The dashed line indicates a gain of one and error bars show ± one standard error of the mean.*

These results suggest that good stereoscopic continuation is a critical determinant of depth magnitude. For the same 5-dot stimulus the amount of estimated depth was dependent on the profile through depth. That this is related to disparity-grouping is shown in the subjective ratings for these patterns, shown in Figure 5.6. For all 5-dot paths, the monocular image rated very high for all observers (ratings >9), suggesting that a strong object percept was created. However, when the stimuli were viewed binocularly, the ratings were markedly different across conditions. The rating for the stimulus possessing good stereo continuation remained high (9.8). However, jittering the depth of the intermediate dots resulted in a marked decrease in ratings. This suggests the percept of a unitary object was weakened by the addition of disparity. These results confirm that perceptual grouping for the dot-contour is dependent on the level of good stereoscopic continuation.

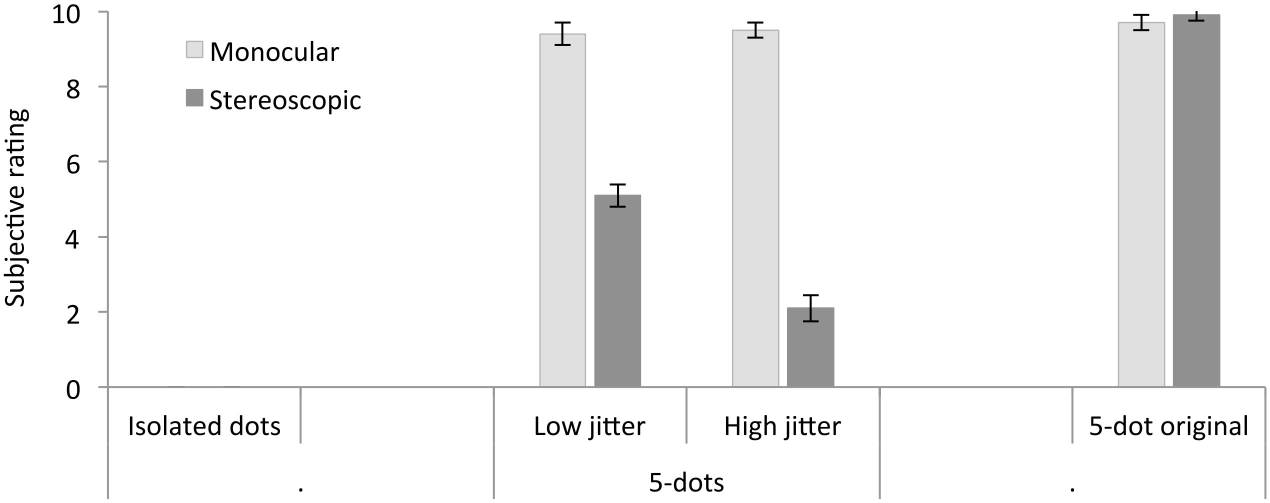

*Figure 5.6. Subjective ratings for the stimulus configurations presented in Experiment 5.1.2. 11 observers rated stimuli presented in a stereoscope. Each condition was presented at the largest test disparity (0.4°) and ratings were made for monocular and binocular viewing. Error bars represent ± one SEM.*

As shown in Figure 5.5, maximum depth was perceived for the high jitter condition and was overestimated compared to veridical (and the two dot version). This overestimation may be related to local depth differences between neighbouring dots. At the high level of jitter, the intermediate dots were positioned at or near the disparity of the outer test dots. While observers did not report this, it is possible that in this large jitter condition, the stimulus appeared more like two pseudo-transparent planes. If so, observers may have judged the separation in depth of these two planes rather than the outer elements. This may have produced exaggerated depth estimates as previous studies have shown that larger disparity gradients enhance perceived depth (Bülthoff, et al., 1991).

The separation of two dots in depth was underestimated when intervening elements were added on a linear disparity gradient; in this experiment I showed that displacing dots away from the linearly interpolated path in depth resulted in systematic increases in perceived depth. The effects of these manipulations on perceived depth were not related to local disparity relationships (such as disparity gradient) that are presumed to occur at early stages of visual processing. Because the outer dots remain at a fixed relative disparity for all disparity profiles, the disparity gradient between them did not change when jitter was applied to the intermediate dots. While the application of large amounts of jitter may introduce disparity gradient violations between neighbouring intermediate elements, the resultant diplopia should degrade, not enhance, perceived depth (Wilcox & Allison, 2009). In addition, given that other potential local influences, such as conflicting depth cues (perspective), assimilation, or normalization to the fronto-parallel plane should occur in all three configurations, these explanations cannot account for these results. Instead, these data are consistent with the previous results in showing that perceptual grouping, in this case via good stereoscopic continuation, is responsible for the reduction in depth percepts shown here.

## 5.4 Experiment 5.2

### 5.4.1 Introduction

The results presented in Experiment 5.1 suggested that visual information in a scene is perceptually organized in terms of the distribution of relative depth. I have proposed that this organization can be formalized as a disparity-based grouping cue called ‘good stereoscopic continuation.’ As outlined above, the depth magnitude and corresponding subjective rating data showed that the presence of this grouping cue promoted object cohesion at the cost of perceived depth. If good 3-D continuation is a cue to object perception, then the visual system could exploit this information to aid object detection. There is evidence that a continuous binocular disparity cue can improve detection of contours in random element displays (Uttal, 1983; Hess, et al., 1997) and that stereoscopic slanted surfaces are detected pre-attentively among fronto-parallel planes (Epstein & Babler, 1990; Holliday & Braddick, 1991; Sousa, Brenner & Smeets, 2009). While these studies provided important information regarding slant detection, they did not address the role of stereoscopic grouping nor evaluate suprathreshold depth percepts. To my knowledge, a visual search task has not been conducted presenting a continuously slanted stimulus in competition with a discontinuous version. The aim of Experiment 5.2 was to directly assess the impact of good stereoscopic continuation on the time required to detect a target array. A search experiment was designed to measure performance for displays where detection was based on the presence or absence of good stereoscopic continuation. I used the same 5-dot contour employed in Experiment 5.1.2 and care was taken to control for other potential cues to search; element density, disparity gradient/span, proximity and collinearity were all held constant. Based on the hypothesis that 'good stereoscopic continuation' is differentiable from 2-D grouping, the prediction was that this 3-D continuation would act as a stimulus feature that was capable of guiding rapid visual search.

### 5.4.2 Methods

**Observers and Apparatus** Nine of the observers that participated in Experiments 5.1.1 and 5.1.2 also participated in Experiment 5.2. The apparatus was the same as that described in Experiment 5.1.

**Stimuli** The test display contained one target and 5, 9 or 13 distractor stimuli (display sizes of 6, 10 or 14 stimuli), as shown in Figure 5.7. The stimulus array consisted of multiple sets of 5 elements arranged in horizontal rows. As in Experiment 5.1, each circular dot was 0.15° in diameter and the separation between the outer dots in each path was fixed at 2.6° and each intermediate dot was laterally separated from its neighbor by 0.65°. In these experiments the relative disparity of each outer dot was fixed at 0.24°. Recall from the results of Experiments 5.1.1 and 5.1.2 that at this disparity, observers consistently reported less depth when the stimulus had continuous disparity but made veridical settings when discontinuities were introduced (Figure 5.2, 5.5).

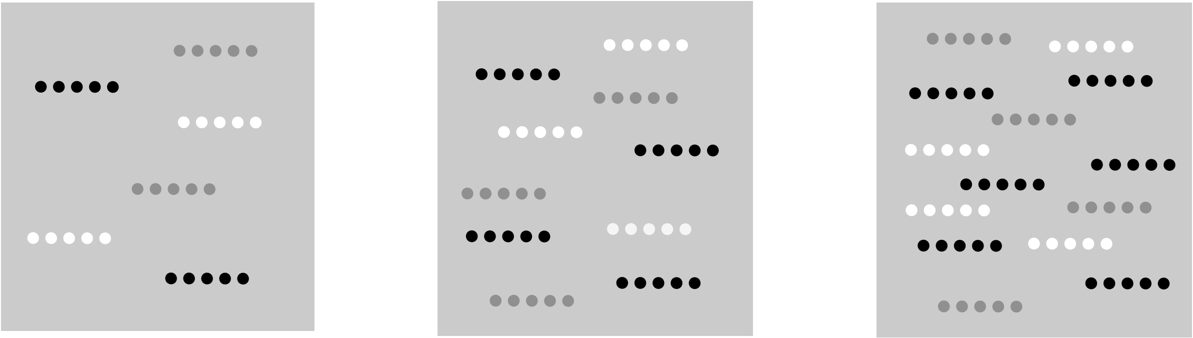

*Figure 5.7 Examples of visual search displays for 6, 10 and 14 items. All contours were composed of 5-dots with equal lateral separation (images are not to scale). Search displays were presented within a fixed distribution zone centred on the mid-point of the screen.*

The target and distractor paths were distinguished by the disparity profiles. In one condition, the target had a continuous disparity gradient and distractors were jittered in depth (discontinuous). For discontinuous paths, the outer dots were positioned in depth the same as the target (same depth range), but the disparities of the intermediate dots were jittered in depth. Disparity jittering was achieved by displacing intermediate dots on the z-axis according to the 'high jitter' described in Experiment 5.1.2. In a separate block of trials, the disparity profile of the target and distractors were reversed: the target was jittered in depth (discontinuous) and the distractors all traversed smoothly through depth (continuous).

Contours were randomly presented at one of three contrasts (Michelson contrasts of 0%, 29%, and 58%) to prevent grouping (via similarity) with nearby contours. Targets were presented on a mid-grey background (15.6 cd/m$^2$). The contrast of the target was counterbalanced to avoid biases and was assigned under the constraint that neighbouring paths did not match. The stimuli were displayed in a 7° x 7° area centred at the mid-point of the display. Each horizontal path was pseudo-randomly positioned within this region, with the constraint that adjacent paths were separated by at least 0.32° in all directions. Path locations were refreshed on every trial.

In a follow-up experiment, I verified that the search task used here was not influenced by any 2-D interpretation of, or information in, the stimuli. To do so, I presented the continuous target in an array of discontinuous distractors and asked observers to perform the task monocularly (one eye viewing the left image of the stereo-pair with disparity) and binocularly (both eyes viewing a stereo-pair at zero disparity). I tested ten trials for each condition at each array size. The results shown in Figure 5.8 confirmed that the task was impossible in the absence of binocular disparity.

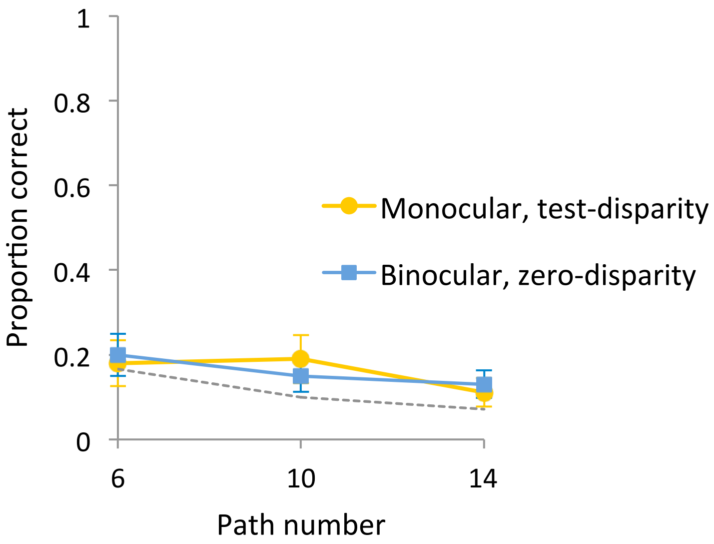

*Figure 5.8 Search accuracy as a function of number of paths for two viewing conditions: the left-eye image of the disparate stereo pair (yellow circles) and binocular viewing with all stimuli at zero-disparity (blue squares). Chance performance is depicted by the grey dashed line. Errors bars show ± one SEM.*

**Procedure** The experimental procedure closely followed that used by Elder and Zucker (1993). This method differs from traditional approaches, because the target is always present. In typical visual search experiments only half of the trials contain the target, and the observer indicates whether it is present or absent. The Elder and Zucker (1993) paradigm is preferable because it yields relatively low error rates (total average of 2.1% and were < 4.5% for all observers in this experiment), is efficient and avoids negative biases introduced by long search times on difficult trials. At the beginning of each trial, an example of the search target was shown for 3 seconds. A fixation cross (2° x 2°) was then presented in the centre of the screen (750ms) and was replaced by the search array. When the observer located the target, they clicked a mouse button and their response time was recorded. The search display was replaced by a validation display in which each dot stimulus was replaced by a small reference square (0.13° x 0.13°) at the position of that stimulus' mid-point. The observer was asked to identify the position of the target by clicking on the appropriate square. If an error was made, the trial was rejected, and subsequently re-tested at random (subsequent analysis showed that there was no stimulus dependency in the probability of an error). Each condition (continuous vs. discontinuous target depth profile) was run in separate blocks, and was tested 20 times for a total of 60 trials (20 trials for 3 display sizes). Each observer initially completed a practice block of 18 trials with feedback.

### 5.4.3 Results and Discussion

Search times for the 5-element stimuli are shown in Figure 5.9A. The results showed that the targets with good stereoscopic continuation were more readily detected (among disparity-jittered paths) than targets containing depth jitter (among continuous disparity paths). While search speed for the discontinuous target depended strongly on the number of stimuli in the display (slope = 0.81sec/item, intercept = 1.1 secs), search for the continuous target was independent of distractor number (slope = 0.095sec/item, intercept = 2.36sec). A two-way ANOVA confirmed that there was a significant effect of disparity profile (continuous vs discontinuous) ($F(1, 48) = 42.13$, $p < 0.001$; $\eta^2 = 0.47$). In the continuous disparity target condition, there was no significant effect of the number of distractors ($F(2,24) = 0.58$, $p > 0.250$) confirming that distractor number had no effect on search speed. However, there was a significant effect of distractor number in the discontinuous target condition ($F(2,24) = 4.80$, $p = 0.018$).

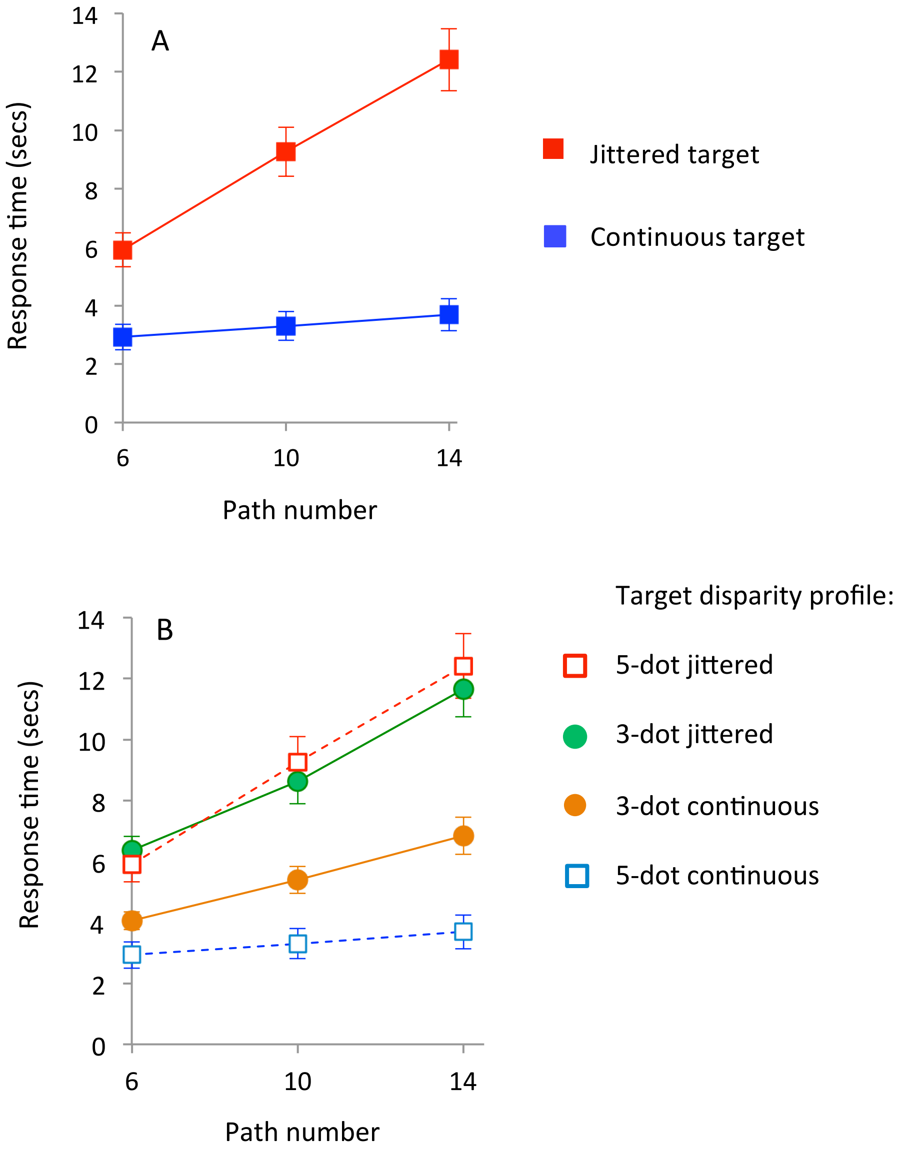

*Figure 5.9. Search results (reaction time) for Experiment 5.2. (A) Search performance for 5-dot contours with a continuous (blue squares) or jittered/discontinuous (red squares) depth profile. (B) 3-dot contours with a continuous depth profile (orange circles) or jittered profile (green circles), with the 5-dot results re-plotted for comparison. Error bars represent ± one standard error of the mean.*

The results depicted in Figure 5.9A suggest that the visual system can take advantage of good stereoscopic continuation in a visual search paradigm. In Experiment 5.1.1, I reported that perceived depth for the end dots was enhanced (the effect of grouping disrupted) by reducing the number of dots in the path from 5 to 3. If the advantage provided in the visual search task is also tied to the effects of grouping, then a reduction in element number should reduce or eliminate the difference between search performance in the continuous and discontinuous target conditions described here. To evaluate this prediction, I repeated Experiment 5.2 with the same observers using targets and distractors defined by only 3 dots. The results are illustrated in Figure 5.9B (along with the results initially obtained with the 5 dot paths) and show that search efficiency for a continuous contour has deteriorated relative to the 5-dot contour, and now depends on the number of distractors (slope = 0.38sec/item, intercept = 1.62secs). While there was some effect of distractors in the 3-element condition, the impact on search time was not as severe as seen in the 5-dot jittered target condition. This was consistent with the fact that perceived depth for the 3-dot stimulus was not fully restored to veridical (see Figure 5.5). On the other hand, performance for the 3-dot discontinuous path was similar to performance in the 5-dot discontinuous condition ($F(1, 48) = 0.035$, $p > 0.250$). This suggested that in the absence of good continuation, performance was consistently poor regardless of the spatial support.

The results presented here demonstrated a search asymmetry. Detection was consistently more efficient for a path with a linear disparity gradient embedded in distractors with jittered disparity, compared to when the jittered path formed the target. By isolating and modulating the disparity cue, I was able to verify that stereopsis alone does not benefit search. Rather it is the local disparity relationships between path components that are important. Observers were unable to perform the task by identifying the path with the unique depth profile; instead, good continuation through depth was the critical feature. However, this type of grouping was dependent on sufficient spatial support, as a continuous path with lower dot density (3 dots) was not as well detected. This is consistent with findings that higher dot density results in stronger perceptual grouping across a contour (eg: Uttal, 1983; Kovacs, 1996; Hess & Field, 1999), and suggests that sufficient 2-D cues (ie: good continuation) are necessary to initially define a contour. Note that 2-D cues alone (such as collinearity, or spacing) were not sufficient to allow detection for either type of target (Figure 5.8). Rather, the additional cues provided by stereopsis – specifically, good continuation in depth – benefits visual search.

### 5.5 General Discussion for Experiments 5.1 – 5.2

I propose that the classic Gestalt principle of good continuation has a disparity-based 3-D counterpart. This disparity-based grouping cue negatively influences perceived depth (Experiment 5.1) but, for similar stimuli, enhances detectability (Experiment 5.2). As reported in Experiment 5.1.1, adding intermediate dots between two isolated elements along a linear disparity gradient resulted in a systematic reduction in relative depth percepts (Figure 5.2). The disruptive effect of intermediate dots critically depended on the presence of a continuous transition through depth: in Experiment 5.1.2, I showed that displacing the intermediate dots in depth was sufficient to break the good stereoscopic continuation and return depth estimates to the level for the two-dot condition (Figure 5.5). It is important to note that in Experiments 5.1.2, the 2-D cues present in the objects were constant across conditions, thus it was the differences in disparity relationships that modulated the perceived depth in these stimuli. I concluded that some minimal amount of 2-D contour representation was necessary, but not sufficient, to yield the reduction in depth illustrated here. Also, the results excluded explanations based on low-level factors such as cue conflict, disparity gradient limits, and re-definition of the fronto-parallel plane. In the second experiment, I confirmed the hypothesis that the visual system can capitalize on good continuation in depth to enhance performance in a visual search task. The stimuli were deliberately designed to eliminate all other cues to the target location; search was virtually impossible when viewed monocularly or binocularly at zero disparity. Targets and distractors could only be distinguished by their disparity profiles. A path of dots that traverses depth along a linear trajectory was detected more rapidly than a target that had sharp disparity discontinuities. The search results provided strong support for the critical role of disparity gradient smoothness in path detection. When taken together, the results suggest that the visual system can take advantage of good continuation though depth to detect contours but that this advantage comes at the cost of perceived depth magnitude.

As reported in the visual search experiment, dot-paths that contain good stereoscopic continuation were easier to detect among stimuli that contain disparity discontinuities. The search results were asymmetric; that is, a discontinuous target was hard to find among distractors with smooth paths. This was the first experiment to directly compare search performance for 3-D stimuli defined by continuous versus discontinuous disparity gradients. However, the result that smooth targets 'pop-out' is not consistent with findings in the 2-D search literature. It is well documented that contours possessing 2-D curvature discontinuities 'pop-out' among straight distractors, whereas straight contours are more difficult to detect among curved distractors (Treisman & Gormican, 1988; Zucker, Dobbins & Iverson, 1989; Wolfe, Yee & Friedman-Hill, 1992; Kristjansson & Tse, 2001). For example, an 'L' or a 'C' shape will pop-out among straight-line segments regardless of the number of distractors (Treisman & Gormican, 1988). This has been taken as evidence that that 2-D curvature is a basic feature for visual search tasks (Wolfe et al. 1992) and provides the most efficient route to recovering 3-D structure from the image (eg, Gibson, 1950). However, this notion is inconsistent with the present results showing that continuous disparity segments pop-out. I propose that the disparity continuation in the path leads to the formation of an object and it is this 'object-ness' that is rapidly detected. There is evidence that collinearity is a critical feature in the detection of a stimulus. Uttal et al. (2000) showed that the perception of motion is enhanced for coherently moving collinear dots embedded in random noise, but is diminished when the dots are non-collinear. It appears that collinearity binds the dots together, enhancing motion perception of a single object. The same logic may apply here. As already described, most coherent 3-D structures are assumed to vary smoothly through depth (Marr & Poggio, 1976, 1979; Marr, 1982). Accordingly, a coherent object would be more likely to 'pop-out' in a cluttered visual search display than an unstable or ambiguous stimulus. However, in trials where the display contains many 'coherent' objects, it would be more difficult to find the discontinuity. Thus, the rapid search rates for coherent objects in this visual search paradigm can be taken as evidence that the visual system is predisposed to rapidly detecting smooth surfaces, using good continuation as one indicator. Previous visual search studies have shown that component parts may be grouped prior to the engagement of attention (Rensink & Enns, 1995; Moore & Egeth, 1997), and that search may be more effectively guided by integrated shapes than by corresponding local features (Found & Müller, 1997). Moreover, early visual processes likely operate on the basis of a variety of grouping principles, such as closure (Elder & Zucker, 1993; Han, Humphreys, & Chen, 1999; Kovács & Julesz, 1993), similarity (Duncan, 1984; Duncan & Humphreys, 1989; Humphreys, Quinlan, & Riddoch, 1989), and proximity (Tversky, Geisler & Perry, 2004). In this manner, good continuation (2-D and 3-D) may be processed pre-attentively in order to supply critical information for shape extraction, ultimately supporting shape identification and segmentation in a cluttered visual world.

# CHAPTER 6

## General Discussion

### 6.1 Summary

This dissertation has addressed the role that perceptual grouping plays in the recovery of depth in simple stereoscopic objects. The research was motivated by a diverse body of work beginning with the work of Kopfermann (1930), who suggested that there is an influence of figural grouping on depth perception. The majority of previous research has focused on the impact of configuration on stereoacuity; increases in depth discrimination thresholds are observed when a closed object is created or implied (Westheimer, 1979; McKee, 1983; Mitchison & Westheimer, 1984; Fahle & Westheimer, 1988; Zalevski et al., 2007). In these investigations, the role of grouping in the perception of suprathreshold depth magnitude had not been assessed. As part of this investigation, I characterized the 2-D and 3-D properties that contribute to 3-D closure and highlighted the perceptual consequences (both costs and benefits) of a new disparity-based grouping cue that interacts with perceived closure.

In Chapter 2, I showed that the amount of suprathreshold depth perceived within a closed rectangular object was consistently reduced relative to its vertical edges presented in isolation or when they formed the edges of two segmented objects. These results extended existing work that showed stereoacuity is degraded by similar manipulations (Westheimer, 1979; McKee, 1983; Mitchison & Westheimer, 1984) and clearly demonstrated that the figural interpretation of a stimulus influences perception of its 3-D form. The effects of the configural manipulations on perceived depth shown here were not due to low-level spatial influences, as shown by the depth estimates obtained when the target contours formed the outer edges of two objects. In this condition, the reduction in depth was not observed. Moreover, conflict with linear perspective cues was shown to not account for the reduction in perceived depth (Experiment 2.2). Instead, these results provide strong evidence that the observed reduction of depth percepts is contingent on mid-level figural grouping, in this case via perceived closure.

In Chapters 3 and 4, I systematically assessed the effect of specific 2-D and 3-D grouping properties – including similarity, connectedness, and proximity – to characterize their role in perceptual closure in depth. The rationale was that perceived depth is maximally degraded when perceptual closure is strong. Therefore, any increase in perceived depth by manipulating individual cues can be used as an indicator of the importance of this property to perceptual closure. The impact of my experimental manipulations on the interpretation of the figure was evaluated by comparing observers' depth estimates to their subjective ratings of closure. Consideration of the rating results in tandem with the corresponding depth estimates confirmed that perceptual closure is not all-or-nothing, but can degrade systematically. I concluded that the dramatic effects of closure on perceived depth do not require physical connections, but critically depend on the presence of aligned L-junctions. Moreover, in Chapter 4, I proposed that the Gestalt cue of good continuation has a stereoscopic counterpart, 'good stereoscopic continuation.' This cue influences perceived depth when the relative disparity of neighboring features varies smoothly. Importantly, good stereoscopic continuation was a necessary component for the grouping phenomenon shown here, suggesting that this 3-D cue supplies critical information for the extraction of shape.

The experiments in Chapter 5 were designed to isolate the 3-D cue of 'good stereoscopic continuation' and assess its impact on depth percepts for rows of dots. The hypothesis was that 'good stereoscopic continuation' is separable from 2-D grouping and can impact performance as a grouping cue in its own right. The results showed that the perceived depth for a row of collinear dots that varied smoothly through depth was reduced compared to the outer dots in isolation. Importantly, when the disparity of the intermediate dots was jittered to disrupt good continuation in the array, veridical depth magnitude estimates were restored. Thus, for a figure with the same 2-D interpretation, the depth profile and interpretation in 3-D modulated depth magnitude. Subsequently, in a visual search task, good stereoscopic continuation was isolated to determine if it has a positive impact on search time. Observers searched for a target, defined by good stereoscopic continuation, among distractors which contained depth jitter, or vice versa. The results showed that detection was dramatically more efficient when the target path contained a continuous disparity gradient. From this, I concluded that 3-D continuation can serve as a stimulus feature that is capable of facilitating visual search. Taken together, the results in Chapter 5 suggest that the visual system can take advantage of good continuation though depth to detect contours, but that this advantage comes at the cost of perceived depth magnitude.

## 6.2 Figural grouping and Closure

As outlined in Chapters 2 – 4, I evaluated the role of closure in 3-D grouping and the image properties essential to perceptual closure. The results of these experiments show that perceptual closure is not simply tied to the presence of connecting lines between the target pair, but is determined by the distribution of components and the configural properties that are generated as a result. The interaction of specific cues necessary for closure suggests that grouping mechanisms are more complex than first described by classic Gestalt psychology (eg, Wertheimer, 1923; Kohler, 1930; Koffka, 1935). The Gestaltists suggested that closure operates as a binary property; groupings either possess closure or they do not. In contrast, the present findings are evidence of the graded nature of closure (Gillam, 1975; Elder & Zucker, 1993; see also Marino & Scholl, 2005). This continuum is reflected in the scatter plot shown in Figure 4.10, which shows there is a strong correlation between the interpretation of a closed object and estimated depth. This pattern of results is evidence of the causal link between perceptual closure and perceived depth magnitude. The strongest percepts of a closed object – and the resultant reduction in depth percepts – were obtained when the horizontal lines were coincident with the end-points of the vertical target lines, creating collinear L-junctions. When the horizontal contours were repositioned vertically so they formed T-junctions, the disruptive effect of grouping was largely eliminated, even though the lines still intersected. Similarly, perceived depth was gradually restored as the L-junction components were vertically offset to remove collinearity.

The importance of L-junctions in human vision has been well documented. Most relevant to the work presented here are psychophysical studies that have suggested that much of the information about the 3-D shape of an object is localized at the intersections of contours. This was first suggested by Attneave (1954), who argued that most of the information along a contour is concentrated in parts of high curvature; a proposal that has received considerable support from psychophysical studies (Koenderink, 1984; Nakayama et al. 1989; Tse & Albert, 1998; Tse, 1999; Kristjannson & Tse, 2001;

Norman, Phillips & Ross, 2001; Rubin, 2001; Feldman & Singh, 2005; Conci et al. 2007). For example, Saarinen and Levi (1999) showed that contour closure enhances the accuracy of shape perception. They showed that recognition of square figures was more accurate when the corners (orientation discontinuities) were intact; presenting the stimulus with only the straight edges resulted in degraded recognition. They proposed that curvature discontinuities – including L-junctions – are important features used to analyze the 2-D image in order to recover 3-D world structure (see also Tse & Albert, 1998; Kristjannson & Tse, 2001). However, closure is not a local property and L-junctions in isolation are not sufficient for the recovery of a global object. In Experiment 4.1, I reported that disconnected, but collinear, L-junction components produced depth magnitude estimates similar to those obtained for continuous horizontal connectors. The fact that both of these stimuli were rated as having a high degree of closure suggests that collinearity between the horizontal line fragments is able to support the grouping of a closed object. This echoes work by Conci et al. (2007) who investigated the impact of grouping by closure on search for target configurations made up of four L-junctions. Their results showed that search performance was optimal when the target was a closed form (composed of collinear corners) and the distractors were open configurations. Notably, search was as efficient when the L-junction components were shortened, indicating that, as I have found, collinearity of the L-junctions was the critical factor in generating closure. They suggested that the visual system is highly tuned to detecting curvature discontinuities (see also Koenderink, 1984; Tse, 1999; Kristjannson & Tse, 2001), and that L-junctions are used to initiate interpolation processes between neighbouring junctions in order to detect a global form (see also Rubin, 2001). This explanation can be extended to the experiments reported here. The detection of L-junctions may have provided initial evidence regarding the shape of the rectangular object, and the collinearity between components was necessary to connect the fragments.

In addition to these 2-D cues, I have shown that 3-D collinearity is also a determinant of figural grouping. In Experiment 4.4, I showed that the alignment of fragments was critical to perceptual closure, as disrupting collinearity was shown to eliminate perceptual closure and increase depth magnitude. I related this perception-based degradation to disparity-based grouping and in Experiment 5.1.2, I showed that ‘good stereoscopic continuation’ is critical to modulating grouping strength. I argue that this disparity-based cue represents more than a continuous disparity gradient between stimulus features, and is contingent upon explicit representation of 2-D continuation between image fragments. This representation need only be a fraction of the stimulus size: gaps of at least 75% between fragments can maintain a high level of closure (Experiment 4.1). In all experiments where 2-D and 3-D continuation were both present between fragments, the depth magnitude estimates consistently matched those obtained for a connected rectangle (Experiment 4.1 – 4.4). This suggests that interpolated stereoscopic contours are treated in the same manner as real contours, in that they are subject to the same reduced depth effect. This proposal is consistent with both behavioural and physiological evidence that have found illusory 2-D boundaries are processed by the visual system in the same manner as luminance-defined edges (eg: von der Heydt, Peterhans, Baumgartner, 1984; Grosof, Shapley & Hawken, 1993; Mendola et al., 1999; Gold et al. 2000). In the disparity domain, Wilcox & Duke (2003) showed that stereoscopically defined illusory surfaces are subject to the same brightness illusions as physical surfaces (assuming Lambertian lighting). In that experiment, like the ones presented here, the percept of a continuous surface was critical to the observed phenomenon. This supports the hypothesis that disparity-based interpolation is driven by the outcome of figural grouping in both 2-D and 3-D.

It is clear that the gradient of disparity in a stimulus can critically influence the percept of depth and should be an important consideration in studying stereoscopic processing. Failure to consider the impact of 3-D good continuation along with 2-D grouping could unwittingly introduce biases toward the fronto-parallel plane and distort depth percepts. Subsequent changes to the configuration (in 2-D or 3-D) that reduce this grouping might then be misattributed or over estimated. This concept may also underlie a number of observed effects of degraded depth for fragmented stimuli (McKee, 1983; Mitchison & Westheimer, 1984; Fahle & Westheimer, 1988; Liu, et al., 1999; Yin, et al., 2000; Vreven et al. 2002; Lu, et al., 2006; Zalevski, et al., 2007). For example, the threshold-based studies described in the Chapter 1 document reduced stereoacuity when an object is created or implied. Those stimulus configurations could all be grouped in 2-D via different cues (including collinear L-junctions) and also possess good stereoscopic continuation. Therefore, it is possible that degraded depth from figural grouping offers a unifying theory for a set of experiments that have previously been considered in isolation.

While the focus of the preceding experiment was on collinear stereoscopic stimuli, it is likely that good stereoscopic continuation is not isolated to linear paths through. Non-planar surfaces (eg, a sinusoidal corrugation) may also exhibit good stereoscopic continuation, provided the stimulus also possesses the attributes required for 2-D good continuation. In such cases, the figure may be prone to the same reduction in depth magnitude observed in these experiments. Good stereoscopic continuation is likely governed by its own set of constraints; I predict these constraints would centre around 'smoothness' through depth. The limits of good stereoscopic continuation would be an interesting avenue for future research.

### 6.3 Object-based smoothing operations

While it is clear that figural grouping impacts percepts of suprathreshold depth, it is not obvious *why* figural grouping results in degraded depth within an object (at threshold and suprathreshold), but enhanced detectability. My working model is based on disparity-based surface encoding and smoothing operations in the visual cortex.

The effects of configuration on stereopsis at detection threshold have been most commonly explained by a form of disparity pooling or averaging (McKee, 1983; Mitchison & Westheimer, 1984; Vreven, et al., 2002). While the neural processing underlying disparity pooling is typically unspecified, a simple hierarchical processing model is often assumed. That is, binocular neurons in later visual processing areas extract global shape information by pooling information from disparity selective neurons in V1. As receptive field size increases and information is pooled, resolution is lost. Disparity pooling or averaging may also be responsible for the results presented here; however, the results are not consistent with simple feed-forward hierarchical processing models. I have shown that 2-D and 3-D grouping cues, thought to be implemented in extrastriate cortical areas, play a critical role in modulating the reduction in perceived depth. Thus, the results are more consistent with recurrent feedback models of neural processing in which high-level inferences concerning visual input drive subsequent processing at earlier levels in the visual system (Lee & Mumford, 2003). This type of model is consistent with the observation that depth judgments are directly linked to the object-based interpretation of the stimuli, as if the disparity information provided by the isolated stimulus components is suppressed if the configuration is perceived as a single object. While still speculative, it is likely that the disparity-based depth discontinuities are initially encoded in area V2 (von der Heydt, Zhou & Friedman, 2000) but are subject to representations of surface-based disparity variation in extra-striate areas. These areas are primarily located in the ventral stream (associated with object recognition), including V4, MT, IT, but slant processing has also been identified in the anterior intraparietal (AIP) in the dorsal stream (eg: Janssen, Vogels & Orban, 1999, 2000; Nguyenkim & DeAngelis, 2003; Liu et al., 2004; Hegde & Van Essen 2005; Verhoef et al. 2010; Rosenburg, Cowen & Angelaki, 2013).

One advantage of an object-based recurrent feedback model for disparity processing is in guiding putative smoothing operations. Since the early computational models of stereopsis (Marr & Poggio, 1976, 1979; Marr, 1982), disparity-based smoothing operations have been employed to help resolve the correspondence problem (eg: Yüille, Geiger & Bülthoff, 1991; Read, 2002; Goutcher & Mamassian, 2005). These assumptions bias matches towards those resulting in constant disparity and bias depth towards the fronto-parallel plane (Marr & Poggio, 1976; 1979). Such a constraint would only be useful for biological systems if the proposed smoothing operations were performed on regions that were organized as common surfaces or objects; otherwise potentially important discontinuities might not be detected, or objects may be fragmented. As noted above, the degraded depth magnitude and improved visual search results were *only* obtained when there was spatial support between the isolated elements which promoted grouping via good continuation. The hypothetical outcome of smoothing operations is that a coherent surface is formed. The results of Experiment 5.2, that showed rapid detection for continuous objects in the presence of discontinuous objects, suggests that the visual system is predisposed to rapid detection of smooth objects – objects that may be differentiated by good stereoscopic continuation. This proposal is supported by computational and neural models. For example, Stanley and Rubin (2003) described ‘salient regions’ as basic image descriptors generated to rapidly guide selective visual processing to candidate target configurations (see also Itti & Koch, 2001).

They showed neurophysiological evidence of responses to salient regions (in that case, closure) and proposed a biologically plausible model for the extraction of salient regions from images (Stanley & Rubin, 2003; Zhaoping, 2005). Within this framework, good stereoscopic continuation may be regarded as one specific property aiding region segmentation, in order to index particular groupings (or objects) for prioritized attentional processing.

The configuration-dependent reduction in depth observed here echoes the results of experiments by Bülthoff and colleagues, who investigated the effects of the disparity gradient between two elements on their perceived location in depth (Bülthoff & Yüille 1991; Bülthoff et al., 1991). Their results showed that suprathreshold depth percepts between two similar components decreased as the disparity gradient increased. Notably, when targets were dissimilar (eg: estimating between a dot and a line), there was no disruption in perceived depth. According to the Bayesian field-based computational model proposed by Yüille, Geiger and Bülthoff (1991), the matching ambiguity between similar components forces the visual system to rely on *a priori* assumptions, such as smoothness, to solve the correspondence problem (this is not necessary with dissimilar components). Thus, the increase in smoothing is responsible for the reduced depth percepts. A fundamental assumption of this approach is the link between matching ambiguity, smoothing operations and degraded depth percepts. However, the results for both line- and dot-based stimuli show that reduction of perceived depth can be observed in the absence of increased matching ambiguity. It is possible that while matching ambiguity may contribute to reduced depth percepts under some conditions, a more significant contributor to smoothing operations may be the salience of the object representation. In the case of Bülthoff et al's (1991) results, manipulation of the disparity gradient by changing element proximity may have also changed the degree to which the neighbouring stimuli were perceptually grouped as single object. This interpretation is supported by the fact that the degraded depth percepts in their study were eliminated when differently shaped elements were used. Similarly, the results shown in Chapters 3 and 4 show that depth percepts can be restored when the degree of object closure is varied, but in those cases the disparity gradient did not vary. Moreover, in the experiments using dot-paths (Experiment 5.1.2), I was able to eliminate the depth disruption by adding disparity jitter, a manipulation that should have increased, not decreased, the correspondence problem.

I propose that object-based disparity smoothing operations are responsible for the degraded depth percepts reported in this dissertation. A modified version of Yüille et al's (1991) computational model could account for the results, if the degree of object-based disparity smoothing were contingent on the strength of the object interpretation. Yüille and colleagues do not rule out such a contingency, but assume that an edge representation has been performed prior to the stereoscopic surface computations. I argue that, as outlined by Lee and Mumford (2003), an object inference from extra-striate cortical areas, which incorporates perceptual grouping information, is used to constrain and modulate disparity smoothing. The result of this inference-based processing is, reduced depth percepts for complete or closed objects and, in addition, enhanced detectability of coherent objects.

### 6.4 Considerations

***6.4.1. Slant perception*** There is evidence from the slant estimation literature that the disparity distribution within stimuli has a significant impact on their suprathreshold appearance. In these experiments, observers are usually asked to match the slant of a display stimulus by manually adjusting a comparator or an on-screen probe. The results typically show that the slant of stereoscopic planar surfaces rotated about the vertical axis are significantly underestimated and slow to develop (eg: Gilliam, 1968; Gillam, et al., 1984; Gillam et al., 1988; Rogers & Cagenello, 1989; Mitchison & McKee, 1990; Ryan & Gillam, 1994; van Ee & Erkelens, 1996). Explanations typically refer to the presence of depth cue conflicts (perspective, accommodation) (Ryan & Gillam, 1994; van Ee & Erkelens, 1996) or the putative insensitivity of the stereoscopic system to smooth disparity gradients (Gillam, et al., 1984; Brookes & Stevens, 1989). Investigations of this phenomenon have not explicitly evaluated the role of perceptual organization. I propose that figural grouping may have played an important role in the reported stereoscopic slant underestimation. That is, the attenuation of slant was not simply tied to the presence of a disparity gradient between features; instead, the reports of degraded depth percepts for continuous disparity profiles are only obtained for stimuli that possess bounding contours or surface texture which create the interpretation of a common surface. In most instances in the literature, this support is explicit, as closed rectangular grids, frames or random dot surfaces are used as stimuli (eg: Gillam et al., 1984; Stevens & Brookes, 1989). Furthermore, these closed figures are slanted in depth and, by definition, contain smooth disparity gradients which, as I have shown, contributes to the perceptual grouping and the resultant loss of perceived depth. In Experiments 4.4 and 5.1.2, I showed that when the smooth disparity gradient is disrupted, depth percepts return to those obtained for isolated targets. Similarly, Gillam et al (1984) used patterns of rectangular grids bounded by a closed contour to assess percepts of slant in depth and found that the addition of a depth discontinuity in the centre of the surface substantially improved slant estimates. Moreover, Stevens and Brookes (1988) showed that perceived depth between two isolated points is closer to veridical when they are placed within a random-dot surface with a 'saw-tooth' profile, compared to a smooth surface. As outlined above, these researchers argued that the visual system is simply more sensitive to disparity discontinuities. Instead, I propose that slant is only underestimated when the stimuli contain 2-D and 3-D good continuation, otherwise veridical depth can be readily extracted across a continuous disparity gradient. This hypothesis offers a converging theory for a set of experiments that have not considered the impact of configuration.

**6.4.2.** ***Stereopsis in isolation*** Interacting with the natural environment often requires that we make precise and accurate depth estimates; from reaching for a coffee cup at arm's length to shooting a hockey puck across the rink into the net. The reduction in depth magnitude between parts of a single object should, if it occurs in natural stimuli, disrupt our ability to interact with objects; spilling the coffee or missing the net. However, there is an abundance of evidence that the perception of depth in natural stimuli is near-veridical despite the complexity and ambiguities of natural scenes (Allison, Gillam, & Vecellio, 2009; Buckley & Frisby, 1993; Durgin, Profitt, Olson, & Reinke, 1995; Frisby, Buckley & Duke, 1996; Loomis, Philbeck, & Zahorik, 2002; Kersten, Mammassian & Yüille, 2004). It has been proposed that humans can compensate for systematic depth errors by relying on multiple redundant cues to depth, particularly in performing well-practiced movements in familiar environments. For example, Loomis et al. (1996) asked observers to match a depth interval (z-axis) to a lateral extent (x-axis) in reduced-cue environment and found that the depth intervals were generally underestimated. However, when they asked observers to walk across the same interval in full-cue real world environments, their motor performance showed no such bias. It is possible that the same result would hold for the phenomenon described here, and that the abundance of depth cues in the natural world would help compensate for the difficulty in extracting disparity in isolation.

To integrate information provided by multiple cues in an efficient manner, observers must assess the degree to which each cue provides reliable versus unreliable information (see Landy, et al., 1995; Ernst & Banks, 2002; Hillis, et al., 2004; Knill & Saunders, 2003; Svarverud, Gilson, & Glennerster, 2010; for review see Jacobs, 2002). Estimating cue reliabilities is a highly complex process, owing to context dependencies. For example, observers may rely heavily on depth-from-stereopsis for very near judgements (eg: threading a needle) but not at all when viewing distant objects. The experiments described here and by others have used restricted cue paradigms to understand stereoscopic mechanisms. While the disruptive effects shown here may not occur in full-cue environments, they do provide potentially important insight into the mechanisms which underlie disparity processing and the constraints that govern the visual recovery of a simple three-dimensional object.

**6.5 Concluding remarks** In this dissertation, I have outlined the constraints that govern the visual recovery of a simple three-dimensional object, as well as highlighted the perceptual consequences of grouping and speculated on the underlying neural machinery. I conclude by supporting the view that closure – and ‘object-ness’ in general – is not an all-or-nothing phenomenon, rather depth-based grouping effects can be independently strengthened or weakened by multiple cues that support the interpretation of a coherent object or surface. Overall, these results highlight the complex interactions between depth from disparity and object recovery: the recovery of a coherent three-dimensional object comes at the cost of veridical depth perception. It is important to understand the broader impact of 3-D grouping in order to develop our knowledge of 3-D object perception. This series of experiments serves as an example of how research on configural effects in depth must extend beyond low-level explanations. Future research must consider the evidence that high-level inferences from visual input drive subsequent processing at earlier levels in the visual system.

# REFERENCES

Allison, R. S., Gillam, B. J., & Vecellio, E. (2009). Binocular depth discrimination and estimation beyond interaction space. *Journal of Vision*, *9*(1), 10.

Allison, R. S., & Howard, I. P. (2000). Temporal dependencies in resolving monocular and binocular cue conflict in slant perception. *Vision research*,*40* (14), 1869-1885.

Anderson, B. L., Singh, M., & Fleming, R. W. (2002). The interpolation of object and surface structure. *Cognitive psychology*, *44*(2), 148-190

Andrews, T. J., Glennerster, A., & Parker, A. J. (2001). Stereoacuity thresholds in the presence of a reference surface. *Vision Research*, *41*(23), 3051-3061.

Attneave, F. (1954). Some informational aspects of visual perception. *Psychological Review*, *61*(3), 183.

Badcock, Haley & Dickinson, 2015. Cues to closed-contour shape revealed using visual search. *Journal of Vision, Abstract supplement.*

Banks, W. P., & Prinzmetal, W. (1976). Configurational effects in visual information processing. *Perception & Psychophysics*, *19*(4), 361-367.

Banks, W. P., & White, H. (1984). Lateral interference and perceptual grouping in detection. *Perception & Psychophysics*, *36*(3), 285-295.

Barenholtz, E., Cohen, E. H., Feldman, J., & Singh, M. (2003). Detection of change in shape: An advantage for concavities. *Cognition*, *89*(1), 1-9.

Bedell, H. E., Gantz, L., & Jackson, D. N. (2012). Perceived Suprathreshold Depth Under Conditions that Elevate the Stereothreshold. *Optom Vis Sci*. 89(12), 1768-73.

Bertamini, M., & Wagemans, J. (2013). Processing convexity and concavity along a 2-D contour: figure–ground, structural shape, and attention. *Psychonomic bulletin & review*, *20*(2), 191-207.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, 94, 115-147.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*(4), 433-436.

Braun, J. (1999). On the detection of salient contours. *Spatial Vision*, *12*(2), 211-225.

Brookes, A., & Stevens, K. A. (1989). Binocular depth from surfaces versus volumes. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(3), 479-484.

Buckley, D., & Frisby, J. P. (1993). Interaction of stereo, texture and outline cues in the shape perception of three-dimensional ridges. *Vision research*,*33*(7), 919-933.

Bülthoff, T, H., Fahle, M., & Wegmann, M. (1991). Perceived depth scales with disparity gradient. *Perception*, *20*, 145-153.

Burge, J., Peterson, M. A., & Palmer, S. E. (2005). Ordinal configural cues combine with metric disparity in depth perception. *Journal of Vision*, *5*(6), 5.

Burt, P., & Julesz, B. (1980). A disparity gradient limit for binocular fusion. *Science*, *208*(4444), 615-617.

Clowes, M. B. (1971). On seeing things. *Artificial intelligence*, *2*(1), 79-116.

Conci, M., Müller, H. J., & Elliott, M. A. (2007). Closure of salient regions determines search for a collinear target configuration. *Perception & psychophysics*, *69*(1), 32-47

Deas L., & Wilcox, L. M. (2012). The role of stereopsis in figural grouping versus segmentation. *Perception*, 41, ECVP Abstract Supplement, 18.

Deco, G., & Lee, T. S. (2002). A unified model of spatial and object attention based on inter-cortical biased competition. *Neurocomputing*, *44*, 775-781.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual review of neuroscience*, *18*(1), 193-222.

Donnelly, N., Humphreys, G. W., & Riddoch, M. J. (1991). Parallel computation of primitive shape descriptions. *Journal of Experimental Psychology: Human Perception and Performance*, *17*(2), 561.

Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General*, *113*(4), 501.

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological review*, *96*(3), 433.

Durgin, F. H., Proffitt, D. R., Olson, T. J., & Reinke, K. S. (1995). Comparing depth from motion with depth from binocular disparity. *Journal of Experimental Psychology: Human Perception and*

*Performance*, *21*(3), 679

Elder, J. H., & Goldberg, R. M. (2002). Ecological statistics of Gestalt laws for the perceptual organization of contours. *Journal of Vision*, *2*(4), 5.

Elder, J., & Zucker, S. (1993). The effect of contour closure on the rapid discrimination of two-dimensional shapes. *Vision Research, 33*(7), 981-991.

Elder, J., & Zucker, S. (1994). A measure of closure. *Vision research*, *34*(24), 3361-3369.

Elder, J. H., & Zucker, S. W. (1998). Evidence for boundary-specific grouping. *Vision research*, *38*(1), 143-152.

Ellis, W.D. (1938). A source book of Gestalt psychology. London: K. Paul, Trench, Trubner & Co., Ltd.

Epstein, W., & Babler, T. (1990). In search of depth. *Perception & psychophysics*, *48*(1), 68-76.

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*(6870), 429-433.

Fahle, M., & Westheimer, G. (1988). Local and global factors in disparity detection of rows of points. *Vision Research, 28*(1), 171-178.

Fanton, C., Gerbino, W., & Kellman, P. J. (2004). Approximation, torsion, and amodally-unified surfaces. *Journal of Vision*, *4*(8), 726-726.

Feldman, J. (2001). Bayesian contour integration. *Perception & Psychophysics*, *63*, 1171-1182.

Feldman, J., & Singh, M. (2005). Information along contours and object boundaries. *Psychological Review*, *112*(1), 243.

Field, D. J., Hayes, A., & Hess, R. F. (1993). Contour integration by the human visual system: Evidence for a local “association field”. *Vision Research*, *33*(2), 173-193.

Foley, J. M. (1980). Binocular distance perception. *Psychological review*, *87*(5), 411.

Frisby, J. P., Buckley, D., & Duke, P. A. (1996). Evidence for good recovery of lengths of real objects seen with natural stereo viewing. *Perception*, *25*, 129-154.

Geisler, W. S., Perry, J. S., Super, B. J., & Gallogly, D. P. (2001). Edge co-occurrence in natural images predicts contour grouping performance. *Vision research*, *41*(6), 711-724.

Gibson, J. J. (1950). *The Perception of the Visual World.* Boston: Houghton Mifflin.

Gilbert, C. D., Das, A., Ito, M., Kapadia, M., & Westheimer, G. (1996). Spatial integration and cortical dynamics. *Proceedings of the National Academy of Sciences*, *93*(2), 615-622.

Gilbert, C. D., & Wiesel, T. N. (1983). Clustered intrinsic connections in cat visual cortex. *The Journal of Neuroscience*, *3*(5), 1116-1133.

Gillam, B. J. (1968). Perception of slant when perspective and stereopsis conflict: experiments with aniseikonic lenses. *Journal of Experimental Psychology, 78*(2), 299-305.

Gillam, B. J., Anderson, B. L., & Rizwi, F. (2009). Failure of facial configural cues to alter metric stereoscopic depth. *Journal of vision*, *9*(1), 3.

Gillam, B., Chambers, D., & Russo, T. (1988). Postfusional latency in stereoscopic slant perception and the primitives of stereopsis. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(2), 163.

Gillam, B., Flagg, T., & Finlay, D. (1984). Evidence for disparity change as the primary stimulus for stereoscopic processing. *Perception & Psychophysics, 36*(6), 559-564.

Gillam, B., & Grant Jr, T. (1984). Aggregation and unit formation in the perception of moving collinear lines. *Perception*, *13*(65), 4-664.

Gillam, B. J., Grove, P. M., & Layden, J. (2010). The role of remote closure in the perception of occlusion at junctions and illusory contours. *Perception*, *39*(2), 145.

Gillam, B., & Ryan, C. (1992). Perspective, orientation disparity, and anisotropy in stereoscopic slant perception. *Perception, 21*(4), 427-439.

Gogel, W. C. (1963). The visual perception of size and distance. *Vision Research*, *3*(3), 101-120.

Goutcher, R., & Hibbard, P. B. (2010). Evidence for relative disparity matching in the perception of an ambiguous stereogram. *Journal of Vision*, *10*(12), 35.

Goutcher, R., & Mamassian, P. (2005). Selective biasing of stereo correspondence in an ambiguous stereogram. Vision Research, 45, 469–483.

Grossberg, S., & Mingolla, E. (1985). Neural dynamics of form perception: boundary completion, illusory figures, and neon color spreading. *Psychological review*, *92*(2), 173.

Guttman, S. E., Sekuler, A. B., & Kellman, P. J. (2003). Temporal variations in visual completion: A reflection of spatial limits? *Journal of Experimental Psychology: Human Perception and Performance*, *29*(6), 1211.

Hadad, B. S., & Kimchi, R. (2008). Time course of grouping of shape by perceptual closure: Effects of spatial proximity and collinearity. *Perception & psychophysics*, *70*(5), 818-827.

Han, S., Humphreys, G. W., & Chen, L. (1999). Uniform connectedness and classical Gestalt principles of perceptual grouping. *Perception & psychophysics*, *61*(4), 661-674

Hartle, B. A., & Wilcox, L. M. (2014). Depth Magnitude From Stereopsis: Assessment Techniques and the Role of Experience. *Canadian Journal of Experimental Psychology,* 68(4), 302-302.

Hartle, B. A., & Wilcox, L. M. (2015). Assessment of depth magnitude from binocular disparity. *Journal of Vision, Abstract supplement*.

Hartmann, G. W. (1935). *Gestalt psychology: a survey of facts and principles*. New York: Ronald Press.

Hegdé, J., & Van Essen, D. C. (2005). Role of primate visual area V4 in the processing of 3-D shape characteristics defined by disparity. *Journal of neurophysiology*, *94*(4), 2856-2866.

Heitger, F., von der Heydt, R., (1993) A computational model of neural contour processing: figure-ground segregation and illusory contours. *Proc. 4th Int. Conf. Computer Vision, 32-40.*

Heitger, F., von der Heydt, R., Peterhans, E., Rosenthaler, L., & Kübler, O. (1998). Simulation of neural contour mechanisms: representing anomalous contours. *Image and Vision Computing*, *16*(6), 407-421.

Herzog, M. H., & Manassi, M. (2015). Uncorking the bottleneck of crowding: A fresh look at object recognition. *Current Opinion in Behavioral Sciences*, *1*, 86-93.

Hess, R., & Field, D. (1999). Integration of contours: new insights. *Trends in cognitive sciences*, *3*(12), 80-486.

Hess, R. F., Hayes, A., & Kingdom, F. A. (1997). Integrating contours within and through depth. *Vision research*, *37*(6), 691-696.

Hillis, J. M., Watt, S. J., Landy, M. S., & Banks, M. S. (2004). Slant from texture and disparity cues: Optimal cue combination. *Journal of vision*, *4*(12), 1.

Hirsch, J., DeLaPaz, R. L., Relkin, N. R., Victor, J., Kim, K., Li, T., ... & Shapley, R. (1995). Illusory contours activate specific regions in human visual cortex: evidence from functional magnetic resonance imaging. *Proceedings of the National Academy of Sciences*, *92*(14), 6469-6473.

Holliday, I. E., & Braddick, O. J. (1991). Pre-attentive detection of a target defined by stereoscopic slant. *Perception*, *20*(33), 355-362.

Hou, F., Lu, H., Zhou, Y., & Liu, Z. (2006). Amodal completion impairs stereoacuity discrimination. *Vision research*, *46*(13), 2061-2068.

Howard, I. P & Rogers, B. J. (2012). *Perceiving in Depth. Vol. 2: Stereoscopic Vision*. New York: Oxford University Press.

Howe, C. Q., & Purves, D. (2005). Natural-scene geometry predicts the perception of angles and line orientation. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(4), 1228-1233.

Huberle, E., & Karnath, H. O. (2012). The role of temporo-parietal junction (TPJ) in global Gestalt perception. *Brain Structure & Function, 217*(3), 735-746.

Humphreys, G. W., Quinlan, P. T., & Riddoch, M. J. (1989). Grouping processes in visual search: Effects with single-and combined-feature targets. *Journal of Experimental Psychology: General*, *118*(3), 258.

Jacobs, R. A. (2002). What determines visual cue reliability? *Trends in cognitive sciences*, *6*(8), 345-350.

Janssen, P., Vogels, R., & Orban, G. A. (1999). Macaque inferior temporal neurons are selective for disparity-defined three-dimensional shapes. *Proceedings of the National Academy of Sciences*, *96*(14), 8217-8222.

Janssen, P., Vogels, R., & Orban, G. A. (2000). Three-dimensional shape coding in inferior temporal cortex. *Neuron*, *27*(2), 385-397.

Johnston, E. B., Cumming, B. G., & Landy, M. S. (1994). Integration of stereopsis and motion shape cues. *Vision research*, *34*(17), 2259-2275.

Johnston, E. B., Cumming, B. G., & Parker, A. J. (1993). Integration of depth modules: Stereopsis and texture. *Vision research*, *33*(5), 813-826.

Kanizsa, G. (1976). Subjective contours. *Scientific American, 234*(4), 48-52.

Kellman, P. J., Garrigan, P., Shipley, T. F., Yin, C., & Machado, L. (2005). 3-d interpolation in object perception: evidence from an objective performance paradigm. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(3), 558.

Kellman, P. J., & Shipley, T. F. (1991). A theory of visual interpolation in object perception. *Cognitive psychology*, *23*(2), 141-221.

Kellman, P. J., Yin, C., & Shipley, T. F. (1998). A common mechanism for illusory and occluded object completion. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(3), 859.

Kersten, D., Mamassian, P., & Yüille, A. (2004). Object perception as Bayesian inference. *Annu. Rev. Psychol.*, *55*, 271-304.

Kimchi, R. (2000). The perceptual organization of visual objects: A microgenetic analysis. *Vision research*, *40*(10), 1333-1347

Knill, D. C., & Saunders, J. A. (2003). Do humans optimally integrate stereo and texture information for judgments of surface slant? *Vision research*, *43*(24), 2539-2558.

Kohler, W. (1930). *Psychologies of 1930*. C. A. Murchison (*Ed.*). London: Oxford University Press.

Kohler, W. (1926). *Psychologies of 1925*. C. A. Murchison (*Ed.*). London: Oxford University Press.

Kovács, I. (1996). Gestalten of today: early processing of visual contours and surfaces. *Behavioural brain research*, *82*(1), 1-11.

Kovács, I., & Julesz, B. (1993). A closed curve is much more than an incomplete one: Effect of closure in figure-ground segmentation. *Proc Natl Acad Sci USA*, *90*(16), 7495-7497.

Koffka, K. (1935). *Principles of Gestalt psychology.* New York: Harcourt, Brace, & World.

Kopfermann, H. (1930). Psychologische Untersuchungen Ïber die Wirkung zweidimensionaler Darstellungen körperlicher Gebilde. *Psychologische Forschung,* 13, 293 - 364.

Kristjánsson, Á., & Tse, P. U. (2001). Curvature discontinuities are cues for rapid shape analysis. *Perception & Psychophysics*, *63*(3), 390-403.

Kubilius, J., Wagemans, J., & Op de Beeck, H. P. (2011). Emergence of perceptual Gestalts in the human visual cortex: the case of the configural-superiority effect. *Psychological Science, 22*(10), 1296-

1303.

Landy, M. S., Maloney, L. T., Johnston, E. B., & Young, M. (1995). Measurement and modeling of depth cue combination: In defense of weak fusion. *Vision research*, *35*(3), 389-412.

Lanze, M., Weisstein, N., & Harris, J. R. (1982). Perceived depth vs. structural relevance in the object-superiority effect. *Perception & Psychophysics, 31*(4), 376-382.

Lee, T. S., & Mumford, D. (2003). Hierarchical Bayesian inference in the visual cortex. *JOSA A*, *20*(7), 1434-1448.

Lehky, S. R., & Sejnowski, T. J. (1990). Neural model of stereoacuity and depth interpolation based on a distributed representation of stereo disparity. *The Journal of Neuroscience*, *10*(7), 2281-2299.

Loomis, J. M., Philbeck, J. W., & Zahorik, P. (2002). Dissociation between location and shape in visual space. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(5), 1202.

Loomis, J. M., Silva, J. A. D., Philbeck, J. W., & Fukusima, S. S. (1996). Visual perception of location and distance. *Current Directions in Psychological Science*, 72-77.

Lowe, D. G. (1987). Three-dimensional object recognition from single two-dimensional images. *Artificial intelligence*, *31*(3), 355-395.

Lindblom, B., & Westheimer, G. (1992). Spatial uncertainty in stereoacuity tests: implications for clinical vision test design. *Acta ophthalmologica*, *70*(1), 60-65.

Liu, Z., Jacobs, D. W., & Basri, R. (1999). The role of convexity in perceptual completion: beyond good continuation. *Vision Research, 39*(25), 4244-4257.

Lu, H., Tjan, B. S., & Liu, Z. (2006). Shape recognition alters sensitivity in stereoscopic depth discrimination. *Journal of Vision*, *6*(1), 75-86.

Malania, M., Herzog, M. H., & Westheimer, G. (2007). Grouping of contextual elements that affect vernier thresholds. *Journal of Vision, 7*(2), 1-7.

Mamassian, P. (2008). Depth, but not surface orientation, from binocular disparities. *Journal of Vision, 8*(6), 89.

Manassi, M., Sayim, B., & Herzog, M. H. (2012). Grouping, pooling, and when bigger is better in visual crowding. *Journal of Vision*, *12*(10), 13, 1-14.

Markov, N. T., & Kennedy, H. (2013). The importance of being hierarchical. *Current opinion in neurobiology*, *23*(2), 187-194.

Markov, N. T., Vezoli, J., Chameau, P., Falchier, A., Quilodran, R., Huissoud, C., ... & Kennedy, H. (2014). Anatomy of hierarchy: feedforward and feedback pathways in macaque visual cortex. *Journal of Comparative Neurology*, *522*(1), 225-259.

Marks, L. E., & Gescheider, G. A. (2002). Psychophysical scaling. In *Stevens' handbook of experimental psychology*. Pashler, H. & Wixted, J. (Eds). Wiley Publsihing.

Marr, D. (1982). Vision: A computational investigation into the human representation and the processing of visual information. New York, NY: Freeman & Company.

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. *Proc Royal Soc London B*, *204*(1156), 301-328.

Mathes, B., & Fahle, M. (2007). Closure facilitates contour integration. *Vision research*, *47*(6), 818-827.

Mathews & Howell (2009). Complex analysis for mathematics and engineering, 6th Ed. Jones & Bartlett Learning

McKee, S. P. (1983). The spatial requirements for fine stereoacuity. *Vision Research, 23*(2), 191-198.

McKee, S. P., Levi, D. M., & Bowne, S. F. (1990). The imprecision of stereopsis. *Vision research*, *30*(11), 1763-1779.

Mendola, J. D., Dale, A. M., Fischl, B., Liu, A. K., & Tootell, R. B. (1999). The representation of illusory and real contours in human cortical visual areas revealed by functional magnetic resonance imaging. *The Journal of Neuroscience*, *19*(19), 8560-8572.

Mitchison, G. J. (1988). Planarity and segmentation in stereoscopic matching. *Perception, 17*(6), 753-782.

Mitchison, G. J., & McKee, S. P. (1987a). Interpolation and the detection of fine structure in stereoscopic matching. *Vision Research, 27*(2), 295-302.

Mitchison, G. J., & McKee, S. P. (1987b). The resolution of ambiguous stereoscopic matches by interpolation. *Vision Research, 27*(2), 285-294.

Mitchison, G. J., & McKee, S. P. (1990). Mechanisms underlying the anisotropy of stereoscopic tilt perception. *Vision Research*, *30*(11), 1781-1791.

Mitchison, G. J., & Westheimer, G. (1984). The perception of depth in simple figures. *Vision Research, 24*(9), 1063-1073.

Moore, C. M., & Egeth, H. (1997). Perception without attention: evidence of grouping under conditions of inattention. *Journal of Experimental Psychology: Human Perception and Performance*, *23*(2), 339.

Nakayama, K., Shimojo, S., & Silverman, G. H. (1989). Stereoscopic depth: its relation to image segmentation, grouping, and the recognition of occluded objects. *Perception*, *18*(1), 55-68.

Nguyenkim, J. D., & DeAngelis, G. C. (2003). Disparity-based coding of three-dimensional surface orientation by macaque middle temporal neurons. *The Journal of neuroscience*, *23*(18), 7117-7128.

Norman, J. F., Phillips, F., & Ross, H. E. (2001). Information concentration along the boundary contours of naturally shaped solid objects. *Perception*, *30*(11), 1285-1294.

Palmer, S., & Rock, I. (1994). Rethinking perceptual organization: The role of uniform connectedness. *Psychonomic Bulletin & Review*, *1*(1), 29-55.

Patel, S. S., Bedell, H. E., Tsang, D. K., & Ukwade, M. T. (2009). Relationship between threshold and suprathreshold perception of position and stereoscopic depth. JOSA A, 26(4), 847-861.

Pelli, D. G. (1985). Uncertainty explains many aspects of visual contrast detection and discrimination. *JOSA A, 2*(9), 1508-1532.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision, 10*(4), 437-442.

Peterhans, E., & von der Heydt, R. (1989). Mechanisms of contour perception in monkey visual cortex. II. Contours bridging gaps. *Journal of Neuroscience*, *9*(5), 1749-1763.

Peterson, M. A., & Gibson, B. S. (1993). Shape recognition inputs to figure-ground organization in three-dimensional displays. *Cognitive Psychology*, *25*(3), 383-429.

Pettet, M. W., McKee, S. P., & Grzywacz, N. M. (1998). Constraints on long range interactions mediating contour detection. *Vision Research*, *38*(6), 865-879.

Pollard, S. B., Mayhew, J. E., & Frisby, J. P. (1985). PMF: A stereo correspondence algorithm using. *Perception*, *4*, 449-470.

Pomerantz, J. R., & Portillo, M. C. (2011). Grouping and emergent features in vision: toward a theory of basic Gestalts. *Journal of Experimental Psychology: Human Perception and Performance, 37*(5), 1331-1349.

Pomerantz, J. R., Sager, L. C., & Stoever, R. J. (1977). Perception of wholes and of their component parts: some configural superiority effects. *Journal of Experimental Psychology: Human Perception and Performance, 3*(3), 422-435.

Prazdny, K. (1985). Detection of binocular disparities. *Biological Cybernetics*, *52*(2), 93-99.

Purcell, D. G., & Stewart, A. L. (1991). The object-detection effect: configuration enhances perception. *Perception & Psychophysics, 50*(3), 215-224.

Qian, N., & Zhu, Y. (1997). Physiological computation of binocular disparity. *Vision research*, *37*(13), 1811-1827.

Rensink, R. A., & Enns, J. T. (1995). Preemption effects in visual search: evidence for low-level grouping. *Psychological review*, *102*(1), 101.

Rock, I., & Brosgole, L. (1964). Grouping based on phenomenal proximity. .*Journal of Experimental Psychology*, *67*(6), 531.

Rogers, B., & Cagenello, R. (1989). Disparity curvature and the perception of three-dimensional surfaces. *Nature,* 339, 135 – 137.

Rosenberg, A., Cowan, N. J., & Angelaki, D. E. (2013). The visual representation of 3D object orientation in parietal cortex. *The Journal of Neuroscience*, *33*(49), 19352-19361.

Rubin, N. (2001). The role of junctions in surface completion and contour matching. *Perception*, *30*(3), 339-366.

Ryan, C., & Gillam, B. (1994). Cue conflict and stereoscopic surface slant about horizontal and vertical axes. *Perception, 23*(6), 645-658.

Saarela, T. P., Sayim, B., Westheimer, G., & Herzog, M. H. (2009). Global stimulus configuration modulates crowding. *Journal of Vision*, 9(2):5, 1–11.

Samonds, J. M., Potetz, B. R., Tyler, C. W., & Lee, T. S. (2013). Recurrent connectivity can account for the dynamics of disparity processing in V1. *The Journal of Neuroscience*, *33*(7), 2934-2946.

Sato, M., & Howard, I. P. (2001). Effects of disparity–perspective cue conflict on depth contrast. *Vision Research*, *41*(4), 415-426.

Sayim, B., Westheimer, G., & Herzog, M. H. (2008). Contrast polarity, chromaticity, and stereoscopic depth modulate contextual interactions in vernier acuity. *Journal of Vision*, 8 (8):12, 1– 9.

Shipley, T. F., & Kellman, P. J. (2003). Boundary completion in illusory contours: Interpolation or extrapolation?. *Perception*, *32*(8), 985-1000.

Sousa, R., Brenner, E., & Smeets, J. B. (2009). Slant cue are combined early in visual processing: Evidence from visual search. *Vision research*, *49*(2), 257-261.

Spehar, B. (2002). The role of contrast polarity in perceptual closure. *Vision Research*, *42*(3), 343-350.

Stanley, D. A., & Rubin, N. (2003). fMRI activation in response to illusory contours and salient regions in the human lateral occipital complex. *Neuron*, *37*(2), 323-331.

Stevens, K. A., & Brookes, A. (1988). Integrating stereopsis with monocular interpretations of planar surfaces. *Vision Research, 28*(3), 371-386.

Stevens, S. S. (1955). The measurement of loudness. *The Journal of the Acoustical Society of America*, *27*(5), 815-829.

Stevens, S. S. (1975). *Psychophysics*. Transaction Publishers.

Svarverud, E., Gilson, S. J., & Glennerster, A. (2010). Cue combination for 3D location judgements. *Journal of vision*, *10*(1), 5.

Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: evidence from search asymmetries. *Psychological Review, 95*(1), 15-48.

Treisman, A., & Paterson, R. (1984). Emergent features, attention, and object perception. *Journal of Experimental Psychology: Human Perception and Performance, 10*(1), 12-31.

Tse, P. U., & Albert, M. K. (1998). Amodal completion in the absence of image tangent discontinuities. *Perception*, *27*, 455-464.

Tversky, T., Geisler, W. S., & Perry, J. S. (2004). Contour grouping: closure effects are explained by good continuation and proximity. *Vision Research*, *44* (24), 2769-2777.

Uttal, W. R. (1983). *Visual form detection in 3-dimensional space*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Uttal, W. R., Spillmann, L., Stürzel, F., & Sekuler, A. B. (2000). Motion and shape in common fate. *Vision Research*, *40*(3), 301-310.

van Ee, R., & Erkelens, C. J. (1996). Temporal aspects of binocular slant perception. *Vision Research, 36*(1), 43-51.

van Ee, R., van Dam, L. C., & Erkelens, C. J. (2002). Bi-stability in perceived slant when binocular disparity and monocular perspective specify different slants. *Journal of Vision*, *2*(9), 2.

van Lier, R. (1999). Investigating global effects in visual occlusion: from a partly occluded square to the back of a tree-trunk. *Acta Psychologica*, *102*(2), 203-220.

Verhoef, B. E., Vogels, R., & Janssen, P. (2010). Contribution of inferior temporal and posterior parietal activity to three-dimensional shape perception. *Current Biology*, *20*(10), 909-913.

Volcic, R., Fantoni, C., Caudek, C., Assad, J. A., & Domini, F. (2013). Visuomotor adaptation changes stereoscopic depth perception and tactile discrimination. *The Journal of Neuroscience*, *33*(43), 17081-17088.

Von der Heydt, R., Peterhans, E., & Baumgartner, G. (1984). Illusory contours and cortical neuron responses. *Science*, *224*(4654), 1260-1262.

von der Heydt, R., Zhou, H., & Friedman, H.S. (2000). Representation of stereoscopic edges in monkey visual cortex. *Vision research*, *40*(15), 1955-1967.

Vreven, D., McKee, S. P., & Verghese, P. (2002). Contour completion through depth interferes with stereoacuity. *Vision research*, *42*(18), 2153-2162.

Wagemans, J., Feldman, J., Gepshtein, S., Kimchi, R., Pomerantz, J. R., van der Helm, P. A., & van Leeuwen, C. (2012a). A century of Gestalt psychology in visual perception: II. Conceptual and theoretical foundations. *Psychological Bulletin*, *138*(6), 1218.

Weisstein, N., & Harris, C. S. (1974). Visual detection of line segments: an object-superiority effect.

*Science, 186*(4165), 752-755.

Wertheimer, M. (2012). *On perceived motion and figural organization*. L. Spillman (Ed.). Cambridge, MA: MIT Press.

Westheimer, G. (1979). The spatial sense of the eye. Proctor lecture. *Investigative Ophthalmology & Visual Science, 18*(9), 893-912.

Westheimer, G., & McKee, S. P. (1977). Spatial configurations for visual hyperacuity. *Vision Research, 17*(8), 941-947.

Wilcox, L. M., & Allison, R. S. (2009). Coarse-fine dichotomies in human stereopsis. *Vision Research, 49*(22), 2653-2665.

Wolfe, J. M., Yee, A., & Friedman-Hill, S. R. (1992). Curvature is a basic feature for visual search tasks. *Perception*, *21*, 465-465.

Yen, S. C., & Finkel, L. H. (1998). Extraction of perceptually salient contours by striate cortical networks. *Vision research*, *38*(5), 719-741

Yin, C., Kellman, P. J., & Shipley, T. F. (2000). Surface integration influences depth discrimination. *Vision Research, 40*(15), 1969-1978.

Yüille, A. L., Geiger, D., & Bülthoff, H. H. (1991). Stereo integration, mean field theory and psychophysics. *Network: Computation in Neural Systems*, *2*(4), 423-442.

Zalevski, A. M., Henning, G. B., & Hill, N. J. (2007). Cue combination and the effect of horizontal disparity and perspective on stereoacuity. *Spatial Vision, 20*(1-2), 107-138.

Zhou, H., Friedman, H. S., & von der Heydt, R. (2000). Coding of border ownership in monkey visual cortex. *Journal of Neuroscience, 20*(17), 6594-6611.

Zucker, S. W., Dobbins, A., & Iverson, L. (1989). Two stages of curve detection suggest two styles of visual computation. *Neural Computation*, *1*(1), 68-81.