Title
Promiscuity and Pruning: Investigations of a Two-Stage Theory of Contour Interpolation

Author
Carrigan, Susan Burns

Publication Date
2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
Los Angeles

Promiscuity and Pruning:
Investigations of a Two-Stage Theory of Contour Interpolation

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Psychology

by

Susan Burns Carrigan

2019

ABSTRACT OF THE DISSERTATION

Promiscuity and Pruning:
Investigations of a Two-Stage Theory of Contour Interpolation

by

Susan Burns Carrigan
Doctor of Philosophy in Psychology
University of California, Los Angeles, 2019
Professor Philip J. Kellman, Chair

One of the most important functions of the human visual system is the construction of representations of whole objects and surfaces. Contour interpolation, the process that connects edges across gaps in the input, is critical to achieving these representations. In this dissertation, I will propose, test, and model a two-stage theory of the contour interpolation process. The first stage of this process interpolates connections indiscriminately across all pairs of edge fragments that meet the geometric relatability constraints defined by Kellman & Shipley (1991). These connections are maintained in an intermediate representation. The second stage of the contour interpolation process receives these promiscuously interpolated contours, and the physically specified edges that gave rise to them, as input. At this stage, a variety of other sources of information in the scene are taken into account to determine the modal or amodal appearance of each interpolated contour as well as its ultimate perceptual strength. The information that impacts an interpolation's perceptual strength includes cues as to border ownership, luminance contrast polarity, equiluminant color contrast, and the corroboration of interpolated contours across multiple spatial scales. The outputs of this second stage determine the subjective appearance of interpolated contours in a scene. Because the outputs of the first stage exist in an intermediate representation, some contours interpolated in the first stage

will not appear in our phenomenological perception. However, because they exist at some level of the perceptual process, they have some measurable impacts on perceptual performance tasks. Chapter 2 introduces a set of experiments that reveal the intermediate outputs of the first stage of the contour interpolation process. Chapter 3 introduces a set of experiments that reveals the effect of a variety of scene cues on the perceptual strengths of interpolated contours, as determined in stage 2. Manipulations of cues as to border ownership, amodal surface spreading (including luminance contrast polarity and equiluminant color contrast differences across inducing fragments), and spatial frequency differences across inducing edges are shown to impact the perception of interpolated contours exiting the second stage. Chapter 4 combines manipulations to scene cues to measure the combined magnitudes of their effects. The results reveal that cue combination in the second stage is well described by an additive model but also show that certain combinations of cues lead to reductions in strength that go beyond additivity. A theory as to what is special about these combinations is discussed in Chapter 5.

Keywords: contour interpolation, promiscuous interpolation, amodal completion, modal completion, multi-stage theory, two-stage theory, border ownership, color contrast polarity, luminance contrast polarity, spatial frequency channels

The dissertation of Susan Burns Carrigan is approved.

Hongjing Lu

Zili Liu

Ladan Shams

Philip J. Kellman, Committee Chair

University of California, Los Angeles

2019

To my advisor and lab mates

TABLE OF CONTENTS

LIST OF TABLES & FIGURES

ACKNOWLEDGEMENTS

This work grew out of several years I spent in a place that has become a second home to me with many people I have grown to care deeply about – some of whom have become like family, and many who are undoubtedly among the greatest minds and mentors I have ever known. To all of them, I am deeply, eternally grateful. Looking back, to cut out of my life the moments I spent with them would mean cutting off of myself many of the most beautiful and precious ideas and qualities that I have managed to grow. This is because many of the best parts of me I have grown largely as a result of the influence of the deeply beautiful, brilliant, and good people I have met here at UCLA.

To all of the following, I send the purest thoughts of gratitude, and the most sincere hope that they will have many more years ahead of them, and that there will be many more like me fortunate enough to meet and be touched by them.

To Dr. Ted Wright – for introducing me to research in vision science.

To Dr. Charlie Chubb – for exciting me, for believing in me, for assuaging my fears, for convincing me that I could do this, and for his continued friendship and support.

To Dr. Ladan Shams – for five opportunities to TA my favorite class; for allowing me space to think, design, and make suggestions – and for taking my ideas seriously; for her contributions as a member of my dissertation committee, and for her kindness, approachability, and enthusiasm.

To Dr. Iris Firstenberg – for eight opportunities to TA a class I enjoy very much, including 3 quarters as a master TA; for her consistent commitment to designing and redesigning a fantastic course - especially when, for unavoidable reasons, it is a much more difficult one to teach; for always welcoming my feedback; and for valuing my contributions.

To Dr. Zili Liu – for always being ready to entertain and evaluate a thought, for his willingness to engage in thoughtful conversations about complicated topics via email, for his contributions as a member of my dissertation committee, and for his rigor, honesty, and good-naturedness.

To Dr. Hongjing Lu – for her readiness to provide explanations, advice, and blessings; for her engaged, open-minded, and interested presence in lab meetings; for her contributions as a member of my dissertation committee; for her computational brilliance and savvy, and for her kindness.

To Dr. Christine Massey – who came relatively recently to our lab and has been a most-welcomed and fantastic addition; for her kind voice and peaceful presence; for her sincere interest, and smart, thoughtful comments in lab meetings; for the friendship she has given my dog, Trex (who thanks her also); and for her warmth and kindness.

To Dr. Gennady Erlikhman – for years of invaluable friendship and tutelage. You, my friend, are a wizard with a beyond-beautiful soul. For every answer, explanation, discussion, script, kind thought, barbecue, and D&D game; and for his sincerity and pure goodness.

To Rachel – who so readily accepted this runt as her friend, for her soul so good and so intact.

To my old friend, Erica Stafford – for being open, for being honest, for being someone I can relate to in so many ways, for being hardworking, for being tenacious, for being my friend, and for being on this journey with me.

To my lab mates: Phil, Christine, Genna, Rachel, Nick, Carolyn, Victoria, Everett, Tim, Matthew, and Dara. For being fun, for being smart, for being a gang, for being friends, for being family.

To my students: who thrill me with their praise, form me with their criticisms, and shape me ever into a better person for having been touched by their joys, sorrows, stresses, fears, and many thoughts about the present and future.

To my sister, Cindy – who is as smart, and as good, but much braver and stronger than I; who has reconnected and stayed in touch; who has shaped me as I have watched her raise her two children; and whose wisdom and character continues to cast a shadow upon my own: for her

unmatchable understanding and empathy, for her toughness, for the deep beauty in her soul, and for her friendship.

To my brother, Bobby – who is also as smart, and as good as I, but much funnier: for his kindness, for his laughter, for his sincerity; for his smile; for his good-naturedness; for our shared dreams about the brewery; for his eagerness to engage in deep conversation; for his many eagerly extended kindnesses and invitations; and for always believing in me.

To my brother, Aaron – who is at least as smart, perhaps not as good, but is in many ways more impressive than myself: for being someone to whom I can always relate, for being someone whom I am always eager to see, for being someone who is always eager to see me, for visiting me in Los Angeles, for many dungeons and not enough dragons, for being a young kid who managed to set a fine example and be a damn good dad, for his perseverance, for his strength, for our many similarities, and for our many more dungeons and many more dragons.

To my sisters-in-law, Dani and Trish – for their kindness, beauty, and sisterhood.

To Daniel Carrigan – for his support, love, and encouragement

To Dr. Kenneth Dean Burns – who took me in at a moment's notice, introduced me to Sinatra, and opened the door to another life.

To my father, David – who has taught me much, in good days and bad: for his strength of conviction, his independence of thought, and for his tenacity.

To my mother, Carol – for her indomitable love, for her easy-going nature, and for being the kind of smart, tough, strong woman that I am glad to be descended of.

To UCLA – a place that might not have existed, but does, a place where people like me can go to grow and become better, and to every taxpayer who ever helped make it.

Above all, to Dr. Philip J. Kellman – who gave me the gift of this opportunity, who continues to invest in me, who has broken barriers and opened doors, who is the most upstanding of mentors,

and from whom I have been learning every moment that I have spent with him: for teaching me so much about the kind of person I want to be; for being a mentor whose concern for his students blurs the line between advisor and dad; for believing in me; for giving me hope; for knowing so much, and teaching so much; for his deep and sincere interests in perception; for his focus on objective methods, and logical processes and mechanisms; for introducing me to Kanizsa; for being patient and forgiving; for being approachable; for being likeable; for belonging to a tradition that I am proud to inherit; for making me into a scientist, and more than that – a better, more considerate, more critical, carefully-shaped mind.

Thank you, thank you, thank you.

Other acknowledgements: Chapters 1 and 2 are a version of Carrigan, Erlikhman, & Kellman, 'Path Detection and Illusory Contours: Evidence for an Intermediate Representation in Visual Contour Interpolation' (in preparation for publication). Susan Carrigan conducted the experiments, completed the final analysis of the results, created all figures except Figure 6, and wrote the first and final drafts of the manuscript except where noted. Gennady Erlikhman wrote the code to create the stimulus displays, created Figure 6, completed the initial analysis of the results, contributed the first draft of the results section, and assisted with editing. Philip J. Kellman revised sections of the introduction and assisted with editing.

VITA

2011-2012	Undergraduate Research Assistant, The Emotion Lab
2012-2013	Undergraduate Research Assistant, The Chubb-Wright Lab
2012	B.A. in Psychology, *Summa Cum Laude*, University of California, Irvine
2013-2019	Graduate Researcher, The Human Perception Lab
2014	M.A. in Psychology, University of California, Los Angeles
2014-2019	Teaching Fellow, University of California, Los Angeles
2014-2017	UCLA Women's Graduate Student Association Co-Founder
2014, 2015	UCLA Psychology-in-Action Symposium Planning Committee Co-Chair
2015	UCLA Underrepresented Graduate Students in Psychology Social Events Committee Chair
2015-2016	UCLA Psychology Graduate Student Association President

SELECTED PUBLICATIONS

Carrigan, S. B., Palmer, E., & Kellman, P. J. (2016) Differentiating global and local contour completion using a dot localization paradigm. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(12) 1928-1946.

Carrigan, S. B., Erlikhman, G., & Kellman, P. J. (in prep). Path detection and illusory contours: Evidence for an intermediate representation in visual contour interpolation.

Guttman, S. E., **Carrigan, S. B.**, & Kellman, P. J. (in prep). Promiscuous interpolation: Contour completion without perception of completed contours.

SELECTED PRESENTATIONS AT SCIENTIFIC MEETINGS

Carrigan, S. B., & Kellman, P. J. (2019). From early contour linking to perception of continuous objects: Specifying scene constraints in a two-stage model of amodal and modal completion.

Poster presented at the Annual Meeting of the Vision Sciences Society in St. Pete Beach, Florida on May 21st.

Carrigan, S. B., Palmer, E., & Kellman, P. J. (2017). Reliable precision and accuracy differences between local and global amodal completion processes: Results from experiments examining symmetry, familiarity, and regularity completions. Poster presented at the Annual Meeting of the Vision Sciences Society in St. Pete Beach, Florida on May 24th.

Kellman, P. J., **Carrigan, S. B.**, & Erlikhman, G. (2017). Path integration and illusory contours: Evidence for an intermediate representation in visual contour interpolation. Talk presented at the Forty-Second Annual Interdisciplinary Conference in Breckenridge, CO, on January 29th.

Kellman, P. J., **Carrigan, S. B.**, & Erlikhman, G. (2016). Path integration and illusory contours: Evidence for an intermediate representation in visual contour interpolation. Talk presented at the annual workshop of the Configural Processing Consortium, Boston, MA, on November 16th.

Carrigan, S. B., Kellman, P. J. (2016). Local and global amodal completion: Revealing separable processes using a dot localization method. Poster presented at the European Conference on Visual Perception in Barcelona, Spain on August 29th.

Carrigan, S. B., Palmer, E., & Kellman, P. J. (2015). Differentiating local and global processes in amodal completion through dot localization. Talk presented at the Annual Meeting of the Vision Sciences Society, St. Pete Beach, FL on May 19th.

Carrigan, S. B., Palmer, E., & Kellman, P. J. (2014). Differentiating processes in amodal completion using a dot localization method. Talk presented at the annual workshop of the Configural Processing Consortium, Long Beach, CA on November 19t

CHAPTER 1

Introduction

Among the most important functions of visual perception is the construction of representations of complete objects and continuous surfaces from information in reflected light. In the guidance of thought and action, it would be hard to overestimate the importance of such representations. Matching their importance is their complexity. Coherent objects and continuous surfaces are not "given" in visual input; they result from extracting particular features and relations from the input and performing particular transformations that produce the objects of our experience. Recovering objects and the continuous contours that define them from visual stimulation requires elaborate routines to overcome what has been termed the *fragmentation problem* in vision (Kellman, 2003c).

Consider the scene in Figure 1. At a basic level of description, what aspects of the visual input define visible areas that are part of a single object? We might think that areas in which spectral and intensity characteristics of light are relatively homogeneous, or smoothly changing, are often part of the same object or surface, and Figure 1 contains many such areas, but these areas very rarely comprise the physical objects in the scene. The following facts: that most objects are opaque, that light moves in straight lines, and that our retinas are flat, combine to mean that most objects and surfaces will be partly occluded in our retinal projections. This is the fragmentation problem: what is connected and continuous in the world is not so on the retina. Far from being occasional or exceptional, partial occlusion is pervasive in viewing ordinary scenes. In Figure 1, for example, the sparrow is occluded by the branch and leaves. Our description of this image as being a scene with a sparrow, branch, and leaves reflects the successful operation of object formation processes. We have

little difficulty in seeing a single sparrow and recovering its shape, for example, despite the fact that the sparrow projects to the eyes in approximately 15 discontinuous patches, with projections of nearer objects separating them.

Figure 1. Illustration of the fragmentation problem in visual object perception. We see the sparrow as a complete object with well-defined boundaries despite occlusion by the branch and leaves. Interpolation processes produce representations of complete objects despite the fact that a single object may project many spatially separated regions to each retina, separated by projections of nearer objects. (See text.)

How does the visual system construct unitary representations of objects and surfaces despite many areas of occlusion? Figure 2 illustrates the two separate and complementary mechanisms that contribute to this process: *contour interpolation* (Ben-Yosef & Ben-Shahar, 2012; Geisler & Perry, 2009; Grossberg & Mingolla, 1985; Heitger, von der Heydt et al, 1998; Kanizsa & Gerbino, 1982; Kalar et al, 2010; Kellman & Shipley, 1991; Maertens & Shapley, 2008; Michotte, Thines & Crabbe, 1991/1964; Keane, Lu, & Kellman, 2007; Kellman, Garrigan, & Shipley, 2005; Palmer, Kellman &

Shipley, 2006) and *surface interpolation* (Yin, Kellman & Shipley, 1997, 2000; He & Ooi, 1998a, 1998b; Su, He & Ooi, 2010a, 2010b).

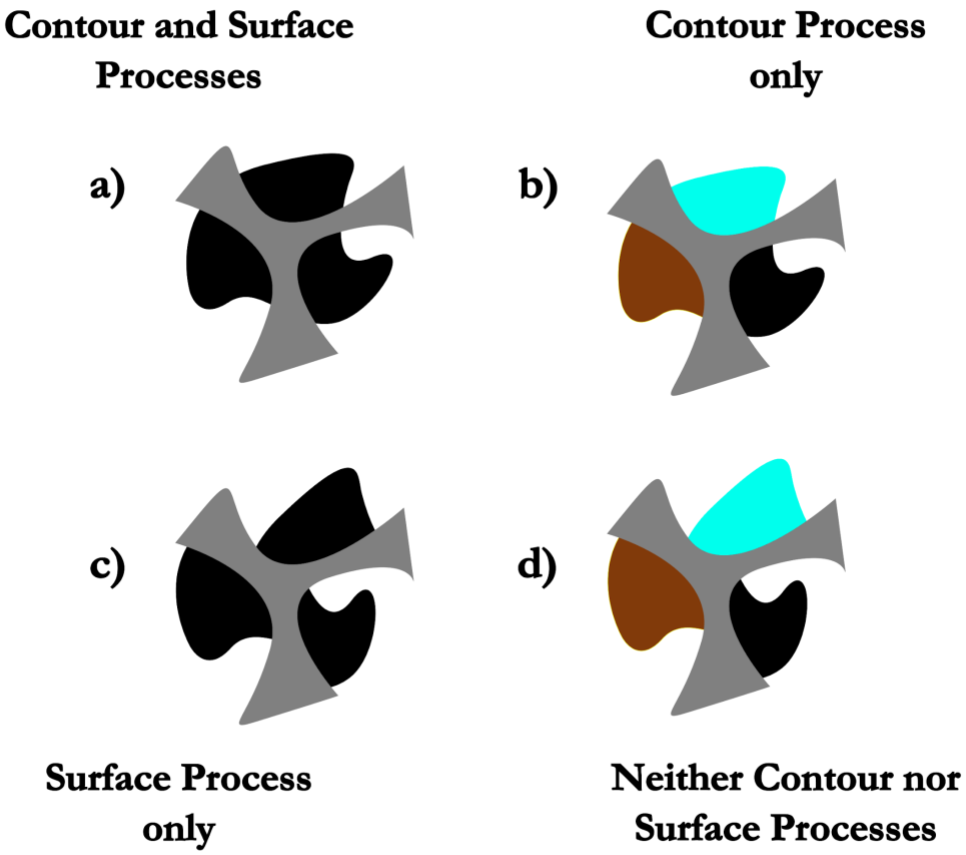

Figure 2. Contour and surface interpolation processes. a) The three black surface regions are connected by contour interpolation, due to relatable edges, and surface interpolation, due to matching surface properties. A single black object is seen as continuous behind the occluder. b) The black, brown, and teal regions are connected by contour interpolation, due to relatable edges, but surface interpolation is blocked due to the different surface colors of the three regions. c) The three black regions are not connected by contour interpolation, due to disruption of contour relatability, but they connect behind the occluder, with indefinite edges, due to surface interpolation. d) Both contour and surface interpolation have been disrupted (relative to (a)). Although each of the black, brown, and teal regions is seen to continue behind the occluder, there is little or no impression that they form a connected object.

Determining an object's boundaries and shape depends foremost on contour interpolation. Considerable evidence has shown that contour interpolation depends on a specific set of spatial relationships between the orientations of spatially separated edge fragments and the positions of the points where they terminate in contour junctions (Kellman & Shipley, 1991; Shipley & Kellman, 1990). Such junctions are an important invariant in vision, as they will always occur at locations where an occluding surface passes in front of a partly occluded object's contours (Hoffman & Richards, 1984; Kellman & Shipley, 1991). The process of contour interpolation may be thought of as a transformation that takes as input the positions and orientations of a pair of edges that conform to certain geometric relations and produces as its output a representation of a continuous contour connection between them. The edge relations that enter into this transformation have been formalized in the notion of contour *relatability* (Kellman & Shipley, 1991; Kellman, Garrigan & Shipley, 2005; see below). Contour relatability specifies that, in order to enter into this transformation, a pair of edges must form an obtuse or right angle at the point of intersection of their linear extensions (the 90 degree constraint), and must be able to be connected by a smooth, monotonic curve (the monotonicity constraint). Edges that violate these constraints, such as those of the black object fragments in Figure 2c, are not perceived to be connected.

Surface interpolation has a complementary relationship to contour interpolation. Under occlusion, surface qualities "spread" unless confined by real or interpolated edges. Spatially separate areas can be perceived to be connected if they are similar in surface properties, including lightness, color, and texture (Kellman & Shipley, 1991; Yin, Kellman & Shipley, 1997, 2000). Where clear 3D information is available, surface interpolation depends on geometric relations between surface patches; these 3D relations are closely related to contour relatability (Fantoni, Hilger, Gerbino & Kellman, 2008; Kellman, Garrigan, & Shipley 2005; Kellman, Garrigan, Shipley, Yin, & Machado, 2005).

Figure 2 illustrated contour and surface interpolation processes with partly occluded objects (amodal completion). These same processes operate in the closely related phenomena of illusory contours and objects (modal completion) (Albert, 2007; Kalar et al., 2010; Kellman & Shipley, 1991; Kellman, Yin, & Shipley, 1998; Murray, Foxe, Javitt, & Foxe, 2004; Regolin & Vallortigara, 1995; Ringach & Shapley, 1996). Figure 3 shows illusory and partly occluded object displays that have corresponding physically-specified contours and gaps between them. For both modal and amodal completion, the result of a case of contour interpolation is a continuous contour connecting two input edges. In modal completion, these connections appear as illusory contours, whereas in modal completion, they appear as partly-occluded contours.

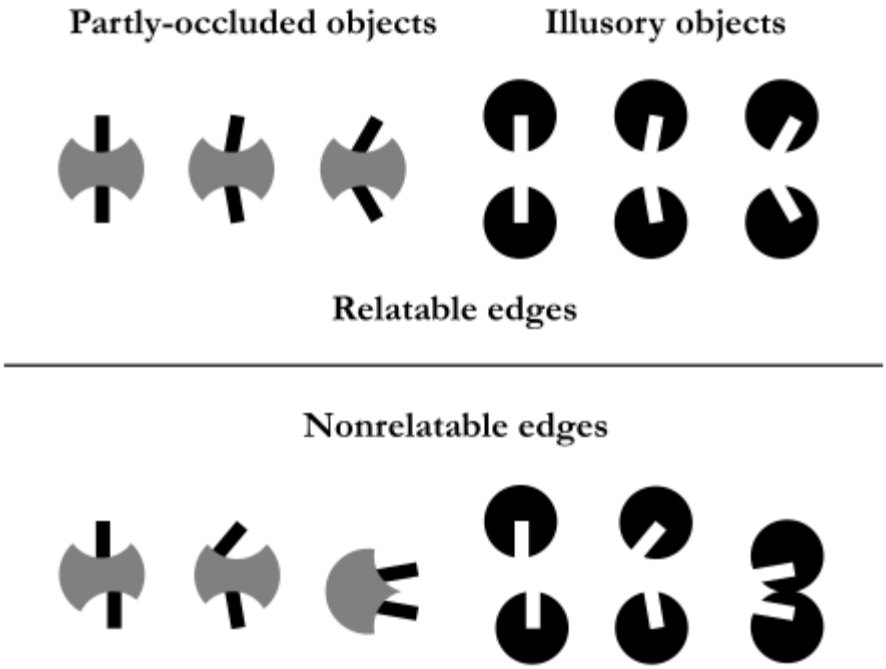

Figure 3. Contour interpolation in partly occluded objects (amodal completion) and illusory objects (modal completion). Top left: Each of the three displays contains two separated black regions, the edges of which satisfy the geometry of contour relatability. In each case, the two black regions are seen as belonging to a single object, amodally completed behind an occluder. Top right: Illusory contour inducing elements containing identical physically specified, relatable edges as in the partly occluded object displays to the left. In each case, the top and bottom cut-out areas are seen as pieces of illusory objects, with perceived continuous contours connecting

the top and bottom pieces across a gap. Bottom left: Identical black regions as in the top left displays, now with contour relatability disrupted. In each display, two disconnected black objects are seen. Bottom right: Circular elements with the cutout areas as in the top right, but with contour relatability disrupted by spatial relations identical to the displays on the left. No illusory contours are perceived. Redrawn from Kellman, P. J. (2003a). Visual perception of objects and boundaries: A four-dimensional approach. In R. Kimchi, M. Behrmann, & C. R. Olson (Eds.), Perceptual organization in vision: Behavioral and neural perspectives. Mahwah, NJ: Lawrence Erlbaum Associates.

A variety of evidence and argument supports the idea that modal and amodal completion share, at least in part, a common interpolation process (Kellman & Shipley, 1991; Kellman, Yin, & Shipley, 1998; Ringach & Shapley, 1996; Murray, Foxe, Javitt, & Foxe 2004; Stevens, 1976; Regolin & Vallortigara, 1995; Albert, 2007; Kalar et al., 2010). Although differing in appearance (i.e., in front of or behind some other surface) both of these outputs of contour interpolation are equally "real" in that it has been shown experimentally that these contour completions are represented in the brain, and these representations support precise spatial localization (Carrigan, Palmer, & Kellman, 2016; Ringach & Shapley, 1996; Gold, Murray, Bennett, & Sekuler, 2000; Guttman & Kellman, 2004; Kellman, 2003b; Keane, Lu & Kellman, 2007). Contour interpolation is crucial in biological vision as a bridge between early, spatially localized, subsymbolic encoding of stimulus input and the representation of contour tokens and object boundaries that define unified objects and receive shape descriptions (Kellman, Garrigan & Palmer, 2010).

1.1 A two-stage theory of contour interpolation

The preceding background on interpolation processes in object formation provides the backdrop for the focus of the present paper: a two-stage theory of the contour interpolation process. Some existing evidence suggests that contour interpolation is a multistage process, with

early interpolated contours appearing in an intermediate representation but not necessarily appearing in our phenomenological perception. Guttman & Kellman (2001) reported results from a series of psychophysical experiments that utilized a fat-thin classification task. In this task, individual inducing "pacman" shaped elements are rotated slightly, either clockwise or counter-clockwise, from trial to trial. When elements are relatable, the percept of a "fat" or a "thin" square results. When the elements are nonrelatable, observers can attempt to classify the images based on whether individual inducers are rotated clockwise versus counter-clockwise. These experiments utilized both filled inducers, which give rise to the perception of an illusory square, as well as outline inducers, for which no interpolated contours are consciously perceived (see Figure 4). In addition, the experiments utilized stimuli with relatable inducers, and stimuli with nonrelatable inducers.

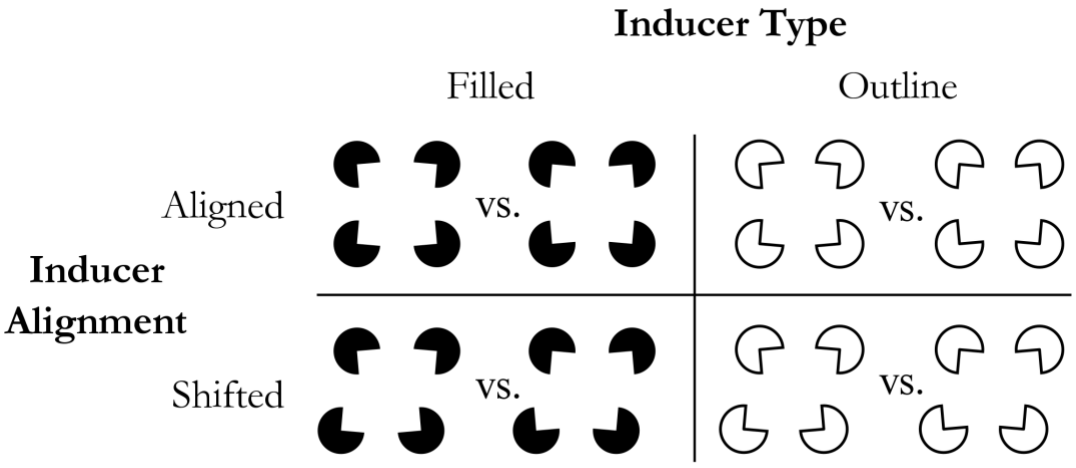

Figure 4. Sample of stimuli from one experiment in Guttman & Kellman (2001). Participants viewed modified Kanizsa square images with either filled or outline inducers. Evidence from control conditions suggests that good performance depends on judging the overall shape of the illusory figure as opposed to the orientations of individual elements. In different manipulations, relatability was disrupted by shifting the inducers, flipping the inducers outward or rounding the tangent discontinuities. Similar to previous experiments, the disruptions to relatability drastically impaired performance. However, performance (both speed and accuracy) was generally unimpaired by the outline manipulation, suggesting that interpolation had occurred for those stimuli, despite the absence of any perceived illusory edge.

The results were well in line with those of previous experiments in that disruptions to relatability drastically impaired performance. However, performance was generally unimpaired by the outline manipulation, suggesting that interpolation had occurred in the presence of relatable edge fragments, *even for those stimuli for which interpolations were not subjectively perceived*, and that those interpolated curves did exist at some stage of the perceptual process, available to be used by observers engaged in the task. In subsequent experiments, the authors implemented several other controls to eliminate a variety of alternative explanations, including the use of local orientation cues from individual pac-man elements. Other work suggests that outline Kanizsa square stimuli produce the same visual search performance as filled-in versions, even though the outline stimuli do not support the subjective perception of illusory contours (Gurnsey, Poirier, & Gascon, 1996).

These results provide some evidence that interpolations are carried out in an early stage, that interpolation proceeds across all relatable edges, that these interpolated edges are maintained in an intermediate representation, and that some of them will not, ultimately, appear in our phenomenological perception. That interpolations are accessible to be utilized in the fat-thin task, even for stimuli for which they are not consciously perceived, indicates that there may well be other paradigms that can access this intermediate representation.

What might this look like? That is, how might the output of the intermediate representation evidence itself in the absence of perceived interpolated contours? It is worth taking a moment to think about the impacts of interpolations across relatable edge fragments. The work from the fat-thin task suggests that, even if an interpolation is not consciously perceived, its trajectory or path (e.g. a convex or concave contour) exists at some stage as a link between the inducing fragments. Even if the interpolated edge is not perceived, its impacts on the grouping of fragments, and the perception of paths or trajectories may be.

CHAPTER 2

Path detection and illusory contours: Evidence for an intermediate representation in visual contour interpolation

2.1 Synopsis

Field, Hayes, & Hess (1993) introduced the phenomenon of path detection: certain position and orientation relations among spatially separated Gabor elements facilitate detection of a group or path. These relations appear to be the same as those of the *contour relatability* geometry that governs contour interpolation in modal and amodal completion. However, interpolated contours are not perceived to connect the Gabor elements in path detection, and the relation between path detection and contour interpolation has remained unclear. Here, I propose that path detection accesses an intermediate representation in a two-stage process of object formation (Kellman, Guttman & Wickens, 2001). An initial contour linking stage joins all oriented edges, terminating in contour junctions, that satisfy contour relatability; however, additional constraints in a second stage, such as consistency of border ownership and surface continuity, must be realized for visible contour connections to appear in final scene representations. A set of experiments tests human observers' path detection and illusory contour perception with paths made from Gabor elements and several types of modified elements that supported surface continuity within linked contours. Misalignment of parallel elements, which has been found to disrupt modal and amodal completion at about 15 arcmin of retinal misalignment (Hilger, 2009), was also manipulated. All four element types produced similar performance in path detection, with path detection falling to chance around 15 arcminutes of element misalignment. Only paths made of elements fulfilling the additional constraint of surface continuity gave rise to reliable perception of illusory contours; for those displays, illusory contour perception also disappeared by 15 arcminutes of misalignment. The results

support a two-stage model of object formation, in which linkage in an intermediate representation governs path detection performance, but perceived contour connections require both the initial linkage and satisfaction of additional constraints.

Keywords: promiscuous interpolation, contour interpolation, contour integration, illusory figures, amodal completion, modal completion, path detection

The previous chapter concluded by saying that there may be other paradigms, besides the fat-thin task, that reveal the intermediate output of the first stage. It was also noted that the phenomenological perception might be something like the sensitivity to a path or trajectory connecting individual edge elements, without the subjective perception of an interpolated contour connecting them.

This consideration brings to mind work from another well known paradigm. In the *path detection* literature, introduced by Field, Hayes & Hess (1993), spatially separated Gabor elements arranged according to certain geometric constraints are perceived to be unified, and achieve perceptual "pop-out" from randomly arranged background elements. These investigators studied arrays of Gabor patches such as those shown in Figure 5.

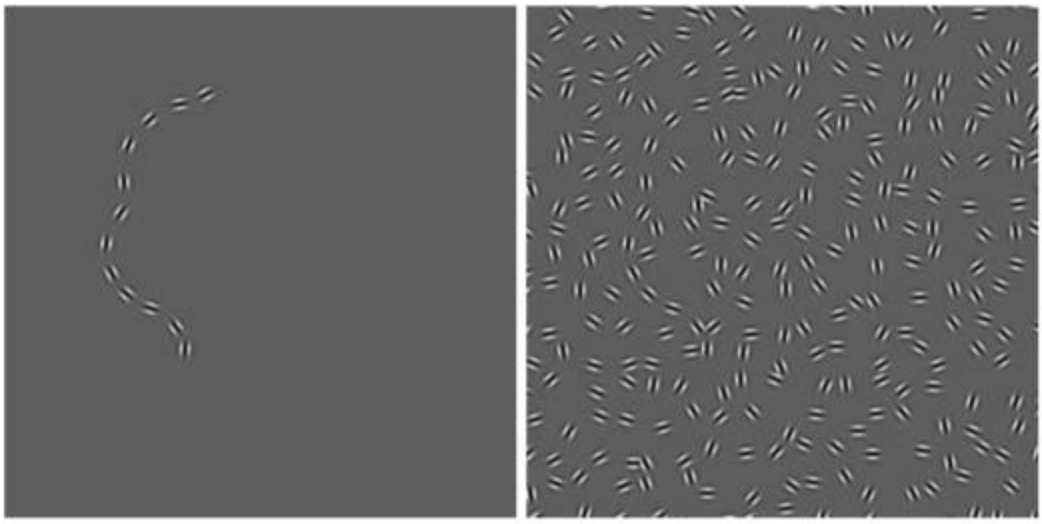

Figure 5. Illustration of the path detection paradigm. At left is a target composed of Gabor elements whose positions and orientations define a path. At right, the target is placed amid a background of randomly arranged Gabor elements. (Reprinted from Field, Hayes & Hess, 1993).

In the "path paradigm," a two interval forced-choice procedure is typically used. The display in both intervals is filled with randomly placed elements; in one of the displays, hidden among the randomly placed elements is a set of elements that defines a path. If certain geometric relations

between pairs of Gabor elements within a path are satisfied, the path is more easily detected. The observer is tasked with identifying which of the two arrays of randomly arranged Gabor elements, presented at different time intervals, contained a hidden path.

Remarkably, the geometric relations upon which path detection depends appear to be the same as those described by contour relatability (Field, Hayes, & Hess, 1993; Hess & Field, 1999; Kellman, Garrigan, & Shipley, 2005; Kellman, Guttman, & Wickens, 2001). Specifically, performance (in terms of proportion correct) in path detection tasks is best when the elements making up the path are close to collinear; performance worsens with increasing inter-element orientation differences, reaching chance performance somewhere around ninety degrees, and is also drastically reduced for paths that require an inflection point between some elements (Field, Hayes, & Hess, 1993). These findings are very suggestive of the contour relatability geometry that constrains modal and amodal completion (Kellman & Shipley, 1991; Field, Hayes, & Hess, 1993; Hess & Field, 1999; Kellman, Garrigan, & Shipley, 2005; Kellman, Guttman, & Wickens, 2001). Inter-element orientation differences beyond 90 degrees violate the ninety degree constraint, and the introduction of inflection points violates the monotonicity constraint (Kellman & Shipley, 1991).

Given that contour interpolation and path detection appear to rely on the same geometric relations, it would be parsimonious to hypothesize that they depend on the same process. There is a problem with this idea, however. As can be seen in the figure, no contours are perceived to connect pairs of Gabor elements in a path. Since there is no occluder between elements, continuous connections, if perceived, would appear as illusory contours connecting each pair of elements along the path. No such illusory contours are visible in path displays made of Gabor elements on grey backgrounds.

Nevertheless, there is some kind of relation among elements of a path, allowing it to facilitate detection in noise. At the same time, the units that make up the path appear as discrete, bounded

elements. So, what is the function of path formation? If a "path" is not a bounded object, what is it? One possibility is that, instead of relating to the perception of discrete objects, paths represent some kind of *grouping* phenomenon. Grouping is a term that has been sometimes applied to the phenomenal experience of seeing collections of separable, discrete objects as somehow related (Koffka, 1915/1938; Kubovy & van den Berg, 2008; Wertheimer, 1923/1938, 1958; see Wagemans, Elder, Kubovy, Palmer, Peterson, Singh, & von der Heydt (2012) for a review). What grouping implies varies in different cases where it has been discussed, and it is often a somewhat vague concept. However, phenomena such as seeing a flock of birds or a collection of dots as being all part of the same group, suggest psychological reality for some notions of grouping. Along these lines, Field, Hayes & Hess (1993) discussed possible applications of their findings to texture perception.

A different idea is that, despite the absence of illusory contours, path formation has some connection to contour interpolation and object perception, in some way that has not been well-specified. Indeed, this suggestion seems implicit in some work. For example, Field, Hayes & Hess (1993) show a picture of a tree whose branches are partly occluded by a grid in a window as a suggestive example (their Figure 1), but they did not connect their results to modal or amodal completion. Paradoxically, much of the literature on the path paradigm can be accessed by searching under the term "contour integration", which is seldom given any explicit definition. The term seems to have originated with Field, Hayes, & Hess (1993), as it appears in the title of their paper. No definition is given there, however, and the phrase in fact does not appear anywhere in the body of the paper. If there is a connection between path phenomena and object perception, some more specific connecting ideas are needed. The primary puzzle is obvious: objects have continuous boundaries, and interpolation processes involved in deriving objects from fragments result in representations of continuous contours (amodal or modal). The lack of perceived contour

connections between elements is the difficulty for relating path phenomena to known phenomena and processes of contour interpolation.

As a result of the ambiguity of what paths really indicate, the interesting phenomenon and elegant paradigm of path detection has grown into a sizeable research literature, without a clear connection to other phenomena or processes. The original Field, Hayes, & Hess (1993) has been cited 1,611 times according to Google Scholar as of this writing. Among these, I can find few that suggest a connection to object perception and none that explain how that connection might operate. In fact, some make an explicit distinction between contour integration and completion phenomena:

> "Whereas contour integration refers to the integration of discrete elements such as dots or oriented elements, contour completion refers to the integration of smooth extended contours, separated by gaps due to occlusion or camouflage." (Wagemans et al., 2012; p. 1191).

In this first set of experiments, I aim to make clear the relation of path detection to contour interpolation and object perception processes. I argue that the second of the two ideas above is correct: *Path detection phenomena derive from (and reveal) processes of contour interpolation in object perception*. I make this connection by presenting some theoretical background and using it to motivate two experiments that test clear predictions about the relation of path phenomena and illusory contour perception.

2.2 The two-stage theory and path detection

The mystery in relating path detection to contour interpolation is that successive elements in perceived paths are not seen as connected by illusory contours. The answer I propose is that contour interpolation in object formation is a two-stage process. An initial linking stage receives

pairs of relatable edge fragments as input and *promiscuously* interpolates a connection between all such pairs. However, the output of this initial stage exists in an intermediate representation that is not phenomenologically accessible nor identical to final scene representations. Perceived continuous contour connections in final scene representations are determined both by this initial contour linking stage and a second stage in which additional scene constraints are implemented. This second stage involves constraints relating to factors such as border ownership, continuity of surface luminance and color, and corroboration across multiple spatial frequency channels. The second stage has the effect of sustaining, weakening, or even deleting contour linkages generated in the first stage. The first stage is automatic, based on neural mechanisms wired up to respond to particular geometric relations, with a minimum of other influences. The second stage is more about constraint satisfaction in using contour and surface information to produce self-consistent, sensible, representations of objects, surfaces, and their arrangements in the physical world.

The idea of an intermediate representation is not new; it has been around for at least as long as Marr's classic work (1982; Marr & Hildreth, 1980). The logic behind it is clear. If, at an early stage, activations corresponding to edges are obtained, it will still need to be determined 1) which edges are real and which are spurious, and 2) which edges belong to the same surface or texture. Any interpolation process will have to address these same concerns. Because the truth of the objects cannot be known at the time interpolation must proceed, it will inevitably be the case that some interpolations produced turn out to connect edges that are not connected in the real world. In order to determine which interpolations are good ones, other scene cues will have to be taken into consideration.

The two-stage theory thus implies that, at some level of the perceptual process, a representation of the interpolated contours exists, even when they are not visible, and therefore they may still have some impact on perception. This prediction of the two-stage theory accounts for both the sensitivity

to paths in the path detection literature as well as the sensitivity to interpolated trajectories in the fat-thin task, in the absence of any perceived interpolated contours.

No interpolated contours are perceived in the path detection stimuli. However, simple modifications to these stimuli can produce the perception of illusory contours. An example of a path detection stimulus is shown in Figure 5. Examining the figure, one may note the absence of any occluder. One may also note that the interiors of the Gabors are black while the background upon which the Gabors lie is grey. In the absence of an occluder, interpolated contours are bound to be spurious unless these interpolated edges complete an object whose surface color and brightness match that of the background. In the path displays used by Field, Hayes, & Hess (1993), however, surface discontinuities exist between the Gabor interiors and the background, in the absence of any occluder.

I propose that it is this cue of surface discontinuity in the absence of an occluder that prevents interpolated contours from appearing in the final scene representation. If this hypothesis is correct, then altering the black centers of the Gabors to match the grey background will cause the physically specified edges to be joined in our phenomenological perception by illusory edges that form the boundaries of a grey bar, which runs across and partially occludes the target elements. This manipulation should have no impact on path detection performance if the path detection task relies on interpolated edges that appear in the intermediate stage rather than our phenomenological perception. However, it should be clear to observers that interpolated contours occur for the modified Gabor elements, but not for the traditional elements that give rise to surface discontinuity. In addition, as variables that impact relatability are manipulated, accuracy performance for both kinds of displays (those that use traditional Gabor elements as well as those using modified grey-centered elements) should follow the same pattern of deterioration.

2.3 Specific, testable predictions

The paragraph above contains three testable predictions: 1) variables that impact relatability should have equal effects on accuracy performance for modified (grey-centered) and traditional Gabors, 2) it should be clear to observers that interpolated contours are visible for the modified elements, but not for the traditional Gabor elements, and 3) manipulations to variables known to impact subjective perception of illusory contours should not impact objective path detection performance. Experiments 1 and 2 will test these predictions.

Experiment 1 was designed to test the first and third predictions given above. That is, path detection performance should be identical for both traditional as well as modified Gabor elements, and, as a known constraint on contour interpolation is manipulated, performance for both kinds of displays should deteriorate in an identical fashion (that is, the manipulation of the centers of the Gabors, which results in the perception of illusory contours, should not impact objective path detection performance). One constraint on contour interpolation that has not yet been examined in path detection is that of an absolute retinal tolerance for misalignments of about 15 arcminutes. Absolute (as opposed to relative) tolerances are fairly unusual in visual perception, as, in most cases, perception of objects tends not to change as an object gets nearer or further away. However, in the case of contour interpolation, slightly misaligned inducers can give rise to the perception of whole objects at further distances (i.e., at distances for which the misalignment is less than 15 arcminutes) while appearing to be nothing more than separate fragments at nearer distances (for which the misalignment of the inducing edges is 15 arcminutes or more) (Hilger, 2009).

As discussed above, research by Field, Hayes, and Hess (1993) has demonstrated that constraints on path detection performance coincide with both the 90 degree constraint and the monotonicity constraint of relatability. Demonstrating that path detection performance also coincides with the

lesser known, even more unusual constraint of a retinal tolerance for misalignments of 15 arcminutes will provide even stronger evidence that path detection phenomena are really revealing the intermediate output of the promiscuous stage of contour interpolation.

Stimuli to be used in Experiments 1 and 2 are shown in Figure 6. The misalignment of target elements was manipulated, utilizing misalignment levels of 0, 4, 8, 12, 16, 20, 24, and 28 arcminutes. Based on previous research that suggests an absolute retinal tolerance of around 15 arcmin for interpolation phenomena (Hilger, 2009), I expected that accuracy performance in the path detection task would decrease with increasing misalignment, reaching chance somewhere around 15 arcminutes of misalignment. In addition, I hypothesized that this pattern of results would hold true for both modified and traditional Gabor elements.

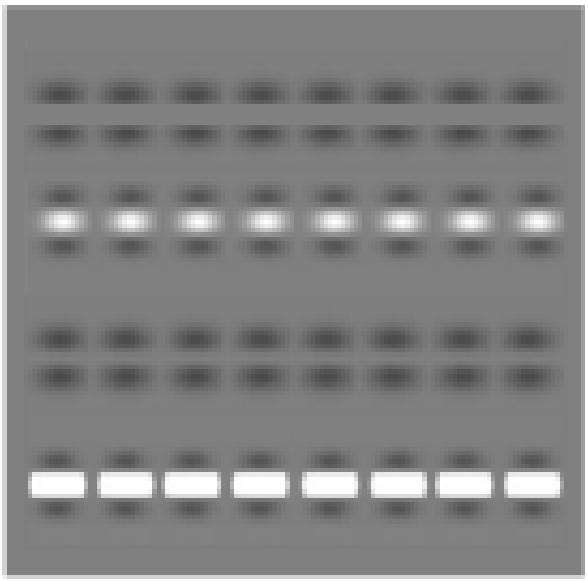

Figure 6. Element types used in Experiments 1 and 2. In the first row are grey-centered step-edge elements. These are Gabor elements that have been modified by cutting the centers out to produce a crisp edge. In the second row are classic Gabor elements ("white-centered Gabors"). In the third row are Gabor elements that have been modified by changing the pixel intensity of the white pixels to an intensity that matches the grey background ("grey-centered Gabors"). On the bottom row are step edge elements that have had a crisp strip of white inserted into their centers ("white-centered step-edges").

Experiment 2 takes the targets out of the path detection fields, presenting them as horizontal paths for an unlimited duration while participants are asked to give a subjective rating of illusory contour clarity. The same stimulus elements and misalignment levels were used as in Experiment 1. This achieves two goals. First, the subjective ratings confirm that participants do report seeing illusory contours for the Gabor elements with grey centers, but not for the Gabor elements with white centers. Second, for those element types that do give rise to the perception of illusory contours, the deterioration of illusory contour clarity ratings mirrors the deterioration of accuracy ratings in the objective path detection task: the subjective ratings decrease with increasing levels of misalignment, bottoming out (reaching baseline) at the same point (around 15 arcminutes of misalignment) that path detection performance reaches chance levels.

2.4 Experiment 1: Objective path detection performance for paths of varying element types and varying misalignment

Four kinds of elements were used to construct the paths as shown in Figure 6 : (1) standard Gabor elements with white centers ("white-centered Gabors"), (2) Gabor elements that have had their central region replaced with a crisp swatch of color that matches the grey background; the result is a crisp "step-edge" transition between the dark surround and the interior grey regions (I call these "grey-centered step-edges"), (3) Gabor elements with a grey center and a gradual, rather than crisp transition from the outer dark to central grey regions ("grey-centered Gabors"), and (4) step-edge elements with white centers ("white-centered step-edges"): these elements are identical to the step-edges with grey centers, except that their interior regions have been replaced with a crisp swatch of white rather than grey. All elements were placed on a grey background.

Because the white-centered Gabors and white-centered step-edges had white centers that did not match the background, no illusory contours were expected to form between the elements. The grey-centered Gabors and grey-centered step-edges, however, had grey centers that matched the background, which was expected to enable the formation of either sharp or fuzzy illusory contours between the elements. For these elements, the boundary between the grey center and the dark surround produces an L-junction at either end of the element. The illusory contours that form between elements begin and end at these L-junctions. For the grey Gabor, this L-junction is fuzzy, and thus the illusory contours that form are fuzzy or diffuse (see Figure 33). For the step-edge, this junction is quite crisp, and thus the illusory contours that form have similarly sharp edges.

The white-centered Gabors and white-centered step-edges also produce junctions at either end of the element - in this case, T-junctions. The T-junction is the result of the element's white center abutting the grey surround. This additional edge produced between the element's center and its background suggests an interpretation of the scene that runs contrary to the formation of illusory contours unless an occluder is present. However, the other aspects of the scene are not consistent with the presence of an occluder. The consideration of these scene cues was predicted to have the effect, at the second stage, of causing any interpolated edges produced in the first stage to not ultimately, appear in the final scene representation.

2.4.1 Method

Participants. Twenty-four participants (19 female, ages 18-38, M_{age} 20.96 years) completed the white Gabor and grey step-edge conditions, and 19 participants (14 female, ages 17-36, M_{age}: 21.89 years) completed the grey Gabor and white step-edge conditions (see below for details). All participants provided informed consent, were students attending the University of California, Los

Angeles, and reported having normal or corrected-to-normal vision. Participants received partial course credit for their participation.

Apparatus. Stimuli were created and displayed using the MATLAB programming language and the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). Stimuli were presented on a Viewsonic G250 CRT monitor powered by a MacPro 4 with a 2.66 GHz Quad-Core Intel Xeon processor and an NVidia GeForce GT120 graphics card. The monitor has a height of approximately 30 cm and a width of approximately 40 cm. The resolution of the monitor was set to 1024 x 768 pixels and a refresh rate of 60 Hz.

Stimulus. Displays consisted of fields of Gabor or Gabor-like elements. Fields contained on average approximately 100 individual elements arranged in a 10x10 grid. Elements were arranged in such a manner so as to maintain a constant density throughout the display, even when paths were added. Displays were created using the GERT toolbox version 1.20 (Demeyer & Machilsen, 2012) which employs methods for controlling the density of the display.

Four types of elements were used to create the fields of elements, as described above. All elements were modifications of Gabor patches, which are oriented sinusoidal patterns multiplied by a circular Gaussian window (see, e.g. Field, Hayes, & Hess, 1993). A sine phase was used so that the center of all elements was white, unless otherwise modified. Elements had a peak spatial frequency of 4.62 cycles/deg and a σ (SD) of 4.0 arcmin. The contrast of the elements was set to maximum. The background of the screen was grey. The display area within which the Gabors could appear was restricted to a 12.5º x 12.5º square region centered on the screen.

Grey-centered step-edges were created by replacing the pixel intensity values of all pixels in the middle of the Gabor patch with the pixel intensity value of the background. The central region was defined as the region between the two darkest bands of the Gabor patch, i.e. from the center of each band. The grey-centered Gabor was created in a similar manner, except the central region was

defined as the region between the two bands containing pixels brighter than the background. This left in place a strip of pixels which faded gradually from the darkest center of each band to the background grey color. This region of gradually increasing pixel intensity was entirely removed in the grey step-edge elements. The white-centered step-edge elements were created in the exact same manner as the grey step-edge elements, except that, instead of replacing the central region with pixels the same intensity as the background, white pixels were used.

On each trial, two fields of elements were shown sequentially. Each field remained on the screen for one second, separated by a 500ms inter-stimulus interval which contained a blank screen. Both fields were made up of the same type of element. One field contained, embedded within it, seven elements placed along a straight path. The length of the path was 6.67 degrees of visual angle from the center of the first to the center of the last element. There was therefore a distance of approximately 1.11 degrees between the centers of each element. The GERT toolbox allowed for the embedding of the path in a field of other elements while maintaining approximately the same total amount of elements and the same density as a display which did not contain the path. The path's position in the field of other elements and its overall orientation were randomized on each trial. Apart from the elements composing the path, all elements in both displays had random orientations. Whether the first or second display contained the path was randomized across trials.

Elements comprising a path were offset from the path by one of eight levels of misalignment: 0, 4, 8, 12, 16, 20, 24, and 28 arcminutes. Misalignment was applied by shifting every other element in the path, starting with the second, by the specified number of arcminutes. To prevent alignment of every other element, element positions were further jittered. The amount of jitter was determined independently for each element and was drawn from a normal distribution with a mean of zero and a standard deviation equal to half of the maximum possible misalignment. Position jitter was applied to all elements along the path, even those that were not shifted away from the path.

Each display lasted for one second, separated by a 500 ms inter-stimulus interval which contained a blank screen. Each trial was repeated 20 times, for a total of 2 (element types per experiment) x 8 (misalignments) x 20 (repetitions): 320 trials (excluding the 10 practice trials). Trial order was randomized in the experiment. The experiment lasted approximately 40 minutes.

Procedure. The four different Gabor types were split in two and were run between-subjects. One group of subjects was shown white-centered Gabors and gray-centered step-edges, while the other group was shown gray-centered Gabors and white-centered step-edges. This was done to shorten the length of the experiment for each participant.

Participants sat at a table at a viewing distance of 134.3 cm from the monitor. Their head was stabilized by a chinrest and forehead guard. The room was dark with the only illumination coming from the monitor. After being given verbal instructions, participants were shown several sample stimuli on the computer screen. They were first shown examples of each kind of stimulus path (i.e., specific to their group, as described above) arranged horizontally on the screen with no other elements. They were told that they would be asked to detect such paths in displays where they would be "hidden". After pressing a key, they were shown a display with a path embedded in a field of Gabors with randomly oriented elements. A second key press highlighted the path within the display. They were then told that, in the course of the experiment, they would see two such images containing fields of Gabors in sequence. They were instructed that one of the images would contain the hidden path, and that the other would not, and that their task was to correctly indicate which of the two images had contained the hidden path. They then completed 10 practice trials in which they received feedback. All paths in the examples and practice trials had 0 misalignment. After the practice trials, participants were told that they would no longer receive feedback and that, moving forward, the paths would not be perfectly aligned. Along with this instruction, they were shown an example of a horizontally oriented path whose elements were misaligned. They then continued on

with the study until completion. Each trial type was repeated 20 times over the course of the experimental session, for a total of 2 (element types per group) x 8 (misalignments) x 20 (repetitions): 320 trials (excluding the 10 practice trials). Trial order was randomized in the experiment. A break was given every 100 trials during which participants were encouraged to rest. The experiment lasted approximately 40 minutes.

2.4.2 Results

Path detection accuracy on the two-interval forced-choice task is shown in Figure 7, separated by each of the element types. Path detection accuracy declined gradually as a function of element misalignment for all element types up until approximately 16 arcminutes of misalignment, at which point performance plateaued at slightly above chance levels.

Because the data were collected between-subjects (white-centered Gabors and gray-centered step-edges for one group, and gray-centered Gabors and white-centered step-edges for the other), the term "element type" is used to describe whether Gabors or step-edges were used, regardless of whether the elements had grey or white centers. The term "element set" is used to indicate whether the first set or second set of elements were used. The accuracy data were submitted to a 2 (element types: Gabors or step-edges) x 2 (element set: first or second between-subjects experimental group) x 8 (misalignments) mixed ANOVA. All F statistics reported below have been Greenhouse-Geisser corrected. There was a main effect of misalignment ($F(4.67, 191.31) = 162.73$, $p < 0.001$, $\eta_{p2} = 0.80$), reflecting the fact that accuracy decreased as a function misalignment. There was a main effect of element type ($F(1, 41) = 18.77$, $p < 0.001$, $\eta_{p2} = 0.31$), but no significant between-subjects effect of element set ($F(1, 41) = 0.77$, $p = 0.39$, $\eta_{p2} = 0.02$), indicating that there was a difference in accuracy as a function of whether the elements were Gabors or step-edges. There was also a significant interaction between element and set type with a marginal effect size ($F(1, 41) = 6.88$, $p = 0.01$, $\eta_{p2} =$

0.14). From visual inspection of Figure 8, one can see that this effect reflects the overall worse performance in the grey-centered Gabor condition relative to the other stimulus types. There was no significant interaction between element type and misalignment ($F(5.11, 209.35) = 1.44$, $p = 0.21$, $\eta_{p2} = 0.03$) or between element set and misalignment ($F(4.67, 191.31) = 0.57$, $p = 0.71$, $\eta_{p2} = 0.01$). There was no significant three-way interaction between element type, element set, and misalignment ($F(5.11, 209.35) = 1.12$, $p = 0.35$, $\eta_{p2} = 0.03$).

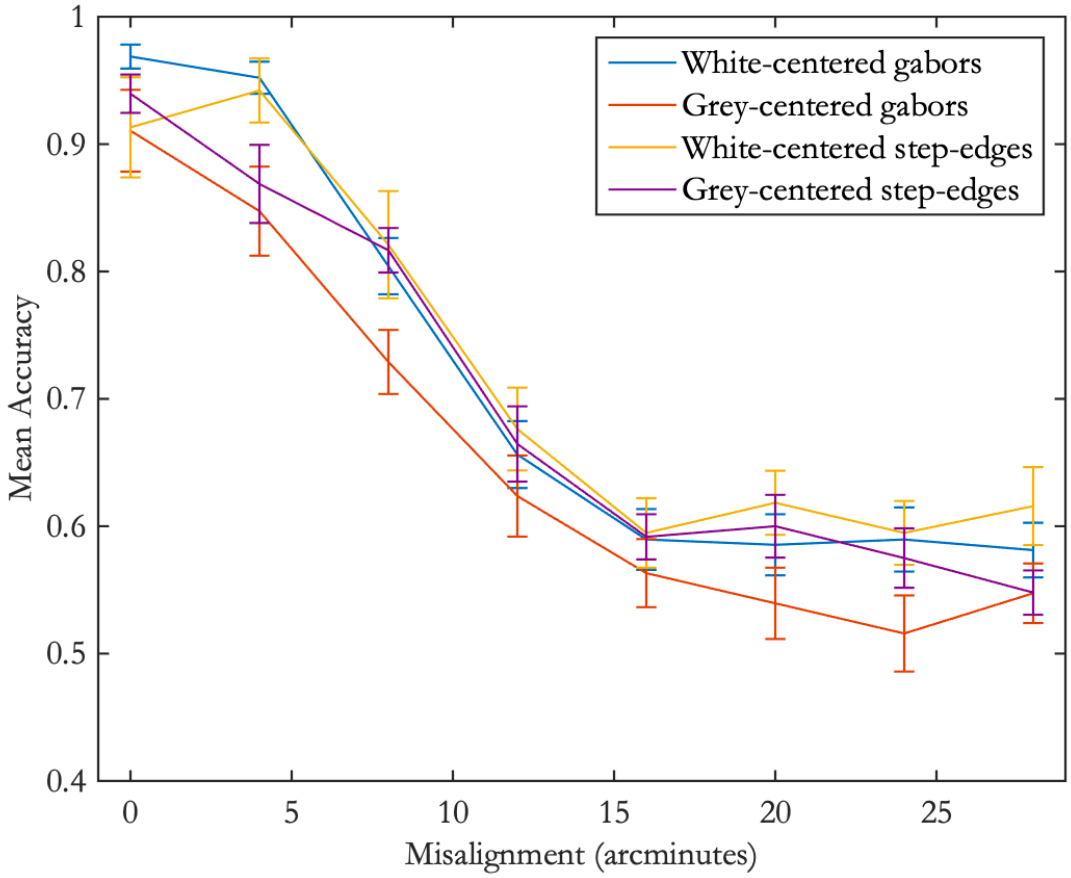

Figure 7. Accuracy in the 2IFC path detection task as a function of element misalignment. Error bars show 1 standard error above and below the condition mean. For all element types, accuracy decreases as a function of misalignment, appearing to reach a plateau at 16 arcminutes of misalignment.

2.4.3 Discussion

The results of Experiment 1 reveal that, for all element types, path detection accuracy declines gradually as a function of element misalignment, up until 16 arcminutes of misalignment, at which point performance plateaus, reaching chance levels. The same performance pattern is evident for all element types, regardless of whether the elements give rise to illusory contours or not.

These results are highly consistent with the predictions. The predictions were made because 1) if path detection depends on contour interpolation, it should be subject to the same constraints that contour interpolation is subject to, including a retinal misalignment of 15 arcminutes, and 2) if path detection depends on interpolated contours in the intermediate representation, it should not be affected by second stage cues, such as surface discontinuity, that eliminate, strengthen, or decrease the phenomenological perception of interpolated contours.

This work adds to the results of Field, Hayes, & Hess (1993) that show that path detection performance is contingent on two other constraints on relatability: the ninety degree constraint (path detection performance worsens for increasing inter-element orientation differences, reaching chance at roughly 90 degrees), and the monotonicity constraint (path detection performance worsens for paths that require an inflection point between some path elements). The results of Experiment 1 demonstrate that yet another constraint on contour interpolation, that of a retinal tolerance for misalignment of inducing edges of about 15 arcminutes, is also true of path detection. This confluence of constraints furthers our confidence that path detection is really revealing the outputs of the first stage of the interpolation process: contours that are interpolated promiscuously between all pairs of relatable edge fragments, and which do not necessarily appear in the final scene representation.

If path detection depends on the output of the first stage, which lies in an intermediate representation, then path detection performance should not be affected by cues that are taken into consideration in the second stage, and which have an impact on the phenomenological perception of interpolated contours. Therefore, I predicted that the decrease in performance, and the eventual plateau at 15 arcminutes of misalignment, would be the same across all element types. This prediction is borne out in the data, with one slight exception: performance was slightly worse for the grey-centered (fuzzy as opposed to step-edge) Gabors than for the other conditions. This can be explained by the fact that, as a result of how the elements were constructed, these elements were the smallest and also produced the least contrast.

Overall, the results strongly support the two-stage theory. Yet another constraint on contour interpolation is also a constraint on path detection, the pattern of decreasing performance for increasing levels of misalignment is exactly as predicted based on previous research on contour interpolation, and factors that are taken into consideration in the second stage have no bearing on performance: the pattern in performance is the same across all element types, regardless of the phenomenological perception of interpolated contours.

2.5 Experiment 2: Subjective ratings of illusory contour clarity

Experiment 1 demonstrated that objective performance in the path detection paradigm degrades in a nearly identical fashion, reaching chance performance around 15 arcminutes of misalignment, for all four element types, regardless of whether the elements were expected to give rise to the perception of illusory contours.

In Experiment 2, we examined whether this critical level of misalignment at which paths could no longer be detected also corresponds to the degree of misalignment at which illusory contours are

no longer seen between those elements that support their formation (i.e., ones with centers that matched the background). A new set of naïve observers provided subjective ratings of perceived illusory contour strength for paths made of the same four element types as in Experiment 1. The paths were shown one at a time on a blank, gray background, without being embedded in a larger array. The same misalignments were used as in Experiment 1.

This experiment achieves two goals. First, the subjective ratings confirm that participants do report seeing illusory contours for the modified Gabor elements with grey centers, but not for the classic and modified Gabor elements with white centers. Second, for those element types that do give rise to the perception of illusory contours, the subjective ratings demonstrate that decreases in the perception of illusory contours mirror the decreases in objective accuracy in the path detection task, also reaching a plateau at 15 arcminutes of misalignment.

Taken together along with the results from Experiment 1, these results suggest that the rich literature on *contour integration* is really revealing the first stage of an interpolation process that interpolates connections indiscriminately between all relatable edge fragments, and that it is surface discontinuities present in the traditional path detection displays that prevents interpolated contours from being perceived.

2.5.1 Method

Participants. Eighteen participants (15 female, ages 18-23, M_{age}: 20.5 years) completed the white-centered Gabor and grey-centered step-edge conditions, and seventeen participants (13 female, ages 19-24, mean age: 20.72 years) completed the grey-centered Gabor and white-centered step-edge conditions. All participants provided informed consent, were students attending the University of California, Los Angeles, reported having normal or corrected-to-normal vision, were offered course credit for their participation, and were naïve to the purposes of the study and had not

participated in Experiment 1. Just as in Experiment 1, the four element types were broken into two sets, presented between-subjects in order to reduce the experiment running time for each subject. One group of participants saw white-centered Gabors and gray-centered step-edges, while the other group saw gray-centered Gabors and white-centered step-edges. Some participants participated in both groups. Three participants who saw the white-centered Gabors and grey-centered step-edges and two who saw the grey-centered Gabors and white-centered step-edges were excluded from the analysis based on a post-experiment questionnaire.

Apparatus. The testing environment and computer were the same as for Experiment 1.

Stimulus. On each trial, participants were shown a path presented alone without any background elements, and were asked to rate the illusory contour clarity on a scale of 0 to 7 by pressing a number key on the keyboard. A 0 rating meant that no contour was visible and 7 indicated the perception of a strong illusory contour. The path was made up of seven elements, just as in Experiment 1. The path was always oriented horizontally and centered on the screen. The same four element types and eight misalignment levels were used as in Experiment 1. Each stimulus was presented 10 times over the course of the experiment. There were therefore a total of 2 (number of element types seen by an individual subject) x 8 (misalignments) x 10 (repetitions): 160 trials. The trial order was randomized.

Procedure. At the beginning of the experiment, participants were familiarized with illusory contours. They were shown an example of a Kanizsa square and were told that the boundaries of the square were illusory contours. They were then shown four circles in the same positions as the inducing elements of the Kanizsa square and were told that while they could imagine the four circles being connected by lines to make a square, there were no illusory contours in this display. Both images, the Kanizsa square and the four circles, were then shown simultaneously so that participants could compare the two side-by-side. They were then told that, in the experiment, they would be

rating the strength of such illusory contours. The Kanizsa square was given a rating of 7. The circles, when there were no illusory contours, were given a rating of 0. They were instructed to use intermediate ratings (1-6) for contours that were of intermediate strength and clarity. At that point they could begin the experiment. On each trial, a single path appeared and remained on the screen until a response was made. Above each element path, text was displayed reminding participants of the meaning of the rating scale: 0 corresponding to no illusory contour and 7 corresponding to strong illusory contours. Once a response was made, the next trial began immediately. Participants were asked to take a break every 80 trials.

At the end of the experiment, participants were asked to complete a post-experiment questionnaire to double check that they understood the instructions, what illusory contours were, and that they were reporting accurately. Participants were given a piece of paper on which were printed two displays. They were asked to rate the illusory contour strength, on the same 0 to 7 scale they used in the experiment. The first display was composed of a rectangular region filled with vertical black bar segments that faded to the background color midway along their height. (see Appendix, Figure A1). The fading of these bars should have produced the perception of a fuzzy white illusory figure traversing an undulating, horizontal path across the bars, as the L-junctions produced by the boundaries of the bars were relatable. We determined a priori that a medium to strong response on the scale would be a reasonable rating for this stimulus if the subject understood the nature of their task. The second display was also composed of black bars, with sections missing. The L-junctions produced by the bar boundaries in this case were not relatable and thus should not have produced illusory contours. We determined a priori that any rating above a 0 for this stimulus would indicate that the subject did not understand the nature of the task, and that their data should be subject to exclusion. Following this post-experiment questionnaire, participants were debriefed.

The experiment lasted approximately 20 minutes followed by less than 5 minutes for the post-experiment questionnaire and debriefing.

2.5.2 Results

The post-experiment questionnaire was examined to make sure that observers understood the task. Three participants who saw the white Gabors and grey step-edges and two who saw the grey Gabors and white step-edges gave contour clarity ratings of 6 or 7 for the stimulus in the questionnaire which had no illusory contours. These participants were excluded from the analysis.

The subjective rating data are presented in Figure 8. Average responses were in the range of 0 to 5, even though the maximum possible response on the scale was 7. The largest illusory contour ratings were given to grey-centered step-edges and grey-centered Gabors when there was no misalignment. Ratings of illusory strength decreased rapidly as a function of misalignment until about 8-16 arcmin. At maximum misalignment, grey-centered step-edges had slightly higher ratings (1.5) than grey-centered Gabors (1). White-centered Gabors and white-centered step-edge displays had very low average ratings (between 0 and 1.5) for all misalignments.

To confirm these observations, the data were submitted to a 2 (element types: Gabors or step-edges) x 2 (element set: first or second between-subjects experimental group) x 8 (misalignments) repeated measures ANOVA. All F statistics were Greenhouse-Geisser corrected. There was a main effect of element type ($F(1, 28) = 10.50$, $p = 0.003$, $\eta p2 = 0.27$) and misalignment ($F(2.32, 64.83) = 17.72$, $p < 0.001$, $\eta p2 = 0.39$), but there was no significant between-subjects effect of experiment type ($F(1, 28) = 0.52$, $p = 0.47$, $\eta p2 = 0.02$). This indicates that, while there was a difference in contour clarity ratings between white-centered Gabors and grey-centered step-edges and between grey-centered Gabors and white-centered step-edges, there was no significant difference between white-centered Gabors and white-centered step-edges or between grey-centered Gabors and grey-

centered step-edges. This is reflected in the absence of an interaction between misalignment and element set ($F(2.32, 64.83) = 2.81$, $p = 0.06$, $\eta p2 = 0.09$) or between element type and element set ($F(1, 28) = 0.18$, $p = 0.92$, $\eta p2 < 0.001$). There was a significant interaction between element type and misalignment ($F(2.66, 74.55) = 13.87$, $p < 0.001$, $\eta p2 = 0.33$). The three-way interaction between element type, element set, and misalignment was not significant ($F(2.66, 74.55) = 0.65$, $p = 0.57$, $\eta p2 = 0.02$).

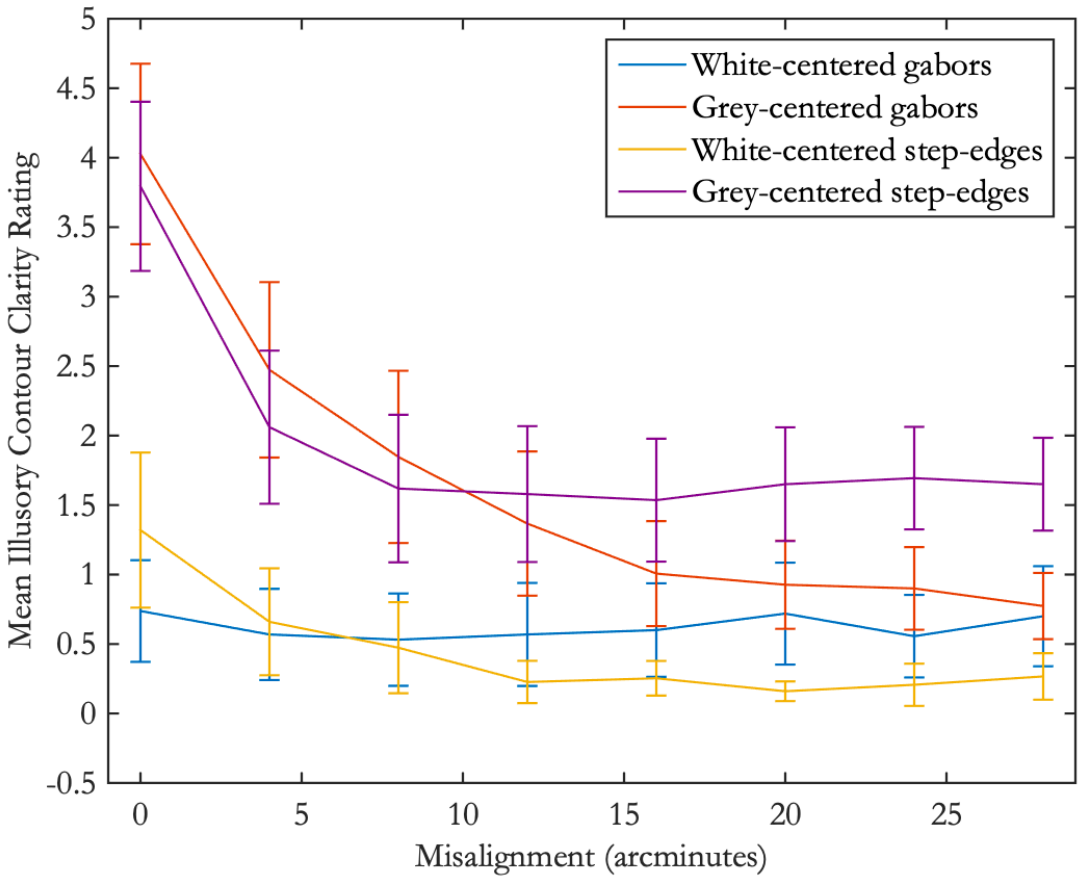

Figure 8. Ratings of illusory contour clarity as a function of element misalignment. The ratings for white-centered element types are roughly at baseline for all misalignment levels. For the grey-centered elements, the ratings of illusory contour clarity reach a plateau at 16 arcminutes of misalignment. Similar to Experiment 1, the grey-centered (fuzzy) Gabors, which were the smallest of all element types and also had the least contrast with the background, produced slightly different responses - here, in the form of slightly different illusory contour clarity ratings at plateau. In

Experiment 1, this element type produced slightly lower accuracy across all misalignment levels. Error bars show 1 standard error above and below the condition mean.

For ease of comparison between the objective and subjective task, the data from both experiments are re-plotted in Figure 9 for the white-centered Gabor and grey-centered step edge conditions. The highest average contour clarity rating for either stimulus type was set to the highest accuracy for any stimulus type. The figure clearly illustrates that contour clarity ratings for grey-centered step-edge displays (dashed orange lines) decrease as a function of element misalignment at qualitatively the same rate as performance in the objective task (solid orange lines). This pattern of results was obtained despite the fact that data were collected from two distinct groups of participants. In contrast, the ratings are flat and near zero (i.e. no visible contour) for paths made of white-centered Gabor elements irrespective of misalignment, while the accuracy data for the same stimulus follow the same decreasing trend as for grey-centered step-edge displays.

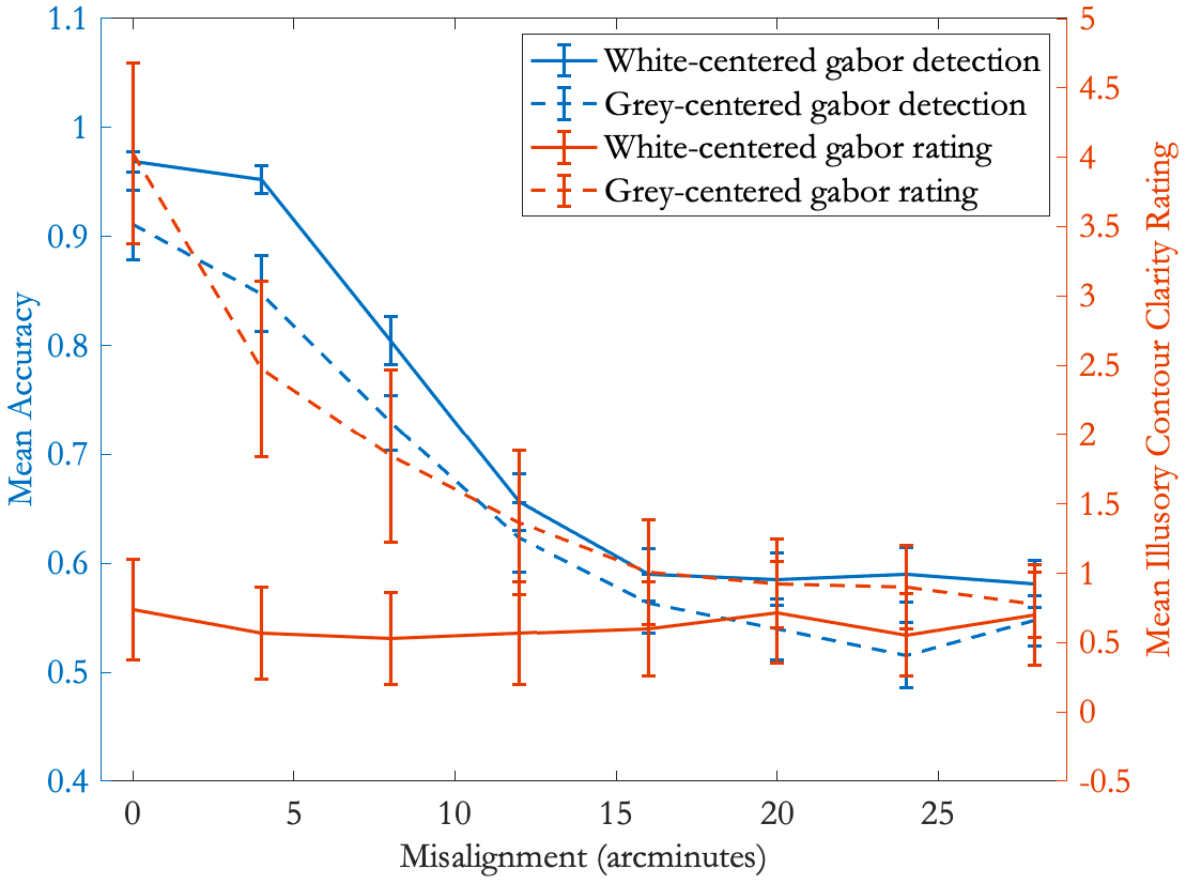

Figure 9. Combination of data from the objective and subjective tasks. Blue axes and lines: mean accuracy data for both white- and grey-centered Gabor conditions from the objective task. Orange axis and lines: mean illusory contour clarity ratings for both white- and grey-centered Gabor conditions from the subjective ratings task. Solid lines: paths made of Gabor elements with white centers. Dotted lines: paths made of Gabor elements with gray centers. The figure reveals that, for grey-centered elements, subjective ratings of illusory contour clarity follow the same pattern as objective performance in the path detection task. White-centered elements, however, for which illusory contours should not be perceived, receive low ratings of illusory contour clarity across all misalignment levels, even though the objective data for these stimuli follow the same trend as that of the grey-centered elements in the path detection task. Error bars show 1 standard error above and below the condition mean.

2.5.3 Discussion

In Experiment 1, path detection accuracy decreased as a function of element misalignment up to 15 arcmin, at which point performance reached an asymptote. Results showed this pattern whether the path element type supported the formation of an illusory contour or not. Experiment 2 had two purposes. First, it confirmed that displays with white centers did not produce illusory contours, as they were rated with the lowest contour clarity irrespective of misalignment level. In contrast, paths made of collinear elements whose centers match the background received high clarity ratings. Second, it confirmed that, for paths made of these elements, rating strength decreased as a function of misalignment, closely matching the pattern of accuracy data from Experiment 1, such that ratings steadily decreased up to misalignments of approximately 15 arcminutes and then remained constant. Critically, this cutoff value is the same for constraints on relatability in contour interpolation.

Taken together, these results strongly support the two-stage theory of interpolation. The phenomenological pop-out seen in path detection displays is really revealing the result of the intermediate output of the first stage of the contour interpolation process, in which contours are interpolated indiscriminately across all pairs of relatable edge fragments. Though these interpolated contours reside in an intermediate representation, their effects on perception are visible both in the sensitivity to paths of Gabor elements in the path paradigm as well as this sensitivity to "fat" versus "thin" interpolated trajectories in the fat-thin task (Guttman & Kellman, 2001, in prep) as well as in results from the path detection literature. Interpolated contours are not perceived in traditional path detection displays because surface discontinuities between the Gabor interiors and the background result in the elimination of these interpolated contours in the second stage. However, altering the Gabor centers to match the background results in the perception of interpolated contours traversing

the gaps between the modified Gabor elements to create the perception of an illusory grey bar that partially occludes the black surrounds of the modified Gabors.

2.6 General Discussion

The term "contour integration" was introduced by Field, Hayes, & Hess as a label for the perceptual mechanism involved in path detection (1993). Though the term has enjoyed wide use, the process so-named has yet to be well defined. Although Field, Hayes, and Hess (1993) proposed the term "association field", suggesting interactions between appropriately located orientation-sensitive units (e.g., V1 cortical cells), they did not propose an actual formal or quantitative model. They suggest a localized linking process and a region around each element within which other elements may be grouped with the first, according to a specific set of rules. What the rules are, the authors do not say. Yen & Finkel (1998) proposed a network model for the relevant relationships of edge fragments. Further, it incorporated temporal synchrony of firing as a mechanism for coding the contour fragments linked into a single path (in the case of elements separated by gaps) or contour (for adjacent elements) (for detailed discussion, see Kellman, Guttman & Wickens, 2001).

Figure 10 shows the association field model as it is typically illustrated. The figure shows a central, horizontally oriented Gabor spatially separated at both the left and right ends from three flanking Gabors. At one end, smooth connections can be made from the central Gabor to its flankers. At the other end, elements are shown for which the orientations and positions violate the rules that govern path detection performance. The relationships among Gabor elements illustrated here - and in other published diagrams of the association field model - are, in fact, perfectly described by relatability. The good connections are those for which the elements can be connected by a smooth, monotonic curve that bends through less than 90 degrees. The violations are cases

where the connection would require a turning angle greater than 90 degrees or the introduction of an inflection point in the connection between elements.

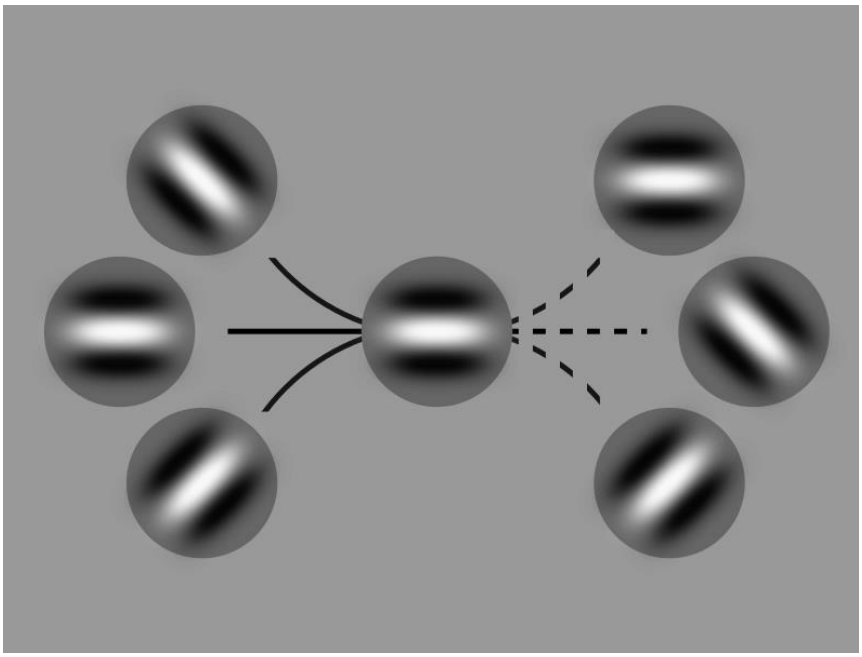

Figure 10. Association field model of Field, Hayes, & Hess, 1993. At left, smooth connections, monotonic connections that bend through less than 90 degrees can be made from the central Gabor to each of its flankers. At right, the orientation and position of the flankers prevents smooth, monotonic connections from being made. Redrawn from Field, Hayes, & Hess, 1993.

Although Field, Hayes, & Hess (1993; see also Hess & Field, 1999) note that the relationships in the association field notion agree with the geometry of relatability (e.g., both the ninety degree constraint and the monotonicity constraint are included), they did not suggest that contour integration had anything to do with contour interpolation, nor did they make any attempt to account for the obvious similarities. This omission was not due to a failure to consider how contour integration might relate to other perceptual phenomena or processes, as the authors did discuss, at length, a possible relation to texture segregation.

In hindsight, reviewing the original discussion in the 1993 paper, one might note the discussion of perceptual continuity; the mention of good continuation; the figure (their Figure 1) that illustrates

the principle of good continuation in the presence of occlusion; the emphasis on local, rather than global, associations; the claim that a specific set of rules governs these local associations, and, finally, results that reveal worsening performance with increasing inter-element orientation differences (with a fall to chance extrapolated for turning angles of 90 degrees - their Figure 7) and the introduction of inflection points (incorporated in the original paper through elements that were offset from the smooth path - their Experiment 3) and be surprised that a connection to contour interpolation was not immediately assumed. To do so, however, is to underestimate the importance that any observer is practically bound to place on phenomenological perception - and the obvious lack of any interpolated contours in the original path displays.

Here, I proposed a multistage theory of contour interpolation that accounts for the similarities in the association field model and the relatability geometry, explains the lack of perceived interpolated contours in the original path displays, and gives rise to several testable predictions. Path detection performance demonstrates a 90 degree limit and a monotonicity limit because the perceptual pop-out observed in path displays is really revealing the intermediate output of an early stage of the contour interpolation process. In that stage, contours are interpolated promiscuously and indiscriminately across all pairs of relatable edge fragments. However, the output of that stage exists in an intermediate representation. Only some of the interpolations existing in the intermediate output will make it onto the final scene representation. In a second stage of processing, other information - for example, cues as to border ownership and surface continuity - are taken into account to determine which interpolations should be sustained, which should be weakened, and which should be deleted entirely from the final scene representation.

Based on this theory, I proposed that the presence of surface discontinuities across the Gabor interiors and the background of traditional path detection displays results in the elimination of interpolated contours at the second stage. Because these interpolated contours don't appear in the

final scene representation, illusory contours are not subjectively perceived. However, the intermediate output - which contains these interpolated contours - still exists at some stage of the perceptual process. We proposed that it is this intermediate output that is responsible for the perceptual pop-out seen in path displays.

We thus, hypothesized, first, that modifying the interiors of Gabors to match the grey background of path detection stimuli would result in the perception of interpolated contours traversing the gaps between the Gabor elements. In addition, we hypothesized that, because path detection performance depends on the intermediate output, it should not depend on the subjective perception of interpolated contours, and thus should not be affected by the manipulation of the centers of the Gabors. In addition, we sought to demonstrate that yet another known constraint on contour interpolation - that of a somewhat unusual 15 arcminute retinal tolerance for the misalignment of inducing edges (relatively few constraints on perceptual processes involve retinal, rather than relative, tolerances) - could be shown to also be a constraint on contour integration. To this end, we manipulated the misalignment of Gabor elements in the target paths, predicting that performance would decrease for increasing levels of misalignment, reaching a plateau at 15 arcminutes of misalignment, and that this pattern of deterioration would be identical for all stimulus types. Finally, we predicted that subjective ratings of illusory contour clarity would confirm that interpolated contours are perceived only for the Gabors with grey centers, and that deterioration in these ratings would mirror the deterioration of performance in the objective task as misalignment between elements increased.

All of these predictions were borne out in the data. Results from Experiment 1 reveal that objective path detection performance deteriorates as misalignment is increased. As predicted, a plateau is reached at around 15 arcminutes - the known tolerance threshold for misalignment in contour interpolation. Further, performance follows the same pattern of deterioration across all

element types. Experiment 2 revealed that interpolated contours are perceived only when the surface discontinuities present in the original path stimuli are eliminated (by making the white centers grey). In addition, the subjective ratings of illusory contour clarity reveal the same pattern of deterioration as the objective data - ratings decrease as misalignment is increased, reaching a plateau at 15 arcminutes of misalignment. Taken together, these results provide very strong support for the multistage theory of contour interpolation.

2.6.1 Second stage cues: Surface spreading & border ownership

The results of the experiments presented here suggest that variations in surface qualities across inducing edges can play a role in determining the final appearance of interpolated contours. Something like this has, in fact, been suggested by previous work, though the results and observations were not interpreted in terms of a multistage theory. He and colleagues have revealed, through careful experimental work, that differences in luminance contrast polarity across inducing edges affect the perceived strength of interpolations connecting those edge fragments (He & Ooi, 1998a), their perceived unity and cohesive motion (Su, He, & Ooi, 2010a), as well as their perceived unity and depth in stereoscopic stimuli (He, Ooi, & Su, 2018). Their work has produced several compelling figures, one of which is redrawn in Figure 25.

Figure 25 shows a redrawing of He & Ooi's (1998a) "illusory-O". The figure shows six wheel-like stimuli, two each of which are set upon a grey, black, or white background. For all six stimuli, interpolation across relatable, rectangular fragments within a single wheel spoke will be produced in the early, promiscuous interpolation stage of processing. Across the two columns, surface relationships across the two rectangular fragments within a single spoke are manipulated. Only for the first row (the grey background) does this have the effect of reversing the luminance contrast polarity across the two columns. At left, relatable fragments within a single spoke have identical

background luminance contrast polarity. At right, the fragments within a spoke have opposite background luminance contrast polarity. The result is that an illusory-O is perceived in the stimulus at left but not in the stimulus at right. For the second (black background) and third (white background) rows, the illusory-O is preserved across both columns, suggesting that it is background luminance contrast polarity, rather than relative or absolute luminance across inducing edges, that determines the appearance of interpolated contours in the second stage.

He & Ooi (1998a) interpreted the elimination of the illusory-O shown in Figure 25b in terms of the absence of amodal surface interpolation. That is essentially what we are proposing here. Specifically, in the differing outcomes for (a) and (b), border ownership issues are crucial. In the absence of potential surface spreading between rectangles within a single spoke, each rectangle (white or black) retains border ownership of its own edges. One may attain weak illusory contours along the inner circle or outer circle (i.e., "crystalline interpolations"; see Kellman, Garrigan, Shipley & Keane, 2007), but the opposed border ownership assignments of the rectangles' edges work against formation of an intervening grey circle in the second stage.

There is good reason to believe that the kinds of scene cues taken into account in the second stage are somewhat diverse, and extend beyond what is evidenced in surface qualities alone. Figure 17 shows multiple examples of outline manipulations of various kinds of illusory contour stimuli - here again, border ownership cues are implicated as being among the kinds of scene cues used to determine the final appearance of interpolated contours in the second stage. Likewise, Figure 11 below shows several manipulations of a drawing in which two different inducing edges produce competing interpolations to complete a glass cup. In this case, cues as to closure and symmetry seem to help resolve the final percept in favor of one completion or another.

2.6.2 Competing interpolations

Given the indiscriminate nature of promiscuous interpolation in the first stage - that is, interpolations are produced across *all pairs of relatable edge fragments* - it seems obvious that, in addition to resolving cases of impossible interpolations, one of the necessary goals of the second stage will be to resolve instances of competing interpolations as well as interpolations that are inconsistent with each other. Figure 11 illustrates three cases where only one interpolation is possible in the first stage (left column) and three other cases (right column) where two interpolations will be produced in the first stage. In these latter cases, cues as to closure and symmetry may be used to determine the strengths of the individual interpolations (see figure caption).

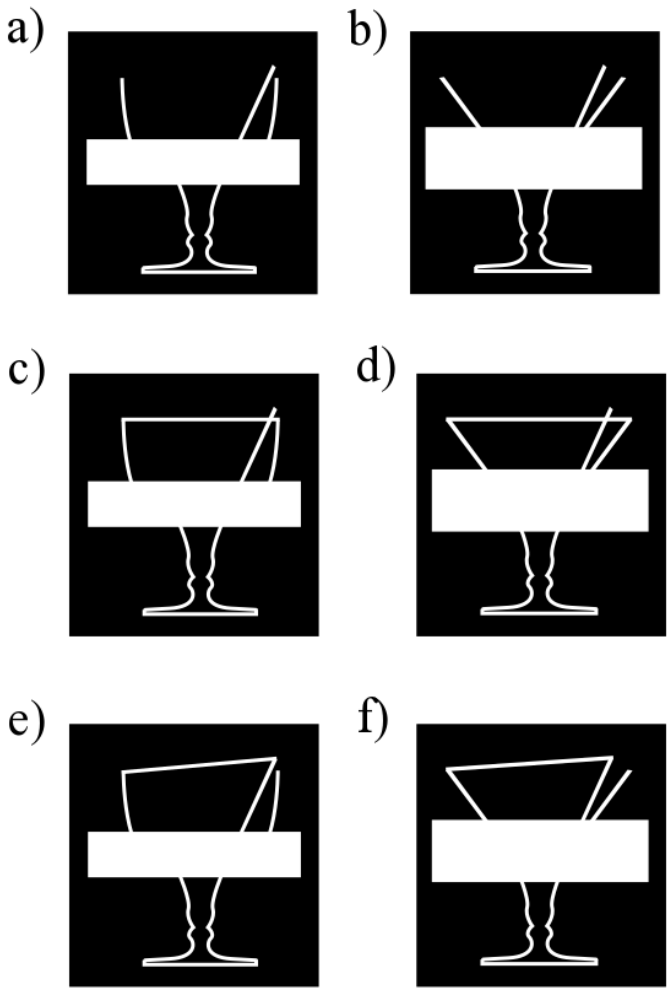

Figure 11. Competing interpolations. For each stimulus at left, the right-most edge of the base of the glass has a relatable connection with only one edge at top. As a result, the right-most inducing edge of the base will be utilized for only one interpolation in the first stage (any other connection would require a violation of the monotonicity constraint). Readers may note the unambiguous perception of the right edge of the base extending upward in a straight line to form an asymmetric cup. This occurs despite the presence of a second, alternative, edge that exhibits reflectional symmetry with the left edge of the cup. Across a), c), and e), the near-horizontal connection at the opening of the glass is manipulated. In a), there is no connection; in c), the two symmetric edges are connected; in e) the two asymmetric edges are connected. One may note the compelling perception that the right edge of the base connects to the straight edge above, even when this interpretation runs counter to (11c) or is unaided by (11a) the connection at the opening of the glass. For each stimulus at right, the edges of the glass have been altered so that two competing interpolations will be produced in the second stage: the bottom right half of the base has a relatable connection to both upper edges at right. Connections at the opening of the glass are manipulated across b), d), and f), in the same way as in the left column. In Figure

11b, symmetry alone suggests one connection over the other. In Figure 11c, symmetry and closure suggest the right-most edge of the base ought to be connected to the right-most upper edge. In Figure 11f, cues to closure and symmetry contradict each other, and one or the other perception must win out.

2.6.3 The known necessity of multiple stages: Determining modal vs. amodal appearance

The theory put forth here is not the first to suggest that interpolation must proceed in at least two stages. The *identity hypothesis* is a part of the unified framework, put forth by Kellman & Shipley (1991), for understanding visual interpolation in object perception; it asserts that modal and amodal completion share a common process. That is, interpolation proceeds first, and afterward, the appearance of an interpolated contour as either modal or amodal is determined by taking into account other scene cues. This hypothesis is supported by experimental work (Shipley & Kellman, 1992a; Kellman, Yin, & Shipley, 1998), as well as compelling logical arguments (Kellman & Shipley, 1991; Shipley & Kellman, 1992; Kellman, Yin, & Shipley, 1998). One of these arguments involves the appearance of crossing interpolations in stimuli known as self-splitting objects. Figure 12 below contains two such stimuli.

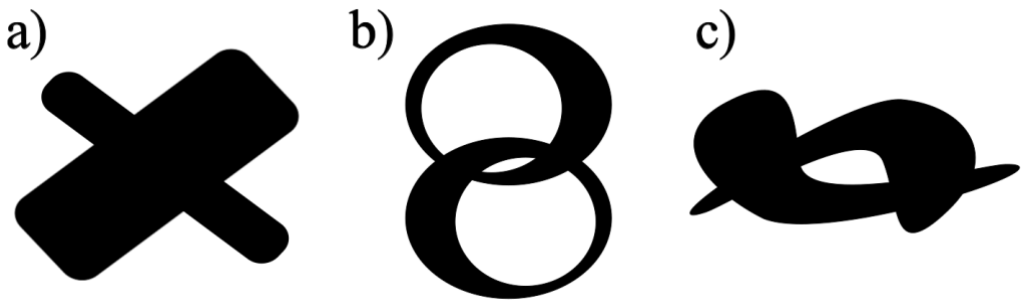

Figure 12. Self-splitting objects. In a), b), and c). homogenous black regions appear to spontaneously split into two overlapping figures. The figures are bi-stable in that either one or the other unit may appear in front, affecting which of the interpolated contours take on a modal or amodal appearance. The "Petter Effect" refers to the known tendency for shorter interpolations to appear in front (taking on a modal appearance). The length of the interpolated edges may be taken as a cue to which edge lies in front, as the shorter interpolated edge is more likely to belong to a nearer object (that appears in front of a more distant object with a smaller retinal

projection). If the length of an interpolated edge determines its appearance as modal or amodal, then interpolation necessarily precedes the determination of the modal or amodal appearance of interpolated edges. Figure redrawn from Kellman et al. (2007) and Kellman et al. (2001).

In the figure above, homogenous black regions appear to spontaneously split into two overlapping figures. The figures are bi-stable in that either one or the other unit may appear in front, affecting which of the interpolated contours take on a modal or amodal appearance. The "Petter Effect" refers to the known tendency for the shorter of two crossing interpolations to appear in front (taking on a modal appearance) (Petter, 1956). This makes some sense, as the shorter interpolated edge is more likely to belong to a nearer object (that appears in front of a more distant object with a smaller retinal projection). The logical argument in favor of a multistage theory is this: if the relative lengths of crossing interpolations determines the modal or amodal appearance of interpolated edges, then the visual system must know the sites and extents of interpolation prior to determination of modal or amodal appearance. This implies a representation of linked contours that is neither modal or amodal and that precedes determination of aspects of the appearance of contours in ultimate scene representations.

Other support for an initial promiscuous interpolation stage and a second stage that determines the final appearance of interpolated contours comes from the existence of so called "quasimodal" displays. In these displays, a single interpolated edge takes on a modal appearance over some of its length and an amodal appearance over the rest of its length. Two quasimodal displays are illustrated in Figure 13a and 13b. These displays, and experimental evidence with them (Kellman, Yin & Shipley, 1998), indicates that a single interpolated edge can have modal appearance along some of its length and amodal (partly occluded) appearance along another part.

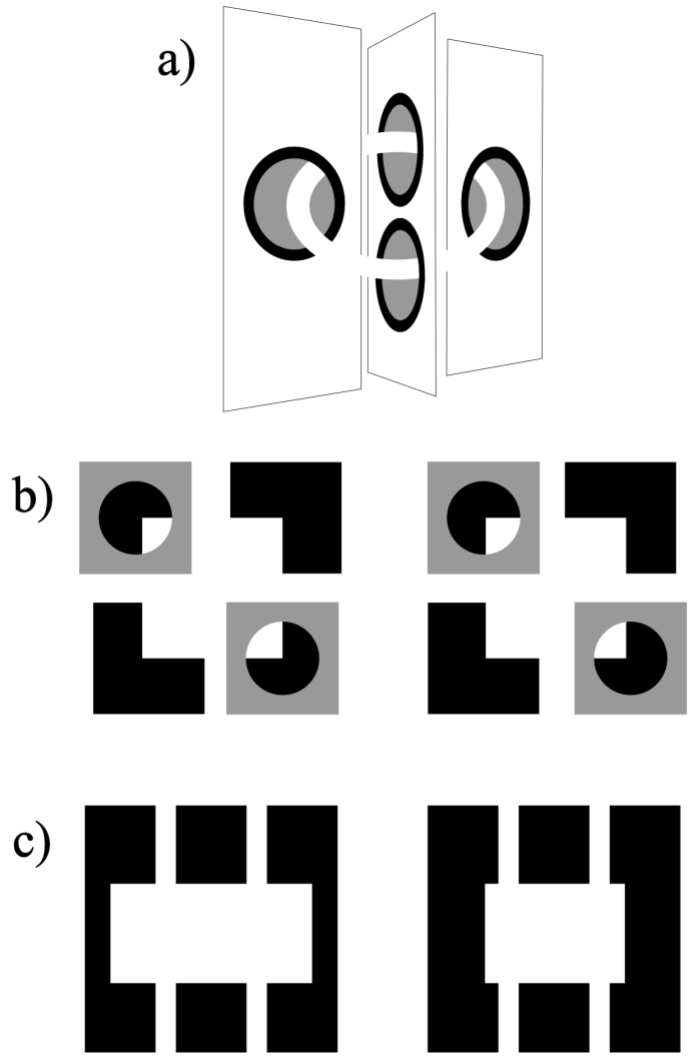

Figure 13. Quasimodal and depth spreading displays. In the first two displays, interpolated edges take on a modal appearance across part of their length, and an amodal appearance across the rest of their length. In a), a white ring appears to be woven through four holes in three different surfaces. The two images in b) can be free fused by crossing the eyes. A white Kanizsa square, which lies atop a black background, is seen through holes in two grey sheets. Each of the four edges of the square must be completed. At bottom, the illusory, modally completed edge at left connects to the amodally completed edge at right. c) shows a display in which interpolation necessarily precedes determination of modal or amodal appearance of contours of surfaces. c) was redrawn from Kellman, P. J., Garrigan, P., & Shipley, T. F. (2005). Object interpolation in three dimensions. Psychological Review, 112(3), 586-609.

Interpolated edges that have both modal and amodal parts and crossing interpolations can be combined to produce phenomena that are especially revealing regarding a contour linking stage that must occur prior to determination of the modal or amodal appearance of interpolated edges, or parts of them. Figure 13c shows an example. The display shows a stereo pair that can be free-fused by crossing or diverging the eyes. (In the description given, I will assume that the reader free-fuses by crossing the eyes; if fusion is accomplished by diverging, all of the depth relations will be opposite to those described, but the main effect will still be apparent.) When fused, the central rectangle appears as the nearest object in the display at its left edge and as the farthest at its right. With reference to the two white vertical "columns" in the display, the rectangle appears to pass in front of the one on the left and behind the one on the right. This simple appearance has profound implications. The top and bottom horizontal edges of the perceived rectangle each arise from interpolation across two gaps (given by the two white vertical columns). Due to the need for matching features for stereoscopic depth, only the endpoints of each horizontal edge have stereoscopic depth specified by the stimulus. Each of these contours (and ultimately, the rectangle they partially bound) appears slanted in depth, due to "depth spreading"; all of the perceived depths of intermediate points along the horizontal edges is inherited from the endpoints. These lie at different depths, and depth spreading between these edges produces the appearance of intermediate positions in depth between these bounding edges. No stimulus information in the middle of the display indicates that the horizontal edges are in front of or behind either vertical "column". Depth spreading occurs along continuous edges or surfaces (Hakkinen Liinasuo, Kojo, & Nyman, 1998); it does not spread unconstrained across a whole scene. Therefore, to be affected by depth spreading, the horizontal edges must be continuous between the endpoints. Due to the gaps in the stimulus, this continuity can only be due to interpolation.

With this in mind, consider the modal or amodal appearances of various contours. Why is one vertical column amodally completed and the other modally completed? These effects are determined by the slant in depth of the horizontal edges of the rectangle. Depth spreading confers depth values all along its boundaries (including the interpolated parts). These depth values of the right column are smaller than that of the left column; thus, the right column appears to lie in front of an occluded white rectangle, and in the area of crossing interpolations, the column therefore has illusory contours. The left vertical column has *occluded* contours in this region of crossing interpolations; it appears behind the slanted rectangle.

The modal–amodal appearances of both the rectangle's contours and the white columns are consequences of the horizontal edges having particular positions in space, given by depth spreading. But the continuous, horizontal contours exist only because of interpolation. The amodal or modal appearances of several contours in this display are consequences of the relative positions in depth of interpolated contours that cross. For the horizontal edges, these positions can only be given by depth spreading; depth spreading presupposes that the horizontal edges are continuous. But the existence of the continuous horizontal edges are results of contour interpolation. Thus, in this type of display, contour interpolation and depth spreading along interpolated contours necessarily precedes determination of the appearance of contours as illusory or occluded. Note that the horizontal edges and the rectangle in the percept have modal and amodal appearances along different parts of their extents in the display. Therefore, when interpolation first occurs, it cannot be amodal or modal; these designations follow from relative depth, which in this case has been derived from contour linking and depth spreading.

2.6.4 Concluding remarks

The multi-stage theory is consistent with empirical work demonstrating the importance of relatability, it explains the reliance of path detection on relatability geometry, it is consistent with empirical work suggesting the promiscuous interpolation of edges, and it explains known effects of some manipulations - that is, outline inducers, and inducing edges with opposite luminance contrast polarities - on the appearance of interpolated contours. In addition, the multi-stage theory provides a much needed framework for incorporating some additional stages of processing - that is, the determination of interpolated contours as either modal or amodal, as well as the determination of the ultimate perceived strength of an interpolation - into a more complete model of the contour interpolation mechanism. Empirical work and logical arguments have provided mounting evidence for these additional stages of processing. Any single-stage proposal is insufficient to account for the evidence.

The evidence suggests that interpolation proceeds first, irrespective of scene cues, taking into account only the relatability of pairs of edge fragments. Having calculated interpolated trajectories for every pair of relatable edge fragments, scene cues are then taken into account, at a later stage of processing, to determine the ultimate appearance of interpolated contours - that is, whether individual interpolations will appear in the final scene representation at all, and, if so, whether their appearance will be modal or amodal, and what their perceived strengths will be.

This work has great implications. The rich literature on *contour integration* is really revealing the intermediate output of the early, promiscuous interpolation stage of the contour interpolation process. What's more, the multistage theory provides a framework that enables us to continue integrating diverse findings to build a more complete picture of the contour interpolation process.

There is, however, a clear weakness in current two-stage accounts. Whereas the first stage of promiscuous contour linking has been mathematically characterized and modeled by several groups

(Grossberg & Mingolla, 1985; Heitger, Von Der Heydt, Peterhans, Rosenthaler, & Kübler, 1998; Kalar Garrigan, Wickens, Hilger, & Kellman, 2010), the second stage has remained vague. Developing a method to evaluate and quantify at least some of the scene constraints that operate in stage 2, and understanding how they interact, is the goal of the experiments in Chapter 3.

CHAPTER 3

Border ownership, luminance contrast polarity, equiluminant color contrast, and spatial frequency effects on the perception of interpolated contours: Evidence for a two-stage theory of visual contour interpolation

3.1 Synopsis

In the introduction to this dissertation, I reviewed some evidence that supports a two-stage theory of contour interpolation. In Chapter 2, I presented evidence that the "contour integration" evident in the path displays of Field, Hayes, & Hess (1993) is really revealing the intermediate output of the first stage of the contour interpolation process – in this stage, contours are interpolated indiscriminately across all pairs of relatable edge fragments. In a second stage, scene cues are taken into account to determine the final perceived strengths of interpolated contours. A weakness of a two-stage theory has been limited evidence that would allow us to clearly specify and quantify the specific constraints, and their interactions, that operate on the stage 1 linked contour representation to determine the final appearance in the final scene representation. We developed an experimental paradigm that we hypothesized would be sensitive to the strength of contours in the output of the second stage. Then, in a series of experiments, I tested a variety of scene cues, including border ownership cues, luminance contrast polarity, equiluminant color contrast, and spatial frequency differences across fragments are varied to observe the effects on the ultimate perceptual strengths of interpolations exiting the second stage. Adapting an intriguing, now classic, display type developed by Bregman (1991), our displays contain a subset of fragments, hidden among a background of randomly arranged fragments, that can be connected to form a whole letter or number. Interpolation is necessary to connect the fragments, thus, recognition accuracy should be worse when the necessary interpolations exit the second stage with relatively weak strengths. Scene cues

and presentation time are manipulated, and recognition accuracy is measured. Results suggest that border ownership cues, luminance contrast polarity, and equiluminant color contrast play a role in determining the strengths of interpolations in the second stage. Results also suggest that interpolation proceeds across multiple spatial frequency channels, and that corroboration across multiple channels plays a role in determining the final perceptual strength of interpolated contours.

Keywords: contour interpolation, object formation, promiscuous interpolation, amodal completion, modal completion, border ownership, luminance contrast polarity, color contrast polarity, spatial frequency channels

The experiments reported in Chapter 2 revealed the intermediate outputs of the first stage of the contour interpolation process. In the first stage, contours are interpolated *promiscuously* across all pairs of relatable edge fragments. In a second stage, scene cues are taken into account to determine the final perceptual strengths of those interpolated contours. Revealing the outputs of second stage processing, and identifying some scene cues that are taken into account in the second stage, are the issues I take up in this chapter.

3.2 Designing a task

The work of Guttman & Kellman (2001) and the path paradigm experiments presented in Chapter 2 made clear that stage 1 output is sufficient to complete some perceptual tasks. Revealing second-stage processing therefore requires some careful consideration to identify a task that selectively targets the outputs of the second stage. Both the fat-thin task and the path detection task, discussed in Chapter 2, can be completed with stage 1 output. We must ask, then, why stage 1 output is sufficient to complete these tasks. The answer may lie in the fact that completing either of these tasks requires only the sensitivity to the length or shape of a single interpolated trajectory. Because interpolated contours produced in stage 1 exist at some stage of the perceptual process, it is not altogether surprising that observers remain sensitive to some aspect of these intermediate representations.

What kind of task then, would be selective for second stage outputs? I have hypothesized that the interpolated contours produced promiscuously in the first stage are then subjected to multiple scene constraints to determine whether they are maintained or deleted as well as their ultimate perceptual strength. In addition, only those interpolations making it out of this second stage are available to be utilized in the building up of object descriptions. Descriptions of complete, coherent,

whole objects must be built up from many real and interpolated edge tokens. In order to target second-stage processing, then, a task must require sensitivity to whole object descriptions that include both real and interpolated edges. That is, because the outputs of stage 2 are utilized to build up descriptions of whole objects from many real and interpolated edges, it should be true that tasks requiring the integration of many interpolated and real edges to form a complete, recognizable object from several fragments rely on stage 2 output.

This is, in some ways, a bold prediction. The experiments presented in Chapter 2 revealed that disruptions of surface spreading across inducers do not affect objective path detection performance, and other work has shown that manipulations to border ownership cues (Guttman & Kellman, 2001) do not affect performance in a fat-thin task, and that differences in luminance contrast polarity across separate inducers do not affect performance in a fat-thin task (Spehar, 2000) or a similar portrait-landscape task (Victor & Conte, 2000). All of these tasks, however, require only sensitivity to a single interpolated trajectory. I therefore predict that these same manipulations that do not affect an individual's sensitivity to individual interpolated trajectories will impact an observer's performance when the task requires generating a complete object description that must be built up out of many real and interpolated edge tokens.

The considerations above suggest that, in order to target second stage output, a task should require the integration of many real and interpolated edges to build up a description of an object that can then be recognized. A simple task that meets this requirement may be suggested by a compelling figure created by Bregman (1990). A redrawing of Bregman's display appears in Figure 14. The figure shows two sets of fragments of multiple letter B's. Bregman observed that the presence of an occluder - which can own the boundaries of the fragments where it lies adjacent to them - supports the grouping of the fragments and makes the letter easier to recognize. The effect is compelling.

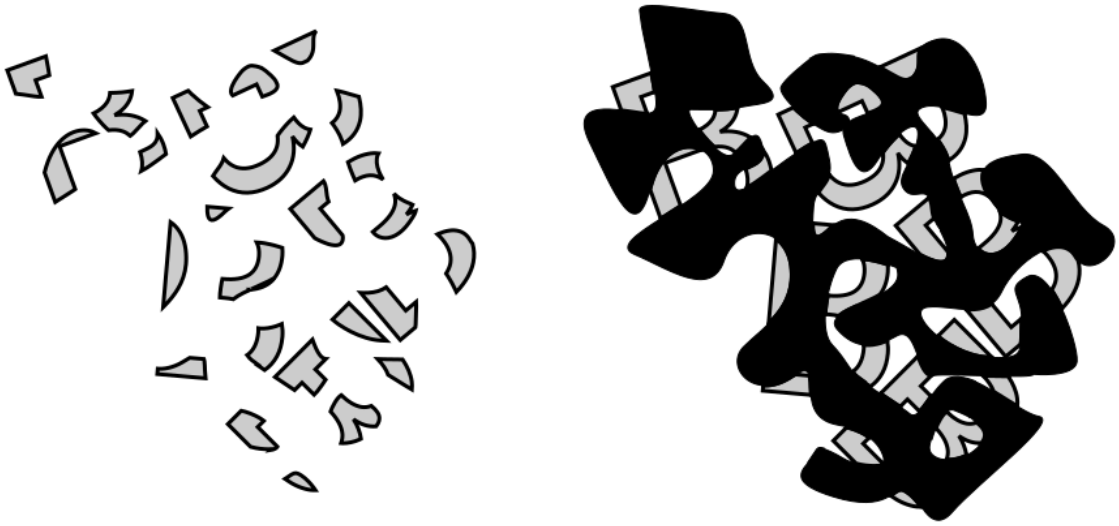

Figure 14. Bregman's Bs. A redrawing of Bregman's original display, from Bregman, A. S. (2002). At left, spatially separated fragments of five block letter B's are difficult to recognize. At right, an occluder accounts for the spatial separation of the fragments, thus supporting the grouping of the fragments into whole B's that are more easily recognized.

Some empirical work has already demonstrated that displays like Bregman's above may be useful in examining the effects of different scene cues on the ability to recognize fragmented alphanumeric characters (Brown & Koch, 1993; Brown & Koch, 2000) and real-world objects (Johnson & Olshausen, 2005). Figure 15 below illustrates some manipulations used by Brown & Koch (1993), who presented participants with stimulus displays like those shown in Figure 15. They had three versions of each of the conditions illustrated in Figure 15 (one containing the letter B, one containing the letter A, and one containing the female symbol, ♀). They showed each participant 3 stimuli (the same condition, one for each of the symbols used) and allowed them one minute to guess at the symbol contained in the stimulus. Participants were allowed an unlimited number of guesses within that minute, and the time taken to get the correct answer was recorded. Their results showed that recognition was much better for displays utilizing "open" fragments (15b and 15c) than

for analogous display using "closed" fragments (15a and 15d). In contrast, the presence of an occluder (15d versus 15a and 15c versus 15b) did not have a large effect on recognition. The authors concluded that closure, but not occlusion, plays a critical role in the recognition of fragmented objects.

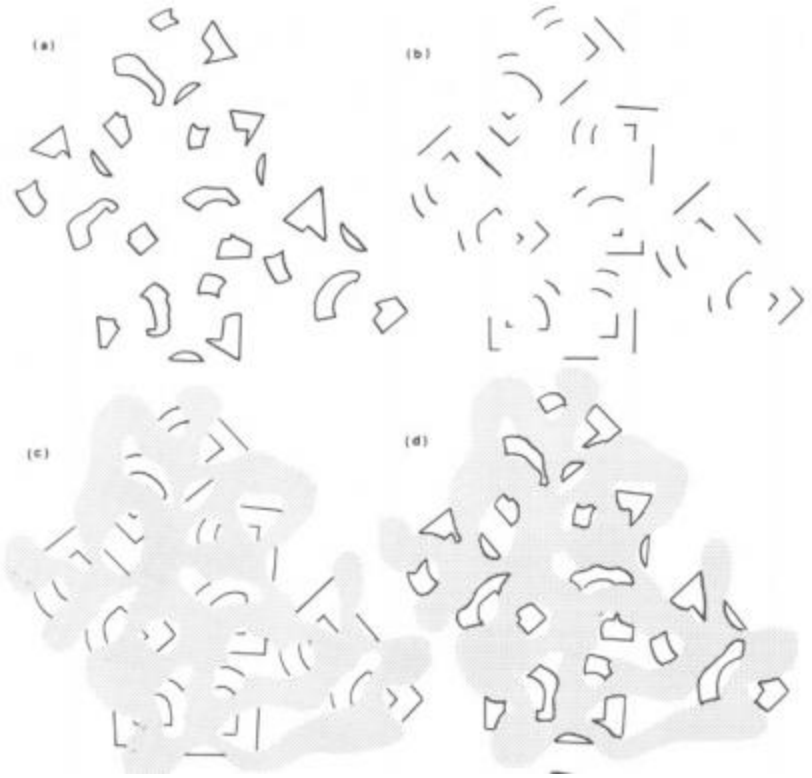

Figure 15. Example displays from Brown & Koch (1993). All four displays show fragments of the letter B. a) Closed, unoccluded fragments. b) Open, unoccluded fragments. c) Open, occluded fragments. d) Closed, occluded fragments. Reprinted from Brown, J. M., & Koch, C. (1993). Influences of closure, occlusion, and size on the perception of fragmented pictures. Perception & Psychophysics, 53(4), 436-442.

Because we will later consider the role that border ownership cues play in the second stage, it is worth mentioning here that an alternative interpretation of Brown & Koch's (1993) results takes into consideration the fact that the closure of the fragments in their displays alters border ownership cues, and that the benefit conferred by an occluder is always one of border ownership. One problem

with Brown & Koch's closed, occluded manipulation (15d) is that the fragment boundaries, where they lie adjacent to the occluder, are hand drawn, are visibly irregular, and, in some places, produce gaps between the occluder and the fragments. An object cannot be considered to be an occluder when it lies adjacent to fragments that clearly end beside it rather than being covered by it. Even if there were no gaps, though, the mere existence of the black outline adjacent to the light grey occluder suggests that the fragments boundaries end there. The light grey occluder is simply unable to own the black outlines of the fragments. It would be extremely unlikely, in real life, that objects would happen to be occluded in just such a way that black outline patterns on their surface happened to lie perfectly adjacent to the occluder. This problem could be resolved by utilizing filled fragments rather than outlines or by setting the color of the occluder to match the outlines of the fragments. In either case, it would also be necessary to ensure that there were no visible gaps between the occluder and adjacent fragment boundaries.

A second paper by Brown & Koch has influenced the design of experiments that will be presented in this paper, so it, too, is worth mentioning here: In 2000, Brown & Koch utilized similar displays as those shown in Figure 15 above except that the occluder was altered to be black. They compared two conditions: one in which the black outline of the fragments was occluded by a black occluder (resolving the problem mentioned in the paragraph above), and one in which the fragment's black outlines were deleted where they would have been occluded (i.e. the black fragments were left "open" in some places, and an illusory white occluder was possible). The occluded displays were noticeably more cluttered, and the authors mentioned, in their discussion, that the longer response times they observed in that condition may have resulted at least partially from the additional time it took observers to segment the occluder from the fragments. I will return to this point in the method for Experiment 3.

Whether Brown & Koch's results are interpreted in terms of border ownership cues or not, the above work certainly demonstrates the usefulness of a character recognition task with fragmented displays. Because the task requires recognition of a fragmented object, and because the outputs of stage 2 go forward to object recognition processes, and because recognition likely depends on complete object descriptions, a modified version of the task used by Brown & Koch is a good means to target the outputs of the second stage.

Drawing inspiration from Brown & Koch's displays, I designed a paradigm in which the recognition of a whole, alphanumeric character is supported by interpolations connecting multiple fragments. In this paradigm, multiple scene cues can be manipulated to affect the strength of interpolations exiting the second stage, and accuracy in recognizing the whole alphanumeric character can be measured. Figure 16 below illustrates a control stimulus in which no scene cues are being manipulated and in which strengths of interpolated connections should generally be strong.

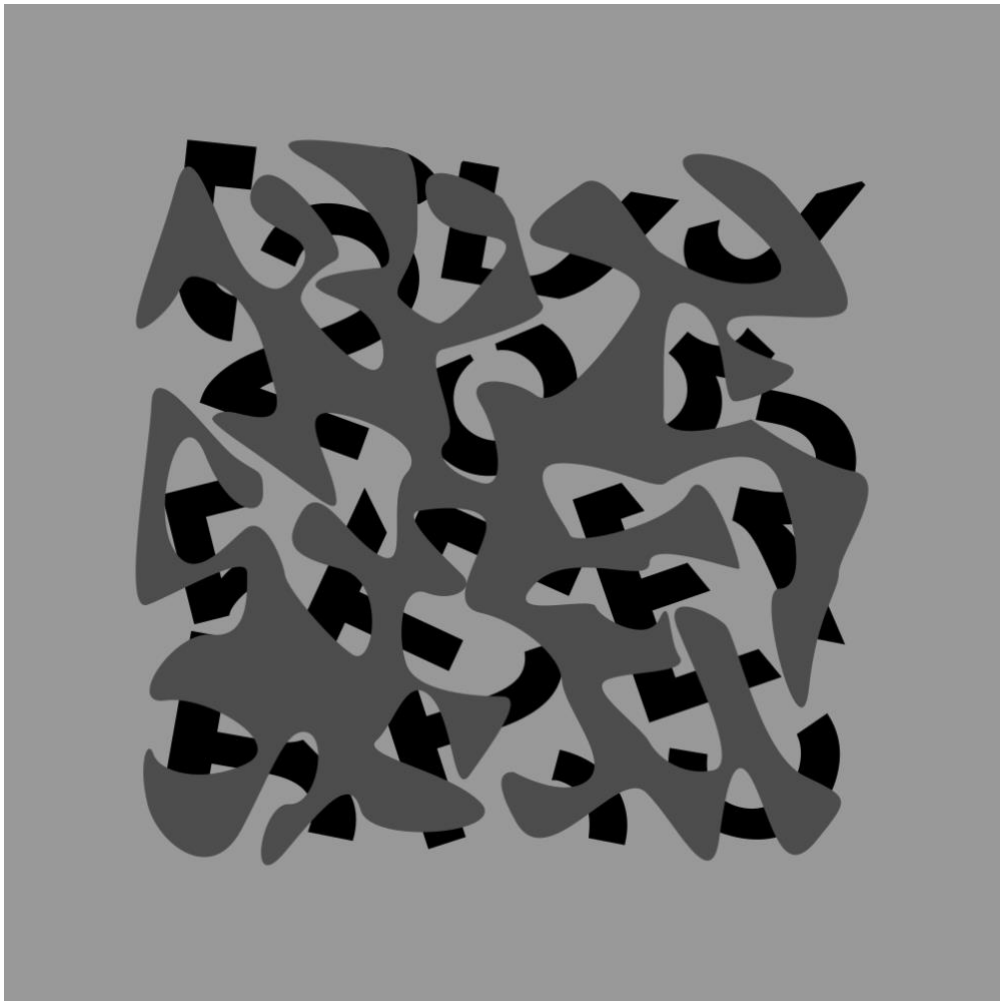

Figure 16. Example control stimulus for character recognition experiments designed to target the outputs of second stage processing. In this particular stimulus, the number '2' is hidden in the upper left. In the experiments reported below, scene cues will be manipulated, and accuracy in recognizing the hidden whole alphanumeric character, relative to control stimuli, will be measured.

The stimulus shown in figure 16 above includes some careful controls that are not present in either Bregman's original display or in the displays of Brown & Koch. The first of these is that all fragments that belong to the same whole are relatable. This is critical to ensuring that the interpolations necessary to complete the whole will be produced in stage 1. In addition, to prevent *recognition from partial information* (or the recognition of a whole merely by the recognition of one of its parts, without perception of a whole), I utilized 10 different alphanumeric characters and ensure that

each stimulus contains at least 1 fragment from each alphanumeric character besides the hidden whole. This ensures that observers cannot arrive at the correct answer merely by recognizing one of the parts in the display. Other controls will be discussed in the method to Experiment 1.

3.3 Scene cues to consider

In the sections that follow, I will focus, in turn, on border ownership cues, surface quality cues (namely, differences in luminance contrast polarity and equiluminant color contrast across fragments), and, finally, spatial frequency differences across fragments. In each section, I will first review the existing evidence that suggests the scene cue in question has an effect on the strengths of interpolations as determined in the second stage. I will then present some experiments that manipulate the cue in the character recognition task briefly described above (and described more thoroughly in the following method section). Through these experiments, I will objectively measure the effects of the cue on the final perception of interpolated contours.

3.4 Border ownership and second stage processing

The role that border ownership cues play in grouping and illusory contour perception is well documented. For example, it has long been known that outline fragments do not support the perception of illusory contours (Kanizsa, 1979; Kellman & Shipley, 1991; Sambin, 1987; He & Ooi, 1998a). However, if the two-stage theory is correct, then interpolation will proceed in the first stage between all pairs of relatable edges terminating at junctions, regardless of the outline nature of fragments. The results of Guttman & Kellman (2001), discussed in the introduction to this dissertation, support this hypothesis. Their work, which utilizes a fat-thin task, shows that observers are equally sensitive to interpolated trajectories connecting outline fragments and interpolated trajectories connecting filled fragments. Yet, just as with the Gabor elements in the path paradigm,

observers do not report seeing interpolated contours connecting outline inducers in the fat-thin task. Figure 17 below contrasts some filled inducer stimuli with their outline inducer versions.

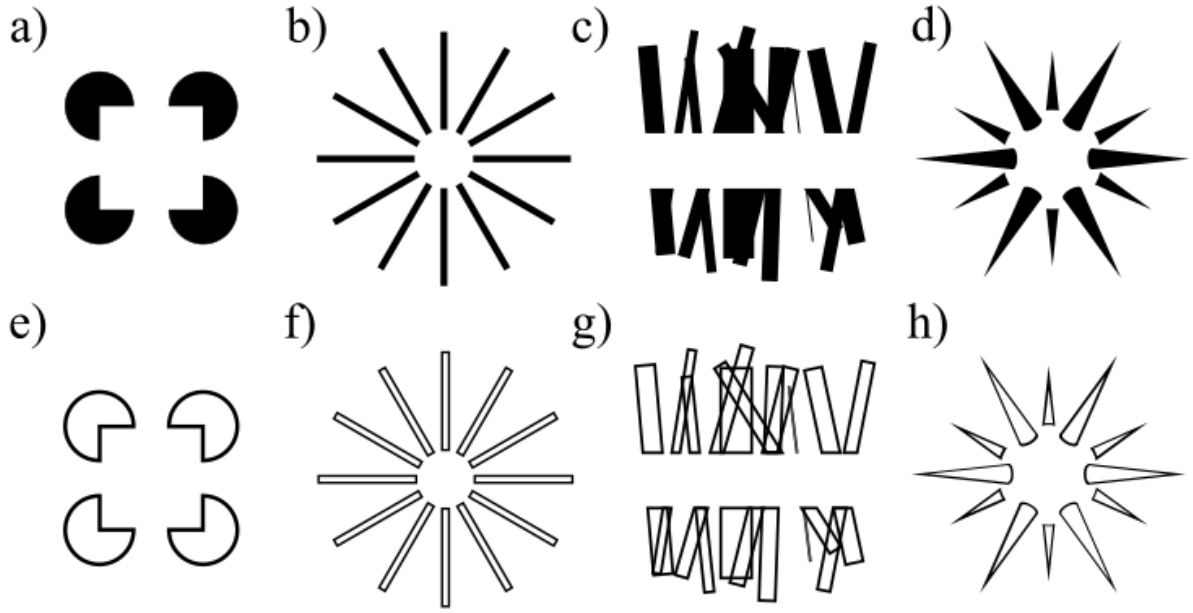

Figure 17. Illusory figures arise from filled but not outline inducers. In a), b), c), and d), filled black inducers give rise to the compelling perception of illusory contours. In a), a white square is seen occluding four black circles. In b, a white circle is seen occluding a set of black lines. In c), a white bar occluders a messy arrangement of black bars. In d), black, conical spikes protrude from a white sphere. In e), f), g), and h), the illusory shapes and contours have disappeared.

What is special about outline fragments that would cause interpolations between them to be interrupted in the second stage? One explanation of the phenomenological absence of perceived illusory contours for outline stimuli is that border ownership cues taken into account in the second stage suggest that the outline fragments own their own borders, resulting in significantly reduced or even eliminated strength of illusory contours for these stimuli. Some work supports the idea that border ownership cues play a role in determining the final appearance of interpolated contours (Albert, 2001; Kogo & Wagemans, 2013; Nakayama, Shimojo, & Silverman, 1989; Johnson & Olshausen, 2005). For example, Nakayama, Shimojo, & Silverman (1989) presented pieces of a face

on slats either in front of or in back of a stereoscopic depth plane. An example of their stimulus is shown in Figure 18. Recognition was better when the face was presented in back of the stereoscopic depth plane. This result makes sense when you consider that the set of black bars are able to own the horizontal borders of the face fragments only when the black bars lie in front. The ownership of the horizontal borders by the black bars confers upon the face fragments the status of part rather than whole, therefore supporting the grouping of the fragments.

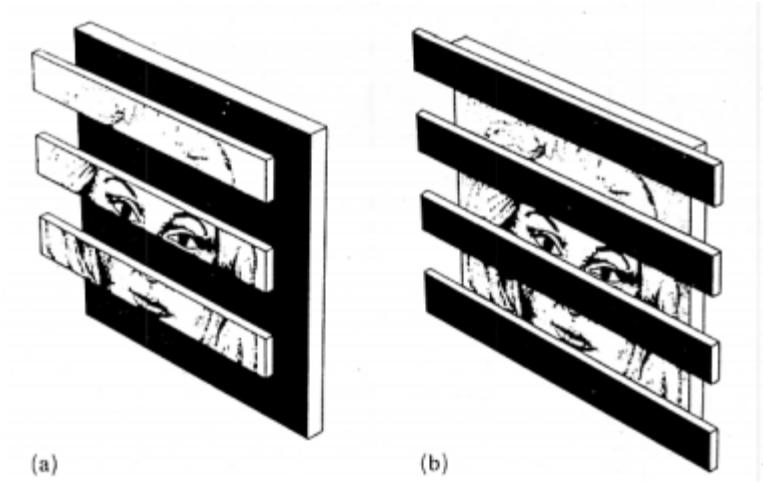

Figure 18. Displays from Nakayama, Shimojo, & Silverman (1989). Parts of faces were presented either at the front or back of a stereoscopic depth plane. Individual error rates for recognizing faces were greatest when the face was presented in front. Reprinted from Nakayama, K., Shimojo, S., & SIlverman, G. H. (1989). Stereoscopic depth: Its relation to image segmentation, grouping, and the recognition of occluded objects. Perception, 18, 55-68.

3.4.1 Specific manipulations

In order to test the possibility that border ownership cues play a role in second stage processing, three manipulations to border ownership cues were utilized in Experiments 3, 4, and 5. In each case, either the occluder or the fragments are manipulated such that the phenomenological perception of

interpolated contours connecting fragments in the manipulated display is expected to be weaker than in the control.

Example stimuli illustrating the manipulations and their respective controls appear in Figures 19, 20, and 21. Figure 19 illustrates the control and experimental stimuli for Experiment 3. For the experimental stimuli, the occluding object was manipulated such that it was set behind the fragments and rotated by 20 degrees so that it could no longer appear to occlude the fragments or own any of their boundaries. This particular manipulation was inspired by a potential issue noted by Brown & Koch (2000). In their displays, the occluder was deleted entirely, leaving gaps between the fragments. Their displays including occluders were noticeably more cluttered, and the authors noted that the increased response times they observed to recognize alphanumeric characters in the occluded case may have been at least partially due to the additional time it took to segment the fragments from the background and the occluder.

For the displays utilized here, in Experiment 3, the grey "occluder" object was present in both the control and the experimental condition - as an occluder in the control condition and as a rotated background object in the experimental condition. Utilizing the occluder as a background object in the experimental displays (rather than deleting it) helped ensure that the amount of clutter, and the need to segment the fragments from the grey object, was roughly similar across the two display types. In addition, there was another benefit of utilizing the grey occluder as a background object in the experimental displays: because the grey object was visible through gaps between the fragments, an illusory occluder, with the same medium grey color of the background, was impossible. The manipulation used here is, therefore, an even stronger manipulation than deleting the occluder. Because no occluder (of any surface color - different from or identical to the background) can possibly exist, the fragments should be much more likely to be perceived to own their own boundaries. The strength of interpolations connecting fragments will likely, therefore, be reduced

even more than if the occluder had simply been deleted (leaving open the possibility of an illusory grey occluder with the same color as the background).

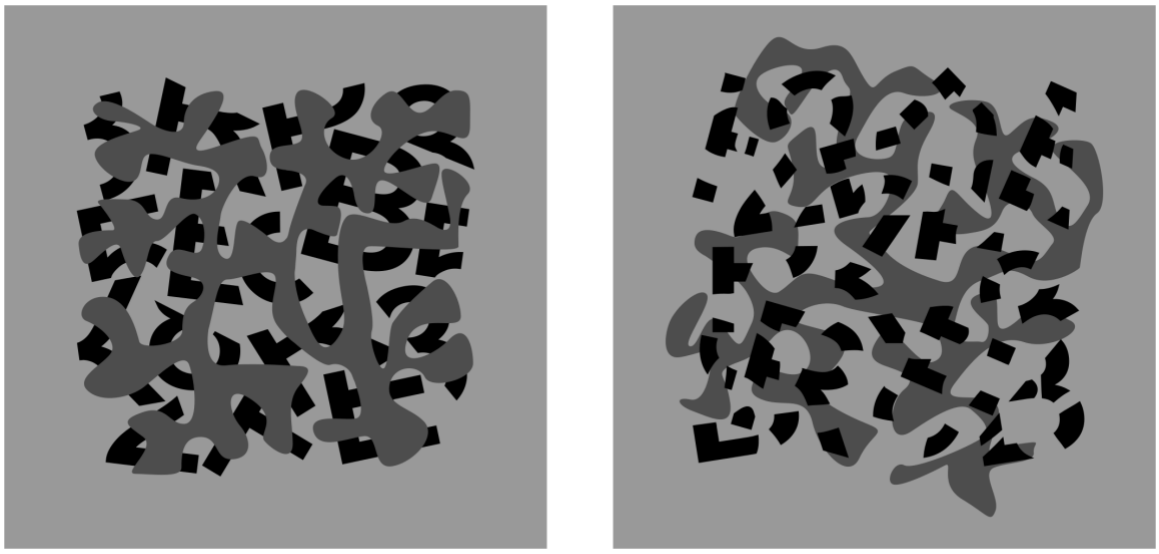

Figure 19. Example stimuli for Experiment 3. At left, a control stimulus is seen in which the letter 'B' is hidden in the upper right. At right, a border ownership cue is manipulated: the occluder has been set behind the fragments and rotated by 20 degrees so that it no longer lies adjacent to the boundaries of the fragments and can no longer own those boundaries. Because this border ownership cue does not support the grouping of fragments into a whole, the 'R' hidden in the lower left is expected to be harder to identify.

Figure 20 illustrates the control and experimental stimuli for Experiment 4. Here, the color of the occluder was altered to match the background (that is, the occluder was deleted, but the gaps between fragments remain). This manipulation is more analogous to the one used by Brown & Koch (1993) except that the problems noted in the introduction to this chapter were solved by utilizing filled, rather than outline fragments. Because there is no occluder to own the fragment boundaries, interpolations connecting fragments are expected to receive reduced strengths in the second stage. However, because an illusory occluder remains possible, the detriment to recognition accuracy may not be as great as for the rotation manipulation depicted in Figure 19 (Experiment 3).

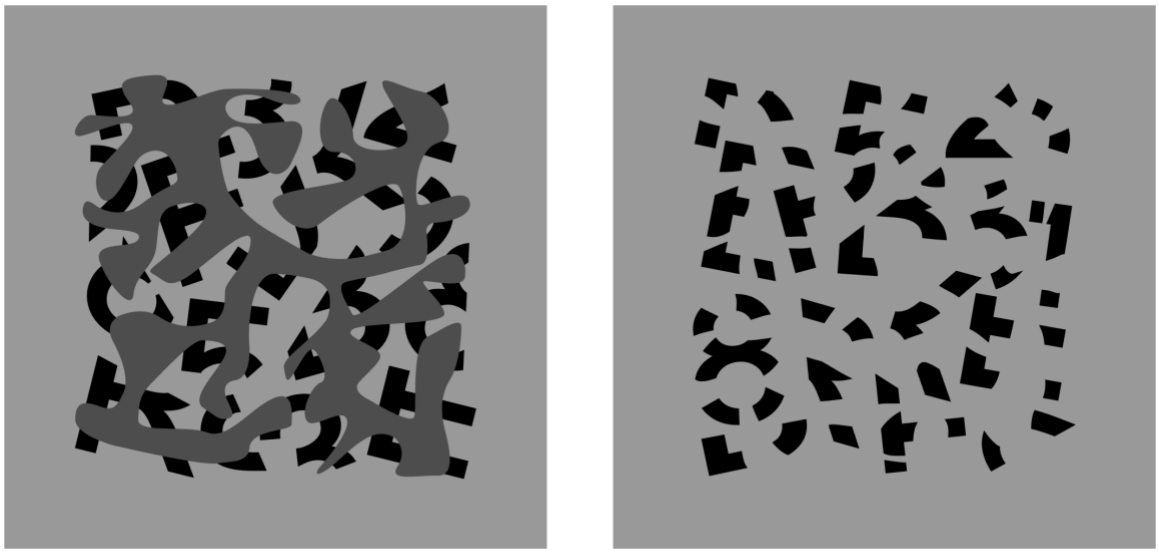

Figure 20. Example stimuli for Experiment 4. At left, a control stimulus is seen in which the letter 'F' is hidden in the upper left. At right, a border ownership cue is manipulated: the occluder has been set to the same color as the background. Because most of the occluder's boundaries are not physically specified, and because the occluder's boundaries are very irregular, the visual system may be less likely to assume that an occluder is present. Without an occluder, the fragments are left to own their boundaries. The '8' hidden in the lower left is therefore expected to be harder to identify.

In Figure 21, the unoccluded, filled fragments used in the control condition for Experiment 5 are contrasted with the experimental stimuli, which utilize outline rather than filled fragments. Because outline fragments tend to own their boundaries (Kanizsa, 1979; Kellman & Shipley, 1991; Sambin, 1987; He & Ooi, 1998a), the strength of interpolated connections, as determined in stage 2, is expected to be weaker in the outline case. Here, because both the control and the experimental stimuli permit the perception of an illusory occluder, the drop in accuracy performance relative to the control might be expected to be smaller than either of the other two border ownership manipulations.

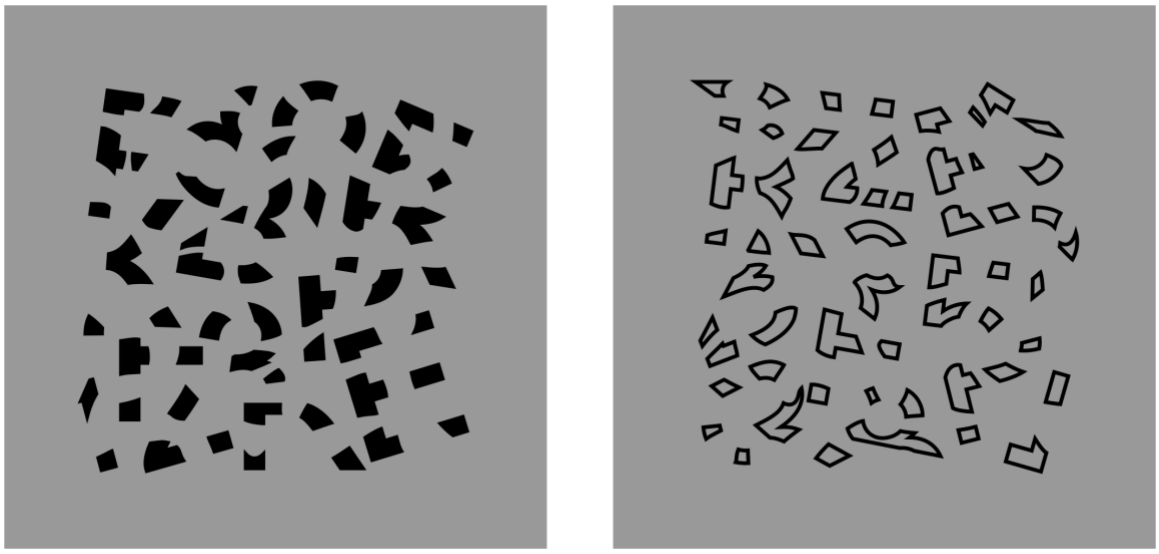

Figure 21. Example stimuli for Experiment 5. At left, a control stimulus is seen in which the letter 'E' is hidden in the bottom right. At right, a border ownership cue is manipulated: the filled fragments have been made into outlines. Because outlines tend to own their own boundaries, the 'K' hidden in the upper left is expected to be harder to identify.

Experiments 3 through 5 have three main goals: 1) to confirm that border ownership cues play a role in determining the perceptual strengths of interpolated contours as determined in the second stage, 2) to identify some specific border ownership manipulations that have an effect, and 3) to determine, for each cue, the magnitude of its effect. These measured magnitudes will be discussed again in Chapter 4, where they will be used in a set of Experiments to test a model of cue combination in the second stage.

3.5 Experiment 3: Occluder set behind fragments & rotated

This experiment utilized the task described in the introduction to this chapter and the border ownership manipulation illustrated in Figure 21. Stimuli consisted of a set of fragments randomly

arranged against a grey background. Hidden within the fragments was a subset of fragments that, if they were grouped together, and if the gaps in between fragments were filled in, formed a whole number or letter. The control condition utilized filled black fragments and a grey occluder. In the manipulation, the grey object was set behind the fragments and rotated twenty degrees so that it could not act as an occluder and was therefore unable to own any of the fragment boundaries.

Because the grey object, in the experimental condition, was unable to own the boundaries of the fragments, and because the fragments were, therefore, expected to own their own boundaries, I predicted that the weights of interpolations connecting fragments in these stimuli would be reduced in the second stage. This was expected to result in lower recognition accuracy for the experimental condition relative to the control.

In addition to the condition manipulation, presentation times were manipulated utilizing times of 50ms, 100ms, 200ms, 400ms, 600ms, 800ms, 1s, and 2s. Because interpolation takes around 150ms to proceed, the hypothesized difference in accuracy performance across the two conditions was expected to materialize somewhere between 100 and 200ms. In addition, because longer presentation times allow more time to scan and search for the hidden alphanumeric character, recognition accuracy in the control condition was expected to increase with increasing presentation time. However, for the experimental condition, increases in presentation times were expected to have less of an impact. This was because the interpolations exiting stage 2 should have relatively weak strengths; as a result, the hidden alphanumeric character was expected to be less apparent, even when the observer was looking directly at the particular fragments that composed it. I therefore expected the difference between the control and experimental conditions to increase with increasing presentation times, as the increasing time to search the array improved performance in the control condition more than in the experimental condition.

3.5.1. Method

Participants. Nineteen University of California, Los Angeles undergraduate students participated in the experiment. One participant, who failed to complete the experiment within the allotted hour, was excluded from the analysis. Thus, eighteen participants (4 male, 14 female, ages 18-32, M_{age}: 20.83 years) were included in the analysis. All participants provided informed consent, were offered partial course credit for their participation, and were naïve to the purposes of the study. Visual acuity was tested using a standardized ETDRS chart as a prerequisite to participation. All participants demonstrated 20/25 or better visual acuity.

Apparatus. Stimuli were presented using the MATLAB programming language and the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). Stimuli were presented on a Viewsonic G250 CRT monitor powered by a MacPro 4 with a 2.66 GHz Quad-Core Intel Xeon processor and an NVidia GeForce GT120 graphics card. The illuminated area of the screen had a height of approximately 29.6 cm and a width of approximately 37 cm. The resolution of the monitor was set to 1080 x 1024 pixels and a refresh rate of 60 Hz.

Stimuli. Some problems with the stimuli used by Brown & Koch (1993, 2000) were taken into consideration in designing the stimuli used here (see the introduction to this chapter above). In each stimulus, only a single whole alphanumeric character appeared, surrounded by a noisy background consisting of parts of ten other different letters and numbers. This was done to ensure that recognition from partial information would not be sufficient to complete the task. That is, an observer could not guarantee a correct response merely by recognizing one of the parts present in the display. In addition, for stimuli that contained a grey occluder (Experiments 3 & 4) or a grey background object (Experiment 3 - rotated occluder set behind fragments), the grey occluder or background object appeared on screen for 500ms before the fragments were added. This was done to ensure that the presence of an additional object would not reduce accuracy performance as a

result of the limited presentation times and the additional time it takes to segment fragments and another object. Brown & Koch (2000) noted that the presence of a physically specified occluder, and the additional time it took to segment the occluder and fragments, may have resulted in longer response times in their experiments relative to conditions with no occluder.

The experiment utilized a set of 40 control stimuli and 40 manipulations of those control stimuli. Example stimuli are illustrated in Figure 19. Stimuli were created in Inkscape. Fragments of the following ten alphanumeric characters were used: 2, 3, 5, 8, B, E, F, K, R, and P. These letters and numbers were chosen for their relative similarity of appearance. This was done to reduce the possibility that correct answers would be achieved merely by recognizing the presence of unique parts in the display. To create these fragments, text of each character (upper case for letters) was first created using Arial font. The '5' and the 'K' were then modified to make their parts more confusable with the parts of the other letters and numbers: the slanted vertical edge of the '5' was made to be perfectly vertical, while the leg of the 'K' was altered so that it intersected the spine halfway up, at the same location as the arm. Two sets of "toolkit" fragments were then created by segmenting each alphanumeric character in two different ways. The fragments in these "toolkit" sets were used to populate the random backgrounds of the stimuli.

Each stimulus was created by randomly arranging and rotating 15 fragments from the "toolkit" along with one whole, complete number or letter on a 2x2 grid. For each of the ten alphanumeric characters used, four different stimuli were created so that a given alphanumeric character appeared once as the hidden whole in each of the four quadrants. In addition to the whole character, each stimulus contained at least one fragment from each of the other possible 9 alphanumeric characters. Fragments were not utilized as part of the noisy background if the alphanumeric character from which they were drawn was the hidden whole in the stimulus. Before adding the occluder (which further segmented fragments), each stimulus contained 15 fragments in addition to the hidden whole

number or letter. Across the entire set of stimuli, each possible alphanumeric character was represented roughly equally in the random fragment backgrounds.

Once the whole characters and fragment backgrounds were in place, a dark grey, blob-like occluder was drawn in to further segment each of the fragments as well as the whole alphanumeric character. The occluder was drawn in such that all resulting fragments belonging to the hidden whole character were relatable. To ensure that occluders were as similar as possible across the entire set of images, they were drawn in a systematic way: a horizontal "spine" traversed the middle region of the stimulus at the intersection of the four quadrants, and four appendages spread off of this spine to partially occlude each of the fragments in each of the four quadrants. Constructing every occluder in this way helped ensure that individual stimuli could not be recognized based on any memorable uniqueness of the occluder.

All fragments and occluders were given a stroke of 7 pixels, which spanned 4.6 arcminutes at presentation. The fragments were set to a filled black with a black stroke; the occluder was set to a filled dark grey with a dark grey stroke. The region covered by the fragments was 750 x 750 pixels, or 8.27 x 8.27 degrees of visual angle at presentation. The stimulus was jittered by up to 1 degree on each trial, but was otherwise presented centered on the screen. Stimuli were presented on a uniform, medium grey background.

In total, 10 (alphanumeric characters) x 4 (quadrants) = 40 stimuli were created in the above manner. Modified versions of this set of 40 control stimuli were used to create the 40 stimuli for the experimental condition. The experimental stimuli were created by first using the occluder to cut sections out of the fragments (regions covered by the occluder in the control stimuli were cut out of the fragments in the experimental stimuli). The occluder was then set behind the fragments and rotated by 20 degrees.

A mask was created by randomly arranging all fragments from one of the "toolkit" sets on a grey background. The mask filled the screen.

Procedure. Participants were first given a visual acuity test using a standardized ETDRS chart to confirm visual acuity was at least 20/25. Participants were then shown two examples of the types of stimuli they would see in the experiment: one version for the control, and one for the experimental condition; both example stimuli contained the number '2' hidden in the upper left corner and were identical apart from the manipulation.

Participants were then instructed that they would see one image on each trial, that the image would look something like one of the example images they were being shown, and that, in either case, there would be some randomly arranged fragments within the image. Participants were then told that, hidden within the randomly arranged fragments, there would be one and only one subset of fragments that, if grouped together, and if the gaps in between the fragments were filled in, could become a whole number or letter. They were instructed that they would be expected to fill in gaps between fragments but would never be asked to fill in end pieces. The experimenter then pointed out one subset of fragments in each image that, if grouped together, and if the gaps between the fragments were filled in, could be completed to form a partial letter 'E'. The experimenter pointed out that the partial 'E' could not be the hidden whole because the ends of two of the prongs of the E were missing. The experimenter then told participants that there was one set of fragments in each of the images that could become a whole number or letter just by filling in the gaps between the fragments - without ever filling in end pieces. The participants were then asked if they saw the hidden whole number or letter in the image. The experimenter then waited until the participant indicated the hidden '2' in either image. The participant was then told that they were correct. The Experimenter then covered the bottom of the two and the top of the two in turn and said that if either the top or the bottom ends of the two were missing, it would not count as the hidden whole.

They then said that, because, as it was, no ends were missing, and only gaps between fragments needed to be filled in, the '2' counted as a hidden whole. The experimenter then ensured the participant saw the hidden '2' in both sets of images by tracing each in turn.

The participant was then told that they would see the images for varying presentation times, and that they would first complete a set of practice trials that would take them through all of the presentation times they would see in the experiment, beginning with the longest and ending with the shortest. Finally, they were told that there would always be a second image (the visual mask) presented after the first image, that this image was there only to prevent their visual system from processing the first image for a longer period of time, that there would never be a hidden whole in the second image, and that they could ignore it.

Participants then completed 65 practice trials utilizing a series of shortening presentation times that was designed to acclimate them to the very fast presentation times that would be utilized in the experiment. The practice trial presentation times utilized were: 5, 4, 3, 2, and 1s, and 800, 600, 500, 400, 300, 200, 100, and 50ms. The participant received 5 trials at each of the practice trial presentation times.

After completing the practice trials, participants began the real experiment. The real experiment utilized presentation times of 50ms, 100ms, 200ms, 400ms, 600s, 800ms, 1s, and 2s. Each trial consisted of the presentation of a fixation cross for 500ms, followed by the presentation of the grey blob object for 500ms (if any such object was included in the stimuli for that Experiment - this does not include Experiment 5, for which both control and experimental stimuli were unoccluded). The fragments were then added (on top of the grey blob object in the experimental condition) to the screen for an amount of time determined by the presentation time for that trial, followed by a mask for 300ms, followed by a screen requesting their response, which was made by keypress. Presentation order was randomized. A progress bar was revealed following every response, and

participants were invited to take a break every 50 trials. The entire experiment involved 10 (alphanumeric characters) x 4 (quadrants) x 2 (conditions) x 8 (presentation times) = 640 trials and took 40-50 minutes to complete.

3.5.2. Results

Figure 22 plots the data from each condition across the range of presentation times used. A clear difference between the control and experimental conditions emerges between 100 and 200ms and increases with longer presentation times. To confirm these observations, the data were submitted to a 2 (conditions) x 8 (presentation times) repeated measures ANOVA. All statistics were Greenhouse-Geisser corrected. There was a main effect of stimulus type ($F(1, 17) = 54.99$, $MSE = .02$, $p < 0.001$, $\eta^2 = .11$, $\eta_p^2 = 0.76$), and presentation time ($F(2.01, 34.09) = 107.26$, $MSE = .03$, $p < 0.001$, $\eta^2 = .66$, $\eta_p^2 = 0.86$), and there was a significant interaction between stimulus type and presentation time ($F(3.74, 63.57) = 24.49$, $MSE = .01$, $p < 0.001$, $\eta^2 = .06$, $\eta_p^2 = 0.59$).

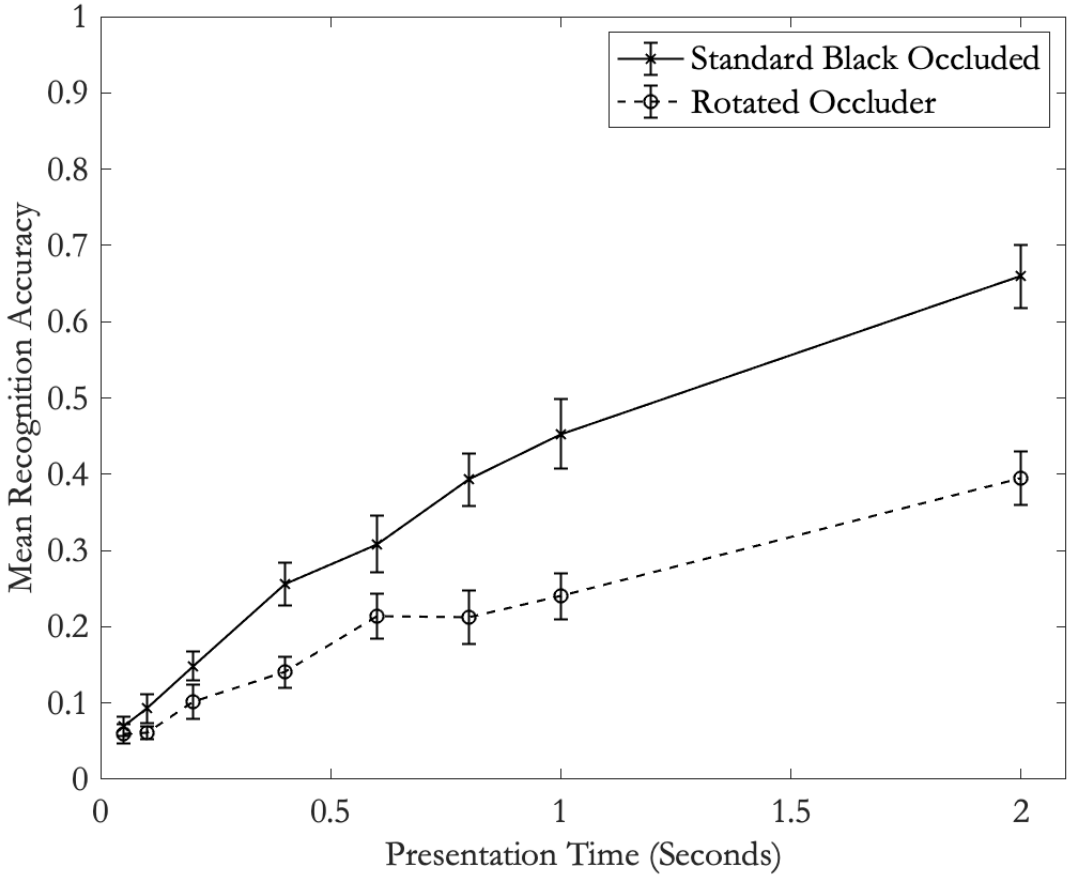

Figure 22. Results from Experiment 3. Performance in the control condition is illustrated by the solid line, while performance in the experimental condition (grey object rotated and set behind fragments) is illustrated by the dotted line. Beginning between 100 and 200ms, a clear difference between the conditions emerges, with the experimental manipulation resulting in substantially worse recognition accuracy. The difference increases with larger presentation times. Error bars show 1 standard error above and below the condition means.

3.5.3. Discussion

It was predicted that recognition accuracy in the control (filled occluded) condition would be greater than recognition accuracy in the experimental condition (occluder object set behind fragments and rotated 20 degrees), that these differences would appear somewhere between 100 and 200ms, and that the difference would get larger with longer presentation times. In addition, it was predicted that the effect of this manipulation would be fairly large - larger than either of the other

two border ownership manipulations (Experiments 4 and 5). All of those predictions are borne out in the data.

A detriment to performance for the experimental condition was predicted because the manipulation caused the grey blob object to be unable to own the boundaries of the fragments. More specifically, when the grey blob object was set behind the fragments and rotated, the physically specified T-junctions of the fragments became L-junctions. L-junctions give rise to the perception of illusory contours only if an occluding object, with the same color as the background, is perceived. Because the grey blob object lied intermediate between the fragments and their grey background in the experimental condition, and was visible through gaps between the fragments, the perception of an illusory occluder, with the same color as the background, was impossible. The fragments were therefore expected to own their own boundaries. In order for the whole number or letter to be recognizable, the observer would need to see the fragments that made up the alphanumeric character as parts of a whole. In order for those fragments to be perceived as parts of a whole, the interpolations necessary to connect the fragments would need to make it out of the second stage with sufficiently high strengths. The results suggest that this particular border ownership manipulation causes the strengths of the interpolations to be sufficiently reduced in the second stage such that recognizing the alphanumeric character is more difficult.

Because all of the to-be-connected fragment edges in the display were relatable, and because the physically specified edges in the two displays were identical, the differences between the two conditions cannot be explained by issues in stage 1. The results of Guttman & Kellman (2001) and Experiments 1 and 2 of this dissertation suggest that identical promiscuous interpolations would have been produced for both the control and experimental displays in stage 1.

These results support the conclusion that border ownership cues are taken into consideration in the second stage to determine which interpolations will appear in the final representation and what

their relative strengths will be, and that this particular border ownership cue - the absence of a physically specified occluder combined with the impossibility of an illusory occluder - has a significant impact on the ultimate perceptual strengths of interpolations. Experiments 4 and 5, will examine two more border ownership cues - the deletion of the occluder (Experiment 4) and the use of outline rather than filled fragments (Experiment 5) - to determine whether and what role they play in stage 2 processing.

3.6 Experiment 4: Unoccluded fragments

Experiment 3 examined the effect of setting the occluder behind the fragments and rotating it by twenty degrees. That manipulation eliminated any physically specified occluder and also ensured that the perception of an illusory occluder, with the same color as the background, was impossible (because the grey blob object lied intermediate between the fragments and their background and was visible between gaps in the fragments). The goal of Experiment 4 was to determine whether the simple deletion of the occluder would be sufficient to cause decreases in the perceptual weights of interpolated contours, determined in stage 2. Examples of the Experiment 4 displays are illustrated in Figure 20. Because the grey blob object is entirely eliminated from the display (rather than appearing rotated in the background), it remains possible, for these displays, that an illusory occluder could be perceived to own some of the fragment boundaries. This manipulation, therefore, might be expected to produce a smaller reduction in performance than was observed in Experiment 3.

In addition to the condition manipulation, presentation time was manipulated using the same levels as in Experiment 3. Differences between the control (filled, occluded fragments) and experimental (filled, unoccluded fragments) conditions were expected to appear somewhere between 100 and 200ms because interpolation is known to take around 150ms to proceed. For the same reasons as in Experiment 3, the difference between conditions was expected to increase with

increasing presentation time. The reasoning was this: the relatively weak strengths of the interpolations for the experimental condition were expected to cause the hidden alphanumeric character to be less obvious, even when the observer was looking directly at the fragments that compose it. Longer presentation times allow more time to scan and search for the hidden alphanumeric character, but this additional time was expected to be less beneficial when the hidden number or letter wasn't as apparent even when looking directly at it.

3.6.1 Method

Participants. Twenty-three University of California, Los Angeles undergraduate students participated in the experiment. One participant, whose performance was at chance at all presentation times in both stimulus conditions, was eliminated from the analysis. Thus twenty-two participants (17 female, 5 male, ages 18-25, M_{age}: 20.50 years) were included in the analysis. All participants provided informed consent, were students attending the University of California, Los Angeles, were offered course credit for their participation, and were naïve to the purposes of the study. Visual acuity was tested using a standardized ETDRS chart as a prerequisite to participation. All participants demonstrated 20/25 or better visual acuity.

Apparatus. The computer and test environment were the same as for Experiment 3.

Stimuli. Stimuli were identical to those used in Experiment 3, except that, in the manipulation condition, the grey occluder was set to the same color as the background (that is, the occluder was eliminated from the display, but the gaps between fragments remained). The physically specified edges of the fragments were, therefore, identical across the two conditions, but the occluder was present in the control stimuli and absent from the experimental stimuli.

The same mask was used as in Experiment 3.

Procedure. The procedure was identical to that of Experiment 3 except that the sample stimuli used in the instructions were modified in the manner specified above.

3.6.2 Results

Figure 23 plots the data from each condition across the range of presentation times used. A clear difference between the control and experimental conditions emerges between 100 and 200ms and increases with longer presentation times. To confirm these observations, the data were submitted to a 2 (conditions) x 8 (presentation times) repeated measures ANOVA. All statistics were Greenhouse-Geisser corrected. There was a main effect of stimulus type ($F(1, 21) = 263.44$, $MSE = .005$, $p < 0.001$, $\eta^2 = .10$, $\eta_{p^2} = .93$), and presentation time ($F(2.07, 43.47) = 238.52$, $MSE = .021$, $p < 0.001$, $\eta^2 = .76$, $\eta_{p^2} = .92$), and a significant interaction between stimulus type and presentation time ($F(5.21, 109.35) = 17.69$, $MSE = .005$, $p < 0.001$, $\eta^2 = .03$, $\eta_{p^2} = .46$).

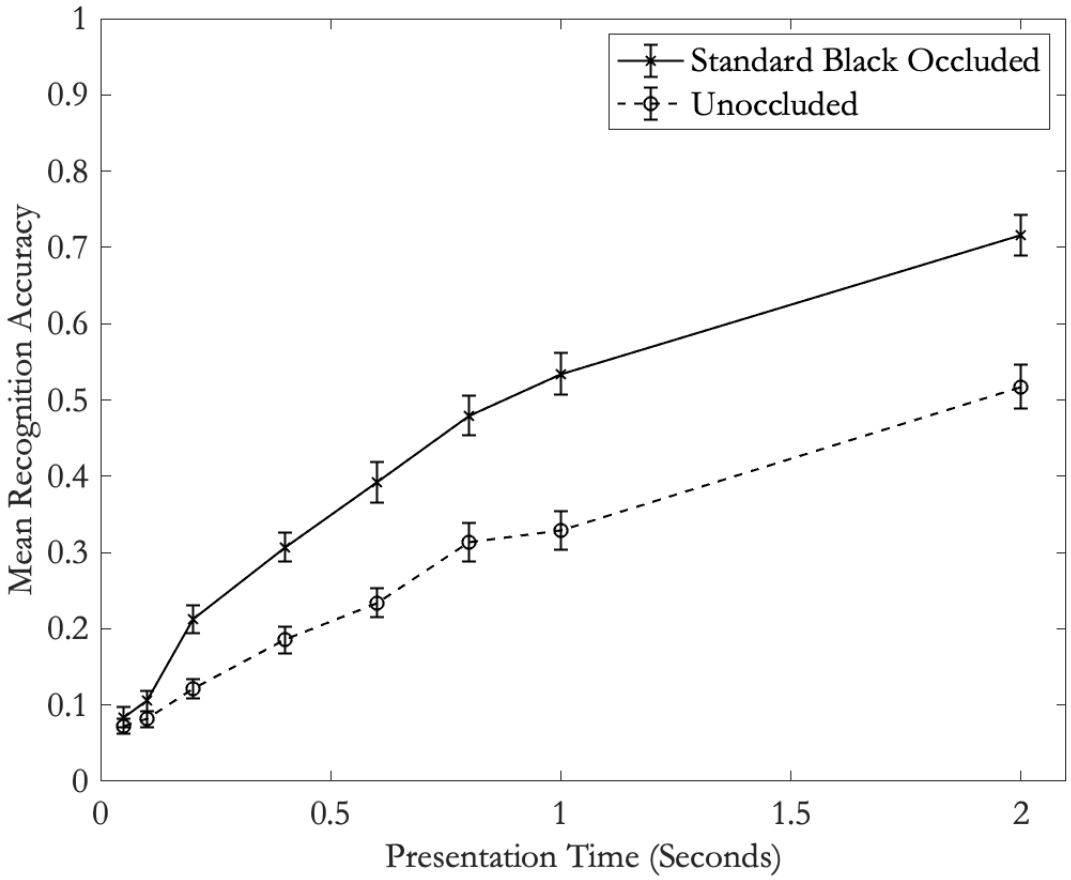

Figure 23. Results from Experiment 4. Performance in the control condition (standard black occluded) is illustrated by the solid line, while performance in the experimental condition (for which the occluder was made the same color as the background) is illustrated by the dotted line. Beginning between 100 and 200ms, a clear difference between the conditions emerges, with the elimination of the occluder resulting in substantially worse recognition accuracy. The difference increases with larger presentation times. Error bars show 1 standard error above and below the condition means.

3.6.3 Discussion

It was predicted that recognition accuracy in the control (filled, occluded fragments) condition would be greater than recognition accuracy in the experimental condition (filled, unoccluded fragments), that these differences would appear somewhere around 200ms, and that the difference would increase with increasing presentation time. In addition, the effect of this manipulation was

expected to be smaller than the effect of the manipulation in Experiment 3, but larger than the manipulation in Experiment 5. The results supported each of these predictions.

The detriment to performance for the unoccluded manipulation was expected because the manipulation eliminates the presence of an occluder that could otherwise own some of the boundaries of the fragments. With the occluder eliminated, the fragments were expected to own their own boundaries unless an illusory occluder was perceived. The perception of an illusory occluder was somewhat unlikely, as many, if not most, of the inducing edges for the occluder were not relatable. In addition, if an illusory occluder were present, its support ratio would have had to be very low. However, the possibility of an illusory occluder was expected to cause the effect of this manipulation to be smaller than that of Experiment 3.

The task, recognizing a hidden alphanumeric character composed of multiple spatially separated fragments, required the observer to see the fragments that made up the alphanumeric character as parts of a whole. In order for the fragments to be perceived as parts of a whole, the interpolations necessary to connect the fragments needed to make it out of the second stage with sufficiently high strengths. The results suggested that this particular border ownership manipulation - deleting the occluder - caused the strengths of the interpolations to be sufficiently reduced in the second stage such that recognizing the alphanumeric character became more difficult.

Because all of the to-be-connected fragment edges in the display were relatable, and because the physically specified edges in the two displays were identical, the differences between the experimental and control conditions cannot be explained by issues in Stage 1. Experiments 1 and 2 of this dissertation, and the earlier work of Guttman & Kellman (2001) suggest that identical promiscuous interpolations were produced for both the control and experimental displays in Stage 1.

These results add to the results of Experiment 3, providing further evidence that border ownership cues are among those cues taken into consideration in the second stage, and that they

play a role in determining which interpolations appear in the final scene representation and what their relative strengths will be. In Experiment 5, I will test a final border ownership cue - the use of outline rather than filled fragments - to determine how making fragments outlines affects the final perceptual strengths of interpolated contours in stage 2.

3.7 Experiment 5: Unocclued filled vs. outline fragments

Experiments 3 and 4 examined the effects of two border ownership manipulations to the occluder: rotating the occluder and setting it behind the fragments, and deleting the occluder entirely from the scene (such that the gaps between fragments remain). Experiment 5 examined a particularly well known border ownership cue that was illustrated in the introduction to this chapter (Figure 17): that of turning filled inducing fragments into outline fragments. Example stimuli for this experiment are illustrated in Figure 21. Because outline fragments are well known to have a tendency to own their own boundaries (Kanizsa, 1979; Kellman & Shipley, 1991; Sambin, 1987), this manipulation was expected to have the effect, in the second stage, of reducing the perceptual strengths of the interpolations necessary to produce a whole alphanumeric character. This, in turn, was expected to result in reduced recognition accuracy for the fragmented alphanumeric character.

In addition to the condition manipulation, presentation time was manipulated using the same levels as in Experiment 3. Interpolation takes around 150ms to proceed, therefore, the differences between the experimental and control conditions were expected to appear somewhere between 100 and 200ms. For the same reasons as in Experiment 3, this difference was expected to increase with longer presentation times. Although longer presentation times allow more time to scan and search for the hidden alphanumeric character, the hidden alphanumeric character was expected to not be very evident, even when observers were looking directly at it, if the perceptual strengths of the interpolations necessary to connect the fragments were low. For this reason, additional time to scan

and search the array for the hidden alphanumeric character was not expected to improve performance in the experimental (outline fragments) condition as much as in the control (filled fragments).

3.7.1 Method

Participants. Seventeen participants (2 male, 15 female ages 18-22, M_{age}: 20.00 years) completed the experiment. All participants provided informed consent, were students attending the University of California, Los Angeles, were offered course credit for their participation, and were naïve to the purposes of the study. Visual acuity was tested using a standardized ETDRS chart as a prerequisite to participation. All participants demonstrated 20/25 or better visual acuity.

Apparatus. The computer and test environment were the same as for Experiment 3.

Stimuli. Stimuli were identical to that of the 'filled unoccluded fragment' experimental condition used in Experiment 4, except that, in the manipulation condition, the filling of the fragments was altered to be the same color as the background. Because the original stimuli were constructed to utilize a 7 pixel stroke (equivalent to 4.6 arcminutes at presentation), the stimuli were otherwise unaltered (the stroke served as the outline for the outline fragments). The physically specified edges of the fragments were, therefore, identical across the two conditions, except that the inducing edges belonged to filled fragments in the control condition, while they belonged to outline fragments in the experimental condition.

The control condition utilized the same mask as was used in Experiment 3. For the experimental condition, the fragments in the mask were modified in the same way as the stimulus (turning the filling to the same color as the background and preserving the black stroke).

Procedure. The procedure was identical to that of Experiment 3 except that the sample stimuli used in the instructions were modified in the manner specified above.

3.7.2 Results

Figure 24 plots the data from each condition across the range of presentation times used. The difference between the filled and outline fragment conditions is small relative to the differences observed in Experiments 3 and 4, and emerges around 200ms. The difference does appear to increase as presentation time increases, but not nearly as much as in Experiments 4 and 5. To confirm these observations, the data were submitted to a 2 (conditions) x 8 (presentation times) repeated measures ANOVA. All statistics were Greenhouse-Geisser corrected. There was a main effect of stimulus type ($F(1, 16) = 19.15$, $MSE = .01$, $p < 0.001$, $\eta^2 = .01$, $\eta_p^2 = .55$), and presentation time ($F(3.48, 55.70) = 80.89$, $MSE = .01$, $p < 0.001$, $\eta^2 = .73$, $\eta_p^2 = .84$), but no interaction between stimulus type and presentation time ($F(5.19, 82.96) = 1.10$, $MSE = .01$, $p = .37$, $\eta^2 = .006$, $\eta_p^2 = .07$).

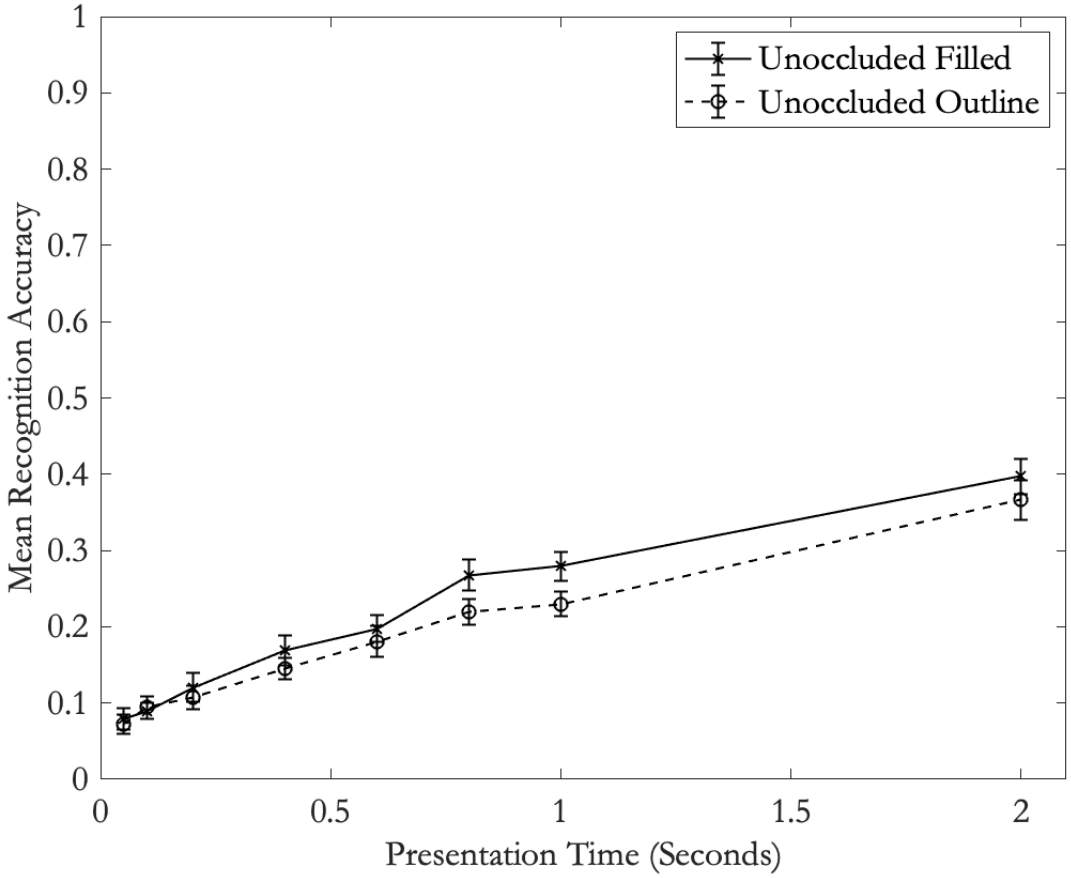

Figure 24. Results from Experiment 5. Performance in the control (filled, unoccluded fragments) condition is illustrated by the solid line, while performance in the experimental (outline, unoccluded fragments) condition is illustrated by the dotted line. A relatively small difference between the two conditions emerges around 200ms and increases up to 1 second. Error bars show 1 standard error above and below the condition mean.

3.7.3 Discussion

It was expected that recognition accuracy in the control (filled, unoccluded fragments) condition would be greater than recognition accuracy in the experimental condition (outline, unoccluded fragments), that these differences would appear somewhere around 200ms, and that the difference would increase as presentation time increased. In addition, the effect of this manipulation was expected to be smaller than the manipulations of Experiments 3 and 4. The results are largely consistent with these predictions. A difference does appear around 200ms, and increases up to 1

second. However, the difference at 2 seconds is not as large as would be expected, and the interaction between stimulus type and presentation time was not significant.

The detriment to performance for the experimental (outline, unoccluded fragments) stimuli was predicted because outlines are known to have a tendency to own their boundaries. If the outline fragments own their boundaries, then interpolations connecting those fragments should receive lower strengths in stage 2. If the interpolations have relatively low strengths, then the hidden whole number or letter, which is formed by the inducing edges and the interpolations that connect them, should be relatively difficult to locate and recognize. These results suggest that this particular border ownership manipulation - utilizing outline rather than filled fragments - does have an effect in stage 2 processing, however, the effect is not as large as the two other border ownership manipulations examined in Experiments 3 (rotating the occluder and setting it behind the fragments) and 4 (deleting the occluder entirely). That the effect of this manipulation would be smaller than either of the manipulations in Experiments 3 and 4 was expected because, in this case, an illusory occluder interpretation is possible for both the control and experimental stimulus. Thus, it was expected that an illusory occluder could be perceived to own the boundaries of the fragments in either case (in Experiments 3 and 4, the control stimulus has a physically specified occluder, which was expected to be even more likely to own the fragment boundaries, causing accuracy performance in that control condition to be even higher).

Because all of the to-be-connected fragment edges in the display are relatable, and because the physically specified edges in the two displays are identical, the effect that was observed cannot be explained by issues in Stage 1. Experiments 1 and 2 of this dissertation, and the work of Guttman & Kellman (2001) suggest that identical promiscuous interpolations will have been produced for both the control and experimental displays in Stage 1.

These results add to the results of Experiments 3 and 4, providing further evidence that border ownership cues are among those cues taken into consideration in the second stage to determine which interpolations will appear in the final representation and what their relative strengths will be.

3.8 Surface qualities and second stage processing

The results of experiments 3 through 5 suggest that border ownership cues play a role in determining the final perceptual strengths of interpolated contours in stage 2. A wealth of evidence suggests that differences in the surface qualities of connected regions also play an important role.

Figure 2 illustrated the separate and complementary processes of contour interpolation and surface interpolation (Yin, Kellman & Shipley, 1997, 2000; He & Ooi, 1998; Su, He & Ooi, 2010). Surface interpolation occurs when two spatially separated surfaces are similar in lightness, color, and texture; it involves the spreading of their surface qualities under occlusion, and the perception of the two spatially separated regions as connected (Kellman & Shipley, 1991; Yin, Kellman, & Shipley, 1997, 2000). Surface interpolation can proceed in the absence of real or interpolated edge information. However, surface interpolation is affected by contour interpolation in that, when either 2D or 3D real or interpolated edge information is available, the spreading of surfaces is confined by those edges (Kellman & Shipley, 1991; Yin, Kellman, & Shipley, 1997, 2000; Fantoni, Hilger, Gerbino, & Kellman, 2008; Kellman, Garrigan, & Shipley, 2005; Kellman, Garrigan, Shipley, Yin, & Machado, 2005). Contour interpolation therefore plays a role in constraining the spreading of surfaces.

But does the surface interpolation process play any role in constraining the contour interpolation process? The work of Guttman & Kellman (2001) combined with the path detection and subjective rating experiments of Chapter 2 have made clear that promiscuous interpolation in the first stage proceeds irrespective of surface differences across inducing fragments. Surface spreading may play a

role, though, in the second stage. It could be the case that cues from surface spreading are one of the types of information taken into consideration in the second stage. If so, then surface spreading that is consistent with an interpolated connection could have the effect of maintaining or increasing the strength of that interpolated contour. Likewise, if the surface qualities across two regions are inconsistent, the fragments may tend to be perceived as owning their boundaries, and the strength of an interpolation may be reduced.

It is worth noting that the hypothesis outlined above would predict that the surface qualities of inducers would affect surface spreading, and therefore the perception of interpolated contours, only for amodally completed surfaces. This is true because, for modally completed (illusory) figures, the surface that spreads within interpolated connections has the same color as the background - it does not have the same color as the inducers. In amodal completion, the surfaces of the inducers will spread under occlusion so long as the surface qualities are similar (Yin, Kellman, & Shipley, 1997, 2000). I will return to this issue in section 3.9.

There is, in fact, some evidence that amodal surface spreading affects the perception of interpolated contours. He and colleagues have demonstrated that both luminance contrast polarity (He & Ooi, 1998a, 1998b; He, Ooi, & Su, 2018; Su, He, & Ooi, 2010a) and equiluminant color contrast (Su, He, & Ooi, 2010b) differences across separate inducers affect the perception of interpolated contours for amodally completed figures.

There is also some physiological basis for reasoning that both luminance and color contrast polarity could potentially affect the perceived border ownership of fragments as a result of surface spreading. It has been demonstrated that some cells that encode border ownership in monkey visual cortex are also sensitive to both luminance and color contrast polarity (Zhou, Friedman, & von der Heydt, 2000).

In the sections below, I will examine both of these manipulations to the surface qualities of inducers in turn. In each section, I will first review the existing evidence and then present an experiment that manipulates that cue in the same character recognition task utilized in Experiments 3 through 5.

3.9 Luminance contrast polarity relative to the background

The work of He and colleagues (He & Ooi, 1998a, 1998b; Su, He, & Ooi, 2010a) strongly suggests that, in cases of amodal completion, lack of surface spreading beneath an occluder due to luminance contrast polarity differences relative to the background results in a severe reduction in the perceptual strength of amodally completed contours. He & Ooi's (1998a) "illusory-O" is redrawn in Figure 25, and makes for a compelling illustration of the effect. In Figure 25a, the surfaces of amodally completed fragments spread beneath an illusory grey ring that is perceived to be partially occluding those fragments. Note that the luminance contrast polarities of the connected fragments are identical: the surfaces of white fragments spread to connect with other white fragments, and the surfaces of black fragments spread to connect with other black fragments. In Figure 25b, the luminances of half of the fragments has been altered. The perception of an illusory ring now depends on perceiving white fragments as connected to black fragments, and vice versa. The phenomenological perception of the "illusory-O" is eliminated. Figures 25c through 25f demonstrate that the relationship of the fragment luminances to the background is critical - simple luminance differences alone cannot account for the disappearance of the illusory-O. A subjective ratings experiment confirmed that observers perceive an illusory-O for Figure 25a but not 25b (He & Ooi, 1998a).

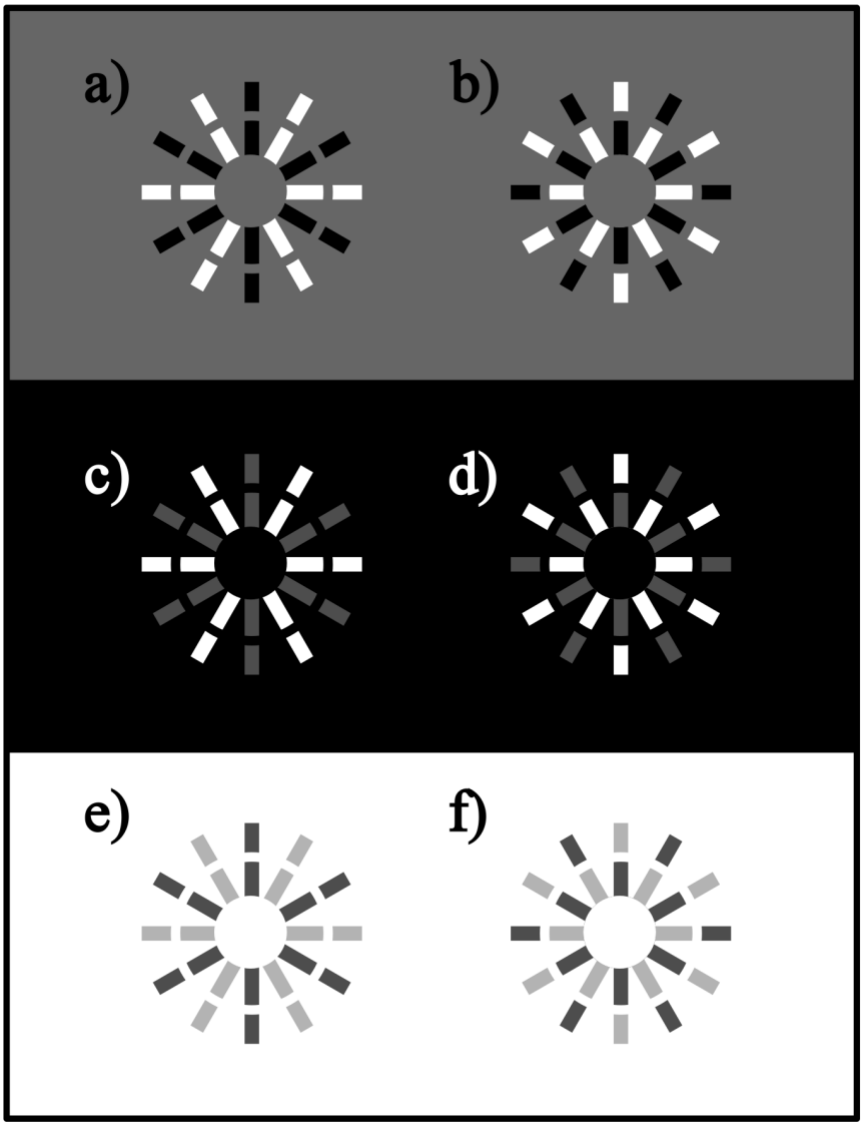

Figure 25. The illusory 'O' of He & Ooi (1998). In a), c), and e), interpolation occurs across gaps to connect fragments of identical luminances. An illusory 'O' is perceived to be occluding parts of the bars. In b), the to-be-connected fragments have opposite luminance contrast polarities with respect to the background. The percept of the illusory 'O' is eliminated. In d), and f), interpolation occurs across gaps to connect fragments that have different luminances, but common background contrast polarity. The percept of an illusory O is maintained.

In 2010(a), Su, He, & Ooi utilized a motion perception task to demonstrate the effect of luminance contrast polarity on the perception of amodally completed surfaces. An illustration of their stimulus is redrawn in Figure 26. The white bars were moved in the direction illustrated by the

arrows. Observers reported which of two percepts resulted: either two bars sliding under and over each other, or four white segments expanding and contracting. For the stimulus illustrated in Figure 26a, observers perceived the motion of two white bars sliding over and under one another within the black diamond. For the stimulus illustrated in Figure 26b, no motion was perceived in the region inside the diamond. Instead, the white bars were perceived to expand and contract.

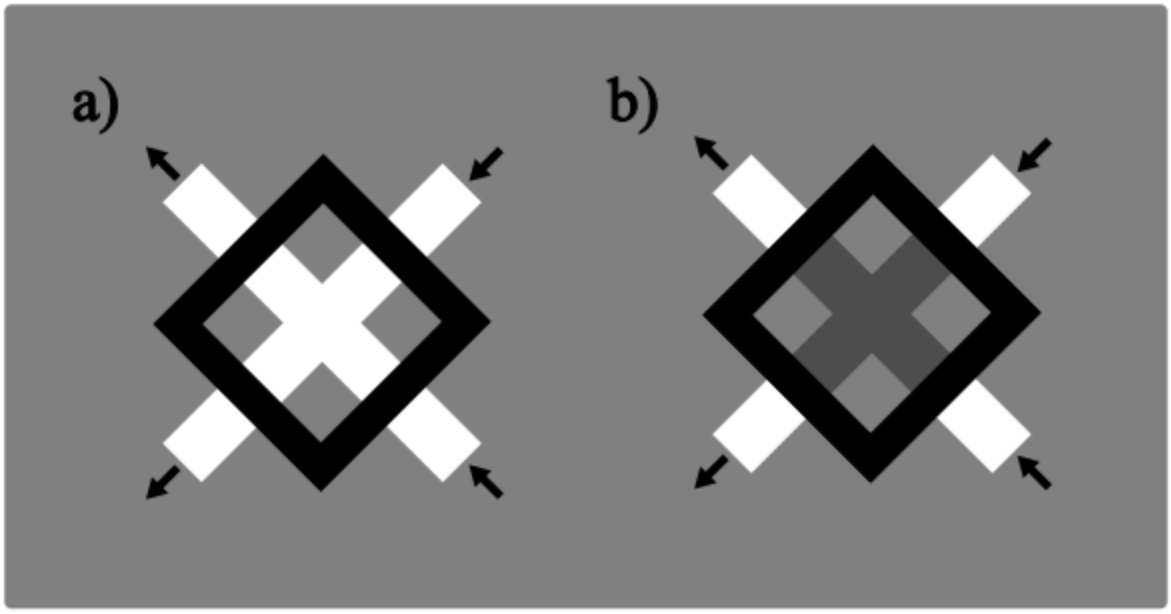

Figure 26. Stimulus from Su, He, & Ooi (2010a). For a), when the outer bar segments moved in the direction illustrated by the arrows, the inner bars (within the black diamond) were seen as also moving and sliding over and under each other. For b), however, no motion was perceived in the inner bars, and the outer bars were perceived to expand and contract.

Others have argued that differences in luminance contrast polarity across inducing elements in illusory figure stimuli do not affect the perception of illusory contours (Shapley & Gordon, 1983, 1985; Grossberg & Mingolla, 1985; Kellman & Cohen, 1984; Prazdny, 1983). It has also been demonstrated empirically that luminance contrast polarity differences across separate inducers do not affect the sensitivity to individual interpolated trajectories for illusory figures in a fat-thin task

(Spehar, 1999, 2000; Spehar & Clifford, 2003) or a similar portrait-landscape task (Victor & Conte, 2000), and that Kanizsa square inducers of opposite luminance contrast polarity facilitate the detection of a subthreshold line flashed upon the illusory contours that connect the inducing edges (Dresp & Grossberg, 1997).

It is important to note that there are two separate issues here. The first is the question of whether luminance contrast polarity differences affect the perception of interpolated contours, and in order to answer that question, we have to first differentiate the effects of inducer surfaces on surface spreading in modal versus amodal completion. As noted in the introduction to this chapter, modifying the inducers for illusory contour stimuli does not affect surface interpolation, as the interpolated surface, in the case of illusory figures, has the same color as the background - it does not share the surface qualities of the inducers. In the case of amodal figures, however, surface spreading depends on the surface qualities of the inducers, as the inducers are parts of the object being interpolated.

If the effect of surface spreading on stage 2 processing is the reason that the perception of interpolated contours is reduced or eliminated when luminance contrast polarity is manipulated for amodal figures, then manipulating the luminance contrast polarity of the inducers for modally completed (illusory) figures *should not* affect the perception of interpolated contours in those stimuli. This prediction, which comes naturally out of the two-stage theory, therefore accounts for some confusing results that otherwise would seem to contradict the work of He and colleagues (He & Ooi, 1998a, 1998b; He, Ooi, & Su, 2018; Su, He, & Ooi, 2010a, 2010b).

One may yet ask, if surface spreading only affects amodally completed surfaces, then why does it affect perception of interpolated contours in the path paradigm with traditional Gabors? Although there is no occluder in the path paradigm, there is also no region within the elements that shares the same color as the background. Surface spreading is prevented by these two facts. Experiment 2 has

already shown that changing the centers of the Gabors to a color that matches the background results in the perception of illusory contours and surfaces spreading across the gaps and through the Gabor centers. It is likewise true that, if occluders are added between the elements, then the surfaces of the elements will be enabled to spread beneath the occluders. Figure 27 illustrates the spreading of Gabor element surfaces beneath occluders that have been placed over the gaps in a traditional path.

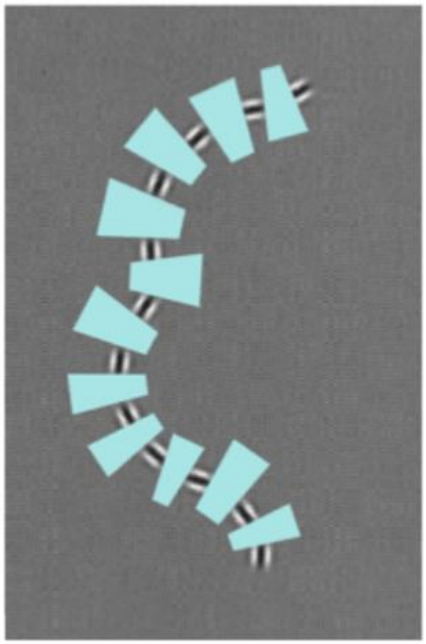

Figure 27. Surface spreading beneath occluders in a path composed of Gabor elements. If occluders are added between Gabor elements in a traditional path display, then the surfaces of the Gabors appear to continue beneath and spread from one element to another.

The second issue is the question of at what perceptual level surface spreading is hypothesized to affect the perception of interpolated contours according to the stage 2 theory. As mentioned above, it has been demonstrated empirically that luminance contrast polarity manipulations across separate inducers does not affect objective performance in a fat-thin task (Spehar 1998, 1999, 2000) or a similar portrait-landscape task (Victor & Conte, 2000). Regardless of the fact that surface spreading

is not affected by manipulations to inducer surfaces for illusory figure stimuli, the two-stage theory would not predict that judgments in either a fat-thin task or a portrait-landscape task would be affected by stage 2 output. The reason is this: both of those tasks merely rely on sensitivity to the length or shape of a single interpolated trajectory. This is entirely similar to the task used by Guttman & Kellman (2000) (who also used a fat-thin task) and the task used by Field, Hayes, & Hess (1993). The work of Guttman & Kellman (2000) and the experiments presented in the second chapter of this dissertation have already demonstrated that stage 1 output is sufficient to complete these kinds of tasks.

In order to target second stage processing, a task needs to target the whole, complete representations that are built up out of many interpolated and real edge tokens. The task used in the character recognition experiments presented in this dissertation requires integrating multiple interpolated edges with many physically specified edges to arrive at a complete description of an otherwise unapparent hidden alphanumeric character. For this reason, the character recognition task is expected to target second stage output. If the two-stage theory is correct, then, performance in the character recognition task should be affected by manipulations to luminance contrast polarity across inducers - even though performance for other kinds of objective tasks is unimpaired by the same manipulation. Experiment 6 tests this hypothesis.

3.10 Experiment 6: Luminance contrast polarity

Figure 28 illustrates two example stimuli from the control and experimental conditions for Experiment 6. The control condition is familiar, as it appeared also in Experiments 3 and 4. The experimental stimuli were manipulated such that half of the fragments were turned white. An effort was made to ensure that fragments connected by interpolations had opposite luminance contrast

polarities relative to the background. That is, if two fragments were connected by interpolations, it was ensured, whenever possible, that one was set to black while the other was set to white.

If luminance contrast polarity differences across surfaces are taken into consideration in the second stage, then interpolations connecting fragments with opposite luminance contrast polarity in the experimental condition should exit the second stage with relatively low perceptual strengths. Because complete representations of the hidden whole number or letter depend on these interpolations, the reduced perceptual strengths of the interpolations should be reflected in accuracy scores for the recognition of the hidden number or letter.

In addition to the condition manipulation, presentation time was manipulated using the same levels as in Experiment 3. Because interpolation takes around 150ms to proceed, differences were expected to appear somewhere between 100 and 200ms. For the same reasons as in Experiment 3, the difference between conditions was expected to increase as presentation time increases. In brief, longer presentation times allow more time to scan and search for the hidden alphanumeric character, but this additional time was expected to benefit performance in the control condition more than the experimental condition. The relatively weak strengths of the interpolations for the experimental condition should cause the hidden alphanumeric character to be less obvious, even when the observer is looking directly at the fragments that compose it. Because the hidden whole number or letter is not clearly visible, even when looking directly at it, additional time to scan or search the array in the experimental condition was not expected to improve performance as much as in the control condition.

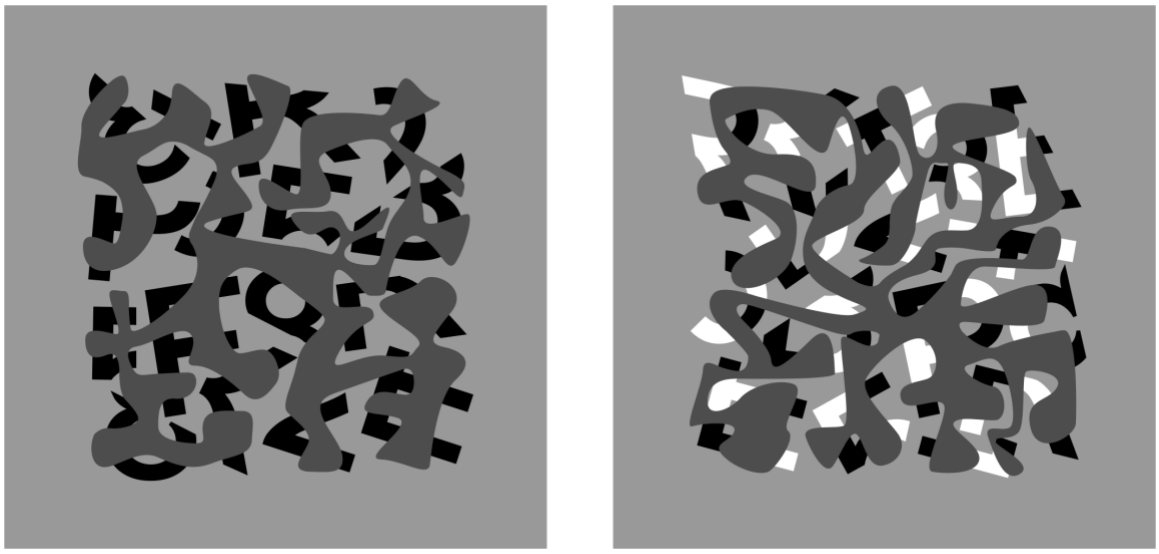

Figure 28. Example stimuli for Experiment 6. At left, a control stimulus is seen in which the letter 'F' is hidden in the lower left. At right, luminance contrast polarity is manipulated such that surface completion does not support the grouping of the fragments into a whole. The 'P' hidden in the lower right may be more difficult to find and recognize as a result.

3.10.1 Method

Participants. Twenty participants 17 female, ages 18-35, M_{age}: 21.10 years) completed the experiment. All participants provided informed consent, were students attending the University of California, Los Angeles, were offered course credit for their participation, and were naïve to the purposes of the study. Visual acuity was tested using a standardized ETDRS chart as a prerequisite to participation. All participants demonstrated 20/25 or better visual acuity.

Apparatus. The computer and test environment were the same as for Experiment 3.

Stimuli. Stimuli were identical to those used in Experiment 3, except that, in the experimental condition, half of the fragments were altered such that their fill and stroke were set to white. An effort was made to ensure that white and black fragments were distributed roughly equally across the stimulus. In addition, an effort was made to ensure that fragments connected by interpolations had

opposite luminance contrast polarity - that is, for two fragments connected by interpolations, whenever possible, one was set to black and the other was set to white.

The control condition utilized the same mask as in Experiment 3. The experimental condition utilized a mask that was identical except that half of the fragments were set to white.

Procedure. The procedure was identical to that of Experiment 3 except that the sample stimuli used in the instructions were modified in the manner specified above.

3.10.2 Results

Figure 29 plots the data from each condition across the range of presentation times used. There is a clear difference, emerging between 100 and 200ms, between the control and experimental conditions, and it appears to increase with increasing presentation times. To confirm these observations, the data were submitted to a 2 (conditions) x 8 (presentation times) repeated measures ANOVA. All statistics were Greenhouse-Geisser corrected. There was a main effect of stimulus type ($F(1, 16) = 102.46$, $MSE = .07$, $p < 0.001$, $\eta^2 = .06$, $\eta_p^2 = .87$), and presentation time ($F(2.40, 38.50) = 231.94$, $MSE = .02$, $p < 0.001$, $\eta^2 = .82$, $\eta_p^2 = .94$), and there was a significant interaction between stimulus type and presentation time ($F(5.30, 84.76) = 9.42$, $MSE = .01$, $p < 0.01$, $\eta^2 = .02$, $\eta_p^2 = .37$).

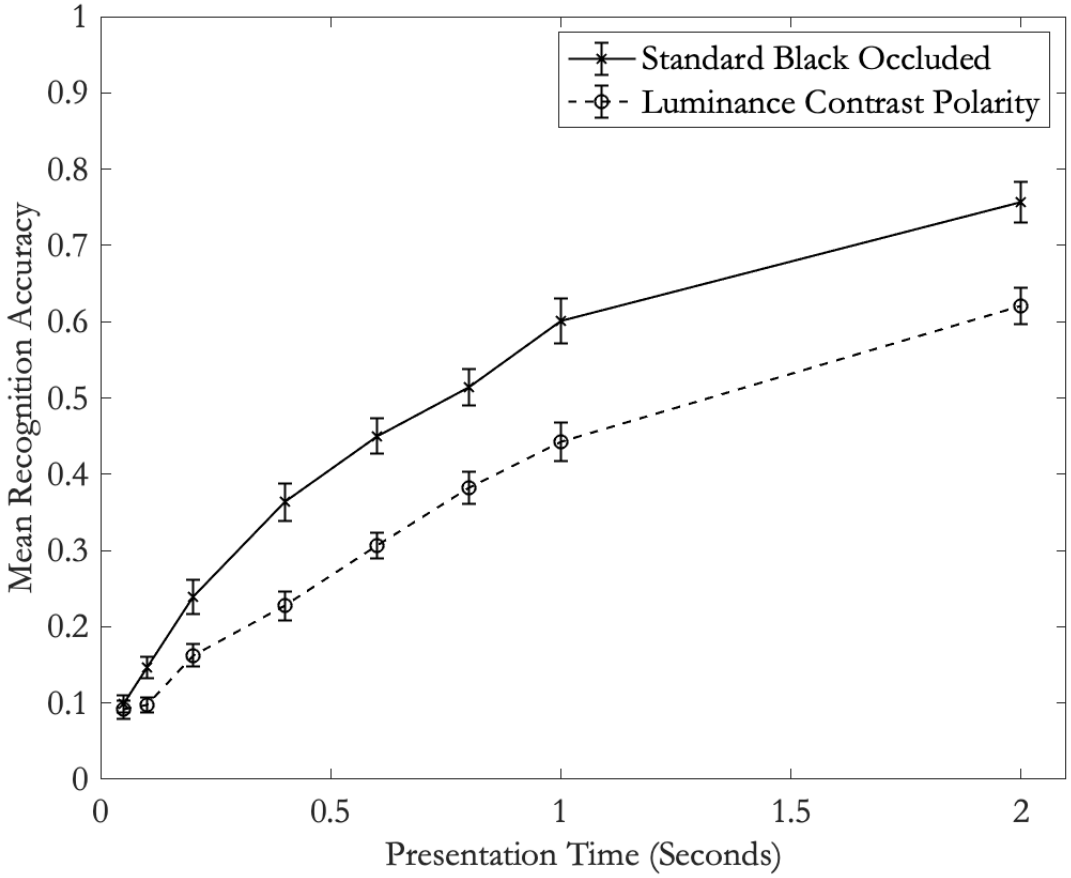

Figure 29. Results from Experiment 6. Performance in the control condition is illustrated by the solid line, while performance in the experimental condition is illustrated by the dotted line. Beginning between 100 and 200ms, a clear difference between the conditions emerges, with the luminance contrast polarity manipulation resulting in substantially worse recognition accuracy. The difference increases with larger presentation times. Error bars show 1 standard error above and below the condition mean.

3.10.3 Discussion

It was predicted that recognition accuracy in the control (filled black, occluded fragments) condition would be greater than recognition accuracy in the experimental condition (mix of filled black and white, occluded fragments), that these differences would appear somewhere around 200ms, and that the difference would increase with increasing presentation times. The results strongly support all of those predictions.

The luminance contrast polarity manipulation was expected to produce a drop in accuracy because it disrupts surface spreading across the fragments connected by interpolated edges. In the absence of potential surface spreading across fragments, each fragment tends to retain border ownership of its own edges. The opposed border ownership assignments of the fragments were expected to be taken into consideration in stage 2, such that the perceptual strengths for interpolations connecting those fragments would be reduced.

The resulting drop in recognition accuracy performance cannot be explained by constraints in stage 1 because all of the to-be-connected fragment edges in the display were relatable, and the physically specified edges in the two displays were identical. The results of Guttman & Kellman (2001), Spehar (1998, 1999, 2000), Victor & Conte (2000), and Experiments 1 and 2 of this dissertation suggest, therefore, that identical promiscuous interpolations were produced for both the control and experimental displays in stage 1. In addition, work from the contour integration literature suggests that disruptions to element phase (which has the effect of causing white edges to align with black edges and vice versa) do not have a strong impact on path detection performance (Field, Hayes, & Hess, 1993). This suggests that differences in the luminance contrast polarity of edge fragments is not a stage 1 constraint.

These results add to the work of He & colleagues (He & Ooi, 1998a, 1998b, He, Ooi, & Su, 2018, Su, He, & Ooi, 2010a, Su, He, & Ooi, 2010b) and Experiments 1 and 2 of this dissertation to indicate that surface spreading plays a role in determining the perceptual strengths of interpolations connecting amodally completed fragments in the second stage. In addition, these results suggest that luminance contrast polarity differences across connected surfaces have the ability to prevent surface spreading, and that this information is taken into consideration in stage 2.

Experiment 7 tests another surface cue - equiluminant color contrast - to determine what role, if any, it plays in second stage processing.

3.11 Equiluminant color contrast

Experiment 6 examined the role that differences in luminance contrast polarity play in the second stage. There is also good reason to believe that equiluminant color contrast has an impact on second stage processing. Su, He, & Ooi (2010b) have shown that the same illusory 'O' discussed in section 3.8 is also affected by equiluminant color contrast. Their compelling illustration is reprinted in Figure 30.

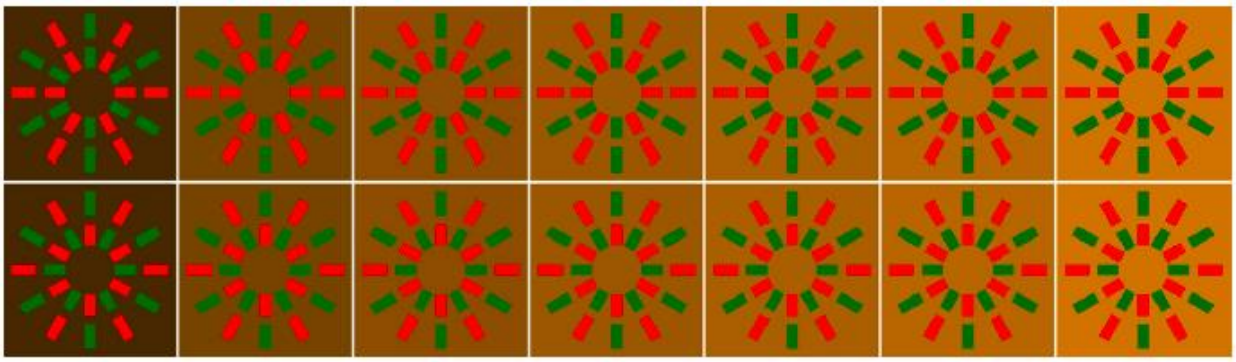

Figure 30. Figure from Su, He, & Ooi (2010b). The red and green rectangles were set to equiluminance for individual observers using heterochromatic flicker photometry (HFP). The luminance level of the yellow background increases across the displays shown from the left to the right. The red and green rectangles in the middle display may be roughly equiluminous with the background for some observers. The top and bottom rows of stimuli have the same and opposite color contrast, respectively.

In addition, as mentioned in section 3.8, there is some physiological basis for reasoning that color contrast, in addition to luminance contrast polarity, could potentially affect the perceived border ownership of fragments through cues to surface spreading. It has been demonstrated that some cells that encode border ownership in monkey visual cortex are also sensitive to both luminance and color contrast polarity (Zhou, Friedman, & von der Heydt, 2000).

In the following experiment, I will test whether equiluminant contrast plays a role in second stage processing by manipulating fragments in the same character recognition task utilized in Experiments 3-6.

3.12 Experiment 7: Equiluminant color contrast

Figure 31 illustrates two example stimuli from the control and experimental conditions for Experiment 7. The stimuli were identical to those used in Experiment 6 except that the white fragments in the experimental stimuli were made red while the black fragments were made an equiluminant green.

If color contrast plays a role in second stage processing, then interpolations connecting fragments with contrasting surface colors should receive relatively low perceptual strengths in the second stage. Because identifying and recognizing the hidden whole number or letter depends on a representation that combines multiple interpolations exiting stage 2, the reduced perceptual strengths of these interpolations should be reflected in reduced accuracy scores for the recognition of the hidden number or letter.

In addition to the stimulus manipulation, presentation time was manipulated using the same levels as in Experiment 3. Because interpolation takes around 150ms to proceed, differences were expected to appear somewhere between 100 and 200ms. In addition, for the same reasons as in Experiment 3, the difference between stimulus conditions was expected to increase as presentation time increased. This was because the relatively weak strengths of the interpolations for the experimental condition were expected to cause the hidden alphanumeric character to be less evident, even when the observer was looking directly at the fragments that made it up. The additional time to scan and search for a hidden alphanumeric character at longer presentation times was expected to be less beneficial when the hidden number or letter was less apparent even when looking directly at it.

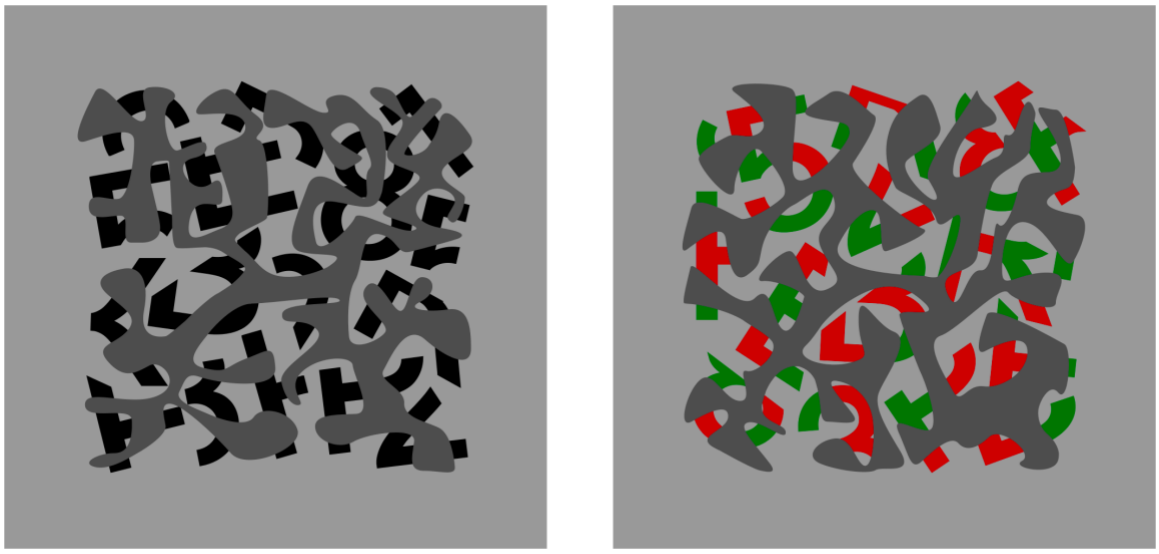

Figure 31. Example stimuli for Experiment 7. At left, a control stimulus is seen in which the number '8' is hidden in the upper right. At right, equiluminant color contrast is manipulated such that surface completion does not support the grouping of the fragments into a whole. The '5' hidden in the upper left may be more difficult to find and recognize as a result.

3.12.1 Method

Participants. Sixteen participants (13 female, ages 19-24, M_{age}: 21.19 years) completed the experiment. All participants provided informed consent, were students attending the University of California, Los Angeles, were offered course credit for their participation, and were naïve to the purposes of the study. Visual acuity was tested using a standardized ETDRS chart as a prerequisite to participation. All participants demonstrated 20/25 or better visual acuity.

Apparatus. The computer and test environment were the same as for Experiment 3.

Stimuli. Stimuli were identical to those used in Experiment 6, except that, in the experimental condition, the white and black fragments were replaced with a given participant's equiluminant red and green, respectively. Equiluminant colors were determined using the modified minimum motion

technique (which uses sinusoidally modulated gratings) described in Cavanagh, MacLeod, and Anstis (1987; for the original technique, see Anstis & Cavanagh, 1983). The masks used were identical to those used in Experiment 6 except that the fragments in the experimental condition mask were modified in the same way as specified above.

Procedure. The procedure was identical to that of Experiment 3 with two exceptions. The first exception is this: after the visual acuity test, participants completed a Dvorine Color Plate test to confirm normal color vision. This test allows a maximum of 7 errors to be made across 15 plates. All subjects read all 15 plates correctly, thus none were excluded from participation based on color blindness. The second exception is this: after completing the color vision test, participants then completed a minimum motion task as described by Cavanagh, MacLeod, and Anstis (1987; for the original technique, see Anstis & Cavanagh, 1983) to identify their equiluminant red and green. In the minimum motion task, observers saw a rotating "pinwheel" type stimulus, in which the peaks of one phosphor coincided with the troughs of the other. Observers perceived either clockwise or counterclockwise motion depending on the relative luminance of the two phosphors. The ratio of the two phosphor intensities was adjusted using the mouse. The isoluminant point was taken as the point where the motion direction was perceived to change and was replaced by the perception of counterphase flicker. The luminance of the red phosphor was adjusted relative to the luminance of the green. The "pinwheel" used spanned 8.27 degrees of visual angle in width and height (a width and height identical to that of the stimuli used in the character recognition task), and the width of the gratings was identical to the thickness (but not the length) of the fragments used (39 pixels, or 25 arcminutes at presentation).

After completing the minimum motion task, participants then began the character recognition experiment utilizing the same procedure as reported in Experiment 3. The isoluminant values

determined by the minimum motion task were utilized to adjust the pixel values of the stimulus fragments in Matlab.

3.12.2 Results

Figure 32 plots the data from each condition across the range of presentation times used. There appears to be a very small difference, emerging between 100 and 200ms, between the control and experimental conditions. The difference does not appear to increase with longer presentation times. To confirm these observations, the data were submitted to a 2 (conditions) x 8 (presentation times) repeated measures ANOVA. All statistics were Greenhouse-Geisser corrected. There was a main effect of stimulus type ($F(1, 15) = 20.00$, $MSE = .01$, $p < 0.001$, $\eta^2 = .004$, $\eta_p^2 = .57$), and presentation time ($F(2.24, 33.60) = 269.88$, $MSE = .02$, $p < 0.001$, $\eta^2 = .92$, $\eta_p^2 = .95$), but no interaction between stimulus type and presentation time ($F(5.12, 76.73) = 1.93$, $MSE = .01$, $p = .10$, $\eta^2 = .002$, $\eta_p^2 = .11$).

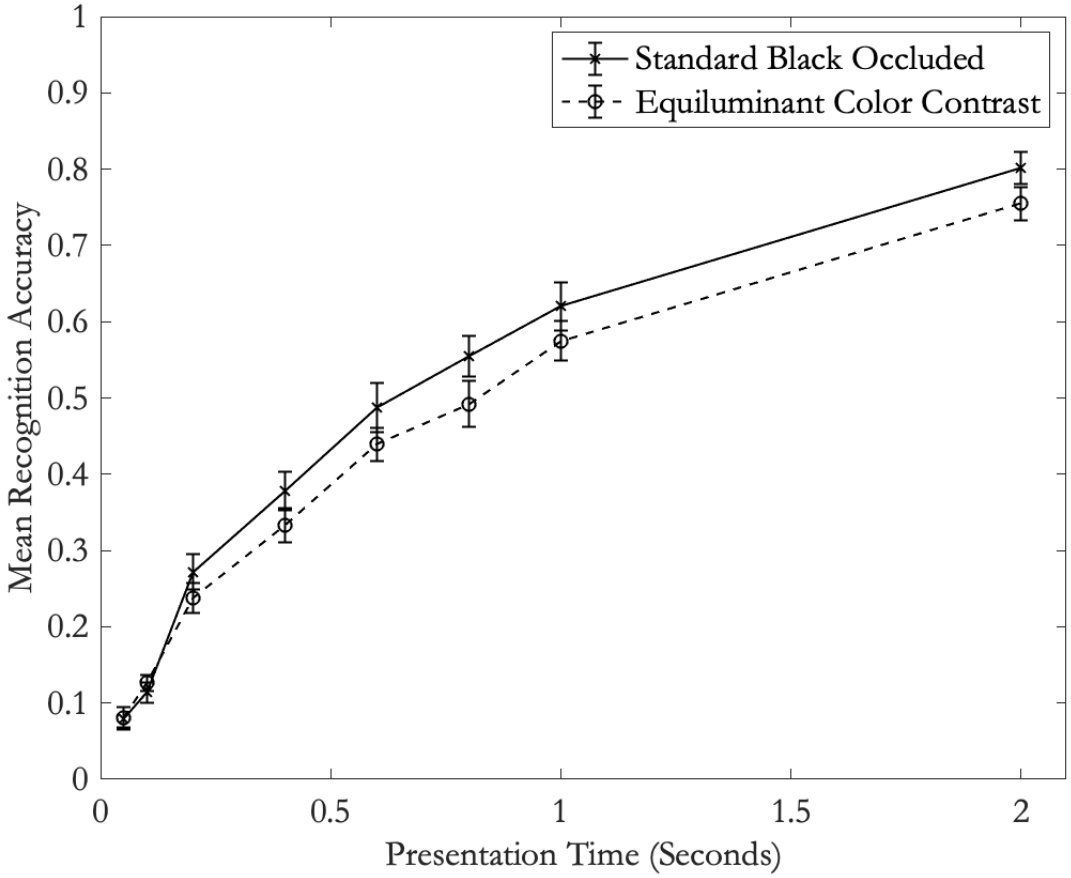

Figure 32. Results from Experiment 7. Performance in the control (black, partially occluded fragments) condition is illustrated by the solid line, while performance in the experimental condition (equiluminant red and green, partially occluded fragments) is illustrated by the dotted line. A very small difference between the conditions emerges at 200ms, with the equiluminant color contrast manipulation resulting in worse recognition accuracy. Beyond 200ms, the size of the difference does not change as presentation time increases. Error bars show 1 standard error above and below the condition mean.

3.12.3 Discussion

It was predicted that recognition accuracy in the control (filled, black partially occluded fragments) condition would be greater than recognition accuracy in the experimental condition (filled, equiluminant red and green partially occluded fragments), that these differences would appear somewhere around 200ms, and that the difference would increase with increasing presentation

times. Only two of these predictions are borne out in the data. There is a difference in recognition accuracy between the two stimulus types, and the difference does appear at 200ms, however, it is quite small (η_2 = .004). Moreover, beyond 200ms, the size of the difference does not change with increasing presentation times.

The small magnitude of the observed difference and its constancy across presentation times is worth noting because it is possible that equiluminance was not perfectly achieved in the displays, and because earlier work suggests that small luminance differences across fragments may produce small decreases in the perceptual strengths of interpolations connecting those fragments (He & Ooi, 2010). It is notoriously difficult to achieve perfectly equiluminant displays for a variety of reasons. Among the problems are the facts that it is very difficult to eliminate luminance artifacts from displays presented on CRT monitors (García-Pérez & Peli, 2000), that equiluminant color matches vary across the retina (Livingstone & Hubel, 1987; Mullen, 1991; Bilodeau & Faubert, 1997), that luminance matches do not hold across different spatial frequencies (Anstis & Cavanagh, 1983), and finally, that chromatic aberration in the eye produces luminance artifacts at any sharp color transitions (Powell, 1981). An effort was made to ensure that the spatial frequencies of the "pinwheel" used in the minimum motion experiment were a good match to the stimuli used in the display (see method), however, the varying sizes of the fragments makes it impossible to ensure a perfect match. In addition, because the following are true: a CRT monitor was used to present our displays, our displays span 8.27 degrees of visual angle, and an achromatizing lens was not available to be used in the study, it is difficult to interpret the fairly small, though reliable, difference observed between these two conditions.

It is possible that the effect is simply a relatively small one. However, in light of the possibility that equiluminance was not perfectly achieved, the results should be interpreted with caution. The

results suggest, at least, that equiluminant color contrast may play a role in determining the perceptual strengths of interpolated contours in the second stage.

3.13 Interpolation across multiple spatial scales

Another possibility, about which relatively little is known, is that interpolation may proceed across multiple spatial scales. The suggestion that visual processing involves the extraction of information in multiple spatial frequency channels is well-established (Campbell & Robson, 1968; Marr & Hildreth, 1980; De Valois & De Valois, 1980; Graham, 1980; Ginsburg, 1984; Morrone & Burr, 1988; Graham, 1989; De Valois & De Valois, 1990; Carandini et al, 2005). The existence of multiple spatial frequency channels is supported by a wealth of psychophysical (Campbell & Robson, 1968; Blakemore & Campbell, 1969; Sachs, Nachmias, & Robson, 1971; Stromeyer & Julesz, 1972; Graham, 1980; Ginsburg, 1986; Julesz & Schumer, 1981) and physiological evidence (Poggio, 1972; De Valois, De Valois, & Yund, 1979; De Valois, Albrecht, & Thorell, 1982; Van Essen & Deyoe, 1995; Bullier, 2001), and is reflected in current influential models of visual perception (Schyns & Oliva, 1994; Bullier, 2001; Bar, 2003; Hegde, 2008).

An interesting question is whether interpolation occurs relatively early in visual processing, proceeding separately in different spatial frequency channels, or, alternatively, proceeds relatively late, following the achievement of symbolic edge tokens based on reconciling the outputs of multiple channels. By "symbolic", I mean to indicate a representation of an edge that is an abstraction from the point-by-point contrast differences in the image that gave rise to it; it is therefore a flexible representation of the relevant structure that discards much irrelevant information (see Kellman, Garrigan, & Erlikhman, 2013 for a discussion on bridging the gap between subsymbolic and symbolic encoding).

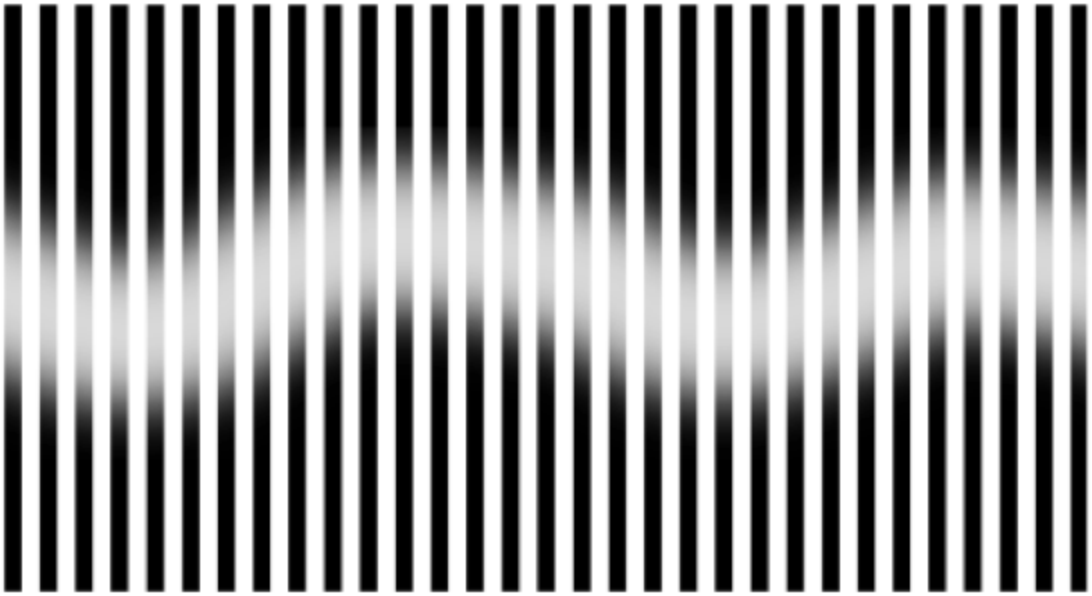

Figure 33. Interpolation across low spatial scales. The edge inputs that induce the illusory contour seen in the figure only exist at relatively low spatial scales.

Relatively few studies have examined multiple spatial frequency channels in the context of interpolation processes. However, some work may indicate that interpolation proceeds across multiple spatial frequency channels (Guttman, Sekuler, & Kellman, 2003; Guttman & Kellman, 2004), and that interpolations existing in relatively few channels appear in the final scene representation with relatively low perceptual strengths (Guttman, Sekuler, & Kellman, 2003). Figure 33 above illustrates an interpolated edge that could only be produced based on edge inputs at a relatively low spatial scale.

In addition, some other work suggests that interpolation in stage 1 proceeds from non-symbolic inputs - that is, the inputs are *not* symbolic, abstracted edge tokens, but rather are contrast differences of a particular spatial frequency (Dakin & Hess, 1998; Dakin & Hess, 1999). In their Figure 1, Dakin & Hess (1998) propose two variants of the association field model: one that takes as input edge tokens that are abstracted from filter outputs across multiple spatial frequency channels, and another in which contour integration occurs separately in different channels. In the second

model, symbolic edge tokens are not achieved until after contour integration has occurred. Utilizing a path paradigm task, they test which of these models is correct by utilizing paths composed of alternating broadband and narrowband elements. They found that observers were somewhat tolerant to spatial frequency differences among element types, however, this was most true for collinear paths - an increasing amount of spatial frequency tuning was evident for paths of higher curvature. The authors concluded that the presence of this tuning makes clear that the first model - in which contour integration proceeds from symbolic edge tokens - cannot be correct. They report that, "the contour-extraction process does not simply extract the local orientation of elements of bandpass path stimuli and operate on some abstracted orientation field". Instead, they concluded, "the visual contour integration system is tuned for the spatial-frequency content of the contour and specifically narrows with increasing contour curvature" (p.1496).

If edges existing in different spatial frequency channels are not integrated prior to stage 1 processing, it may well be the case that information about how many channels an interpolated edge appeared in is available to be utilized in stage 2 to adjust interpolation weights based on the amount of corroboration a given interpolation has across multiple spatial scales. The idea that the visual system utilizes corroboration of edge information across spatial scales is far from new. In his classic book, Marr wrote that, if an edge exists in multiple channels, with the same position and orientation in each channel, it can be taken as indicating that the detected contrast results from the same physical phenomenon (1982). In addition, edges that exist only in relatively low spatial frequency channels may be taken to indicate shadows, while edges existing only in relatively high spatial frequency channels may be taken to indicate scratches or other spurious markings upon a surface.

It seems possible then, that, if interpolation occurs in multiple different channels, the visual system may take advantage of information about the channels an interpolation appears in to determine the final perceptual weights of the interpolated edge. If so, corroboration of interpolated

edges across many channels could have the effect of maintaining or increasing the perceived strength of the interpolated edge. Likewise, interpolations that do not appear in as many channels may have their strengths maintained or reduced. In order to examine this possibility, Experiment 8 will utilize displays that are designed to enable interpolation only within relatively low spatial frequency channels.

If the conclusions of Dakin & Hess (1998, 1999) were wrong, and interpolation does proceed from symbolic edge tokens, then the manipulation described in the method below (which utilizes alternating high and low spatial frequency edge fragments) should not have any effect on recognition accuracy performance. If an effect is observed, it would support the idea that interpolation occurs separately in different spatial frequency channels, and that interpolations existing in relatively few channels receive lower perceptual strengths in stage 2.

3.14 Experiment 8: Interpolation in multiple spatial frequency channels

Experiment 8 was designed to test the possibility that interpolation proceeds across multiple spatial frequency channels, and that the second stage takes into consideration the number of channels that an interpolated contour appears in. Example control and experimental stimuli are illustrated in Figure 34.

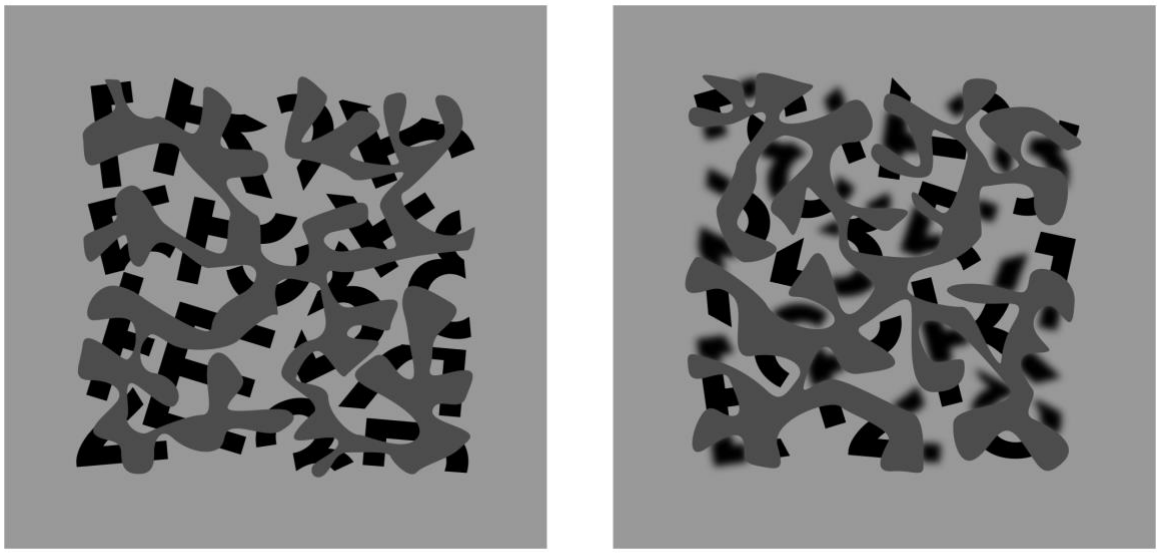

Figure 34. Example stimuli from Experiment 8. At left, a control stimulus is seen in which the letter 'E' is hidden in the lower left. At right, the spatial frequency of some fragments is manipulated such that interpolations across fragments should exist only in relatively low spatial frequency channels. The 'K' hidden in the lower right may be more difficult to find and recognize as a result.

3.14.1 Method

Participants. Twenty-one University of California, Los Angeles undergraduate students participated in the experiment. Two participants, whose performance was at chance at all presentation times for both stimulus conditions, were eliminated from the analysis. Thus, nineteen participants (17 female, ages 18-22, M_{age}: 20.37 years) were included in the analysis. All participants provided informed consent, were offered partial course credit for their participation, and were naïve to the purposes of the study. Visual acuity was tested using a standardized ETDRS chart as a prerequisite to participation. All participants demonstrated 20/25 or better visual acuity.

Apparatus. The computer and test environment were the same as for Experiment 3.

Stimuli. Stimuli were identical to those used in Experiment 6, except that, in the manipulation condition, instead of turning half of the fragments white, half of the fragments were manipulated by applying a Gaussian blur with a standard deviation of 2 pixels (1.32 arcminutes at presentation) to the fragments. Gaussian blur results in low-pass filtering, preferentially removing or attenuating high spatial frequencies (e.g., see https://en.wikipedia.org/wiki/Gaussian_blur). The boundaries of the blur were not limited by any window, and technically encompassed the whole of the stimulus image. However, the standard deviation of the Gaussian was small enough that visual inspection of the displays reveals the visible boundaries of the blur to be limited to an area much more immediate to the original, unblurred fragment edges.

The control condition utilized the same mask as Experiment 3. The experimental condition utilized a mask that was modified in the same way as described above: the same Gaussian blur was applied to half of the fragments in the mask.

Procedure. The procedure was identical to that of Experiment 3.

3.14.2 Results

Figure 35 plots the data from each condition across the range of presentation times used. There is a clear difference, emerging between 100 and 200ms, between the control and experimental conditions. The difference increases with increasing presentation times. To confirm these observations, the data were submitted to a 2 (conditions) x 8 (presentation times) repeated measures ANOVA. All statistics were Greenhouse-Geisser corrected. There was a main effect of stimulus type ($F(1, 15) = 69.98$, $MSE = .01$, $p < 0.001$, $\eta^2 = .03$, $\eta_p^2 = .82$), and presentation time ($F(2.76, 41.36) = 205.38$, $MSE = .02$, $p < 0.001$, $\eta^2 = .87$, $\eta_p^2 = .93$), and there was a significant interaction between stimulus type and presentation time ($F(4.22, 63.32) = 6.79$, $MSE = .01$, $p < 0.001$, $\eta^2 = .01$, $\eta_p^2 = .31$).

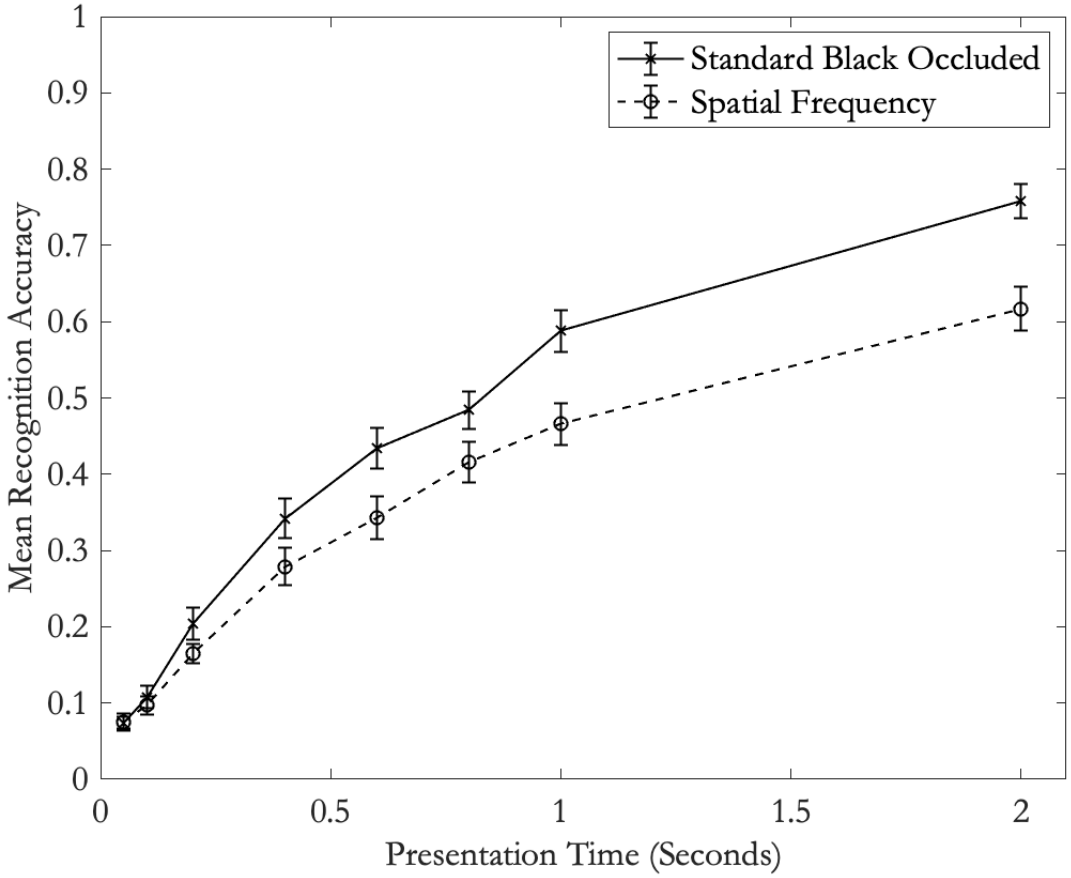

Figure 35. Results from Experiment 8. Performance in the control condition is illustrated by the solid line, while performance in the experimental (spatial frequency manipulation) condition is illustrated by the dotted line. Beginning between 100 and 200ms, a clear difference between the conditions emerges, with the spatial frequency manipulation resulting in substantially worse recognition accuracy. The difference increases with larger presentation times. Error bars show 1 standard error above and below the condition mean.

3.14.3 Discussion

It was predicted that recognition accuracy in the experimental (spatial frequency manipulation) condition would be worse than recognition accuracy in the control (filled occluded) condition, that this difference would appear somewhere between 100 and 200ms, and that the difference would increase with increasing presentation times. All of those predictions are supported by the results.

Introducing spatial frequency differences across fragments was expected to reduce recognition accuracy performance because the manipulation should cause interpolations across fragments to exist only in relatively low spatial frequency channels. In addition, for one of the edge pairs, there is a matching edge in a high spatial frequency channel. Because interpolations in different channels will at some point have to be reconciled, it was hypothesized that corroboration across multiple channels is taken into consideration in stage two to determine the ultimate perceptual strength of an interpolation. For the experimental stimuli of this study, the strength of interpolations would be expected to be reduced in the second stage because 1) the interpolation only appears in relatively low spatial frequency channels and, 2) one of the connected edges appears in relatively high spatial frequency channels but does not appear in relatively low ones. This mismatch of spatial frequency information may be taken as evidence in stage 2 that the two edges in the pair do not really belong to the same object. The data revealed that, indeed, recognition accuracy performance was much worse for the manipulated stimuli. This suggests that information about the spatial frequency channels an interpolated edge appears is does come into play in stage 2.

This finding is remarkable for three reasons: 1) it adds to existing evidence to suggest that interpolation does indeed proceed separately across multiple spatial frequency channels (Guttman, Sekuler, & Kellman, 2003; Guttman & Kellman, 2004), 2) it adds to existing evidence that suggests that promiscuous interpolation proceeds from nonsymbolic edge inputs in separate spatial frequency channels (Dakin & Hess, 1998), and 3) it suggests that the final perceptual strength of interpolations is partially based on the number of channels that an interpolation exists in.

3.15 General discussion

Experiments 3 through 8 were designed to examine the impacts of various cues hypothesized to play a role in determining the final perceptual strengths of interpolated contours in the second stage. Experiments 3 through 5 examined border ownership cues, while Experiments 6 and 7 examined surface qualities that prevent amodal surface spreading (i.e. luminance contrast polarity and equiluminant color contrast, respectively). Finally, Experiment 8 examined the possibility that interpolation proceeds across multiple spatial frequency channels, that these interpolations in different channels are reconciled in the second stage, and that the number of channels an interpolation exists in plays a role in determining that interpolation's final perceptual strength. In separate experiments, each of these cues was modified, and the effects on the recognition of an alphanumeric character, composed of multiple fragments and hidden within a background of random alphanumeric character fragments, was observed.

Because the experiments in Chapter 2 demonstrated that stage 1 output is sufficient to complete some tasks - like the path paradigm (Experiments 3 & 4), the "fat-thin" classification task (Guttman & Kellman, 2001), classification imaging tasks (Gold et al., 2000; Keane, Lu & Kellman, 2007) and a portrait-landscape classification task (Victor & Conte, 2000) - that rely on simple sensitivity to the shape or length of an individual interpolated contour, the character recognition task was designed to target second stage output. It is the final versions of interpolated contours exiting stage 2 that should be sent forward to object recognition processes to be utilized in the building up of object descriptions. For this reason, tasks that require integrating many interpolated and real edges to build a complete, recognizable object shape from several fragments are expected to rely on stage 2 output. The character recognition task used in Experiments 3-8 requires combining many interpolated edges and physically specified edges to arrive at a complete representation of a whole alphanumeric character that, without the perception of wholeness, does not stand out from its random background.

The results reveal that recognition accuracy performance in the character recognition task is strongly affected by manipulations to border ownership cues, amodal surface spreading, and spatial frequency differences across connected fragments. The results strongly support the two stage theory and suggest that both border ownership cues and amodal surface spreading are taken into consideration to determine the perceptual strengths of interpolations in the second stage. In addition, the results support the conclusion that interpolation proceeds across multiple spatial frequency channels, and that the degree of corroboration across multiple channels affects a given interpolation's final perceptual strength. The experiments in Chapter 3 confirm that, among the border ownership cues that play a role in the second stage are: having no physically specified occluder (see Experiment 4); having no physically specified occluder when an illusory occluder is impossible (see Experiment 3); and having outline, rather than filled, fragments (see Experiment 5). In addition, the prevention of amodal surface spreading as a result of manipulations to the luminance contrast polarity of fragments relative to the background can be concluded to play a definite role in the second stage (see Experiment 6). This result concurs well with the results of Geisler & Perry (2009), who examined statistics of contours in natural scenes and found that edge fragments of opposite contrast polarity are rarely connected in the world. In addition, observers who reported whether they perceived isolated edges leading into an occluder to be connected agreed with the scene statistics - their ratings indicated that edges of opposite luminance contrast polarity were not perceived to be connected unless they were collinear or nearly collinear. The data from Experiment 7 suggest that, in addition to luminance contrast polarity, equiluminant color contrast also plays a role. However, because it is very difficult to eliminate all luminance information from displays (see the discussion for Experiment 7 for an explanation), and in light of evidence from other work that suggests that small luminance differences across fragments can produce small

effects (He & Ooi, 2010), the small effect of equiluminant color contrast in Experiment 7 should be interpreted with caution.

These results add to the work of Chapter 2, which revealed the early, *promiscuous* stage of the interpolation process - in this stage, interpolated contours are produced promiscuously across all pairs of relatable edge fragments. These interpolated contours reside in an intermediate representation - whether they will ultimately appear in the final representation of a scene depends on processing that occurs in a later, second stage. Effects of these early interpolated contours are visible in the grouping and perceptual "pop-out" of paths in the path detection tasks of Field, Hayes, & Hess (1993), and in the sensitivity to interpolated trajectories in a fat-thin classification task that utilizes outline fragments that do not support the perception of illusory contours for Kanizsa square stimuli (Guttman & Kellman, 2001), and are also apparent in the results from classification imaging studies (Gold, Murray, Bennett, & Sekuler, 2000; Keane, Lu, & Kellman, 2007) . Experiment 1 confirmed that, in addition to the monotonicity and 90 degree geometric constraints of *relatability*, which governs when two fragments can be connected by interpolated contours (Kellman & Shipley, 1991; Kellman, Garrigan & Shipley, 2005), path detection is also constrained by the same 15 arcminute misalignment tolerance that governs contour interpolation (Hilger, 2009). In addition, Experiment 2 revealed that, when you alter the centers of Gabors in a traditional path to match the color of the background, interpolated contours are perceived to connect path elements. Experiment 1 demonstrated that the perception of interpolated contours in path displays does not affect objective path performance. For both types of displays - those with modified and traditional Gabors - performance degrades in an identical fashion with increasing misalignment, reaching a plateau at 15 arcminutes - the known tolerance for misalignment for contour interpolation. In addition, subjective rankings of interpolated contour clarity for modified path detection stimuli degrade in an identical fashion with increasing misalignment, reaching the same plateau at 15 arcminutes of misalignment.

All of this work strongly supports the two-stage theory by revealing the effects of interpolated contours produced promiscuously in the first stage. The experiments presented in this chapter reveal the outputs of the second stage - which takes into consideration a variety of different cues to determine the ultimate perceptual weights of interpolated contours produced in the first stage.

3.15.1 Other second-stage influences: Symmetry and familiarity

There are some other cues that were not examined in this chapter, but which are very much worth mentioning. Many studies have shown effects of global influences, such as symmetry (Sekuler, Palmer, & Flynn, 1994; van Lier, van der Helm, & Leeuwenberg, 1995; van Lier & Wagemans, 1999) on the perception of interpolated contours, though other work has made clear that reflectional symmetry (Carrigan, Palmer, & Kellman, 2016a), translational symmetry (Carrigan & Kellman, 2017), and familiarity (Carrigan & Kellman, 2016), in absence of relatable edges, are not sufficient to give rise to percepts of precise, interpolated edges. In the absence of relatable edge fragments, observers may rely on symmetry or familiarity information to arrive at an idea of what a whole object looks like - but the process is a higher-level, more cognitive one that involves conscious judgments based on the recognition of parts (*recognition from partial information*). The outputs of this sort of relatively high level process can, and have been differentiated from contour interpolation, which is a relatively low-level process that produces precise interpolated trajectories with specific, spatially localized positions in space (Carrigan, Palmer, & Kellman, 2016).

Most of the work in amodal completion that has suggested a role for symmetry and/or familiarity is explainable by RPI, a process distinct from contour interpolation. That said, both symmetry and familiarity may influence perception in a variety of ways. It is possible that, apart from RPI, these global shape influences could have a strengthening effect on actual contour interpolation in stage 2. We currently do not have any evidence for this idea, but it bears further investigation.

Figure 11, from the discussion of Chapter 2, illustrated some cases where competing interpolations may be resolved in the second stage based on either closure, familiarity, or symmetry. It may also be the case that an interpolated contour that is also symmetrical with other interpolated or real edges in the scene receives a larger perceptual weight exiting stage 2. Likewise, there may be cases where interpolated contours receive greater weights when the interpolations they produce result in a familiar shape. These possibilities could certainly be examined in future work.

3.15.2 Other motivations for a second stage: Reconciling interpolations produced in separate spatial frequency channels

This chapter introduced yet another reason why the interpolation process might include a second stage. That is, it is widely believed that early visual processing involves the extraction of edge information across multiple spatial frequency channels (see section 3.10), and some work has suggested that interpolation, likewise, proceeds separately within different channels (Guttman, Sekuler, & Kellman, 2003; Guttman & Kellman, 2004). In addition, other work suggests that the inputs to stage 1 are nonsymbolic contrast differences existing in separate spatial frequency channels (Dakin & Hess, 1998, 1999). If interpolation proceeds separately in different spatial frequency channels, then it will inevitably be necessary, at a later stage, to reconcile the interpolations existing in different channels. Because earlier work has suggested that the visual system may utilize information about the corroboration of edges existing in different channels (Marr, 1982), it was hypothesized that information about the channels an interpolated edge exists in might be utilized to determine the final perceptual strengths of interpolated edges in stage 2. The results of Experiment 8 combine with earlier work (Guttman, Sekuler, & Kellman, 2003) to suggest that the number of channels an interpolation appears in plays a role in determining its final perceptual strength in stage 2.

This finding is remarkable because it helps to shed some light on the answer to a very interesting question about when, and from what inputs, interpolation proceeds. That is, does interpolation occur relatively early in visual processing, proceeding separately from subsymbolic input in different spatial frequency channels? Or, alternatively, does interpolation proceed relatively late, from inputs that consist of symbolic representations of edges that are an abstraction from the point-by-point contrast differences across many spatial frequency channels? If interpolation proceeds from symbolic edge inputs, it might be expected not to matter which or how many channels originally extracted the inducing edges. Experiment 8 compared performance for two display types: control displays for which all fragments had crisp edges (the inducing edges could therefore be picked up at both relatively high and relatively low spatial frequencies), and experimental displays for which half of the fragments had blurry edges (half of the inducing edges could therefore only be picked up at relatively low spatial frequencies). The results revealed that recognition accuracy (detecting and recognizing a hidden alphanumeric character composed of multiple fragments and interpolations across fragments) was substantially worse for the experimental displays. That is, altering the spatial frequency of half of the fragments, such that interpolations had to connect crisp fragments to blurry fragments, produced a marked decrease in recognition accuracy. The results therefore combine with earlier work (Dakin & Hess, 1998, 1999) to provide some support for the idea that interpolation occurs relatively early and proceeds from information contained in separate spatial frequency channels.

3.15.3 Concluding remarks

The results of Experiments 3 through 8 provide strong support for the multi-stage theory. They also reveal that border ownership cues, amodal surface spreading cues, and information about the

number of spatial frequency channels an interpolation was produced in all contribute to determining an interpolation's final perceptual strength.

The first stage of the multi-stage theory is consistent with a wealth of empirical work that demonstrates the importance of relatability, and it explains the dependence of some tasks, for which interpolated contours are not perceived, on relatability geometry. The second stage of the multi-stage theory explains the effects of some manipulations on the subjective appearance of interpolated contours. The second stage also makes clear how global cues, such as closure, symmetry, and familiarity, can play a role in determining the subjective appearance of individual interpolated contours. In addition, the existence of a second stage is consistent with earlier work that has demonstrated that interpolation proceeds first, and that afterward, the appearance of an interpolated contour as modal or amodal is determined. The existence of a second stage is also consistent with the idea that interpolation proceeds separately across multiple spatial frequency channels.

In summary, the multi-stage theory provides a much needed framework for understanding a variety of findings that simply cannot be accounted for with any single stage model. The evidence suggests that interpolation proceeds first, according to relatability geometry, across separate spatial frequency channels, such that interpolations are produced promiscuously across all pairs of relatable edges, and that afterward, a variety of constraints - which take into consideration border ownership cues, surface spreading, and the number of channels an interpolated was produced in, among other things - are resolved to determine an interpolated contour's final appearance as modal or amodal and also its ultimate perceptual strength. For some tasks, the output of the first, promiscuous interpolation stage is sufficient to complete the task. Others, which are sensitive to the subjective appearance of contours or which require a complete object shape description, rely on the outputs of the second stage.

In the next chapter, some of the cues examined in this chapter will be combined in a series of experiments in order to determine the magnitudes of their combined effects. A model for cue combination in the second stage will be discussed.

CHAPTER 4

Modeling cue combination in the second stage

4.1 Synopsis

In the introduction to this dissertation, and in the discussions of Chapters 2 and 3, evidence that supports a two-stage theory of contour interpolation was presented. In the first stage, interpolations are produced promiscuously across all relatable edge fragments. In the second stage, a variety of scene cues are taken into consideration to determine an interpolated contour's appearance as modal or amodal and its final perceptual strength. Experiments 1 and 2 revealed the intermediate outputs of the first stage of the interpolation process using the path detection paradigm of Field, Hayes, & Hess (1993) and a subjective rating experiment. Experiments 3 through 8 revealed some scene cues that play a role in the second stage. A variety of border ownership cues, amodal surface spreading cues, and spatial frequency differences across inducing edges were all shown to play a role in the second stage. Here, manipulations of one border ownership cue (the deletion of the occluder), one amodal surface spreading cue (luminance contrast polarity differences across inducing fragments), and spatial frequency differences across inducing fragments, will be combined in a set of experiments to measure the magnitude of their combined effects. The results of the experiments presented in Chapter 3 will be used to generate predictions for the combined effects based on an additive model of cue combination in the second stage. The results of the combined cue experiments presented in this chapter will be compared to those predictions. The results suggest that the data are well described by an additive model with the exception that certain combinations of cues result in reductions to perceptual strength that go beyond additivity. The implications are that, at least for the different categories of cues examined here, the second stage integrates information in

a simple, additive manner, but for certain combination of cues, there is an interaction such that interpolated strengths are reduced more than additivity would predict. A theory as to what is special about these combinations will be discussed. The paradigm used is identical to that used in Chapter 3. Stimuli consist of a background of random alphanumeric character fragments within which is hidden a subset of fragments that can be connected to form a whole letter or number. Interpolation is necessary to connect the fragments, and recognition accuracy is worse when the necessary interpolations exit the second stage with relatively weak strengths. Scene cues and presentation time are manipulated, and recognition accuracy is measured.

Keywords: contour interpolation, object formation, promiscuous interpolation, amodal completion, modal completion, border ownership, amodal surface spreading, luminance contrast polarity, spatial frequency channels

In the introduction to this dissertation, I proposed a two-stage theory of the contour interpolation process. The second chapter of this dissertation revealed the intermediate outputs of the first stage of that process. In the first stage, contours are interpolated indiscriminately across all pairs of relatable edge fragments. These *promiscuously* interpolated contours exist in an intermediate representation - whether they appear in the final scene representation depends on processing in a subsequent stage. In the second stage of this process, a variety of scene cues are taken into consideration to determine the final perceptual strengths of interpolated contours along with their modal or amodal appearance. The output of this second stage was revealed by a series of experiments presented in Chapter 3. The results of those experiments reveal that border ownership cues, amodal surface spreading cues (including background luminance contrast polarity and equiluminant color contrast), and spatial frequency differences across fragments all play a role in determining the final perceptual strengths of interpolated contours.

The experiments in Chapter 3 also revealed the magnitudes of the individual effects of the border ownership, amodal surface spreading, and spatial frequency cues that were manipulated. These magnitudes reveal the amount of impact that an individual cue has on the perceptual strength of interpolated contours in the second stage. This final chapter has two main goals. First is to determine the combined magnitudes of some cues when they are manipulated in the same image. This will be done through a series of character recognition experiments (the same paradigm used in the experiments presented in Chapter 3). Second is to arrive at a model of cue combination in the second stage based on these empirical results.

4.2 Combining cues

Chapter 3 demonstrated that multiple scene cues are utilized in determining the final perceptual strengths of interpolated contours in the second stage. However, in those experiments, each cue was presented in an isolated fashion - two cues were never manipulated in the same image. An obvious question is this: what will the brain do when two cues are present, and each provides a different estimate for the strength of an interpolated edge? Numerous possibilities exist. For example, a very simple solution would involve using one of the cues and ignoring the information conveyed by others. The question would then be which cue wins out. Another simple solution would involve simply combining the information conveyed by multiple cues in an additive fashion -- that is, the visual system, in this case, would not take into consideration which cue is providing the information or weight the information from different cues differently.

However, the goal of the output of the second stage is to offer a perceptual strength that reflects the real life existence or nonexistence of an interpolated edge as accurately and reliably as possible. There are, therefore, some characteristics of cues and groups of cues that the second stage might likely utilize in order to determine how information conveyed by those cues should be combined.

One interesting question is whether the brain takes into consideration the degree to which two or more cues are non-independent - that is, the degree to which the information conveyed by those cues is correlated or redundant. It seems possible, for example, that the information conveyed by different border ownership cues overlaps to some extent. Likewise, differences in surface characteristics across fragments, whether in luminance or in color, might be conveying some of the same information. Cues of different categories, too (for example, a border ownership cue and a manipulation to the spatial frequencies of edge fragments) might, when they appear in natural images, be the results of some of the same facts about the objects in the scene. For this reason, computations in the second stage might involve subtracting out some of the influence of a given cue

when the information conveyed by that cue overlaps with the information conveyed by another cue that has already been taken into account.

Another interesting question is whether the brain takes into consideration the degree to which a given cue tends to be accurate or reliable. Some cues may be more variable in the estimates they produce, or they may tend to produce less accurate estimates. For this reason, computation in the second stage may involve placing greater weight in cues that have a tendency to be more accurate or reliable (c.f. Landy, Maloney, Johnston, & Young, 1995). This would have the effect of ensuring that the information conveyed by those cues has a greater impact on the final perceptual strength of an interpolated edge relative to information conveyed by cues that tend to be less accurate or less reliable.

If computation in the second stage takes into consideration either of the possibilities mentioned above, then the combined magnitudes of two cues will not necessarily equal the simple addition of the magnitudes of their individual effects. The goal of the series of experiments presented here was to combine a few cues to arrive at a model of cue combination in the second stage. Three cues, each of which showed a particularly large individual effect in the experiments covered in Chapter 3, were chosen to be combined in a series of experiments. The results reveal that the combined magnitude of the effects is, in general, very well described by the simple addition of their individual effects. An exception, however, is that certain pairs of cues, when combined, result in a much larger reduction in accuracy performance than would be predicted by simple additivity. Two possible interpretations of this will be discussed.

4.3 Cues to be combined

Three cues, one of each of the following types: border ownership, amodal surface spreading, and spatial frequency differences across fragments, were chosen to be implemented in combined

manipulations in the following series of experiments. The cues chosen were: deleting the occluder (Experiment 4 of Chapter 3), background luminance contrast polarity differences across fragments (Experiment 6 of Chapter 3), and spatial frequency differences across fragments (Experiment 8 of Chapter 3). Each of the chosen cues was shown to produce a large individual effect in Chapter 3. Cues of three different types were chosen because it was expected that cues of different types would be more likely to carry independent information. If non-additivity were found with these cues, it would strongly support the hypothesis that cue combination is not simply additive in the second stage.

Four experiments were conducted, combining luminance contrast polarity and spatial frequency differences across fragments (Experiment 9), luminance contrast polarity differences across fragments and the deletion of the occluder (Experiment 10), spatial frequency differences across fragments and the deletion of the occluder (Experiment 11), and all three cues (luminance contrast polarity and spatial frequency differences across fragments and the deletion of the occluder - Experiment 12). The results reveal that luminance contrast polarity differences across fragments combine in a simple additive way with either of the other two manipulated cues. Whenever spatial frequency differences across fragments and the deletion of the occluder are combined, however, accuracy performance is significantly lower than additivity would predict. The discussion of this chapter presents a theory as to what is special about the combination of these two manipulations.

4.4 Experiment 9: Luminance contrast polarity combined with spatial frequency differences across fragments

Experiment 9 combined luminance contrast polarity differences across fragments (manipulated in isolation in Experiment 6) with spatial frequency differences across fragments (manipulated in isolation in Experiment 8). Figure 36 illustrates two example stimuli from the control and

experimental conditions for Experiment 9. The experimental stimuli have been altered in two ways: half of the fragments have been turned white, and a Gaussian blur has been applied to the edges of the black fragments. Presentation time was manipulated using the same levels as in Experiments 3-8.

The individual effects of the luminance contrast polarity manipulation (manipulated in isolation in Experiment 6) and the spatial frequency manipulation (manipulated in isolation in Experiment 8) were used to generate a prediction for the combined magnitude of these cues based on assuming simple additivity. The results of Experiment 9 are tested against the prediction generated based on the empirical data from Experiments 7 and 8.

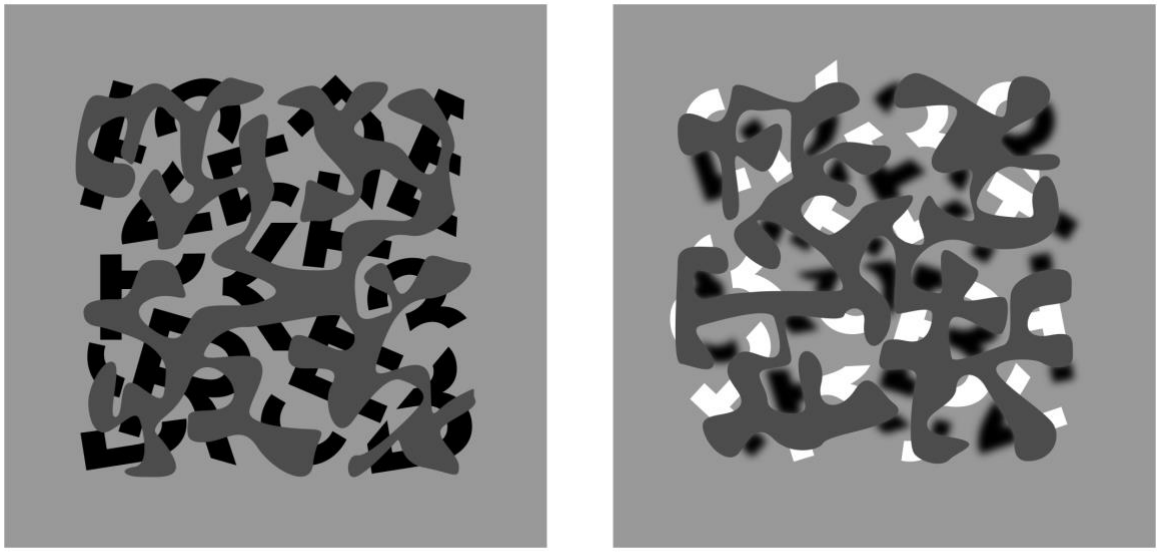

Figure 36. Example stimuli for the combined luminance contrast polarity and spatial frequency manipulation (Experiment 9). At left, a control stimulus is seen in which the letter 'R' is hidden in the lower left. At right, luminance contrast polarity and spatial frequency differences across fragments are manipulated. The 'B' hidden in the lower right may be more difficult to find and recognize as a result.

4.4.1 Method

Participants. Twenty-two University of California, Los Angeles undergraduates (17 female, ages 18-24, M_{age}: 20.21 years) participated in the experiment. One participant, who failed to complete

the experiment in the allotted hour, was excluded from the analysis. All participants were offered course credit for their participation and were naïve to the purposes of the study. Visual acuity was tested using a standardized ETDRS chart as a prerequisite to participation. All participants demonstrated 20/25 or better visual acuity.

Apparatus. The computer and test environment were the same as for Experiments 3-8.

Stimuli. Stimuli were identical to those used in Experiment 6, except that, in the experimental condition, the same Gaussian blur used in Experiment 8 was applied to the black fragments (see Experiment 9, Method). The control condition utilized the same mask as in Experiment 3. The experimental condition utilized a mask that was identical to the mask used in Experiment 6 except that the black fragments were modified in the same way as described above.

Procedure. The procedure was identical to that of Experiment 3 except that the sample stimuli used in the instructions were modified in the manner specified above.

4.4.2 Results

Figure 37 plots the data from the control condition as well as the experimental condition for Experiment 9. Figure 37 makes clear that, unsurprisingly, the combined manipulations result in a decrease in accuracy performance relative to the control.

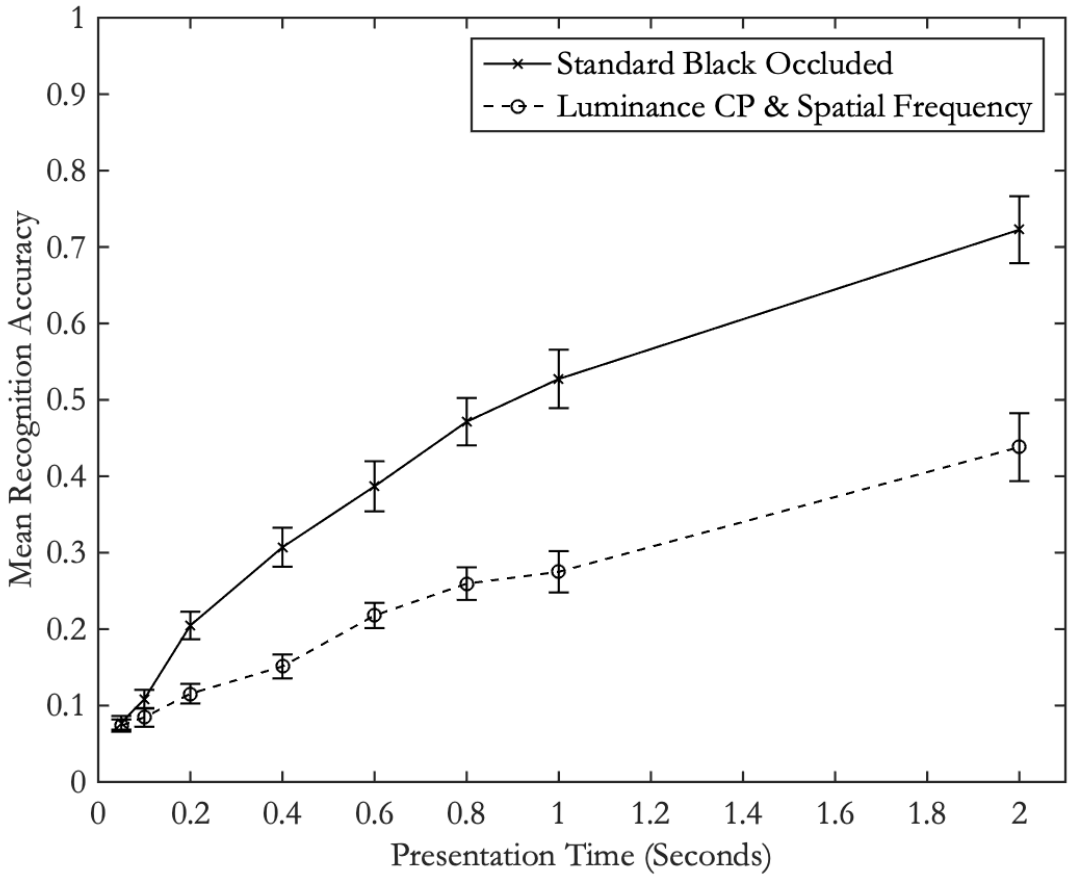

Figure 37. Mean recognition accuracy for the combined luminance contrast polarity and spatial frequency manipulation. Performance in the control condition is illustrated by the solid line, while performance in the experimental condition (combined luminance contrast polarity and spatial frequency manipulations) is illustrated by the dotted line. The combined manipulations result in a clear decrease in performance relative to the control. Error bars show 1 standard error above and below the condition mean.

4.4.2.1 Calculating average control condition performance

In order to generate the predicted additive effects for each of Experiments 9-12, performance for the control condition was first averaged across all experiments that utilized any of the manipulations (luminance contrast polarity, spatial frequency, deletion of the occluder) that are combined in Experiments 9-12. Thus, all of the predictions for Experiments 9 through 12 utilize the

same average control performance as a reference, and this was calculated by averaging the scores of participants in the control conditions of Experiments 4, 6, 8, 9, 10, 11, and12.

4.4.2.2 Calculating predictions based on additivity

In order to calculate predicted scores, the effects of the individual manipulations were first determined by subtracting performance for the individual manipulation from performance in the averaged control condition. That is, because the experimental condition of Experiment 9 utilizes a combination of the luminance contrast polarity and spatial frequency cues, the individual effects of those cues relative to the average control condition performance was first determined. Thus, performance in the experimental condition of Experiment 6 (luminance contrast polarity manipulation) was subtracted from the average control performance to determine the individual effect of the luminance contrast polarity manipulation. Likewise, performance in the experimental condition of Experiment 8 (spatial frequency manipulation) was subtracted from the average control performance to determine the individual effect of the spatial frequency manipulation. These two individual effects were then added together and subtracted from the average control condition performance to generate the predicted scores for Experiment 9. Wherever the predicted scores resulted in a performance prediction that was below chance performance, the predictions were adjusted to be at chance.

Chance was defined as the performance achieved with a 50ms presentation in the average control condition (8%). This was done because, although participants were told that the hidden alphanumeric character could be any letter A to Z and any number 0 to 9 (chance would be 2.7%), only eight alphanumeric characters were ever used (chance would be 12.5%), and participants clearly realized this, as some participants spontaneously inquired after the experiment why certain numbers and letters appeared over and over again, while others were never seen. That is all to say that chance

performance could not simply be calculated based on either the instructions participants were given or the number of possible alphanumeric characters. However, as mentioned in Chapter 3, interpolation is known to take around 100-200ms to proceed. Differences in the control and experimental conditions were expected to appear somewhere between 100ms and 200ms for this reason. This prediction was borne out in each of the 6 experiments presented in Chapter 3. Performance at 50ms - a time too brief for interpolation to occur - was consistently around 8% accuracy in all conditions at the 50ms time point. That is, performance was never worse than around 8% accuracy, even with a very brief 50ms presentation. It may be that participants were affected by the instructions they were given, informing them that any letter or number was possible, but that they also realized certain letters and numbers were common, and so adjusted their bias accordingly, resulting in a chance performance that was somewhere between what we would observe if participants took the instructions seriously or let their bias be totally controlled by their observations about what numbers and letters seemed to be appearing in the experiment. Whether that is true or not, it seems very likely the case that we can expect performance never to go below 8% in any of the experiments presented here. From my many pilot experiences, I can report that it is nearly impossible to find and recognize the hidden alphanumeric character in that amount of time, and that the condition manipulations have little impact. For all of these reasons, I have concluded that the performance observed at the 50ms timepoint very likely reflects the guess rate of participants in the character recognition experiments.

Figure 38 plots the data from the experimental condition of Experiment 9 against the additive prediction based on the data from Experiments 6 and 8.

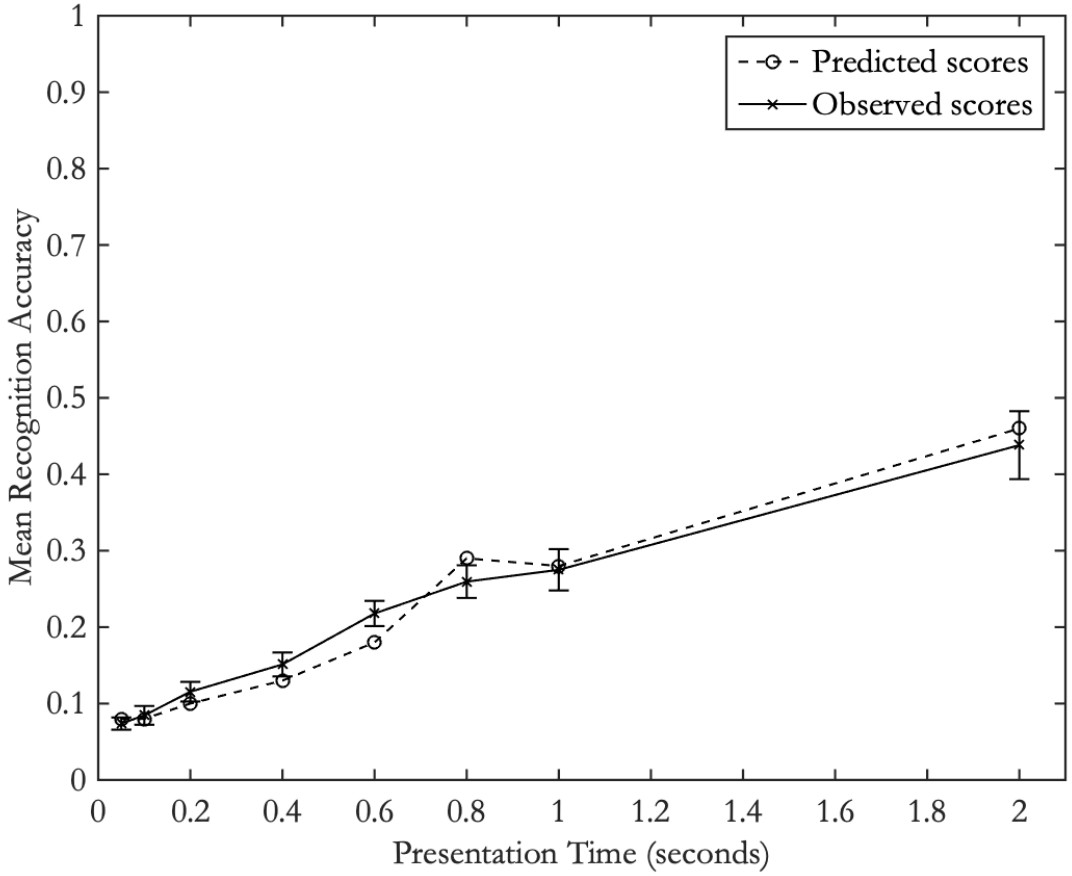

Figure 38. Mean recognition accuracy for the combined luminance contrast polarity and spatial frequency manipulation relative to a prediction based on the simple addition of the individual effects. Performance in the experimental condition is illustrated by the solid line. The prediction based on additivity is illustrated by the dotted line. The prediction was made by adding together the individual effects (measured in Experiments 6 and 8) of the luminance contrast polarity (Experiment 6) and spatial frequency (Experiment 8) manipulations. Error bars show 1 standard error above and below the observed condition means.

4.4.2.3 Fitting functions

As seen in Figure 40, the predicted magnitude of the combined effect, based on the simple addition of the individual effects, appears to be a very good match to the observed data. To test this, we fit cumulative Gaussian functions (cgf's) for every subject individually to their mean performance at all exposure durations in the experimental condition. A function was also fit to the predicted

means. The use of the cumulative Gaussian function in this application is not intended to imply any particular underlying model of the process generating the data; rather it is a mathematical way to summarize and compare the trends in the data. For this use, it was important that the function produced a reasonably good fit across the conditions, which was the case (R_2 values ranged from .83 to .99). All fits assumed a lapse rate (lambda, λ) of .01 (based on the experimenter's lapse rate in a pilot study), and a guess rate (gamma, γ) of .08. Nonlinear regression was used to determine the bias (alpha, α, or function midpoint or threshold) and slope (beta, β, that is, 1/sigma) of the cumulative Gaussian function that minimized squared deviations from the to-be-fitted data points. Illustrations of the fitted functions appear in Figures 39 and 40 below. The slopes of all the functions fit to the means of individual participants served as test means in a one-sample *t*-test against the slope of the function fit to the predicted means. The results reveal that the slope of the function fit to the predicted means (.30, R_2 = .87) is not significantly different from the average slope of functions fit to individual participant means (M = .29, SD = .20, mean R_2 = .83, $t(20) = -0.18$, $p = 0.86$). The slope of the function fit to the predicted means (.30) falls inside the confidence interval on the average slope of functions fit to observed participant means (CI: 0.20 - 0.38).

Figure 39 below illustrates the mean observed scores (solid black line), along with the fit to the predicted scores (dashed and dotted pink line), and a function with slope equal to the average slope of all functions fit to individual subject means (dashed blue line). Note that the actual analysis involved fitting individual upper portions of cumulative Gaussian functions to each individual participant's means. Because a plot with all participant means and fitted functions would be illegible (due to many overlapping lines), a single summary illustration of the average slope was chosen instead. The function for this summary illustration was generated by fitting a function first to the observed means across subjects, and then adjusting the slope as necessary. That is, the bias of this function was determined by the fit to observed across subject means; the slope was determined by

the average slope of functions fit to individual participant means (as discussed previously, gamma (γ), the guess rate, was set to .08, and lambda (λ), the lapse rate, to .01).

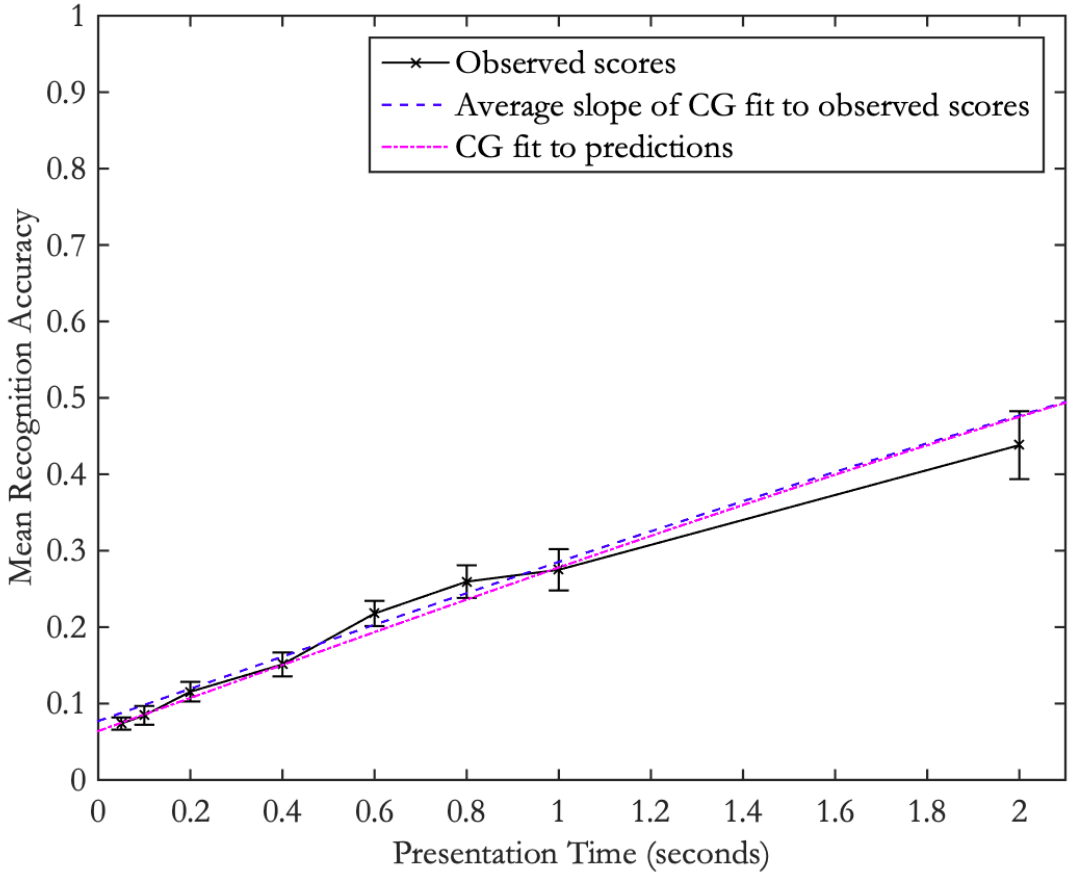

Figure 39. Observed accuracy means plotted with the function fit to predicted accuracy scores and the function with slope equal to the average slope of functions fit to individual participant means for the combined luminance contrast polarity and spatial frequency manipulation. Black line illustrates the observed accuracy means across subjects. The actual analysis involved fitting a function to each participant's individual means. The slopes of these functions, fit to individual participant means, were then submitted to a one sample t-test against the slope of the function fit to the predicted means. Here, individual participant means and the functions fit to them are not shown. Instead, the dashed blue line is a summary that reveals the average slope of functions fit to individual participant means (see text for more information). The slope of this fitted function is very similar to (and not significantly different from) the function fit to the predicted means (dashed and dotted pink line). The predicted means were based on the addition of the individual effects of the luminance contrast polarity (Experiment 6) and spatial frequency (Experiment 8) manipulations. Error bars show 1 standard error above and below the observed condition means.

Figure 40 below illustrates the function fit to the predicted accuracy scores within the confidence interval on the average slope fitted to individual observed participant means. That is, the function fit to the predicted accuracy scores is illustrated by the pink dashed and dotted line. In addition, three other fitted functions are shown: the function with slope equal to the average slope of functions fit to individual participant means (dashed blue line), and the two functions with slope equal to +/- the confidence interval on the average slope of functions fit to individual participant means (upper and lower dashed black lines).

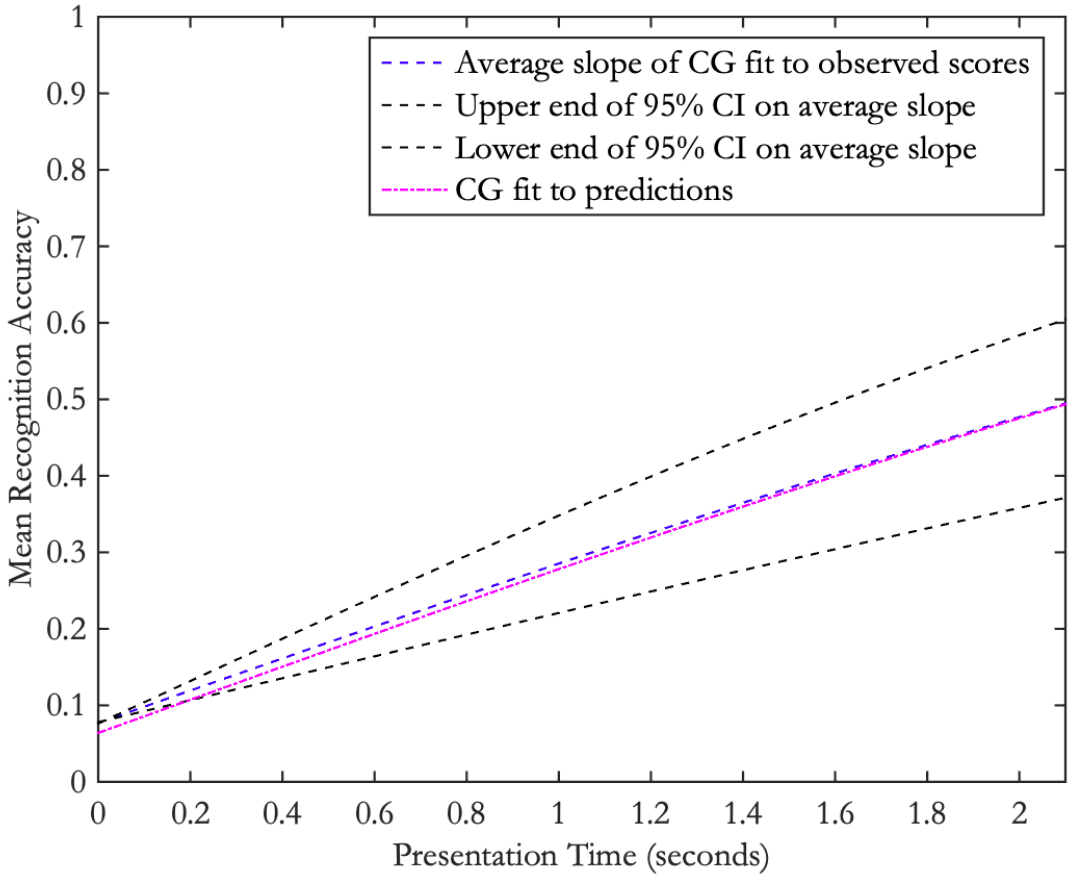

Figure 40. Function fit to the predicted accuracy means plotted within the confidence interval on the average slope of functions fit to observed individual participant mean for the combined luminance contrast polarity and spatial frequency manipulation. The dashed and dotted pink line illustrates the function fit to the predicted accuracy means. The dashed blue line illustrates the average slope of functions fit to individual

participant means. The dashed upper and lower black lines illustrate the 95% confidence interval on the average slope of functions fit to individual participant means.

4.4.3 Discussion

This experiment combined the individual cues manipulated in Experiments 6 (luminance contrast polarity differences across fragments) and 8 (spatial frequency differences across fragments). Recognition accuracy for the combined manipulation was compared to a prediction based on the simple addition of the individual effects observed in Experiments 6 and 8. The results revealed that the prediction based on simple additivity is a very good match to the observed data. These results suggest that information conveyed by cues as to luminance contrast polarity differences across fragments and spatial frequency differences across fragments combine in a simple additive fashion in Stage 2. That is, the strength of interpolated contours is reduced in a predictable fashion based on simply adding together the decrements of each of the individual cues. Experiments 10, 11, and 12 will examine whether predictions based on additivity account well for the observed accuracy decrements when other cues are combined.

4.5 Experiment 10: Unoccluded with luminance contrast polarity differences across fragments

Experiment 10 combined luminance contrast polarity differences across fragments (manipulated in isolation in Experiment 6) with the deletion of the occluder (or, turning the occluder the same color as the background - manipulated in isolation in Experiment 4). Figure 41 illustrates two example stimuli from the control and experimental conditions for Experiment 10. The experimental stimuli have been altered in two ways: half of the fragments have been turned white, and the

occluder has been set to the same color as the background. Presentation time was manipulated using the same levels as in Experiments 3-9.

The individual effects of the luminance contrast polarity manipulation (manipulated in isolation in Experiment 6) and the deletion of the occluder (manipulated in isolation in Experiment 4) were used to generate a prediction for the combined magnitude of these cues based on assuming simple additivity. The results of Experiment 10 are tested against this prediction.

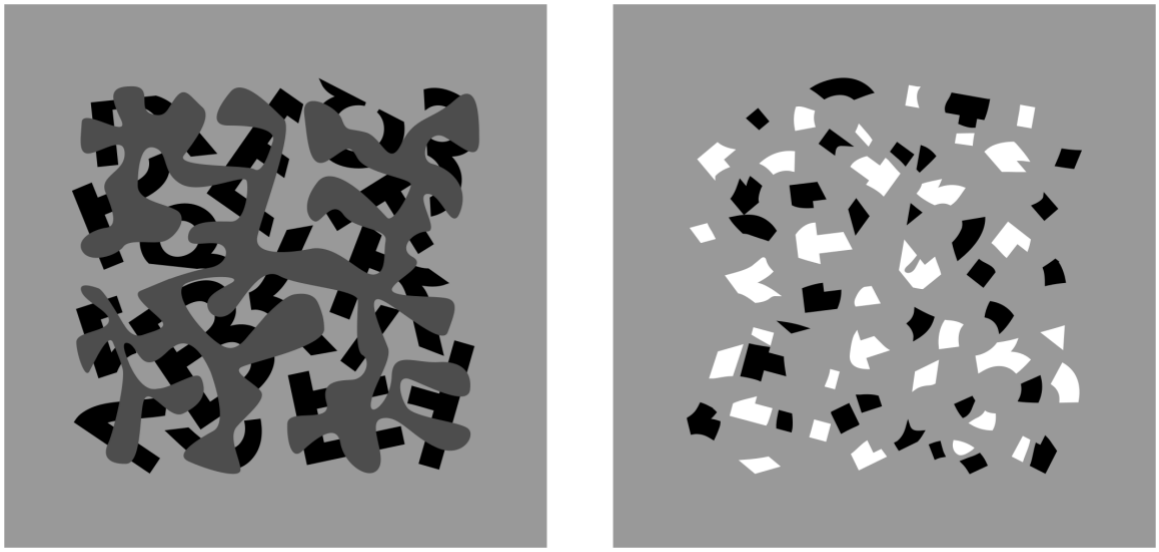

Figure 41. Example stimuli for the combined unoccluded and luminance contrast polarity manipulation (Experiment 10). At left, a control stimulus is seen in which the letter 'B' is hidden in the lower left. At right, luminance contrast polarity (relative to the background) across fragments is manipulated, and the occluder has been set to the same color as the background. The '3' hidden in the lower right may be more difficult to find and recognize as a result.

4.5.1 Method

Participants. Sixteen participants (12 female, 4 male, ages 18-34, M_{age}: 21.13 years) completed the experiment. All participants provided informed consent, were students attending the University of California, Los Angeles, were offered course credit for their participation, and were naïve to the

purposes of the study. Visual acuity was tested using a standardized ETDRS chart as a prerequisite to participation. All participants demonstrated 20/25 or better visual acuity.

Apparatus. The computer and test environment were the same as for Experiment 3.

Stimuli. Stimuli were identical to those used in Experiment 6, except that, in the experimental condition, in addition to utilizing white and black fragments, the occluder was turned the same color as the background. The control condition utilized the same mask as in Experiment 3. The experimental condition utilized a mask that was identical to the mask used in Experiment 6.

Procedure. The procedure was identical to that of Experiment 3 except that the sample stimuli used in the instructions were modified in the manner specified above.

4.5.2 Results

Figure 42 plots the data from the control condition as well as the experimental condition for Experiment 10. Clearly apparent in the plot is the unsurprising fact that the combined manipulations result in a decrease in accuracy performance relative to the control.

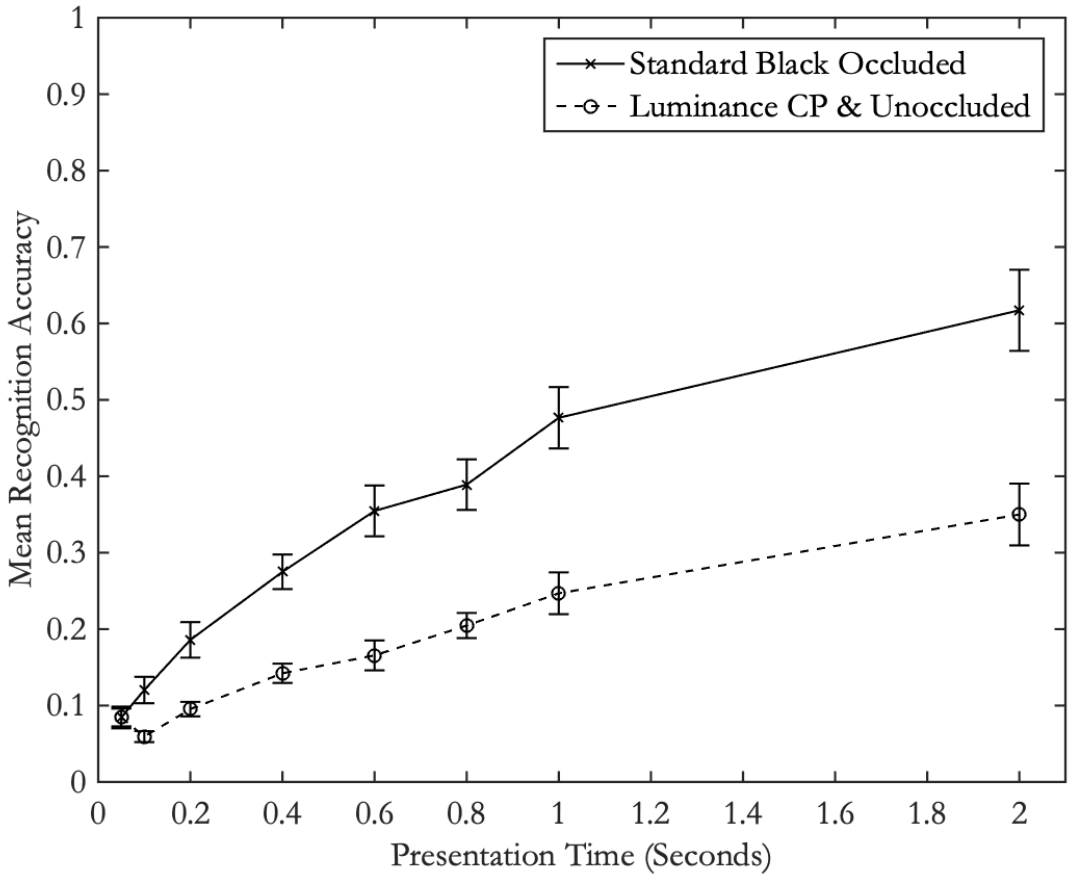

Figure 42. Mean recognition accuracy for the combined unoccluded and luminance contrast polarity manipulation. Performance in the control condition is illustrated by the solid line, while performance in the experimental condition is illustrated by the dotted line. The combined manipulations result in a clear decrease in performance relative to the control. Error bars show 1 standard error above and below the observed condition means.

Figure 43 plots the data from the experimental condition of Experiment 10 against the additive prediction based on the data from Experiments 6 and 4. In order to generate the predicted additive effect, the average control condition performance was calculated in the same manner as described in section 4.4.2.1 of Experiment 19. All predictions across Experiments 9-12 are made relative to this same average control condition performance. The effects of the individual manipulations were calculated in the same manner as described in section 4.4.2.2 of Experiment 9. That is, the effects of

the manipulations of Experiments 6 (luminance contrast polarity) and 4 (deletion of the occluder) were calculated relative to the average control condition performance. This was done by subtracting performance in each experimental condition from the averaged control condition performance. The resulting individual effects were then added together and subtracted from the average control condition performance to obtain the expected decrease in accuracy that the combined manipulation would produce, assuming simple additivity of the individual effects. As before, the predictions were adjusted to chance performance (8%) wherever they fell below chance (see section 4.4.2.2 for a discussion of determining chance performance).

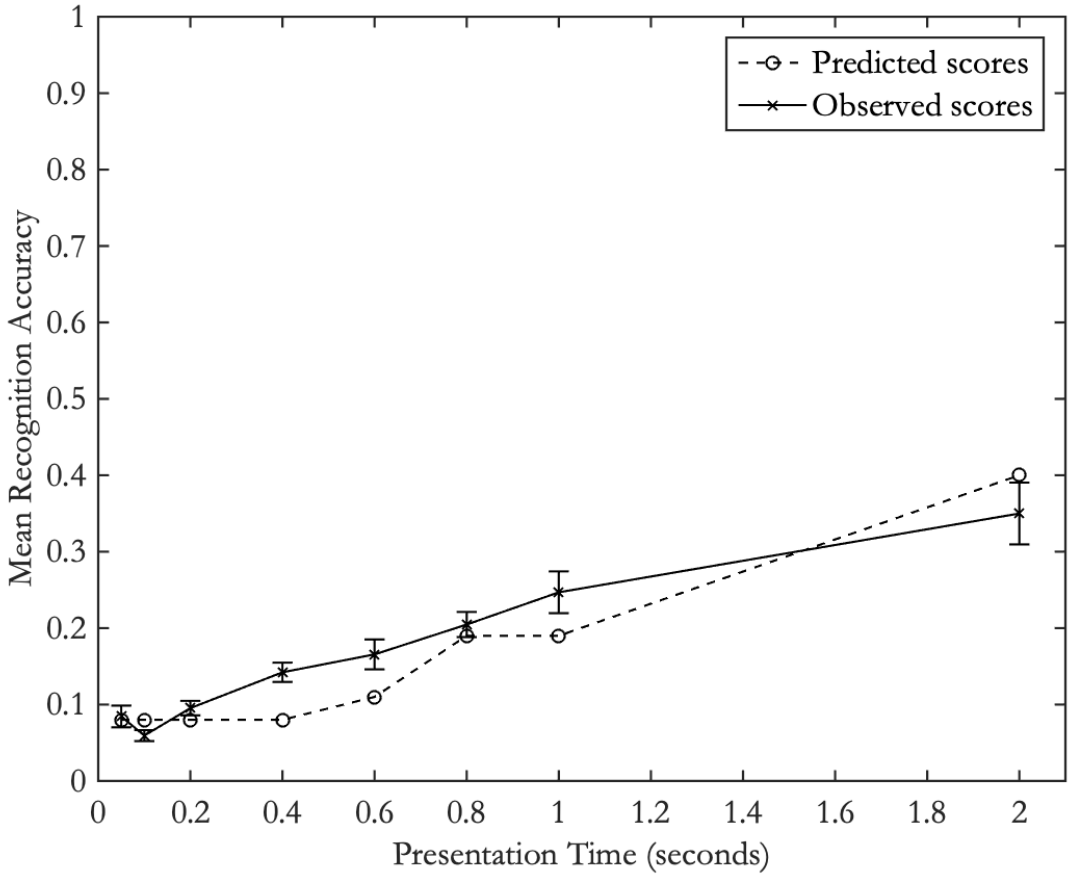

Figure 43. Mean recognition accuracy for the combined luminance contrast polarity and unoccluded manipulation relative to a prediction based on the simple addition of the individual effects. Performance in the experimental condition is illustrated by the solid line. The prediction based on additivity is illustrated by the dotted line. The prediction was made by adding together the individual effects (measured in Experiments 6 and 4) of the luminance contrast polarity (Experiment 6) and

unoccluded (Experiment 4) manipulations. Error bars show 1 standard error above and below the observed condition means.

As seen in Figure 43, the predicted magnitude of the combined effect, based on the simple addition of the individual effects, appears to be a very good match to the observed data. To test this, a the upper half of a cumulative gaussian function was first fit, for every subject individually, to their mean performance at all time points in the experimental condition. The upper half of a cumulative gaussian function was also fit to the predicted means. The fits were determined in the same manner described previously (see Experiment 9, section 4.4.2.3). The slopes of all the functions fit to the data of individual subjects then served as test means in a one-sample *t*-test against the slope of the function fit to the predicted means. The results reveal that the slope of the function fit to the predicted means (.24, R_2 = .91) is not significantly different from the average slope of functions fit to individual participant means (M = .22, SD = .13, mean R_2 = .89, $t(15) = 0.01$, $p = 0.51$). The slope of the function fit to the predicted scores (.24) falls inside the confidence interval on the average slope of functions fit to individual participant means (CI: 0.15 - 0.29). Figure 44 illustrates the average slope of functions fit to individual participant means (dashed blue line) against the function fit to the predicted means (dashed and dotted pink line) and the actual observed means across all participants (solid black line). Figure 45 illustrates the function fit to the predicted means (dashed and dotted pink line) within the confidence interval on the average slope fit to individual participant means (dashed black lines).

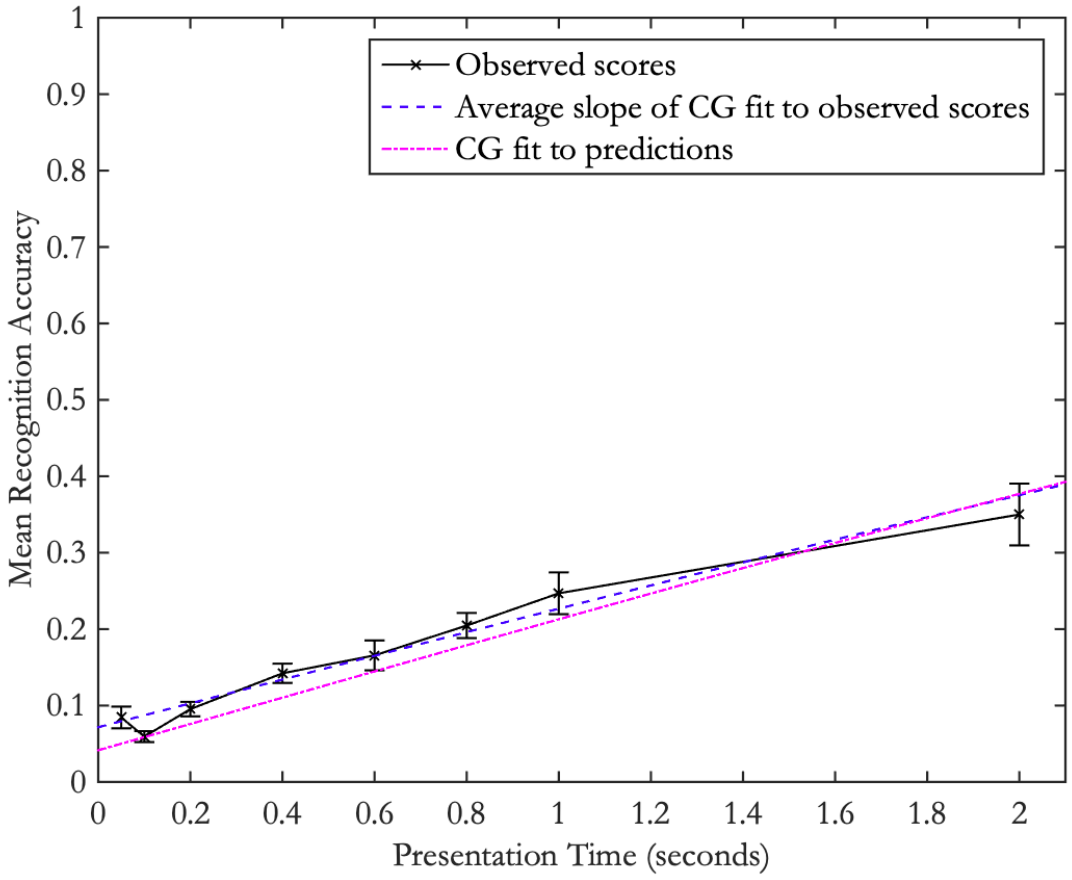

Figure 44. Observed accuracy means plotted with the function fit to predicted accuracy scores and the function with slope equal to the average slope of functions fit to individual participant means for the unoccluded luminance contrast polarity manipulation. Black line illustrates the observed accuracy means across subjects. The actual analysis involved fitting a function to each participant's individual means. The slopes of these functions, fit to individual participant means, were then submitted to a one sample t-test against the slope of the function fit to the predicted means. Here, individual participant means and the functions fit to them are not shown. Instead, the dashed blue line is a summary that reveals the average slope of functions fit to individual participant means (see text for more information). The slope of this fitted function is very similar to (and not significantly different from) the function fit to the predicted means (dashed and dotted pink line). The predicted means were based on the addition of the individual effects of the luminance contrast polarity (Experiment 6) and unoccluded (Experiment 4) manipulations. Error bars show 1 standard error above and below the observed condition means.

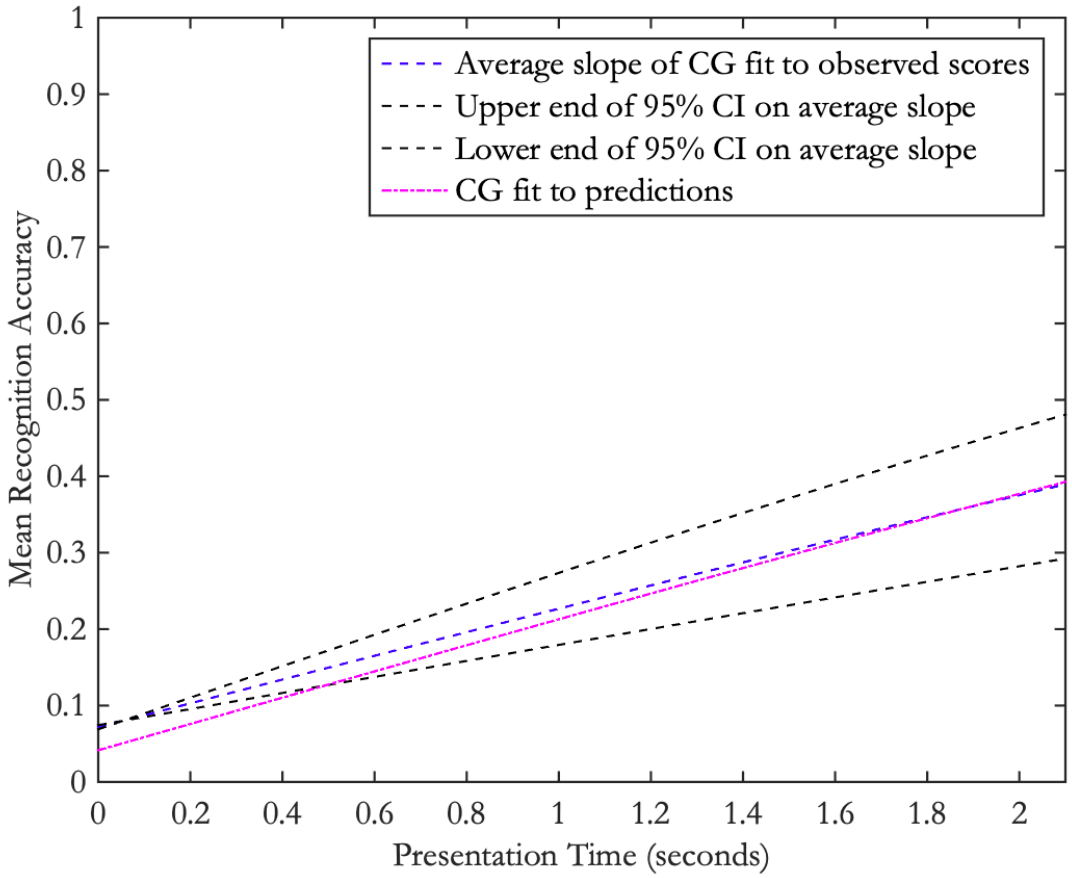

Figure 45. Function fit to the predicted accuracy means plotted within the confidence interval on the average slope of functions fit to observed individual participant mean for the combined luminance contrast polarity and unoccluded manipulation. The dashed and dotted pink line illustrates the function fit to the predicted accuracy means. The dashed blue line illustrates the average slope of functions fit to individual participant means. The dashed upper and lower black lines illustrate the 95% confidence interval on the average slope of functions fit to individual participant means.

4.5.3 Discussion

This experiment combined the individual cues manipulated in Experiments 7 (luminance contrast polarity differences across fragments) and 4 (deletion of the occluder). Recognition accuracy for the combined manipulation was compared to a prediction based on the simple addition of the individual effects observed in Experiments 6 and 4. The results revealed that the prediction

based on simple additivity is a very good match to the observed data. These results suggest that information conveyed by cues as to luminance contrast polarity differences across fragments and the lack of an occluder combine in a simple additive fashion in stage 2. That is, the strength of interpolated contours is reduced in a predictable fashion based on simply adding together the decrements of each of the individual cues. Together, Experiments 9 and 10 reveal that manipulations to luminance contrast polarity differences across fragments combine in an additive way with both spatial frequency differences across fragments and the lack of an occluder (a border ownership cue). Experiment 11 will examine whether spatial frequency differences across fragments and the lack of an occluder combine in an additive fashion with each other, and Experiment 12 will examine whether all 3 cues, when combined, produce an accuracy decrement that is well accounted for by the simple addition of their individual effects.

4.6 Experiment 11: Unoccluded with spatial frequency differences across fragments

Experiment 11 combined spatial frequency differences across fragments (manipulated in isolation in Experiment 8) with the elimination of the occluder (or, setting the occluder to the same color as the background - manipulated in isolation in Experiment 4). Figure 46 illustrates two example stimuli from the control and experimental conditions for Experiment 11. The experimental stimuli have been altered in two ways: a Gaussian blur has been applied to the boundaries of half of the fragments (see Experiment 8, method), and the occluder has been set to the same color as the background. Presentation time was manipulated using the same levels as in Experiments 3-10.

The individual effects of the spatial frequency manipulation (manipulated in isolation in Experiment 8) and the deletion of the occluder (manipulated in isolation in Experiment 4) were used to generate a prediction for the combined magnitude of these cues based on assuming simple

additivity. The results of Experiment 11 are tested against the prediction generated based on the empirical data from Experiments 8 and 4.

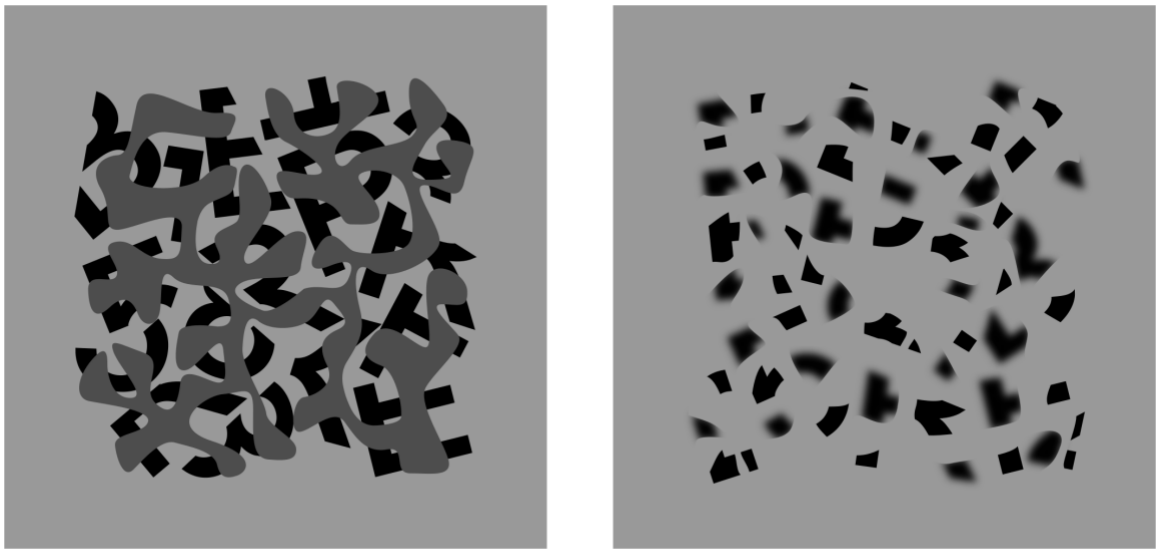

Figure 46. Example stimuli for the combined unoccluded and spatial frequency manipulation (Experiment 11). At left, a control stimulus is seen in which the letter 'P' is hidden in the upper right. At right, spatial frequency differences across fragments are manipulated, and the occluder has been set to the same color as the background. The '5' hidden in the lower left may be more difficult to find and recognize as a result.

4.6.1 Method

Participants. Twenty-seven University of California, Los Angeles undergraduates participated in the experiment. One participant, who failed to complete the experiment within the allotted hour, was excluded from the analysis. Two other participants - one whose performance was at chance at all timepoints in the control condition, and another whose performance revealed a non-monotonic function that was at chance at the two-second time point in the control condition - were also eliminated from the analysis. Thus, twenty-three participants (16 female, 7 male, ages 18-25, M_{age}: 20.50 years) were included in the analysis. All participants provided informed consent, were offered

course credit for their participation, and were naïve to the purposes of the study. Visual acuity was tested using a standardized ETDRS chart as a prerequisite to participation. All participants demonstrated 20/25 or better visual acuity.

Apparatus. The computer and test environment were the same as for Experiment 3.

Stimuli. Stimuli were identical to those used in Experiment 8, except that, in the experimental condition, in addition to utilizing half blurred fragments, the occluder was set to the same color as the background. The control condition utilized the same mask as in Experiment 3. The experimental condition utilized a mask that was identical to the mask used in Experiment 8.

Procedure. The procedure was identical to that of Experiment 3 except that the sample stimuli used in the instructions were modified in the manner specified above.

4.6.2 Results

Figure 47 plots the data from the control and experimental conditions of Experiment 11. In the plot, it can be clearly seen that the combined manipulation results in an unsurprising decrease in mean accuracy scores relative to the control.

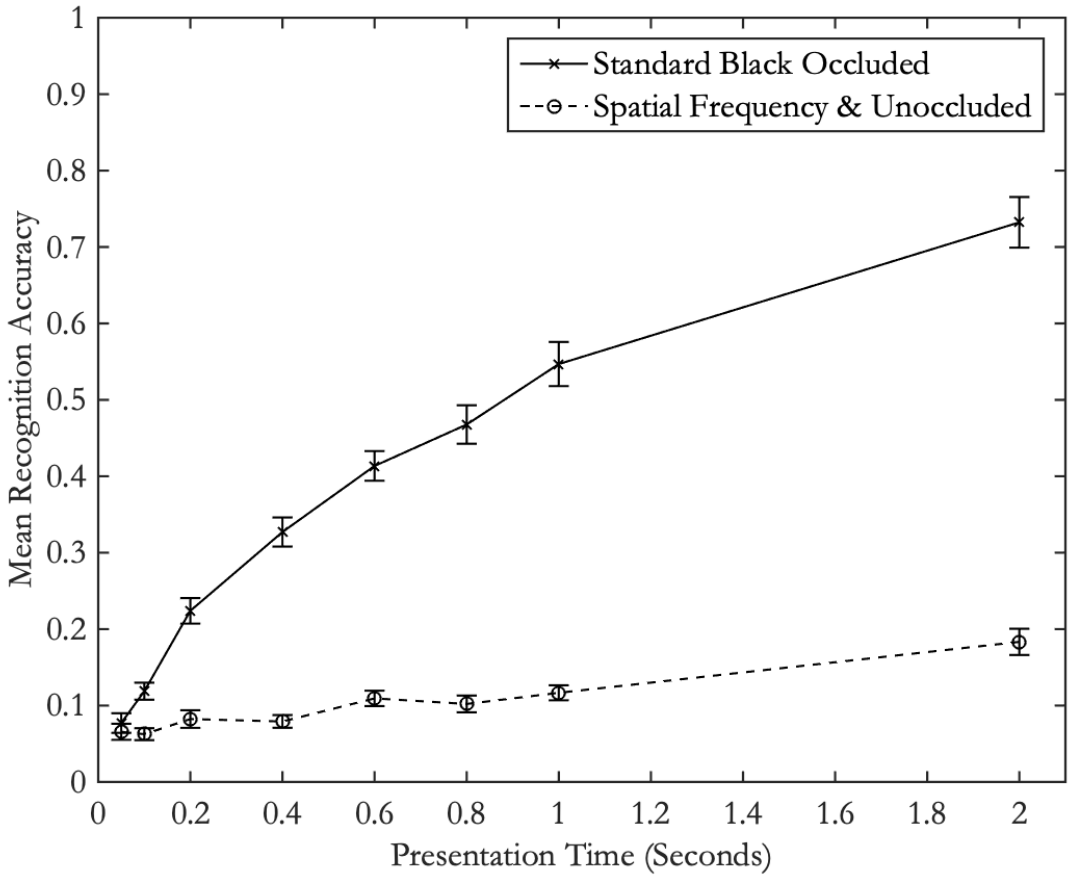

Figure 47. Mean recognition accuracy for the combined unoccluded and spatial frequency manipulation. Performance in the control condition is illustrated by the solid line, while performance in the experimental condition is illustrated by the dotted line. The combined manipulations result in a clear decrease in performance relative to the control. Error bars show 1 standard error above and below the observed condition means.

In order to generate the predicted additive effect, the average control condition performance was calculated in the same manner as described in section 4.4.2.1 of Experiment 9. All predictions across Experiments 9-12 were made relative to this same average control condition performance. The effects of the individual manipulations were calculated in the same manner as described in section 4.4.2.2 of Experiment 9. That is, the effects of the manipulations of Experiments 4 (deletion of the occluder) and 8 (spatial frequency differences across fragments) were calculated relative to the

average control condition performance. This was done by subtracting performance in each experimental condition from the averaged control condition performance. The resulting individual effects were then added together and subtracted from the average control condition performance to obtain the expected decrease in accuracy that the combined manipulation would produce, assuming simple additivity of the individual effects. As before, the predictions were adjusted to chance performance (8%) wherever they fell below chance (see section 4.4.2.2 for a discussion of determining chance performance). Figure 48 plots the data from the experimental condition of Experiment 11 against the additive prediction based on the data from Experiments 4 (unoccluded manipulation) and 8 (spatial frequency manipulation).

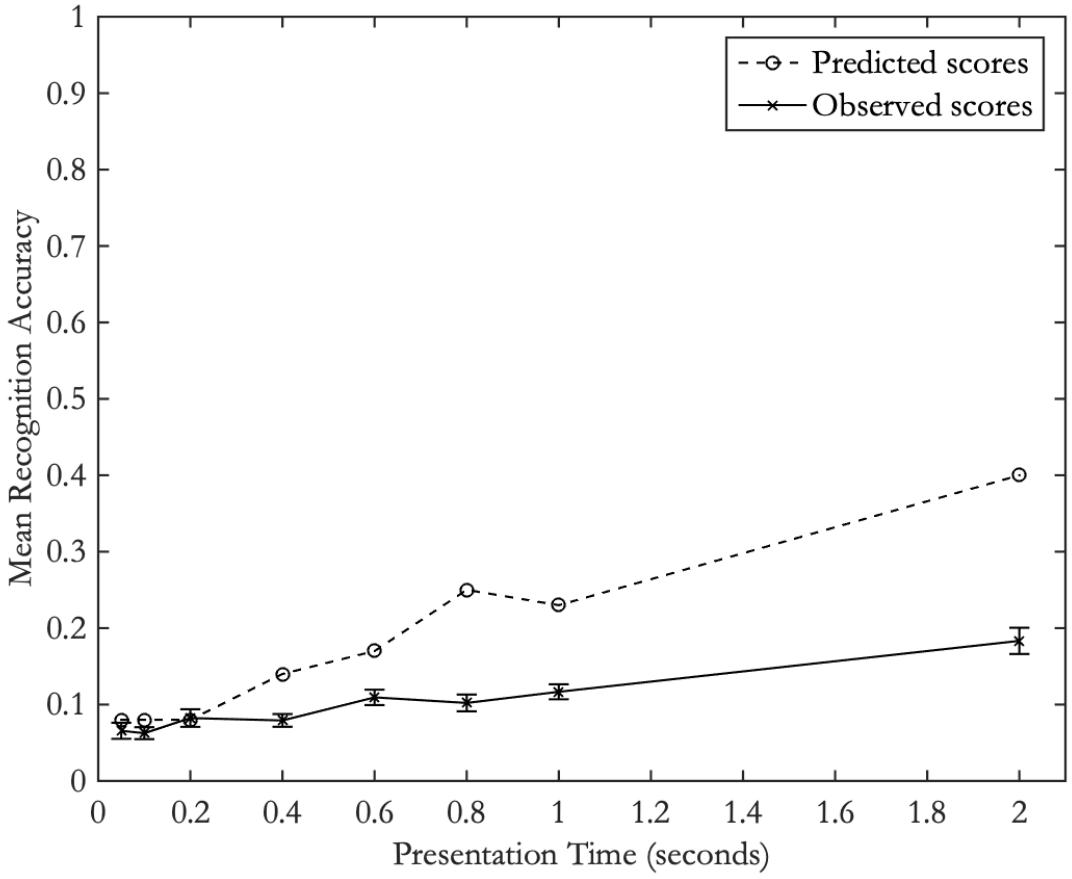

Figure 48. Mean recognition accuracy for the combined luminance contrast polarity and unoccluded manipulation relative to a prediction based on the simple addition of the individual effects. Performance in the experimental condition is illustrated by the

solid line. The prediction based on additivity is illustrated by the dotted line. The prediction was made by adding together the individual effects (measured in Experiments 4 and 8) of the unoccluded (Experiment 4) and spatial frequency (Experiment 8) manipulations. Error bars show 1 standard error above and below the observed condition means.

As seen in Figure 48, the predicted magnitude of the combined effect, based on the simple addition of the individual effects, is far off from the observed data. To test this observation, the upper half of a cumulative gaussian function was first fit, for every subject individually, to their mean performance at all time points in the experimental condition. The upper half of a cumulative gaussian function was also fit to the predicted means. The fits were determined in the same manner described previously (see Experiment 10, section 4.4.2.3). The slopes of all the functions fit to the data of individual subjects then served as test means in a one-sample *t*-test against the slope of the function fit to the predicted means. The results reveal that the slope of the function fit to the predicted means (.24, R_2 = .91) is significantly different from the average slope of functions fit to individual participant means (M = .09, SD = .06, mean R_2 = .97, $t(23) = -12.52$, $p < .01$). The slope of the function fit to the predicted scores (.24) falls outside of the confidence interval on the average slope of functions fit to individual participant means (CI: 0.06 - 0.11). Figure 49 illustrates the average slope of functions fit to individual participant means (dashed blue line) against the function fit to the predictions (dashed and dotted pink line) and the actual observed means across all participants (solid black line). Figure 50 illustrates the function fit to the predicted means (dashed and dotted pink line) within the confidence interval on the average slope fit to individual participant means (dashed black lines).

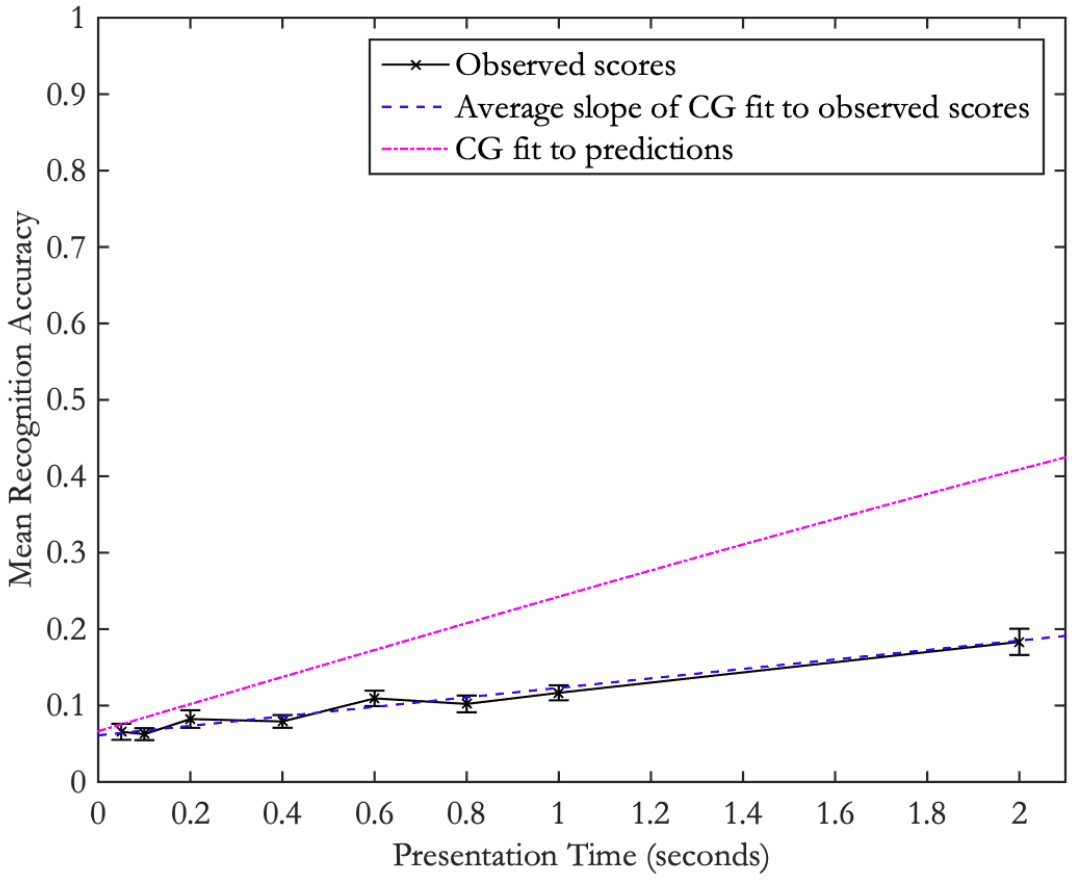

Figure 49. Observed accuracy means plotted with the function fit to predicted accuracy scores and the function with slope equal to the average slope of functions fit to individual participant means for the unoccluded spatial frequency manipulation. Black line illustrates the observed accuracy means across subjects. The actual analysis involved fitting a function to each participant's individual means. The slopes of these functions, fit to individual participant means, were then submitted to a one sample t-test against the slope of the function fit to the predicted means. Here, individual participant means and the functions fit to them are not shown. Instead, the dashed blue line is a summary that reveals the average slope of functions fit to individual participant means (see text for more information). The slope of this fitted function is very similar to (and not significantly different from) the function fit to the predicted means (dashed and dotted pink line). The predicted means were based on the addition of the individual effects of the unoccluded (Experiment 4) and spatial frequency (Experiment 8) manipulations. Error bars show 1 standard error above and below the observed condition means.

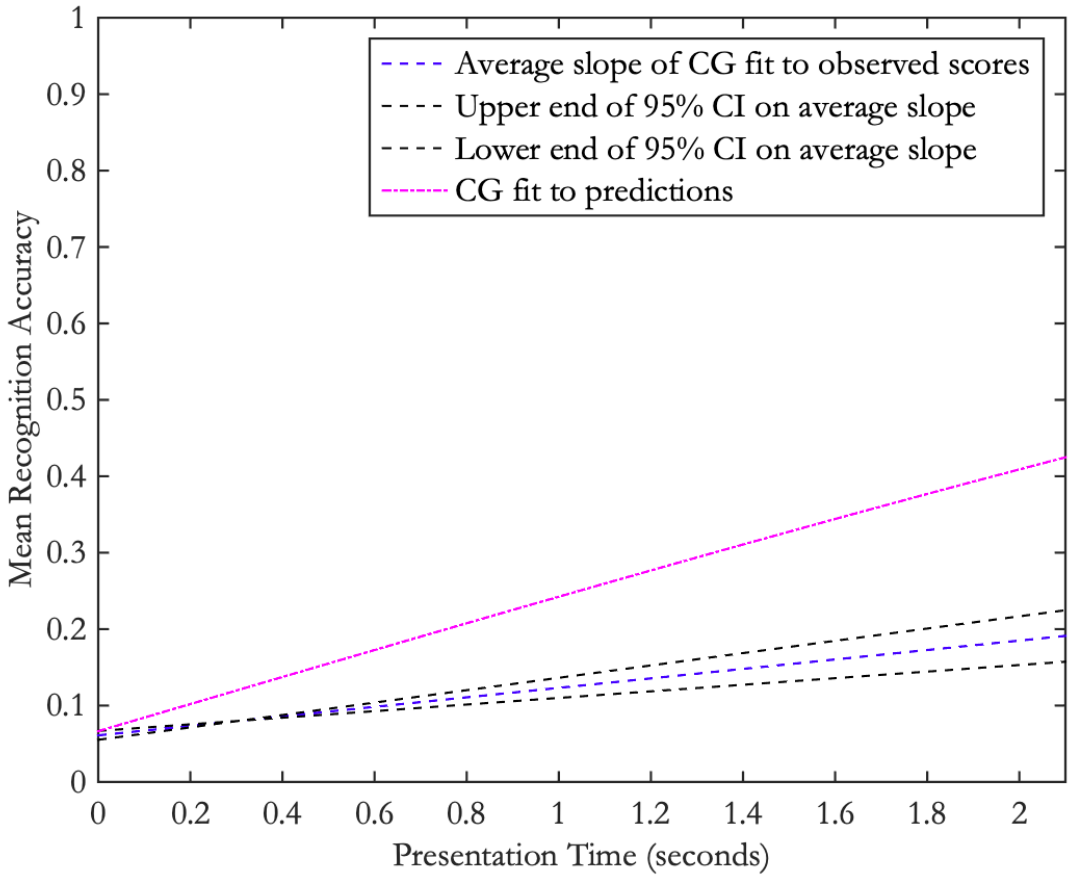

Figure 50. Function fit to the predicted accuracy means plotted within the confidence interval on the average slope of functions fit to observed individual participant mean for the combined unoccluded and spatial frequency manipulation. The dashed and dotted pink line illustrates the function fit to the predicted accuracy means.The dashed blue line illustrates the average slope of functions fit to individual participant means. The dashed upper and lower black lines illustrate the 95% confidence interval on the average slope of functions fit to individual participant means.

4.6.3 Discussion

This experiment combined the individual cues manipulated in Experiments 8 (spatial frequency differences across fragments) and 4 (eliminating the occluder). Recognition accuracy for the combined manipulation was compared to a prediction based on the simple addition of the individual effects observed in Experiments 8 and 4. The results revealed that the prediction based on simple

additivity is a poor match to the observed data. This suggests that the combination of spatial frequency differences across fragments and the deletion of the occluder interact in such a way that the perceptual strengths of interpolated contours in Stage 2 are decreased by an amount that goes beyond the simple addition of the individual strengths of those cues. Experiment 12 will examine the combination of all 3 cues (luminance contrast polarity differences across fragments, spatial frequency differences across fragments, and turning the occluder the same color as the background).

4.7 Experiment 12: Unoccluded with luminance contrast polarity and spatial frequency differences across fragments

Experiment 12 combined all three cues: deletion of the occluder (manipulated in isolation in Experiment 4), background luminance contrast polarity differences across fragments (manipulated in isolation in Experiment 6), and spatial frequency differences across fragments (manipulated in isolation in Experiment 8). Figure 51 illustrates two example stimuli from the control and experimental conditions for Experiment 12. The experimental stimuli have been altered in three ways: half of the fragments have been turned white, a Gaussian blur has been applied to the boundaries of the black fragments (see Experiment 8, method), and the occluder has been set to the same color as the background. Presentation time was manipulated using the same levels as in Experiments 3-11.

The individual effects of the deletion of the occluder (manipulated in isolation in Experiment 4), luminance contrast polarity differences across fragments (manipulated in isolation in Experiment 6), and spatial frequency differences across fragments (manipulated in isolation in Experiment 8) were used to generate a prediction for the combined magnitude of these cues based on assuming simple additivity. The results of Experiment 12 were tested against the prediction generated based on the empirical data from Experiments 4, 6, and 8.

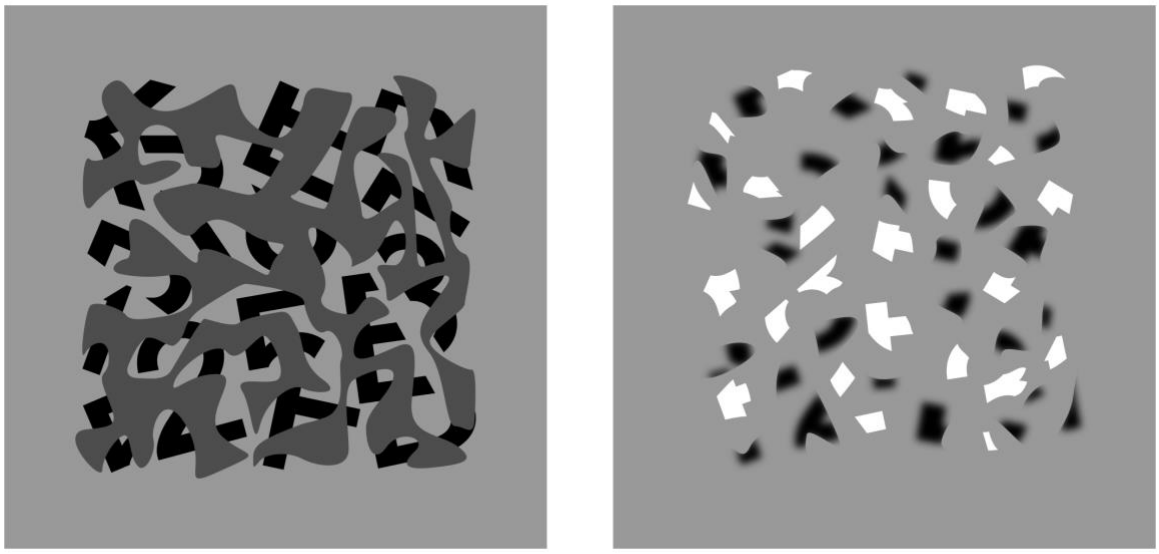

Figure 51. Example stimuli for the combined unoccluded luminance contrast polarity andn spatial frequency manipulation (Experiment 12). At left, a control stimulus is seen in which the number '2' is hidden in the lower left. At right, luminance contrast polarity and spatial frequency differences across fragments are manipulated and the occluder is set to the same color as the background. The '5' hidden in the upper right may be more difficult to find and recognize as a result.

4.7.1 Method

Participants. Twenty-four University of California, Los Angeles undergraduates participated in the experiment. Two participants - one who failed to complete the experiment within the allotted hour, and another whose performance was at chance at all presentation times - were excluded. The analysis thus included twenty-two participants (20 female, 2 male ages 18-23, M_{age}: 20.23 years). All participants provided informed consent, were offered course credit for their participation, and were naïve to the purposes of the study. Visual acuity was tested using a standardized ETDRS chart as a prerequisite to participation. All participants demonstrated 20/25 or better visual acuity.

Apparatus. The computer and test environment were the same as in Experiments 3.

Stimuli. Stimuli were identical to those used in Experiment 8, except that, in the experimental condition, in addition to utilizing half blurred fragments, the occluder was turned the same color as the background. The control condition utilized the same mask as in Experiment 3. The experimental condition utilized a mask that was identical to the mask used in Experiment 8.

Procedure. The procedure was identical to that of Experiment 3 except that the sample stimuli used in the instructions were modified in the manner specified above.

4.7.2 Results

Figure 52 shows mean accuracy scores for the control condition and the combined (luminance contrast polarity, spatial frequency, and unoccluded) manipulation. Figure 52 makes clear that, unsurprisingly, the combined manipulations do result in a decrease in accuracy performance relative to the control.

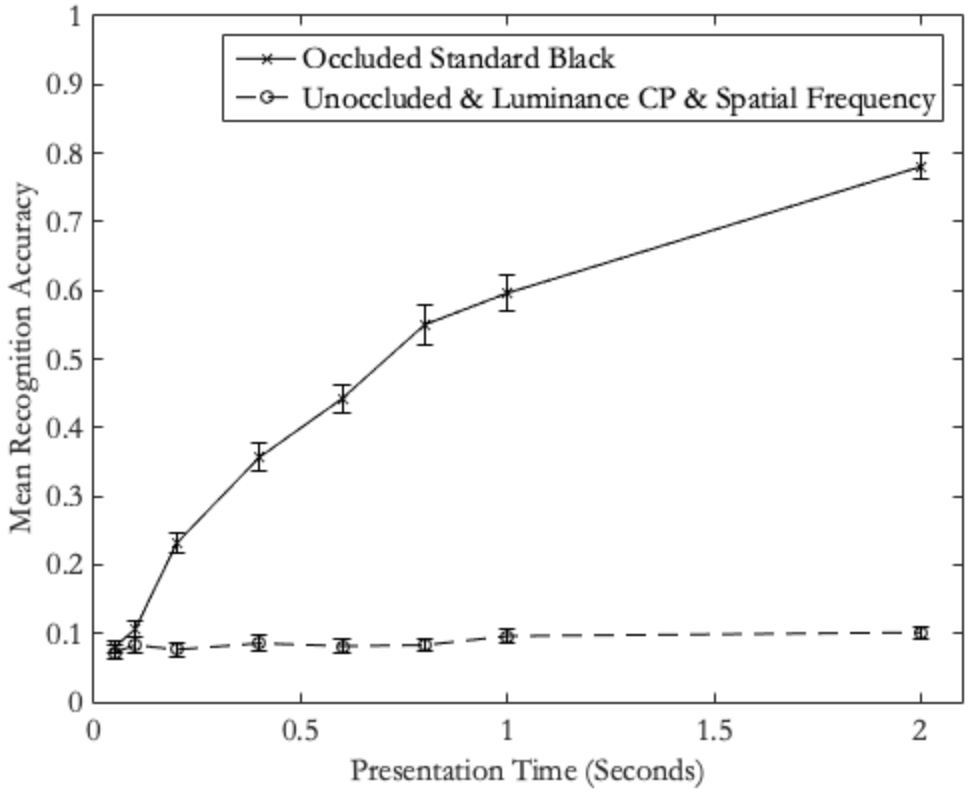

Figure 52. Mean recognition accuracy for the combined unoccluded and luminance contrast polarity and spatial frequency manipulation. Performance in the control condition is illustrated by the solid line, while performance in the experimental condition is illustrated by the dotted line. The combined manipulations result in a clear decrease in performance relative to the control. Error bars show 1 standard error above and below the observed condition means.

Figure 53 plots the data from the experimental condition of Experiment 12 against the additive prediction based on the data from Experiments 4, 6, and 8. In order to generate the predicted additive effect, the average control condition performance was calculated in the same manner as described in section 4.4.2.1 of Experiment 9. All predictions across Experiments 9-12 are made relative to this same average control condition performance. Then the effects of the individual manipulations were calculated in the same manner as described in section 4.4.2.2 of Experiment 9. That is, the effects of the manipulations of Experiments 4 (deletion of the occluder), 6 (luminance contrast polarity), and 8 (spatial frequency manipulation) were calculated relative to this average

control condition performance. This was done by subtracting performance in each experimental condition from the averaged control condition performance. The resulting individual effects were then added together and subtracted from the average control condition performance to obtain the expected decrease in accuracy that the combined manipulation would produce, assuming simple additivity of the individual effects. As before, the predictions were adjusted to chance performance (8%) wherever they fell below chance (see section 4.4.2.2 for a discussion of determining chance performance).

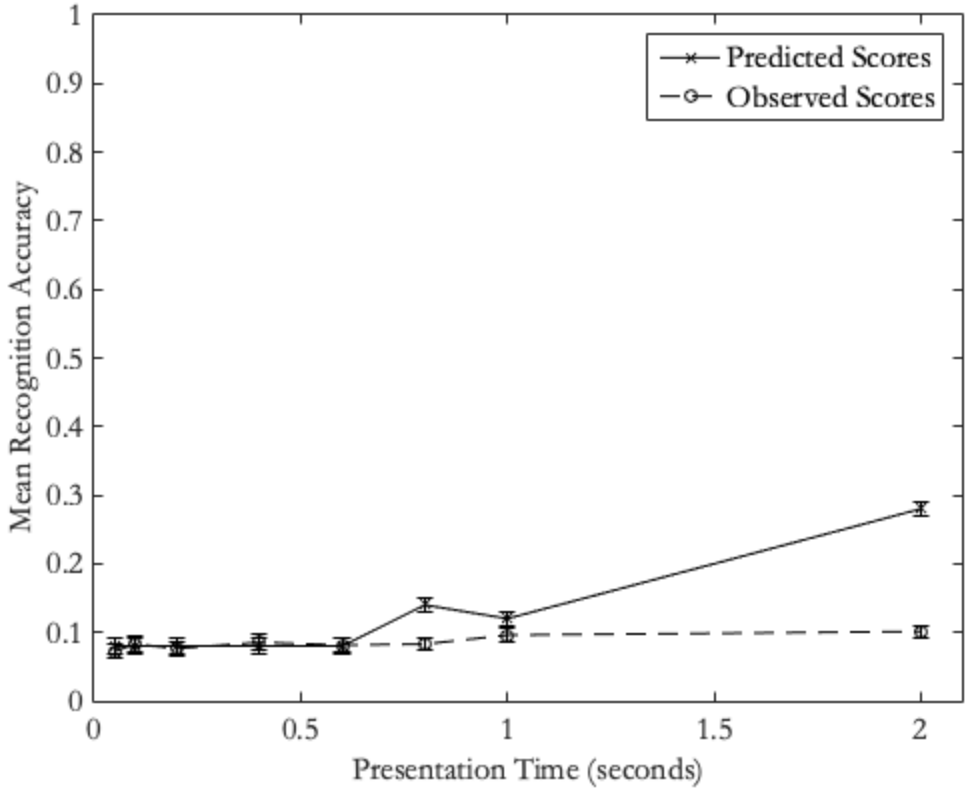

Figure 53. Mean recognition accuracy for the combined luminance contrast polarity, spatial frequency, and unoccluded manipulations relative to a prediction based on the simple addition of the individual effects. Performance in the experimental condition is illustrated by the solid line. The prediction based on additivity is illustrated by the dotted line. The prediction was made by adding together the individual effects (measured in Experiments 4, 6, and 8) of the unoccluded (Experiment 4), luminance contrast polarity (Experiment 6) and spatial frequency (Experiment 8) manipulations. Error bars show 1 standard error above and below the observed condition means.

As seen in Figure 53, the predicted magnitude of the effect based on adding together the individual effects of the combined cues measured in Experiments 4, 6, and 8 is a poor fit to the observed data. To confirm this observation, the same mid-upper portion of a cumulative gaussian function was first fit, for every subject individually, to their mean performance at all time points in the experimental condition. The upper half of a cumulative gaussian was also fit to the predicted means. The fits were determined in the same manner described previously (see Experiment 9, section 4.4.2.3). The slopes of all the functions fit to the data of individual subjects then served as test means in a one-sample *t*-test against the slope of the function fit to the predicted means. The results reveal that the slope of the function fit to the predicted means (.12, R_2 = .97) is not significantly different from the average slope of functions fit to individual participant means (*M* = .02, *SD* = .02, mean R_2 = .98, $t(20) = -24.26$, $p < .01$). The slope of the function fit to the predicted scores (.12) falls outside the confidence interval on the average slope of functions fit to individual participant means (CI: 0.01 - 0.03). Figure 54 illustrates the average slope of functions fit to individual participant means (dashed blue line) against the function fit to the predictions (dashed and dotted pink line) and the actual observed means across all participants (solid black line). Figure 55 shows the function fit to the predicted means (dashed and dotted pink line) outside the confidence interval on the average slope of functions fit to individual participant means.

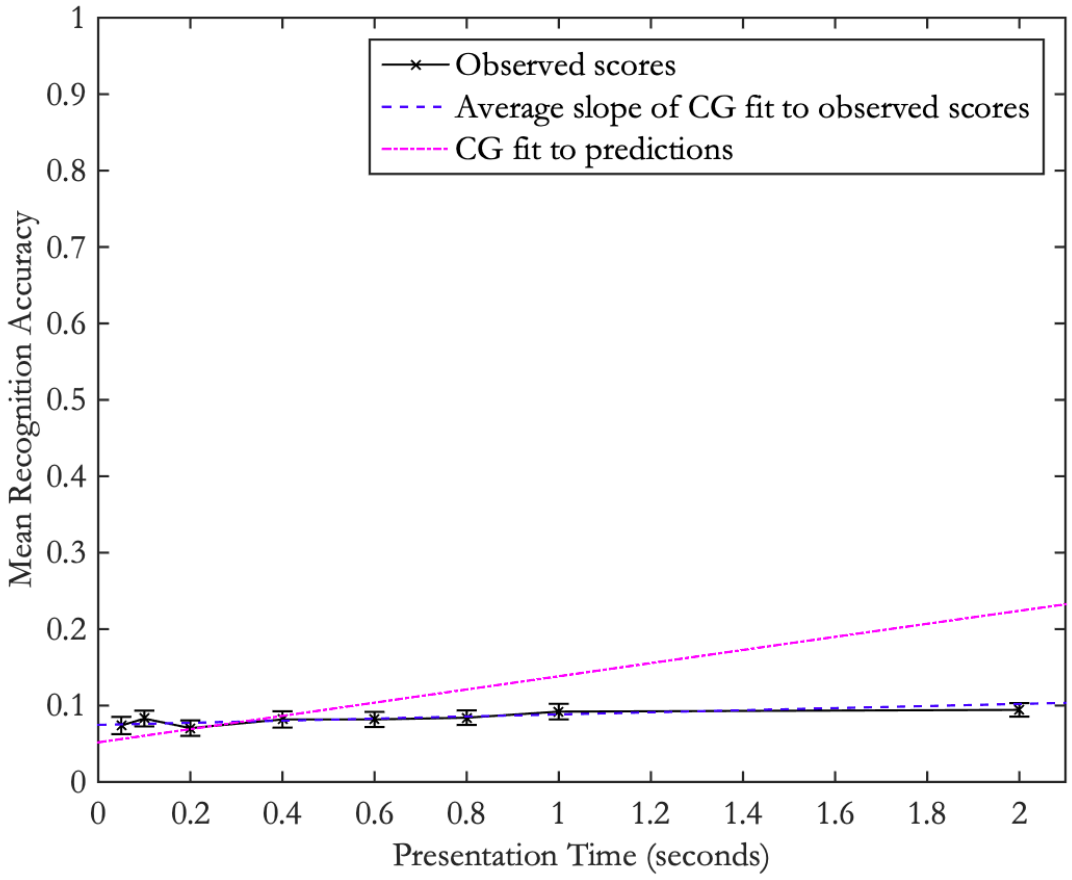

Figure 54. Observed accuracy means plotted with the function fit to predicted accuracy scores and the function with slope equal to the average slope of functions fit to individual participant means for the unoccluded, luminance contrast polarity and spatial frequency manipulation. Black line illustrates the observed accuracy means across subjects. The actual analysis involved fitting a function to each participant's individual means. The slopes of these functions, fit to individual participant means, were then submitted to a one sample t-test against the slope of the function fit to the predicted means. Here, individual participant means and the functions fit to them are not shown. Instead, the dashed blue line is a summary that reveals the average slope of functions fit to individual participant means (see text for more information). The slope of this fitted function is very similar to (and not significantly different from) the function fit to the predicted means (dashed and dotted pink line). The predicted means were based on the addition of the individual effects of the unoccluded (Experiment 4), luminance contrast polarity (Experiment 6), and spatial frequency (Experiment 8) manipulations. Error bars show 1 standard error above and below the observed condition means.

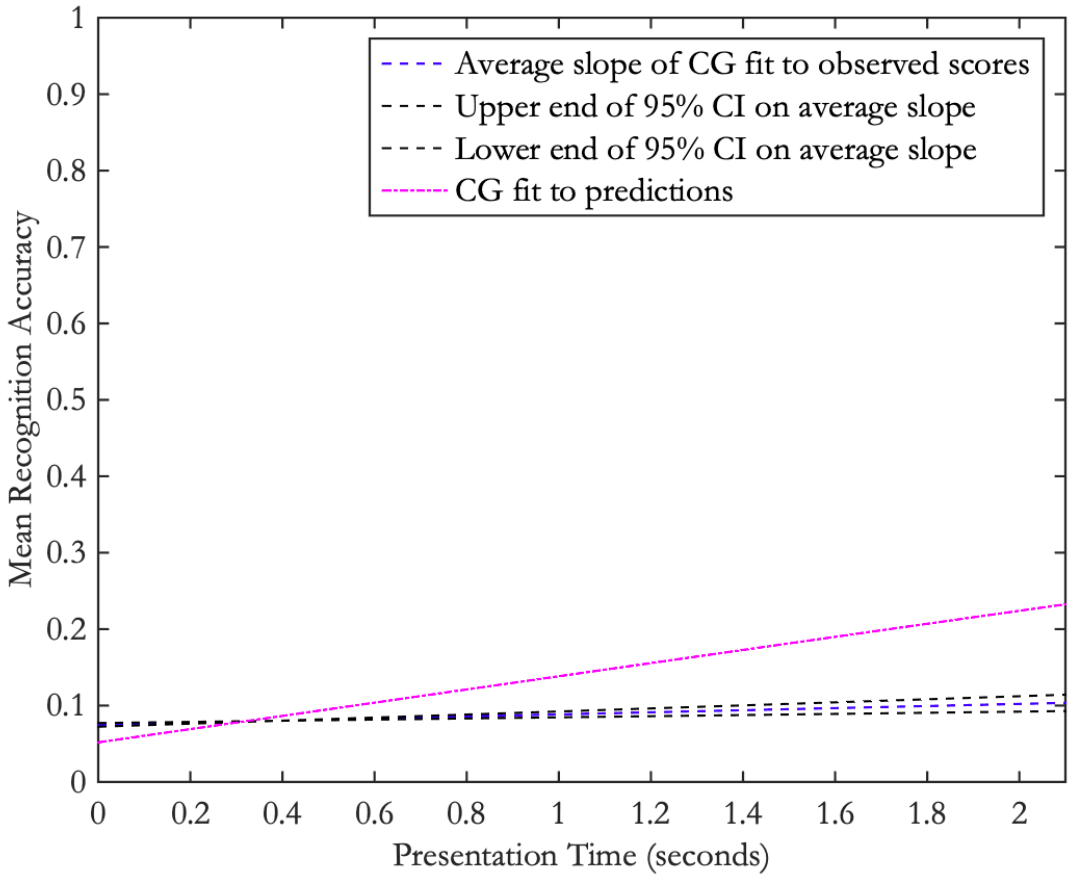

Figure 55. Function fit to the predicted accuracy means plotted within the confidence interval on the average slope of functions fit to observed individual participant mean for the combined unoccluded luminance contrast polarity and spatial frequency manipulation. The dashed and dotted pink line illustrates the function fit to the predicted accuracy means. The dashed blue line illustrates the average slope of functions fit to individual participant means. The dashed upper and lower black lines illustrate the 95% confidence interval on the average slope of functions fit to individual participant means.

4.7.3 Discussion

This experiment combined the individual cues manipulated in Experiments 4 (deletion of the occluder), 6 (luminance contrast polarity), and 8 (spatial frequency differences across fragments). Recognition accuracy for the combined manipulation was compared to a prediction based on the simple addition of the individual effects observed in Experiments 4, 6, and 8. The results revealed

that the prediction based on additivity is a very poor match to the observed data. In fact, recognition accuracy in the experimental condition was at chance levels at all presentation times (~8%). At the two second presentation time, performance was therefore 17% worse than expected. This is somewhat similar to the result observed for Experiment 12 - where only the unoccluded and the spatial frequency manipulation were combined. There, performance was roughly 20% worse than expected at the 2 second time point. At all other time points, performance was near chance. It seems possible, therefore, that anytime the unoccluded and spatial frequency manipulations are combined, an interpolated contour's perceptual strength is reduced to zero. That is, it may be the case that the unoccluded and spatial frequency manipulations, taken together, have the result of causing interpolated contours to be deleted in stage 2. In the following section, I will present a new analysis to examine the deletion hypothesis. I will then present a theory as to why the spatial frequency and unoccluded combinations result in a decrement to performance that goes beyond what additivity predicts.

4.8 General discussion

The goal of Experiments 9-12 was to determine how information from different cues is combined in the second stage. Each experiment compared recognition accuracy performance when multiple cues were combined to a prediction based on the simple additivity of the individual effects measured in Experiments 4, 6, and 8. That is, Experiments 4 (deletion of the occluder), 6 (luminance contrast polarity manipulation), and 8 (spatial frequency manipulation) examined the effects of individual cues and revealed how much performance drops, relative to the control condition, when a single cue is manipulated. The effects of individual cues, obtained in Experiments 4, 6, and 8, were used to obtain an estimate of the amount of performance decrease that would be expected when multiple cues are combined if the effects of the cues are simply added together in stage 2. These

predictions based on simple additivity were compared to actual observed performance in Experiments 9 (luminance contrast polarity & spatial frequency cues combined), 10 (luminance contrast polarity & unoccluded cues combined), 11 (spatial frequency & unoccluded cues combined), and 12 (luminance contrast polarity & spatial frequency & unoccluded cues combined). That is, in Experiments 9-12, multiple cues were combined, and the drop in performance, relative to the control condition (averaged across experiments), was measured. This drop in performance was compared to the predicted drop based on the simple additivity of the individual cue effects. The results revealed that the luminance contrast polarity cue combines in an additive fashion with either of the other (spatial frequency or unoccluded) cues. The other two cues, however, produce a drop in performance that is larger than additivity would predict any time that they are combined. That is, when just spatial frequency differences across fragments and the unoccluded cue are combined, recognition accuracy was lower by 22% than predicted at 2 seconds of presentation time. When all three cues are combined, performance drops to chance (8% accuracy) or very near chance levels at all presentation times, although simple addition of the individual cue effects predicts that recognition accuracy would be at 26% at 2 seconds.

4.8.1 Deletion? Or some beyond-additive drop?

One question that the analyses conducted in Experiments 12 and 13 leave open is how, specifically, the combination of the spatial frequency and unoccluded manipulations impact recognition accuracy performance in the character recognition task. The data reveal that, for both experiments in which these two manipulations are combined, performance drops to chance or very near chance levels. There is one exception: performance with a 2 second presentation time in Experiment 11 resulted in a mean 18% accuracy across subjects - this result is 10% higher than

chance. The question is how to interpret the information from this single data point. Performance at all other time points was near chance in the combined manipulation condition for both experiments.

It may be true that, any time these two manipulations are combined, the visual system simply deletes the interpolated edge entirely from the scene. In that case, performance would be expected to be at chance at all time points. However, it may also be the case that the result at the 2 second timepoint for Experiment 11 is no accident - and that the combination of these two cues does not result in deletion of an interpolated edge, but merely some reduction in the perceptual strength that is much greater than additivity would predict, but less than a drop to zero strength. The results of this set of experiments leave open this question. An examination of the cumulative gaussian fits with slope equal to the average observed slope for individual participant data fits in the two experiments suggests that the mean observed fitted slope in Experiment 12 (M = 0.09, SD = 0.06) is significantly steeper than the mean observed fitted slope in Experiment 13 (M = 0.02, SD = 0.02, $t(44) = 4.83$, $p < .001$, 95% CI: 0.04 - 0.09). A comparison of the fitted slopes is plotted in Figure 56 below. The same analysis, conducted with the data from the 2 second time point removed, suggests that the difference between the mean observed fitted slope in Experiments 12 (M = 0.09, SD = 0.07) and 13 (M = 0.03, SD = 0.05, is still significant ($t(44) = 2.93$, $p < .01$, 95% CI: 0.02 - 0.09). Though the mean observed accuracy scores in Experiment 12 are just slightly above chance (ranging from 10% to 12% for the 600ms, 800ms, and 1s presentation times), this difference relative to the performance at chance in Experiment 13 (8% to 9% at all presentation times) does result in a significant difference in the average slope of functions fitted to individual participant data. This suggests that performance for the spatial frequency and unoccluded combination was better than performance for the luminance contrast polarity, spatial frequency, and unoccluded combination. That is, the results suggest that the interpolations were not totally deleted in stage 2 anytime the spatial frequency and

unoccluded manipulations were combined - rather, the combination resulted in a drop in perceptual strength that was greater than additivity predicted, but was still less than a drop to zero.

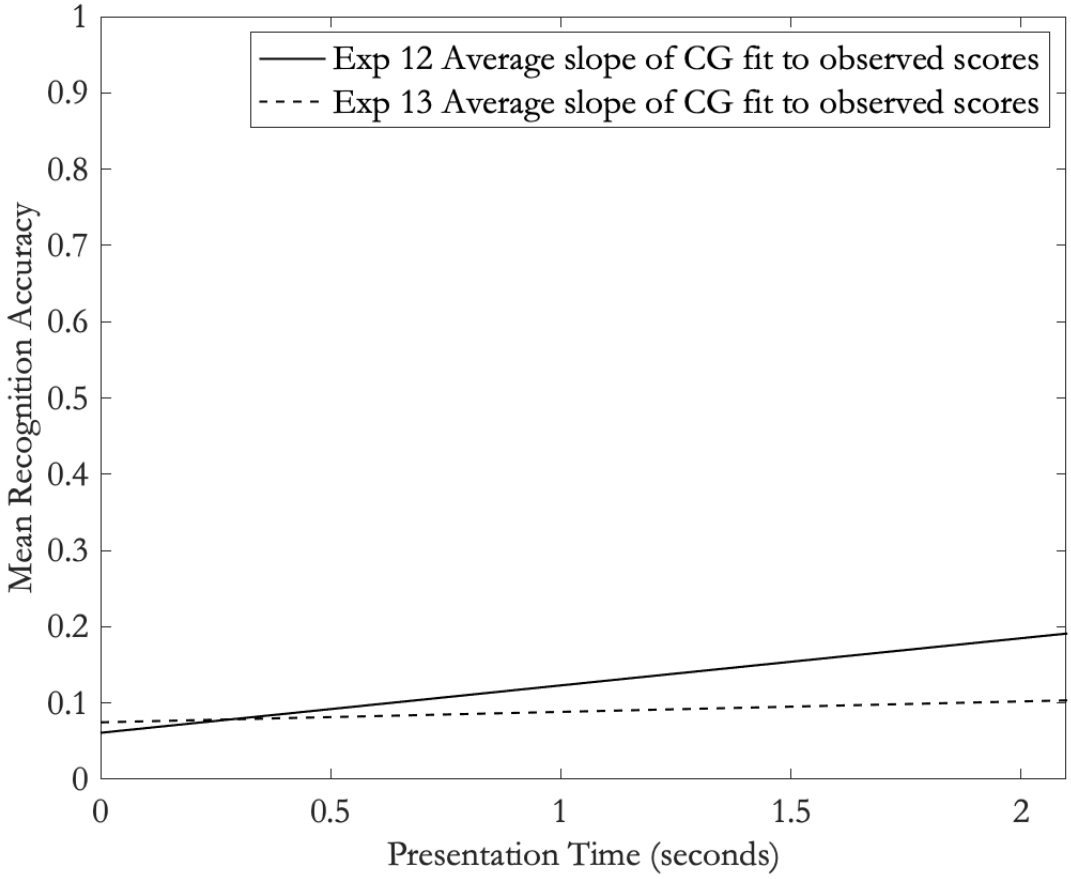

Figure 56. Comparison of average slope fits for Experiments 12 and 13. The function with slope equal to the mean observed slope fitted to individual participant data in Experiment 12 is illustrated by the solid black line. The function with slope equal to the mean observed slope fitted to individual participant data in Experiment 13 is illustrated by the dashed black line. The mean observed slope for Experiment 12 is significantly steeper than that of Experiment 13.

What is special about the combination of the unoccluded and spatial frequency manipulations? It is worth taking another look at the stimulus presented in Figure 46. One might notice that the crisp and blurry fragments tend to group together. However, to this observer, the grouping does not seem any different than what is observed for the crisp and blurry fragments in the occluded case, or for

the black and white fragments in the unoccluded or occluded luminance contrast polarity manipulations, or for the red and green fragments in the occluded equiluminant color contrast manipulation.

There does, however, seem to be some new information in the unoccluded spatial frequency stimulus. To this observer, the blurry fragments appear further away than the crisp ones. Taking a look at the original occluded spatial frequency stimulus (Figure 34), both the blurry and crisp fragments appear to lie immediately behind the occluder. When the occluder is removed, however, the crisp fragments appear to advance into the foreground while the blurry fragments fall behind. It is worth noting that relatability operates in 3-dimensions (Fantoni, Hilger, Gerbino & Kellman, 2008; Kellman, Garrigan, & Shipley 2005; Kellman, Garrigan, Shipley, Yin, & Machado, 2005), so any 3-dimensional depth differences across fragments would prevent interpolations from being produced in the first stage. Here, however, all sets of fragments are relatable in 3-dimensions. However, the scene cues considered in the second stage seem to convey additional information about the depth of the fragments. If this information is taken into consideration in the second stage, it may well be the case that the combination of the unoccluded and spatial frequency manipulations causes the fragments, in the second stage, to be estimated to lie at different depths. This could then cause the strengths of interpolations connecting those fragments to be reduced beyond what would be predicted by the individual strengths of those manipulations when manipulated in isolation.

CHAPTER 5

Pulling it all together

5.1. A Multi-stage theory

In the introduction to this dissertation, I proposed a two-stage account of the contour interpolation process. The processing in the first stage is governed by geometric relationships labeled *contour relatability* described by Kellman & Shipley (1991). These relationships determine whether a pair of edges is linked in an intermediate representation. The relatability geometry is applied only to edges that terminate in contour junctions, and, as a qualitative description, relatability is equivalent to the statement that only those pairs of edges that can be connected by a smooth, monotonic curve that bends through fewer than 90 degrees are linked. In the first stage, contours are interpolated indiscriminately across all candidate pairs (*promiscuous interpolation*). The interpolations across these candidate pairs are then evaluated in a second stage relative to information from a variety of scene cues. The information from the cues considered suggests something about the likelihood that an interpolated connection is connecting a pair of edges that are actually connected in the real world. The information from these cues is therefore used to determine which interpolations should be maintained, which should be deleted, and which should be strengthened or weakened. Figures 57 and 58 illustrate promiscuous interpolation across relatable edge fragments in a simple geometric scene, and a naturalistic scene, respectively.

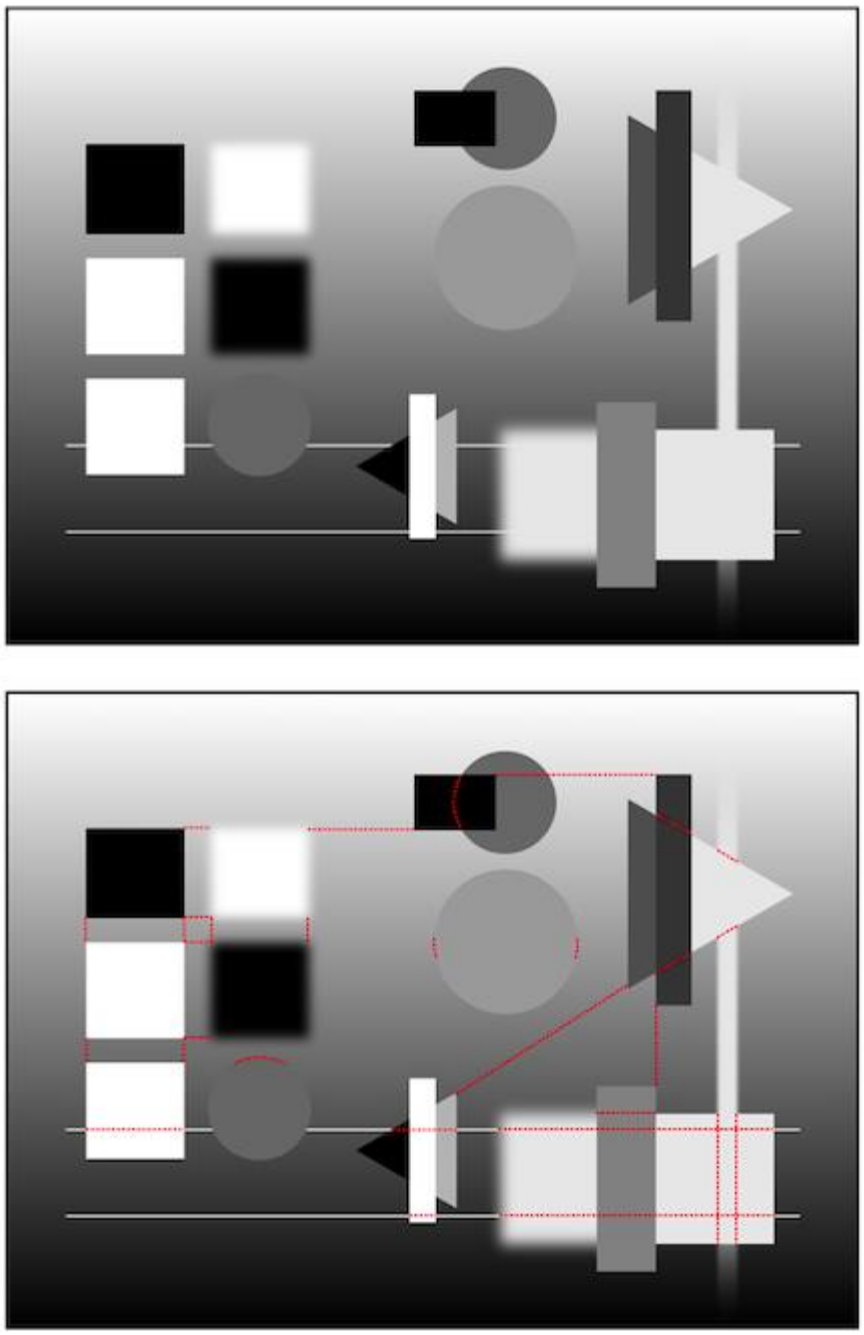

Figure 57. Promiscuous interpolations in a simple geometric scene. The top panel shows a simple geometric scene. The bottom panel shows some of the interpolations that are produced promiscuously in the first stage. Note that some of these interpolated edges connect edge fragments that should not be connected. Consideration of other scene cues in stage 2 will have the effect of pruning and down-weighting the bad interpolations and maintaining and increasing the perceptual strengths of the good ones.

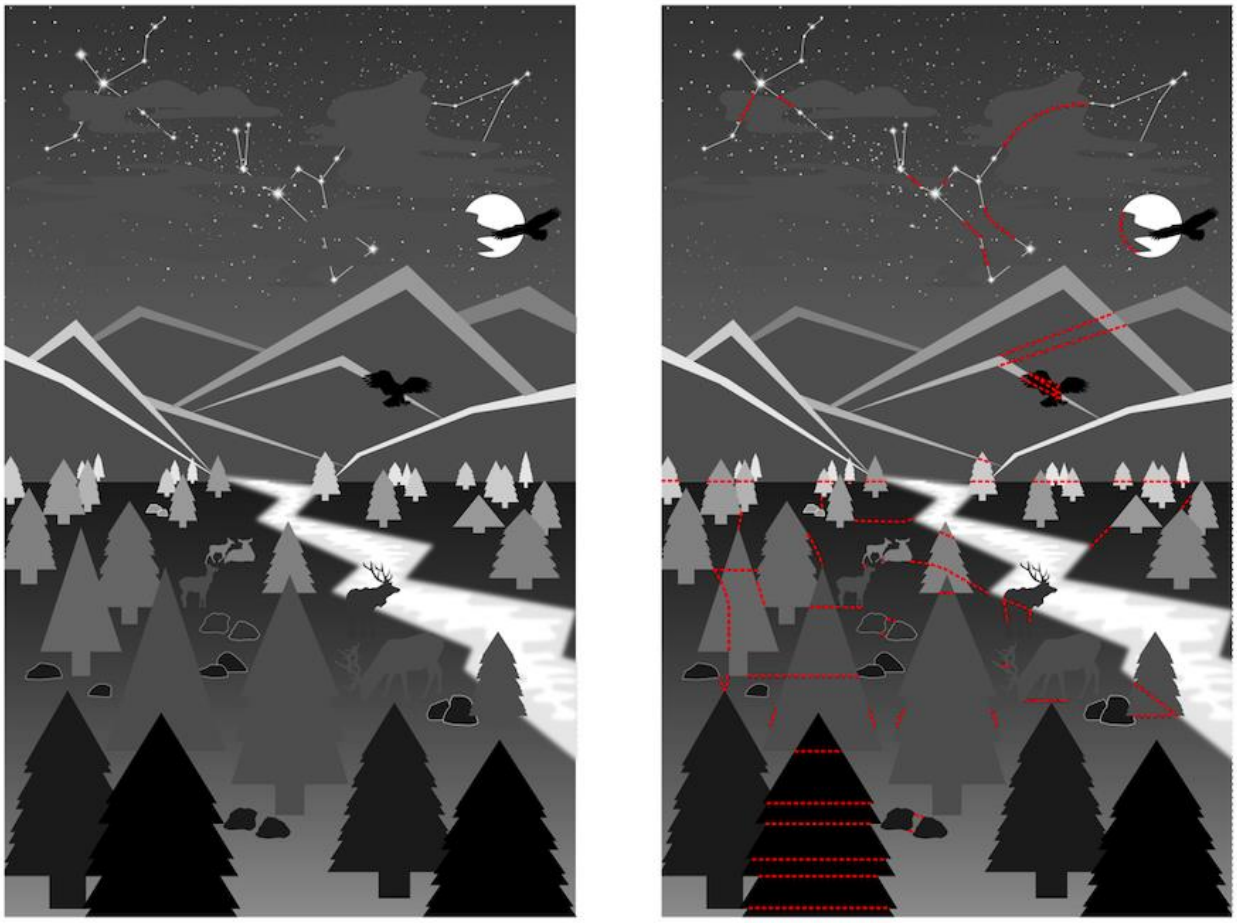

Figure 58. Promiscuous interpolations in a naturalistic scene. The left panel shows a simple naturalistic scene. The right panel illustrates just some of the interpolations that would be produced promiscuously in the first stage. What should be evident here is the fact that natural scenes are quite complex, and many more interpolations will inevitably be produced in the first stage than will ultimately appear in our phenomenological perception. Those that make it into our phenomenological perception will be those that were generally supported by the information from scene cues considered in the second stage. Interpolations that were contradicted by evidence from scene cues will be deleted or weakened in their perceptual strength.

The idea that contours are interpolated indiscriminately at an early stage is consistent with the fact that the truth of the objects in the scene probably cannot be known until the object descriptions have been built out of many real and interpolated edges. Likewise, the truth of which interpolated edges are good ones and which are spurious probably cannot be known until the truth of the objects in the scene is known. The promiscuous interpolation of contours in an early stage is also consistent

with other published work that has revealed the sensitivity to interpolated trajectories in the absence of any phenomenological perception of illusory or occluded contours (Field, Hayes, & Hess, 1993; Guttman & Kellman, 2001; Victor & Conte, 2000; Gold et al., 2000; Keane, Lu & Kellman, 2007). The first of these cited works is especially important to this dissertation.

5.2 Revealing the intermediate output of the first stage

In the path detection paradigm of Field, Hayes, and Hess (1993), paths composed of Gabor elements are embedded within a background of randomly arranged Gabor elements. Observers see two displays at separate intervals for each trial and are tasked with reporting which display contained a path. The original authors observed that detection of paths depended on inter-element orientation differences. In fact, the results of their Experiment 1 reveal that performance reached chance somewhere around 90 degrees of inter-element orientation difference. In addition, they reported results from an experiment that reveals that path detection is also better when paths do not contain inflection points. The authors noted that these results agree with the relatability geometry of Kellman & Shipley (1991), but they never suggested any explicit connection to contour interpolation, nor did they make clear how contour integration might be related to contour interpolation.

Performance in path detection tasks has since been investigated in numerous studies and has grown into a sizable literature. However, this dissertation contains the first explicit proposal of a specific connection to contour interpolation. The connection is this: contour integration in the path paradigm is really revealing the promiscuous interpolation of contours in the first stage of the interpolation process.

In Chapter 1, I reviewed the evidence suggesting that the contour integration evident in path displays is revealing the intermediate outputs of the first stage of the interpolation process. These

promiscuously interpolated connections confer some unity on the connected elements, but do not necessarily appear in our phenomenological perception. Whether they appear in our phenomenological perception depends on of stage 2 processing.

In Chapter 2, I tested the hypothesis that contour integration is the result of promiscuously interpolated connections. I did this by testing two predictions: the first was based on a relatively little known and somewhat unusual constraint on contour interpolation: a retinal tolerance for misalignments of about 15 arcminutes (Hilger, 2009). By "retinal", I mean that the amount of misalignment tolerated is based on the degree of visual angle rather than the size of the misalignment relative to the size of the edge fragments. This particular constraint on contour interpolation is somewhat unusual because retinal tolerances are fairly uncommon in visual perception, as, in most cases, the perception of objects does not change as an object gets nearer or further away. In this case, however, fragments can appear connected at further away distances (i.e. distances at which the misalignment of the fragments is less than 15 arcminutes) and unconnected at nearer distances (i.e. for which the misalignment of the fragments is 15 arcminutes or more). This particular constraint on contour interpolation has never been tested in contour integration. Therefore, one of the goals of Experiment 1 was to test whether contour integration is subject to the same, somewhat unusual retinal tolerance for misalignments of 15 arcminutes.

Another prediction was based on the idea that, if contour integration is revealing promiscuous interpolation, then, because no illusory or occluded edges are typically perceived in path displays, some stage 2 cue must be causing them to be deleted from the final scene representation. In typical path displays, there are no occluders, and there are surface differences between the Gabor elements and their grey background. That is, in order for illusory contours to be perceived, the contours would have to spread across the white regions of the Gabors, through the grey gap, and into the white regions of the next Gabor. Likewise, the black regions of the Gabors would have to spread

through the grey gap and into the black regions of the next element in the path. However, surface spreading is blocked in these displays by the presence of these surface discontinuities and the absence of any occluder to spread beneath. I hypothesized that it is this lack of amodal surface spreading that causes the interpolated connections to be deleted in stage 2. I also hypothesized that, if amodal surface spreading were enabled, illusory contours would be perceived to connect the elements within a path.

Experiment 1 utilized a path detection task with varying misalignment levels across inducing elements and also utilized multiple inducer types: some traditional 'Oreo' cookie-like black and white elements, and some elements for which the white centers had been turned grey to match the grey surround. Experiment 2 utilized a subjective rating task with paths composed of the same element types using the same levels of misalignment: observers were asked to report illusory contour clarity ratings for any illusory contours they perceived to connect the path elements. The results of these two experiments revealed that 1) path detection accuracy performance decreases continuously with increasing misalignment, reaching chance performance at 15 arcminutes of misalignment, 2) performance is nearly identical across all element types, regardless of the colors of their centers, 3) illusory contours are perceived when amodal surface spreading is enabled, and 4) perceptions of illusory contour clarity decrease in an identical fashion with increasing misalignment, also reaching a plateau at 15 arcminutes of misalignment.

These results strongly support the idea that the contour integration evident in path detection displays is the result of interpolated connections residing in an intermediate representation. They are also highly consistent with other work showing that disruptions to amodal surface spreading disrupt the perception of interpolated contours (He & Ooi, 1998a, He & Ooi 1998b; He, Ooi & Su, 2018; Su, He, & Ooi, 2010a; Su, He, & Ooi, 2010b).

5.3 Revealing the final outputs of the second stage

Having revealed the output of the first stage of the interpolation process, the experiments of Chapter 3 were designed to examine second stage outputs. This was done by testing the effects of various scene cue manipulations on the recognition of whole alphanumeric characters that had to be pieced together from physically specified fragments and interpolated edges. This task was expected to target second stage output because it requires the integration of many real and interpolated edges into a complete object description, and it should be the final outputs of the interpolation process that end up being utilized to build object descriptions.

Before presenting my experiments, I reviewed the evidence that certain cues are among those taken into consideration in the second stage. Among the cues I considered were three border ownership cues (moving an occluding object such that it is rotated and set behind fragments - preserving gaps between fragments, deleting an occluder while preserving gaps between fragments, and making unoccluded fragments outlines), disruptions to amodal surface completion (luminance contrast polarity differences across fragments relative to the background, and equiluminant color contrast differences across fragments), and manipulations to the spatial frequency of edge fragments that cause interpolations connecting those fragments to be produced in fewer spatial frequency channels. I then predicted that each of these cues would have an effect on recognition accuracy in my task utilizing displays for which a whole alphanumeric character could be built up out of physically specified fragments and interpolated edges. Observers were presented with control displays (utilizing partially occluded, filled black fragments) and displays in which a single scene cue was manipulated (for which recognition accuracy was expected to be reduced) for varying presentation times and asked to press a key to indicate what whole alphanumeric character had been presented in the stimulus.

The prediction that these manipulations would affect performance in this task was a somewhat bold one, as some of the manipulations discussed had previously been demonstrated not to affect performance in some objective tasks (among these were certain border ownership manipulations utilized in a fat thin task; manipulations to luminance contrast polarity in a fat thin task, portrait-landscape task, or path detection task; and manipulations of spatial frequency differences across inducing edges in path detection). However, the tasks in question were all similar to path detection in that they required sensitivity only to a single interpolated trajectory - the interpolation did not have to be integrated with other edges in the scene to form a complete object description. For this reason, I concluded that performance in the above-mentioned experiments relied on stage 1 outputs - that is, the promiscuous interpolations residing in the intermediate representation were sufficient to complete the task. Thus, the consideration of scene cues in stage 2, which affects our ultimate phenomenological perception, would not affect performance in such cases. I therefore predicted that, because building and recognizing a complete object description for a whole alphanumeric character composed of many interpolated and real edges should rely on stage 2 output, accuracy in recognizing the hidden alphanumeric character would be affected by each of the manipulations mentioned above. I also predicted that differences between the control and manipulated displays would appear between 100 and 200ms, as interpolation is known to take about 150ms to proceed.

The results confirmed the predictions in every case. Across all six experiments, manipulations that affected the phenomenological perception of interpolated contours had significant effects on recognition accuracy in my objective task. The results were also consistent with a hypothesis regarding the relative magnitudes of the effects of the border ownership cues. In Experiment 3, it was expected that the fragments would be most likely to own their own boundaries because no physically specified occluder was visible (an occluder could otherwise own some of the fragment boundaries), and an illusory occluder was impossible owing to the fact that a dark background object

could be seen lying on the medium grey background through the gaps between the black fragments. In Experiment 4, the occluder was deleted in the experimental condition, but an illusory occluder (with the same color as the grey background) could be supposed. The manipulation was therefore expected to produce a drop in performance, but the effect was expected to be smaller than that of the manipulation of Experiment 3. In addition, Experiment 5 manipulated unoccluded black outline fragments relative to unoccluded filled black fragments. In both cases, an illusory occluder interpretation was possible, so the effect of the manipulation relative to this control was expected to be the smallest of all three. The results revealed that the magnitude of the effects was exactly as predicted. Thus, cues that were expected to convey information that was more inconsistent with the possibility of a pair of edges really being connected resulted in a larger reduction in the perceptual strengths of interpolated edges.

All of these results suggest that border ownership cues, amodal surface spreading cues (including luminance contrast polarity differences relative to the background and equiluminant color contrast differences), and information about the spatial frequency channels an interpolated contour appears in, are all taken into consideration in the second stage to determine the ultimate perceptual strengths of promiscuously interpolated contours.

5.4 Clarifying conflicting results of earlier work

These results also help to clarify some conflicting evidence that is otherwise confusing: namely, some studies report that certain manipulations affect performance for tasks involving interpolated contour stimuli, while other studies report that the same manipulations have no effect on performance for a different task. The two-stage theory can explain the discrepancy: some tasks, like path detection, can be completed with the intermediate outputs of the first stage of the contour interpolation process. These tasks require relatively simple sensitivity to a single interpolated

trajectory and do not require the building up of complete object descriptions from numerous real and physically specified edges. In contrast, the task utilized in Chapters 3 and 4 required observers to build the description of a whole alphanumeric character from many physically specified edge fragments and the interpolations necessary to connect them. Likewise, the tasks of other researchers who have reported effects of some of the manipulations above rely on the perception of a whole complete object and depend on the successful integration of numerous real and interpolated edges. For example, the "illusory O" of He & Ooi requires the successful perception of numerous completed edges that are perceived to lie behind the illusory-O as well as the successful integration of numerous interpolated edges that connect the inducing edges to produce the illusory-O. It is also worth noting that some tasks, which have reported no effect of manipulations to luminance contrast polarity, were manipulating the luminances of inducers for illusory figure stimuli. In the case of illusory figures, the surfaces of the inducers do not spread - it is the surface of the illusory figure, whose surface has the same color and luminance properties as the background, that spreads. Manipulations to the surfaces of inducers would only be expected to affect surface spreading in the case of amodally completed figures.

5.5 Returning to Brown & Koch (1993, 2000)

These results also offer a new interpretation of the work of Brown & Koch (1993, 2000) discussed in the introduction to Chapter 3. These authors had previously investigated the effects of closure and the presence of an occluder on recognizing a letter or symbol in stimuli much like the ones used in Experiments 3 and 4 (with certain exceptions - see the careful controls added to the stimuli used here, discussed in section 3.5.1. - the method for Experiment 3). Certain problems with their stimuli were noted in the introduction to that chapter. They interpreted their results as indicating that closure, and not the presence of an occluder, causes the spatially separated fragments

to more easily be perceived as whole letter. However, in their hand-drawn stimuli, when an occluder was present, it did not lie perfectly adjacent to the fragment boundaries (instead, gaps were visible between the fragments and the occluder), so it could not be expected to own the fragment boundaries. The results of the experiments reported here suggest that the benefit conferred by an occluder is one of border ownership: when an occluder is present (or when an illusory occluder can be inferred), the fragments are less likely to be perceived to own their boundaries. If the fragments do not own their boundaries, interpolations connecting the fragments will receive higher perceptual strengths in stage 2.

5.6 Interpolation in multiple spatial frequency channels

The results of one particular experiment reported in Chapter 3 are interesting for another reason. Experiment 8 examined the effect of applying a Gaussian blur to half of the fragment outlines so that their edges would be picked up only in relatively low spatial frequency channels. This was done to investigate the possibility that interpolation proceeds separately in separate channels, and that corroboration across multiple channels is taken into consideration in stage 2. The idea that the visual system utilizes information about the corroboration of edges to existing in multiple channels was proposed by Marr (1982). The idea is that edges existing in a wider range of channels are more likely to result from the boundaries of an object, and are less likely to be, for example, shadows (edges existing only in relatively low spatial frequency channels) or scratches upon a surface (edges existing only in relatively high spatial frequency channels). If the visual system utilizes information about what channels physically specified edges appear in, it seems likely it would also take this information into consideration for interpolated edges to determine which interpolations are good ones and which ones are spurious.

Some work from the path paradigm has already indicated that interpolation in stage 1 proceeds from non-symbolic inputs - that is, the inputs are *not* symbolic, abstracted edge tokens, but rather are contrast differences of a particular spatial frequency (Dakin & Hess, 1998; Dakin & Hess, 1999). If this is true, then the interpolations produced in stage 1 will need to be reconciled at a later stage. Other work has already suggested that interpolation proceeds separately in different channels (Guttman, Sekuler, & Kellman, 2003; Guttman & Kellman, 2004), and that interpolations existing in relatively few channels appear in the final scene representation with relatively low perceptual strengths (Guttman, Sekuler, & Kellman, 2003).

The results of Experiment 8 add to this earlier work to support the idea that interpolation proceeds from non-symbolic input consisting of contrast differences in a given spatial frequency channel, and that corroboration of promiscuously interpolated edges across channels is taken into consideration in the second stage to help determine which interpolations will be maintained, which will be deleted, and which will be strengthened or reduced.

5.7 Not in stage 1

It is worth mentioning that the effects of the manipulations presented in Chapter 3 cannot be explained by first stage considerations. A variety of evidence supports this idea. In the first place, the inducing edges, in all cases, were identical for both control and experimental stimuli - the manipulations to scene cues never altered the inducing edges (except in the spatial frequency case, which will be discussed in a moment). Relatability of the fragments was, therefore, unaffected.

In addition, other experiments suggest that manipulations to border ownership cues do not affect stage 1 processing (Guttman & Kellman, 2001), and that manipulations to luminance contrast polarity relative to the background do not affect stage 1 processing (Field, Hayes, & Hess, 1993 - see Experiment phase manipulation, which causes black edges on a grey background to align with white

edges). Likewise, the experiments presented in Chapter 2 suggest that disruptions to amodal surface completion do not affect stage 1 processing.

In the case of the spatial frequency manipulation, relatability should not have been disrupted at lower spatial frequencies. The inducing edges should have appeared in low spatial frequency channels, and, if interpolation receives symbolic edge tokens that are an abstraction over the edges existing in different spatial frequency channels, then interpolation should have preceded identically for both the control and experimental stimuli of Experiment 8. There should have been no effect of the manipulation, as interpolation would have proceeded from symbolic edges - there would be no opportunity to reconcile interpolations produced in different channels or reduce interpolation strengths based on information about the spatial frequency channels an interpolation appeared in.

Another interpretation is possible: perhaps the spatial frequency manipulation prevented interpolation from occurring in stage 1. Then it would have to be explained why edges existing in low spatial frequency channels cannot give rise to interpolated contours. The undulating, fuzzy interpolation illustrated in Figure 33 suggests that this is not the case. The interpolated edges visible in this figure can only have been detected in relatively low spatial frequency channels.

Yet another possibility remains: perhaps edge information across channels is integrated prior to stage 1 processing. And perhaps information about what channels an edge appeared in determines which edge pairs are candidates for interpolation in stage 1. That is, perhaps the visual system took note of the fact that a pair of relatable edge fragments was detected in relatively low spatial frequency channels, but that only one of these edges appeared in relatively high spatial frequency channels. And perhaps this information was utilized to prevent interpolation across abstractions of those edges from occurring in the first place. Then it would have to be explained why work from the path paradigm indicates that contour integration, and therefore, promiscuous interpolation, does not proceed from symbolic, abstracted edge tokens (Dakin & Hess, 1998; Dakin & Hess, 1999).

5.8 Other cues that may play a role in second stage processing

Other cues, not investigated here, may also play a role in second stage processing. Other border ownership manipulations - for example cues to convexity, and other cues that block surface spreading, for example, differences in texture, might affect the final perceptual weights of interpolated contours. Likewise, information about the potential symmetry of an interpolated edge with another real or interpolated edge may have the effect of strengthening the perceptual strength of that interpolation. Figure 11 makes clear that, while information about symmetry is insufficient to give rise to an interpolation in the first place (this point is also supported by empirical evidence - see Carrigan, Palmer, & Kellman, 2016), information about symmetry can play a role in strengthening or weakening interpolated edges that resulted from promiscuous interpolation across relatable edge fragments. The figure also suggests that closure may play a role in second stage considerations. Other work has suggested that cues as to parallelism have a considerable impact on the perception of interpolated contours (Gillam, 1987; Albert, 1993; Gillam & Chan, 2002). The effects of parallelism, too, would be determined in the second stage.

5.9 Other evidence that supports a multi-stage theory

Although this dissertation offers the first suggested explicit connection between the contour integration literature and contour interpolation, it is not the first time a two-stage theory of contour interpolation has been proposed (Kellman, Guttman, & Wickens, 2001; Kellman, Garrigan, & Shipley, 2005). In fact, much earlier work has revealed the fact that interpolation proceeds first, and afterward, the appearance of an interpolated edge as either modal or amodal - depending on its perceived depth relative to other objects in the scene - is determined (Kellman & Shipley, 1991;

Shipley & Kellman, 1992a; Kellman, Yin, & Shipley, 1998; Shipley & Kellman, 1992; Kellman, Yin, & Shipley, 1998). A variety of empirical evidence and logical arguments support this conclusion and is discussed in detail in section 2.6.3 of the general discussion of Chapter 2.

5.10 Combinations of information from multiple cues in stage 2

One important question that the experiments presented in Chapter 3 leave open is what happens when multiple scene cues weigh in on the final perceptual strength of an interpolated connection in stage 2. That is, how is the information conveyed by multiple cues integrated? Answering this question was the goal of the experiments presented in Chapter 4. Multiple possibilities were discussed. The simplest solution might be to disregard the information from some cues and allow the information from one cue to win out. Another simple solution would be to simply add the information conveyed by the different cues without taking into consideration which cues were providing the information. Alternatively, the visual system might seek to weight information from different cues differently depending on whether the information conveyed by multiple cues is independent (i.e., non-overlapping), or based on which cues are most reliable or most accurate.

Three cues were selected to be combined in a series of four experiments presented in Chapter 4. One border ownership cue (deleting the occluder), one disruption to amodal surface completion (luminance contrast polarity differences across fragments relative to the background), and one manipulation that causes interpolations to be produced only in relatively low spatial frequency channels (spatial frequency differences across fragments). The magnitudes of the effects of the combined manipulations of the experiments in Chapter 4 were calculated and compared to predictions based on the simple additivity of the individual effects measured in the experiments in Chapter 3. Somewhat surprisingly, the results for two of the experiments were very well described by the simple addition of the combined effects of the individual manipulations, as measured in

Chapter 3. The two experiments in question examined the combination of the luminance contrast polarity manipulation with either the unoccluded manipulation or the spatial frequency manipulation. That is, the luminance contrast polarity manipulation combines in a simple additive fashion with either of the other cues.

Something beyond additivity happens when you combine the other two cues with each other, however. Any time the unoccluded and spatial frequency manipulations appeared together, recognition accuracy performance was reduced to near chance levels. There was a single exception: the results for the 2 second presentation time in Experiment 11 revealed recognition accuracy performance that was greater than chance by 10%.

5.11 Deletion? Or something else?

A post-hoc analysis, presented in the general discussion for Chapter 4 compared the average slopes observed for functions fitted to individual participant data in Experiments 12 and 13. The slopes are significantly different. The same analysis conducted without the 2 second timepoint reveals that the average observed slope for functions fitted to individual participant data is still significantly different even without data from the two second time point. This suggests that interpolations were not simply deleted any time the spatial frequency and unoccluded manipulations co-occurred. Rather, the combination of those two manipulations resulted in a decrease in the perceptual strength of interpolated contours that was greater than additivity would have predicted, but which was still less than a drop to zero perceptual strength.

5.12 What is special about the unoccluded and spatial frequency combination?

The above result begs the question: what is special about the unoccluded and spatial frequency combination? That is, why is it the case that the combination of these two cues results in performance decrements that go beyond the simple addition of the individual effects of the cues?

One interpretation is that both of these manipulations produce changes in the depth of the fragments. That is, the presence of an occluder is itself a depth cue - when the occluder is present, both the blurry and crisp fragments appear to lie immediately behind it (see Figure 34 for an illustration of the occluded spatial frequency manipulation). However, when the occluder is deleted (see Figure 47 for an illustration of the unoccluded spatial frequency manipulation), the crisp fragments appear to advance into the foreground while the blurry fragments fall behind. The crisp fragments therefore appear nearer to the observer than the blurry fragments.

It is important to note that relatability operates in 3D ((Fantoni, Hilger, Gerbino & Kellman, 2008; Kellman, Garrigan, & Shipley 2005; Kellman, Garrigan, Shipley, Yin, & Machado, 2005), and that no pairs of edge fragments would make it into the first stage if they were not relatable in 3D. Here, however, both sets of fragments are relatable in 3D, and so interpolation in the first stage can be expected to proceed. However, if 2D scene cues considered in the second stage convey additional information about the depth of fragments, it may be the case that this information is utilized to adjust the perceptual strengths of the interpolations.

5.13 Future directions

The work presented here suggests that interpolation proceeds in multiple stages, that, at an early stage, interpolated contours are produced promiscuously across all pairs of relatable edge fragments, and that, at a later stage, information from multiple scene cues is taken into consideration to adjust the perceptual weights of these interpolated edges. However, many research opportunities remain. Specifically, this work has not addressed the variable strength of stage 1 outputs. It seems very likely

that interpolations exiting stage 1 come packed with initial perceptual strengths based on the support ratio of the interpolated edge (the ratio of physically specified to total edge length) and the nearness of the interpolated edge to violating the 90 degree constraint. The effect of both of these variables on perceptual strengths of interpolated contours residing in the intermediate representation is worth investigating. Here, stimuli were controlled such that variables that affect relatability (such as the 90 degree constraint above) and support ratio were never varying across the stimuli examined. In fact, the inducing edges were identical in all stimulus conditions. Future work should examine stimuli in which these variables are manipulated to determine whether and how they affect initial interpolation strengths. In addition, the effect of many other cues on stage 2 processing (e.g. parallelism, closure, symmetry, differences in surface texture, other border ownership cues such as convexity) have yet to be examined or modeled. Investigating the effects of other cues would be particularly worthwhile in the context of cue combination. Doing so could reveal which cues are combined in an additive fashion, which are not, and help clarify how interpolation strengths are determined when the combined effect is non-additive. Finally, future work should explore the development and testing of theories as to why additivity holds true for some cue combinations but not others.

APPENDIX A

Rating tests for Experiment 1b: Subjective ratings of illusory contours

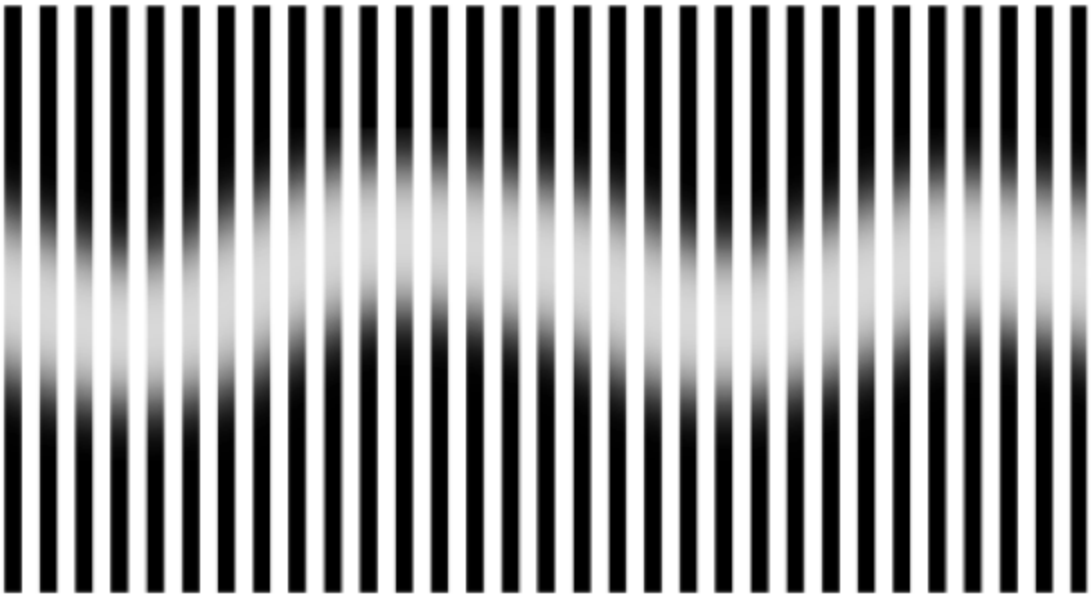

Figure A1: Image with fuzzy illusory contours.

Figure A2: Image with no illusory contours.

REFERENCES

Albert, M. K. (1993) Parallelism and the perception of illusory contours. *Perception, 22*, 589-595.

Albert, M. K. (2001). Cue interactions, border ownership and illusory contours. *Vision Research, 41*(22), 2827-2834.

Albert, M. K. (2007). Mechanisms of modal and amodal interpolation. *Psychological Review, 114,* 455-469.

Anstis, S. M., and Cavanagh, P. C. (1983) A minimum motion technique for judging equiluminance. In *Colour Vision: Physiology and Psychophysics*, J. Mollon and R. T. Sharpe, eds., pp. 155-166, Academic, New York.

Bar, M. (2003). A cortical mechanism for triggering top-down facilitation in visual object recognition. *Journal of Cognitive Neuroscience, 15*, 600–609.

Ben-Yosef, G., & Ben-Shahar, O. (2012). Tangent bundle curve completion with locally connected parallel networks. *Neural Computation, 24*, 3277-3316.

Bilodeau, L., & Faubert, J. (1997). Isoluminance and chromatic motion perception throughout the visual field. *Vision research, 37*(15), 2073-2081.

Blakemore, C., & Campbell, F. W. (1969). On the existence of neurons in the human visual system selectively sensitive to the orientation and size of retinal images. The Journal of Physiology, 203(1), 237-260.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*, 433-436.

Bregman, A. S. (1981). Asking the "what for" question in auditory perception. In M. Kubovy and J. R. Pomerantz (eds.), *Perceptual Organization*. Hillsdale, N.J.: Erlbaum.

Bregman, A. (1990). Auditory Scene Analysis. MIT Press, Cambridge, MA.

Brown, J. M., & Koch, C. (1993). Influences of closure, occlusion, and size on the perception of fragmented pictures. *Perception & Psychophysics, 53*(4), 436-442.

Brown, J. M., & Koch, C. (2000). Influences of occlusion, color, and luminance on the perception of fragmented pictures. *Perceptual and Motor Skills, 90*(3), 1033-1044.

Bullier, J. (2001). Integrated model of visual processing. *Brain Research Reviews, 36*, 96–107. doi: 10.1016/S0165-0173(01)00085-6

Campbell, F. W., & Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *The Journal of Physiology, 197*, 551-566.

Carandini, M., Demb, J. B., Mante, V., Tolhurst, D. J., Dan, Y., Olshausen, B. A., Gallant, J. L., & Rust, N. C. (2005). Do we know what the early visual system does? The Journal of Neuroscience, 25(46).

Carrigan, S. C., Palmer, E. M., & Kellman, P. J. (2016). Differentiating global and local contour completion using a dot localization paradigm. *Journal of Experimental Psychology: Human Perception and Performance, 42*(12), 1928-1946.

Carrigan, S., & Kellman, P. (2016). Differentiating Local and Global Processes in Amodal Completion: Dot Localization with Familiar Logos. *Journal of Vision, 16*(12), 310-310. [abstract]

Carrigan, S., & Kellman, P. (2017). Separating perception and recognition in amodal completion: Dot localization with regular patterns. *Journal of Vision, 17*(10), 1370-1370. [abstract]

Cavanagh, P., MacLeod, D. I A., & Anstis, S. M. (1987). Equiluminance: Spatial and temporal factors and the contribution of blue-sensitive cones. *Journal of the Optical Society of America A, 4*(8), 1428-1438.

Dakin, S. C., & Hess, R. F. (1998). Spatial-frequency tuning of visual contour integration. *Journal of the Optical Society of America A*, *15*(6), 1486-1499.

Dakin, S. C., & Hess, R. F. (1999). Contour integration and scale combination processes in visual edge detection. *Spatial vision, 12*(3), 309-328.

Demeyer, M., & Machilsen, B. (2012). The construction of perceptual grouping displays using GERT. *Behavior Research Methods, 44*(2), 439-446.

De Valois, R. L., Albrecht, D. G., & Thorell, L. G. (1982). Spatial frequency selectivity of cells in macaque visual cortex. Vision Research, 22, 545-559.

De Valois, R. L., & De Valois, K. K. (1980). Spatial vision. *Annual Review of Psychology, 31*, 304-341.

De Valois, R. L., & De Valois, K. K. (1990). *Spatial vision*. Oxford University Press, New York.

De Valois, K. K., De Valois, R. L., & Yund, E. W. (1979). Responses of striate cortex cells to grating and checkerboard patterns. The Journal of Physiology, 291, 483-505.

Dresp, B., & Grossberg, S. (1997). Contour integration across polarities and spatial gaps: From local contrast filtering to global grouping. *Vision Research, 37*(7), 913-924.

Fantoni, C., Hilger, J. D., Gerbino, W., & Kellman, P. J. (2008). Surface interpolation and 3D relatability. *Journal of Vision, 8*(29).

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.

García-Pérez, M. A., & Peli, E. (2000). Luminance artifacts of cathode-ray tube displays for vision research. *Spatial Vision, 14*(2), 201-215.

Geisler, W. S. & Perry, J. S. (2009). Contour Statistics in natural images: Grouping across occlusions. *Vision Neuroscience, 26*(1), 109-121.

Gillam, B. (1987). Perceptual grouping and subjective contours. In: Petry, S., & Meyer, G. E. (Eds). *The Perception of Illusory Contours.* (pp. 268-273).

Gillam, B., & Chan, W. M. (2002). Grouping has a negative effect on both subjective contours and perceived occlusion at T-junctions. *Psychological Science, 13*(3), 279-283.

Ginsburg, A. P. (1984). Visual form perception based on biological filtering. Sensory experience, adaptation, and perception: Festschrift for Ivo Kohler, 53-72.

Ginsburg, A. P. (1986). Spatial filtering and visual form perception. In Boff, K. R., Kauffman, L, & Thomas, J. P. (Eds.), Handbook of Perception and Human Performance. New York: Wiley.

Gold, J. M., Murray, R. F., Bennett, P. J., & Sekuler, A. B. (2000). Deriving behavioral receptive fields for visually completed contours. *Current Biology, 10*(11), 663-666.

Graham, N. (1980). Spatial-frequency channels in human vision: Detecting edges without edge detectors. In C. H. Harris (Ed.), Visual coding and adaptability. Hillsdale, NJ: Lawrence Erlbaum.

Graham, N. V. S. (1989). Visual pattern analyzers (Vol. 16). Oxford University Press.

Grossberg, S., & Mingolla, E. (1985). Neural dynamics of form perception: Boundary completion, illusory figures, and neon color spreading. *Psychological Review, 92*(2), 173-211.

Gurnsey, R., Poirier, F. J., & Gascon, E. (1996). There is no evidence that Kanizsa-type subjective contours can be detected in parallel. *Perception, 25*(7), 861-874.

Guttman, S.E., & Kellman, P. J. (2001). Contour interpolation: Necessary but not sufficient for the perception of interpolated contours [Abstract]. Journal of Vision, 1(3): 384.

Guttman, S. E., & Kellman, P. J. (2004). Contour interpolation revealed by a dot localization paradigm. *Vision Research, 44*(15), 1799-1815.

Guttman, S. E., & Kellman P. J. (in prep). Promiscuous interpolation: Contour completion without perception of completed contours.

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.

Häkkinen, J., Liinasuo, M., Kojo, I., & Nyman, G. (1998). Three-dimensionally slanted illusory contours capture stereopsis. *Vision Research, 38*(20), 3109-3115.

He, Z. J., & Ooi, T. L. (1998a). Illusory-contour formation affected by luminance contrast polarity. *Perception, 27*, 313-335.

He, Z. J., & Ooi, T. L. (1998b). Luminance contrast polarity: a constraint for occlusion and illusory surface perception. *Investigative Ophthalmology and Visual Science, 39*(Suppl.), 848.

He, Z. J., Ooi, T. L., & Su, Y. R. (2018). Perceptual mechanisms underlying amodal surface integration of 3-D stereoscopic stimuli. *Vision Research, 143*, 66-81.

Hegde, J. (2008). Time course of visual perception: Coarse-to-fine processing and beyond. *Progress in Neurobiology, 84*, 405–439.

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*(6-7), 407-421.

Hess R., & Field, D. (1999). Integration of contours: New insights. *Trends in Cognitive Sciences, 3*(12), 480-486.

Hilger, J. D. (2009). Contour integration and interpolation: Geometry, phenomenology, and multiple inputs (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses. (Accession Order No. 3394906).

Hoffman, D. D. & Richards, W. A. (1984). Parts of recognition. *Cognition, 18*, 65-96.

Johnson, J. S., & Olshausen, B. A. (2005). The recognition of partially visible natural objects in the presence and absence of their occluders. *Vision Research, 45*(25-26), 3262-3276.

Julesz, B., & Schumer, R. A. (1981). Early visual perception. Annual Review of Psychology, 32, 575-627.

Kalar, D. J., Garrigan, P., Wickens, T. D., Hilger, J. D., Kellman, P. J. (2010). A unified model of illusory and occluded contour interpolation. *Vision Research, 50*, 284-299.

Kanizsa, G. (1979). *Organization in Vision*. New York, NY: Praeger Publishers.

Kanizsa, G., & Gerbino, W. (1982). Amodal completion: Seeing or thinking? In: *Organization and representation in perception*, ed. B. Beck. Erlbaum.

Keane, B. P., Lu, H., & Kellman, P. J. (2007). Classification images reveal spatiotemporal contour interpolation. *Vision Research, 47*, 3460-3475.

Kellman, P. J. (2003a). Visual perception of objects and boundaries: A four-dimensional approach. In R. Kimchi, M. Behrmann, & C. R. Olson (Eds.), *Perceptual organization in vision: Behavioral and neural perspectives*. Mahwah, NJ: Lawrence Erlbaum Associates.

Kellman, P. J. (2003b). Interpolation processes in the visual perception of objects. *Neural Networks, 16*, 915-923.

Kellman, P. J. (2003c). Perceptual processes that create objects from fragments. *Proceedings of the 2003 IEEE International Joint Conference on Neural Networks, 1*, 2-7.

Kellman, P. J., & Cohen, M. H. (1984). Kinetic subjective contours. *Perception & Psychophysics, 35*(3), 237-244.

Kellman, P. J., Garrigan, P., & Erlikhman, G. (2013). Challenges in understanding visual shape perception and representation: Bridging subsymbolic and symbolic coding. *In Shape Perception in Human and Computer Vision* (pp. 249-274). Springer, London.

Kellman, P. J., Garrigan, P., & Palmer, E. M. (2010). 3D and spatiotemporal interpolation in object and surface formation. In C. W. Tyler (Ed.), *Computer vision: From surfaces to 3D objects* (pp. 183-207), Boca Raton, FL: Chapman & Hall/CRC.

Kellman, P. J., Garrigan, P., & Shipley, T. F. (2005). Object interpolation in three dimensions. *Psychological Review, 112*, 586 – 609.

Kellman, P. J., Garrigan, P., Shipley, T. F., & Keane, B. P. (2007). Interpolation processes in object perception: Reply to Anderson. *Psychological Review, 114*(2), 488-508.

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*.

Kellman, P. J., Guttman, S., & Wickens, T. (2001). Geometric and neural models of contour and surface interpolation in visual object perception. In T. F. Shipley & P. J. Kellman (Eds.), *From fragments to objects*: *Segmentation and grouping in vision* (pp. 193–245). London: Elsevier.

Kellman, P. J., & Shipley, T. F. (1991). A theory of visual interpolation in object perception. *Cognitive Psychology, 23*, 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), 859-869.

Kogo, N., & Wagemans, J. (2013). The emergent property of border-ownership and the perception of illusory surfaces in a dynamic hierarchical system. *Cognitive Neuroscience, 4*(1), 54-61.

Koffka, K. (1938). Beitra¨ge zur Psychologie der Gestalt. III. Zur Grundlegung der Wahrnehmungspsychologie. Eine Auseinandersetzung mit V. Benussi [Contributions to Gestalt psychology: III. Toward a foundation of the psychology of perception—A debate with V. Benussi]. In W. D. Ellis (Ed.), A source book of Gestalt psychology (pp. 371–378). New York: Harcourt, Brace. (Reprinted from Zeitschrift für Psychologie, 73, 11–90, 1915)

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.

Livingstone, M. S., & Hubel, D. H. (1987). Psychophysical evidence for separate channels for the perception of form, color, movement, and depth. *Journal of Neuroscience, 7*(11), 3416-3468.

Maertens, M., & Shapley, R. (2008). Local determinants of contour interpolation. *Journal of Vision, 8*(3).

Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information.* New York, NY: W.H. Freeman and Company.

Marr, D. & Hildreth, E. (1980). Theory of edge detection. *Proceedings of the Royal Society of London B, 207*(1167), 187-217.

Michotte, A., Thinès, G., & Crabbé, G. (1991). Amodal completion of perceptual structures (E. Miles & T. R. Miles, Trans.) In G. Thinès, A. Costall, & G. Butterworth (Eds.), *Michotte's experimental phenomenology of perception* (pp. 140-167). Hillsdale, NJ: Erlbaum. (Original work published 1964).

Morrone, M. C., & Burr, D. C. (1988). *Proceedings of the Royal Society of London, B*, 235, 221-245.

Mullen, K. T. (1991). Colour vision as a post-receptoral specialization of the central visual field. *Vision research, 31*(1), 119-130.

Murray, M. M., Foxe, D. M., Javitt, D. C., & Foxe, J. J., (2004). Setting boundaries: Brain dynamics of modal and amodal illusory shape completion in humans. *Journal of Neuroscience, 31*, 6898-6903.

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

Palmer, E. M., Kellman, P. J., & Shipley, T. F. (2006). A theory of dynamic occluded and illusory object perception. *Journal of Experimental Psychology: General, 135*(4), 513-541.

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

Petter, G. (1956). Nuove ricerche sperimentali sulla totalizzazione percettiva. *Rivista di Psicologia, 50*, 213–227.

Poggio, G. F. (1972). Spatial properties of neurons in striate cortex of unanesthetized macaque monkey. Invest. Ophthalmol. 11, 368–377.

Powell, I. (1981). Lenses for correcting chromatic aberration of the eye. *Applied Optics, 20*(24), 4152-4155.

Prazdny K. (1983). Illusory contours are not caused by simultaneous brightness contrast. *Perception & Psychophysics*, 34, 403–404.

Regolin, L., & Vallortigara, G. (1995). Perception of partly occluded objects by young chicks. *Perception & Psychophysics, 57*(7), 971-976.

Ringach, D. L., & Shapley, R. (1996). Spatial and temporal properties of illusory contours and amodal boundary completion. *Vision Research, 36*(19), 3037-3050.

Sachs, M., Nachmias, J., & Robson, J. G. (1971). Spatial-frequency channels in human vision. Journal of the Optical Society of America, 61(9).

Sambin, M. (1987). A dynamic model of anomalous figures. In: Petry, S., & Meyer, G. E. (Eds). *The Perception of Illusory Contours.* (pp. 131-142).

Schyns, P. G., & Oliva, A. (1994). From blobs to boundary edges: Evidence for time-and spatial-scale-dependent scene recognition. Psychological science, 5(4), 195-200.

Shapley, R., & Gordon, J. (1983). A nonlinear mechanism for the perception of form. *Investigative Ophthalmology and Visual Science*, *24*(Suppl.), 238.

Shapley, R., & Gordon, J. (1985). Nonlinearity in the perception of form. Attention, Perception, & Psychophysics, 37(1), 84-88.

Shipley, T. F., & Kellman, P. J. (1990). The role of discontinuities in the perception of subjective figures. *Perception & Psychophysics, 48*(3), 259-270.

Shipley, T. F., & Kellman, P. J. (1992a). Perception of partly occluded objects and illusory figures: Evidence for an identity hypothesis. *Journal of Experimental Psychology: Human Perception and Performance, 18*(1), 106-120.

Spehar, B. (1999). The role of contrast polarity in illusory contour formation. *Investigative Ophthalmology & Visual Science, 40*(4).

Spehar, B., & Clifford, C. W. G. (2003). When does illusory contour formation depend on contrast polarity? *Vision Research, 43*(18), 1915-1919.

Spehar, B. (2000). Degraded illusory contour formation with non-uniform inducers in Kanizsa configurations: the role of contrast polarity. *Vision Research, 40*(19), 2653-2659.

Stevens, K. A. (1976). Occlusion clues and subjective contours, *MIT A.I. Memo*, 363.

Stromeyer, C. F., & Julesz, B. (1972). Spatial-frequency masking in vision: Critical bands and spread of masking. *Journal of the Optical Society of America*, *62*(10), 1221-1232.

Su, Y. R., He, Z. J., & Ooi, T. L. (2010a). Surface completion affected by luminance contrast polarity and common motion. *Journal of Vision, 10*(3), 1-14.

Su, Y. R., He, Z. J., & Ooi, T. L. (2010b). Boundary contour-based surface integration affected by color. *Vision Research, 50*, 1833-1844.

Van Essen, D. C., & Deyoe, E. A. (1995). Concurrent processing in the primate visual cortex. The cognitive neurosciences, 383-400.

Van Lier, R. J., Van der Helm, P. A., & Leeuwenberg, E. L. J. (1995). Competing global and local completions in visual occlusion. *Journal of Experimental Psychology: Human Perception and Performance, 21*(3), 571.

Van Lier, R., & Wagemans, J. (1999). From images to objects: Global and local completions of self-occluded parts. *Journal of Experimental Psychology: Human Perception and Performance, 25*(6), 1721-1741.

Victor, J. D., & Conte, M. M. (2000). Illusory contour strength does not depend on the dynamics or relative phase of the inducers. *Vision Research, 40*(25), 3475-3483.

Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., & von der Heydt, R. (2012). A century of Gestalt psychology in visual perception I. Perceptual grouping and figure-ground organization. *Psychological Bulletin, 138*(6), 1172-1217.

Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: I. Fitting, sampling, and goodness of fit. Perception & psychophysics, 63(8), 1293-1313.

Wertheimer, M. (1938). Laws of Organization in Perceptual Forms. In Ellis, W. D. (Ed.), *A source book of Gestalt Psychology* (71-88). London, England: Routledge & Kegan Paul. (Original work published in 1923).

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. (1997). Surface completion complements boundary interpolation in the visual integration of partly occluded objects. *Perception, 26*, 1459-1479.

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

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.