Influences of spatial context on dynamic perceptually unstable figures.

Marouane Ouhnana
Integrated Program of Neuroscience
McGill University
November 2017

A thesis submitted in partial fulfillment of the requirements of the degree of Doctor of Philosophy in Neuroscience

Table of Contents

List of figures

List of equations and tables

Abstract

An important part of our understanding of human vision comes from the study of ambiguous figures, a sub-set of which include bi-stable figures, such as the well-known Necker cube, and binocularly rivalrous figures. Key insights into the mechanisms of ambiguous figure perception have come from studying the effects of context on their variable perceptual interpretations. Since a given context tends to result in a predominance of one perceptual interpretation over the other, contextual influences are instructive for our understanding of the process of perceptual binding. In three studies I explore the effects of spatial context on the properties of the following motion-defined ambiguous figures: (1) a moving target whose perceived direction of motion is influenced by the motion trajectory of a secondary object, where the secondary object is either a cast shadow or an object similar to the target; (2) a rotating Necker cube whose perceptual interpretation is influenced by the motion direction of a nearby unambiguous skeleton cube; (3) binocularly rivalrous skeleton cubes whose dominant interpretation is influenced by a non-rivalrous nearby skeleton cube. Results with (1) argue for the influence of not only cast shadows on perceived object trajectories, but the Gestalt grouping principle of common-fate combined with the principle that objects tend to be perceived as moving along a common planar surface. Results with (2) demonstrate that change-synchrony not common-fate underpins the perceptual binding between a rotating unambiguous skeleton figure and a bi-stable Necker cube. Results with (3) demonstrate that an identical contextual figure has a similar influence on both a bi-stable and a binocularly rivalrous rotating skeleton cube. Together, these findings argue that the perceptual binding between a dynamic spatial context and ambiguous target figure, be it ambiguous, bi-stable or binocularly rivalrous, is critically dependent on grouping principles such as common fate, the putative common planar surface constraint and change-synchrony. The results also support the idea that bi-stable and binocularly rivalrous figures share a similar or common dynamical mechanism.

Résumé

Une partie importante de notre compréhension de la vision humaine provient de l'étude de figures ambiguës, dont un sous-ensemble comprend des figures bi-stables, tels que le célèbre cube de Necker et les figures binoculairement rivales. Les idées clés sur les mécanismes de la perception de figure ambiguë proviennent de l'étude des effets d'un contexte sur leurs interprétations perceptuelles. Étant donné qu'un contexte tend à entraîner une prédominance d'une interprétation perceptuelle par rapport à une autre, les influences contextuelles sont instructives pour notre compréhension du processus de liaison perceptuelle. Dans trois études, j' explore les effets d'un contexte spatial sur les propriétés des figures ambiguës définies par le mouvement: (1) une cible mobile dont la direction de mouvement perçue est influencée par la trajectoire de mouvement d'un objet secondaire, où l'objet secondaire est soit une ombre ou un objet similaire à la cible; (2) un cube Necker rotatif dont l'interprétation perceptive est influencée par la direction du mouvement d'un cube de squelette non ambigu voisin; (3) un cubes squelettiques binoculairement rivaux, dont l'interprétation dominante est influencée par un cube de squelette à proximité non rival. Les résultats avec (1) plaident pour l'influence non seulement des ombres sur les trajectoires des objets perçus, mais le principe du regroupement Gestalt *common-fate* combiné avec le principe selon lequel les objets ont tendance à être perçus comme se déplaçant le long d'une surface plane commune. Les résultats avec (2) démontrent que le regroupement Gestal *change* synchrony et non *common-fate* sous-tend la liaison perceptive entre une figure de squelette non ambigu rotative et un cube Necker bi-stable. Les résultats avec (3) démontrent qu'une figure contextuelle identique a une influence similaire à la fois sur un cube de squelette rotatif bi-stable et binoculairement rival. Ensemble, ces résultats font valoir que la liaison perceptuelle entre un contexte spatial dynamique et une cible ambiguë, qu'elle soit ambiguë, bi-stable ou binoculairement rivale, dépend de manière critique des principes de regroupement tels que le *common-fate*, la contrainte de surface planaire putative commune et le regroupement Gestal *change synchrony*. Les résultats confirment également l'idée que des figures bi-stables et binoculairement concurrents partagent un mécanisme dynamique similaire ou commun.

Acknowledgements

I am indebted to many who supported me during the course of my PhD, not all of whom I can list here.

I must express my sincerest gratitude to my supervisor - Professor Frederick Kingdom. Over the years, he has provided me with his insight, guidance, and encouragement. His patience and understanding are unparalleled and I count myself incredibly fortunate to have been able to train under his supervision.

I would also like to thank my lab mates throughout these years - Dr. Jason Bell, Dr. Ben Jennings, as well as Dr. Gunnar Schmidtmann. Their support, advice, and the hours upon hours they have generously given me and my experiments have been paramount to my completion. I have learned a tremendous amount from all you and I am forever grateful. I would also like to thank them for the laughs, annoyances, and for enduring my antics throughout these years.

I am grateful to my friends and family for their support, encouragement, and learning fairly quickly to stop asking when I will be done.

Lastly, I must thank my wife Maria Ene for putting up with my ups and downs throughout all these years. Your support, encouragement, and impeccable sense of humor have allowed me to forge on even when I wasn't sure myself.

Preface

Original contributions

The empirical studies presented in this thesis are all original work investigating the contextual influences that an unambiguous context figure exerts on a simultaneously presented perceptually ambiguous target figure through the perceptual binding between these two figures. The following is a summary of the main distinct contributions this thesis adds to the scientific knowledge on this particular topic.

Chapter 2 presents a novel methodology to extract height and depth spatial estimates of a moving target, building on the paradigm of Kersten, Mamassian, and Knill's ball-in-box animation. With this change in methodology, the results express that a similar object moving at the same speed and direction to a target, exerts a similar, but albeit weaker, spatial modulatory influence to that demonstrated with a cast-shadow. This finding also revealed the perceptual common-plane formed between target and secondary-object is subject to a fronto-parallel bias.

Chapter 3 presents novel methodology to establish the magnitude of contextual influence between an unambiguous context and perceptually ambiguous target. Through this method, we present the first psychophysical evidence of the properties that influence the magnitude of perceptual binding between these two types of figures. Our results revealed that the similarity of motion speed between a dynamic context and ambiguous target figure did not increase the perceptual binding between these two figures, as would be expected through common-fate. Rather we find that the change-synchrony, an alternative Gestalt grouping principle and a property that is agnostic to motion speed, to be the mediating factor in the magnitude of contextual influence between these two figures. Furthermore, our results also suggest, that the magnitude of contextual influence is increased when the unambiguous figure is presented below the target, arguing for prior knowledge of frictional forces.

Chapter 4 presents a novel binocularly rivalrous stimuli that elicits an identical perceptual experience to that of a bi-stable Necker cube. Using the novel methodology introduced in Chapter 3, the similarity of influence of an identical spatial contextual figure in a both bi-stable and binocularly rivalrous figure, equated for perceptual experience, was demonstrated. These results argue that both these class of figures, although being physically very different, share a common feedback mechanism responsible for the perceptual disambiguation of an uncertain visual input.

Organization of the thesis

This thesis consists of three published research papers, Chapters 2 through 4, lead in by an Introduction, Chapter 1, and followed by a Summary and Conclusion section, Chapter 5. As the research papers each contain their own Introduction and Conclusion sections, the Introduction and Conclusion chapters of this thesis provide general background information and an overall framing on the reported results, respectively. The research papers which make up the body of this thesis are the following:

- Chapter 2 - Ouhnana, M., & Kingdom, F. A. A. (2016). Objects Versus Shadows as Influences on Perceived Object Motion. *I-Perception*, *7*(6).
- Chapter 3 - Ouhnana, M., & Kingdom, F. A. A. (2016). Perceptual-binding in a rotating Necker cube: The effect of context motion and position. *Vision Research*, *126*, 59–68.
- Chapter 4 -Ouhnana, M., Jennings, B. J., & Kingdom, F. A. A. (2017). Common contextual influences in ambiguous and rivalrous figures. *PLoS ONE*, *12*(5).

These research papers have also been presented as abstracts and poster presentations in international conferences:

- 2013, May Common fate versus cast shadows as influences on perceived motion direction and depth. Vision Science Society Meeting, St. Pete Beach, Florida
- 2014, May Contextual disambiguation of rotating Necker cubes. Vision Science Society Meeting, St. Pete Beach Florida & CVR Vision Conference, York University Toronto, Ontario
- 2015, May Driving a rotating Necker cube – context position matters. Vision Science Society Meeting, St. Pete Beach, Florida
- 2016, May A binocular context exerts a similar influence on both binocular rivalry and ambiguous figure perception. Vision Science Society Meeting, St. Pete Beach, Florida

Contributions of other authors

I am the first author on all journal articles that make up the body of this thesis, Chapter 2 to 4. All these articles have been co-authored with Dr. Frederick Kingdom who provided his insight and guidance as research and graduate supervisor. Chapter 4 is also co-authored with Dr. Benjamin Jennings, a postdoctoral fellow in the Kingdom lab, who provided assistance during the pilot study and analysis on the project and has provided permission to include the co-authored publication in the main body of this thesis. The introduction and Summary and conclusion, Chapter 1 and 5, were written by myself with guidance from Dr. Frederick Kingdom.

Chapter 1

Introduction

Contextual influences that bias perception

The visual perception of our immediate environment is generally stable and veridical. When viewing an object, our visual system almost always arrives at a single perceptual conclusion. Yet the apparent stability of the visual world belies the fact that a given retinal image can arise from an almost endless number of possible scenes. In short, the proximal image has many possible distal origins. Put another way, the proximal image is often potentially ambiguous.

A major source of potential ambiguity is that the retinal shape of an object is a two-dimensional projection of a three-dimensional object, and therefore may correspond to an almost infinite number of three dimensional shapes. Take the projection illustrated in Figure 1.1, where the retinal image (left) can arise from any of the objects illustrated to the right of the figure. However, we rarely experience fluctuations in our perceptual interpretation. The reason for this is that the visual environment provides contextual information that favors one perceptual interpretation over another.

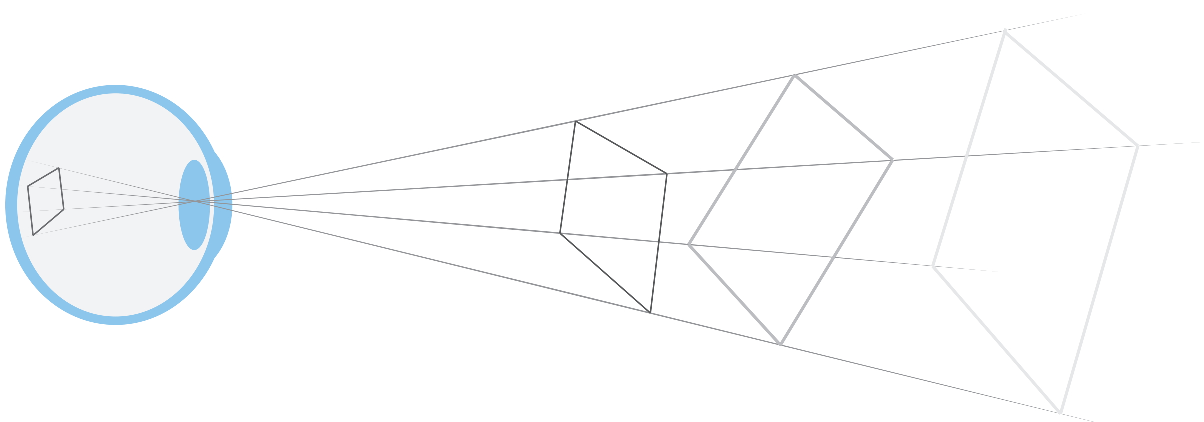


Figure 1.1. Ambiguity of the retinal image. The retinal image of the small square (left) may correspond to any number of objects (right).

Figure 1.2, a reproduction of a demonstration by Bruner and Minturn (1955), is an example of how ambiguity is resolved by context. The same pair of characters in the middle of both the top and bottom figures is perceived either as a **B** or the number **13** depending on its flanking context.

Figure 1.2. Ambiguity and context. In this reproduction of the Bruner & Minturn (1955) figure, the middle figure may be perceived as either the letter B or the number 13 depending on the flanking characters.

Information in the visual environment can also provide an indication of an object's spatial layout within a scene. In an example by Kersten, Knill, Mamassian, and Bülthoff (1996), reproduced in Figure 1.3, a square is presented in front of a checker background where its spatial layout is drastically changed through a manipulation of its immediate environment. In both of the left and right patterns in the figure, the square is presented in the same location and casts a shadow; however, the position of the cast-shadow is different. On the left, the square appears to be close to the background, while on the right it appears farther away. Thus a small change in context has a dramatic impact on our perception, in this case on depth.

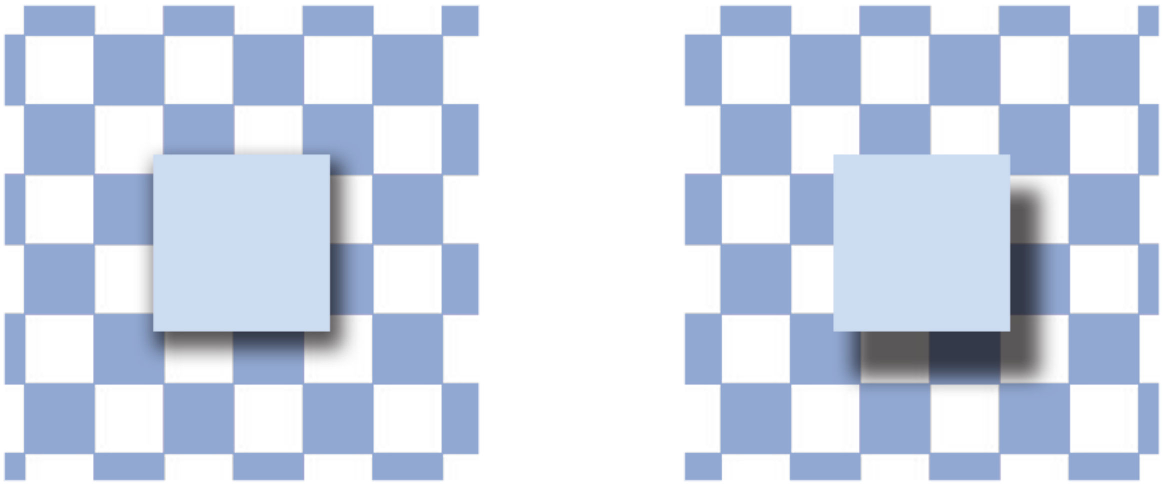


Figure 1.3. Kersten et al.'s (1996) demonstration of how the perceived depth of an object is influenced by a cast-shadow.

Perceptual instability

Although rare in everyday vision, there are cases where the visual system is unable to converge onto a single perceptual interpretation. The result is a multiple of short lived percepts, in which the alternative interpretations switch back and forth. The competing perceptual alternations are often mutually exclusive, i.e. only one of the two or more percepts is experienced at any one time.

Bi-stable figures

In this thesis, I will explore two types of perceptual instability: bi-stable figures and binocularly rivalrous figures. Bi-stable figures are characterized by a single physical stimulus producing two alternating percepts. Such figures are often termed ambiguous, but arguably the term bi-stable better captures the phenomenon, so will be used here. Bi-stable figures have a long history, dating as far back as the early 1800s. The first formal description of a bi-stable figure is found in a communication published in the London and Edinburgh philosophical magazine and journal of science in 1832. Louis Albert Necker, a Swiss crystallographer and geographer described a sudden and involuntary change in the appearance of a rhomboid crystal while inspecting plates of crystalline forms. The perceptual reversal was described as a change in the position of the faces of the now famous Necker cube figure (Necker, 1832), shown in Figure 1.4. The two square faces of the cube alternate in terms of which is perceived in front and which at the back. Figure 1.5 shows three other examples of bi-stable figures: on the left the Schröder staircase (Schröder, 1858), in the middle the chalice and faces figure (Rubin, 1915), and on the right the Necker-like overlapping square figure (Grünau, Wiggin, & Reed, 1984).

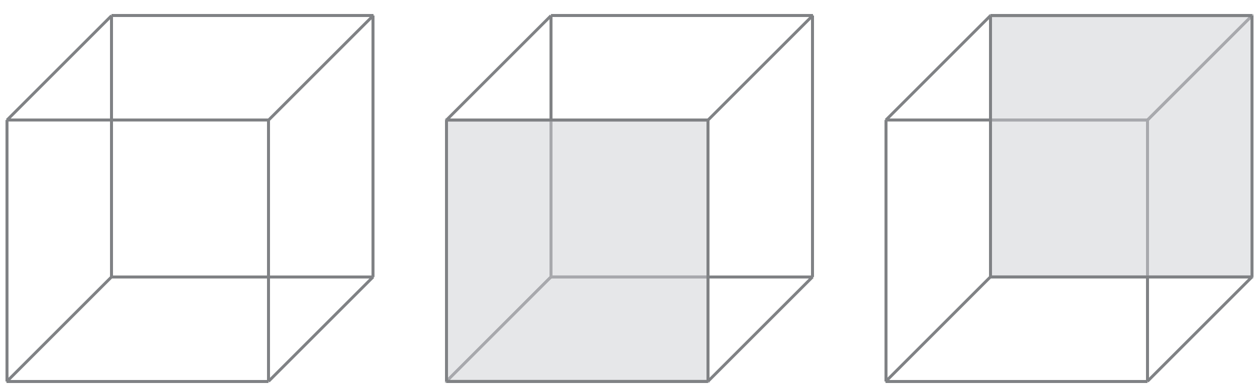


Figure 1.4. Example of the Necker cube (left). The two faces of the cube may be interpreted as either in the front or the back, alternating between the two. The alternations are either of a downward (middle) or upward (right) tilting cube.

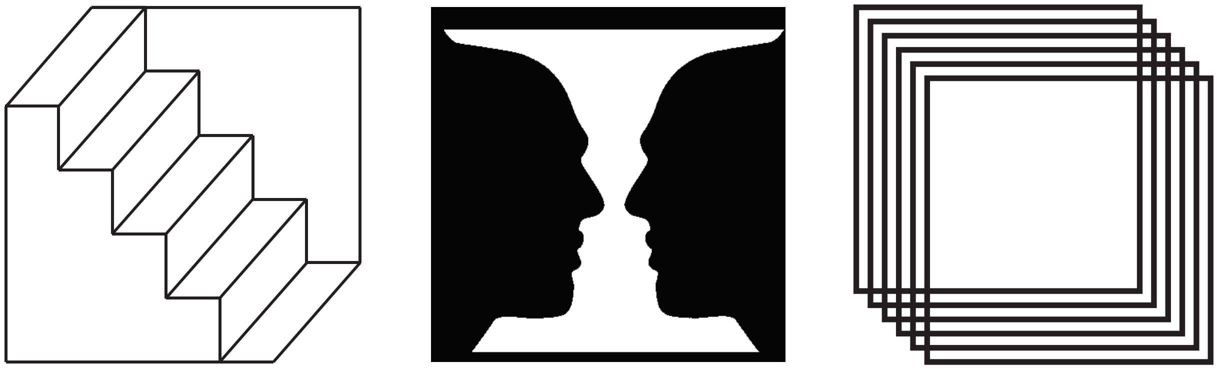


Figure 1.5. Examples of other bi-stable figures. Left: Schröder staircase; middle: chalice versus face; right: overlapping squares.

Theories of perceptual instability

Perceptual instability is believed to be caused by dynamic competitive interactions between aggregates of neurons, or 'channels', that underpin the competing perceptual interpretations (Orbach, Ehrlich, & Heath, 2011). At any point in time the interpretation depends on the channel with the highest level of activity (Köhler, 1940), and the exclusivity of the perceptual interpretation is a result of the channel's 'winner-take-all' behavior. However, there is disagreement as to the relative influences of bottom-up and top-down processes in perceptual instability (Long & Toppino, 2004). As a rule of thumb, bottom-up theories of perceptual instability stress more passive, automatic, and locally adaptable mechanisms at early stages of visual processing, whereas top-down theories favor more active, cognitive processes that embody forms of perceptual hypothesis testing (Kornmeier, Hein, & Bach, 2009).

Bottom-up theory of perceptual instability

Bottom-up theories of perceptual instability stress the importance of specific stimulus characteristics as well as the retinal location to which the stimuli are presented (Long & Toppino, 2004). When one inspects a bi-stable figure for an extended period of time, the shifts in percept - specifically the reversal rate - tend to increase over time

(Köhler, 1940). In other words, the length of time for which a single percept is dominant becomes shorter the longer the figure is inspected. Proponents of the bottom-up theory have argued that this characteristic is consistent with a feedforward theory of perceptual instability. The idea is that as the perceptual interpretation of a bi-stable figure oscillates between each of two percepts, so too the underlying neural channels oscillate between periods of fatigue and recovery. When a particular perceptual interpretation is experienced, its underlying neural channel exhibits the highest level of activity. During this period of high activity, there is a build up of fatigue from self-adaptation, which causes the channel's activity to gradually diminish. Eventually a point is reached where the competing channel takes over and now exhibits the highest level of activity. The acceleration of the reversal rate with time is then explained through the asymmetry between the periods of response and recovery, where the threshold of fatigue is reached at an earlier stage than the time needed to fully recover. The consequence of this asymmetry is that the threshold of fatigue will be reached at shorter and shorter intervals; hence the perceptual reversal rate increases over time.

Interestingly, some manipulations have been shown to alter this increase in reversal rate and in doing so strengthen the bottom-up argument. Spitz and Lipman (1962) reported that with a Necker cube, reversal rates returned to baseline levels if the figure was switched to a different visual field location, showing that the underlying neural mechanisms were retinotopically specific. Reversal rates have also been found to slow down or nearly stop when the bi-stable figure is presented discontinuously, by intermittently removing the stimulus from view (Leopold, Wilke, Maier, & Logothetis, 2002). These manipulations demonstrate the crucial relationships between reversal rates and retinal location, underscoring the importance of bottom-up influences.

Adaptation has also been shown to alter reversal rates, as well as having a significant impact on the predominance of each perceptual interpretation. Grünau et al. (1984) demonstrated that if an unambiguous version of a bi-stable figure was first presented, the perceptual interpretation of a subsequently presented bi-stable figure was opposite to that of the adapted figure. However the adaptation effect was eliminated when the adapting figure and bi-stable target were presented to different retinal locations, in keeping with a bottom-up theory of perceptual instability (Toppino & Long, 1987).

Top down theories of instability

In parallel to the evidence for a bottom-up theory of perceptual instability, a growing body of literature supports a more top-down approach. In this view, the visual system selects the perceptual interpretation based upon spatial and temporal context, past experience, and the preferences of the observer (Long & Toppino, 2004). In support of a top-down view, contextual information has been found to be particularly important. For example, using a bi-stable figure similar to the Bruner and Minturn (1955) figure in Figure 1.1, Intaitė, Šoliūnas, and Gurčinienė (2013) found that if flanked by unambiguous versions of the figure, the dominant interpretation shifted to that of the flanking context. Similarly, the interpretation can be biased towards that of a briefly exposed unambiguous version of the figure (Long, Toppino, & Mondin, 1992).

Further evidence for higher order influences comes from the demonstrations of the effects of prior learned associations and experiences, such as the study by Haijiang, Saunders, Stone, and Backus (2006). In their study, observers underwent a training period in which a rotating, disambiguated skeleton cube was presented together with various cues. For example, the motion direction of the skeleton cube was associated with cues such as presentation location, translation of motion (added to the rotation of motion of the figure), or a specific sound. During the test phase, an ambiguous rotating Necker cube was presented along with one of the previously learned cues. It was found that the perceived motion direction of the bi-stable target was reported in the motion direction associated with the particular cue during the training period.

Prior experience has also been shown to influence the perceptual interpretation of bi-stable figures, for example familiarity with frictional forces. Gilroy and Blake (2004) used a shape-from-motion-defined sphere, which is bi-stable in its perceived direction of rotation. The figure was presented alongside a similar but unambiguous shape-from-motion sphere. The reported motion direction of the bi-stable sphere depended on whether the two figures abutted or were separated by a small gap. When abutting, the perceived motion direction of the bi-stable sphere was in counter-rotation to the unambiguous sphere, however with a gap there was no consistent rotation relationship. Gilroy and Blake (2004) argued that our prior knowledge of physics, more specifically friction, provided the contextual information to disambiguate the rotation direction of the bi-stable sphere.

Binocularly rivalrous figures

Another class of alternating percept is binocular rivalry (Alais, 2012; Brascamp, Klink, & Levelt, 2015). The perceptual instability in this case occurs when the observer is presented with different stimuli to the two eyes. Binocular rivalry occurs when the two inputs are sufficiently different to preclude fusion (Levelt, 1965), and as with bi-stable figures, the percept oscillates back and forth between the two eyes' stimuli. Sometimes the percept is not exclusively one eye's stimulus or the other, but a patchwork of the two (Kovacs, Papathomas, Yang, & Feher, 1996).

The first formal description of binocular rivalry occurred almost two centuries ago from (Dutour, 1760). He noted that when presenting a pair of blue and yellow disks to the two eyes, instead of appearing as a green mixture of the two, the two colours appeared to oscillate back and forth.

As with bi-stable figures, competition has been invoked as the underlying mechanism in binocular rivalry; however, with binocular rivalry particular importance is placed also on the contributory role of interocular inhibition. The competition in the case of binocular rivalry is presumably between monocular neurons that respond to corresponding retinal locations (Levelt, 1965). The monocular channel exhibiting the highest level of activity, a result of inhibition of the other eye's channel, determines the percept at any one time. The perceptual switching occurs because the dominant channel adapts, reducing its own level of activity as well as reducing the amount of inhibition of the other channel. As the monocular channel with the highest level of activity adapts or fatigues, the level of inhibition it exerts on the other monocular channel decreases, while at the same time the adaptation of the suppressed channel itself decays. Eventually the suppressed channel's activity becomes greater and a perceptual switch occurs (Wilson, 2007). The exclusivity of each percept has raised the question of whether it is the other eye's visual input, or visual percept that is suppressed (Alais, 2012). These have in turn produced competing theories of suppression, either low-level, or high-level (Alais, 2012). The former refers to a complete inhibition of the signal coming from one eye, akin to blocking the eye's view. In this view, the dominance and suppression refers to which eye's signal is experienced and which is blocked. The high-level theory of suppression places the competition between the neural representations of each of the alternative interpretations, where dominance and inhibition is specific and based on the feature content of the rivaling stimuli.

Low-level binocular suppression

Evidence for low-level interocular suppression in binocular rivalry receives support from a variety of paradigms. One of these involves the detection of a probe presented to one or other eye during rivalry. Blake and Fox (1974) presented a probe at random to one of the eyes either during a period of dominance, when the eye's input was visible, or suppression, when the eye's input was invisible. They found that regardless of the type of rivalling pattern, either a grating or contour, or the type of probe, either a spots or letter, the detectability of the probe was reduced when presented during periods of inhibition. This was taken to indicate that the suppression in binocular rivalry was wholesale, in other words irrespective of the particular probe or rivaling pattern.

Evidence for wholesale suppression has also found support in neuroimaging studies. Tong, Nakayama, Vaughan, and Kanwisher (1998) used fMRI to measure the brain activity in response to a binocularly-rivalrous house versus face stimulus, in which an image of a house and an image of a face was presented to the two eyes. As the percept shifted from house to face, the brain activity shifted from the Fusiform Face Area (FFA) to the Parahippocampal Place Area (PPA). Interestingly, when Tong et al. (1998), rather than presenting a rivaling stimulus, presented each stimulus in succession to mimic perceptual alternation, they found that the measured brain activity was not significantly different from that measured with the rivalry stimuli.

High-level binocular suppression

There is competing evidence for high-level suppression in binocular rivalry. In a study by Logothetis, Leopold, and Sheinberg (1996), flickering orthogonal gratings were presented to each eye and swapped from one eye to the other at a rate several times faster than that typically found with perceptual reversals. Logothetis et al. (1996) reported that the perceptual alternations did not follow the physical eye swaps of the presented stimuli; rather, typical rates of perceptual reversals were found. As the physical alternations of the eye's input did not match the perceptual alternations reported by the observers, the authors argued that the competition between the percepts could not have arisen at the eye level but instead must have occurred at a later binocular stage.

Hybrid model of binocular rivalry

In light of these competing theories about the site of suppression, a hybrid model of binocular rivalry has been suggested that incorporates both low-level and high-level binocular suppression (Blake & Logothetis, 2002). This view reconciles the evidence for both levels of suppression as well as accounting for contextual effects which modulate the perceptual experience of rivalrous figures through feedback mechanisms (Alais, 2012).

Contextual influences in binocular rivalry

As with ambiguous and bi-stable figures, spatial context influences binocular-rivalry. In an early study by Levelt (1965), one of the components of a rivalrous pair was displayed with a small annular surround. The effect was to increase the predominance of the surrounded stimulus. Kovacs et al. (1996) demonstrated that a grid of red and a grid of similarly positioned green colored dots presented to different eyes alternated in color, as one would expect. However, if one eye's grid comprised a mixture of green and red dots, with the other eye's grid the complementary color mixture, the resulting percept was of alternating single-color grids. This shows that high-level grouping mechanisms influence binocular rivalry.

Global coherence in binocular rivalry has also been demonstrated in dynamic displays. In a study by Sobel and Blake (2002), observers were presented with four apertures displaying gratings drifting in different directions. The gratings either defined a coherent global diamond shape, or were not globally coherent. One of the apertures was presented as a component of a rivaling pair, with a radial checkerboard presented to the other eye. When the drifting gratings did not define a global shape, the rivalling pattern alternated between grating and checkerboard. However, when defining the diamond shape the grating percept dominated.

Gestalt laws and perceptual binding

The aforementioned influences of context on the interpretations of ambiguous, bi-stable and binocularly rivalrous figures, may also be considered within the framework of Gestalt grouping principles. Gestalt grouping principles define ways in which a set of discrete elements take on the appearance of a global percept – ways in which the "whole is greater than the sum of its parts". One can think of Gestalt grouping as an example of perceptual binding. The Gestalt grouping principles include proximity (elements that are closer together are bound together), similarity (of color, size or orientation), common fate (elements that move together are grouped together) and synchronicity (elements that change together are grouped together). In addition, perceptual grouping may occur because of the way elements complement each other, such as symmetry, parallelism, continuity, and closure. Both Köhler (1940) and Wertheimer (1923), in discussing laws of perceptual organization, argued that the global percept critically depended on the relationship between discrete elements and stressed how the resulting perceptual grouping had an impact on all other aspects of visual processing.

As spatial contextual influences favor a perceptual interpretation over another in both bi-stable and binocularly rivalrous stimuli, one may make the parallel between the influence of context and perceptual binding. If one thinks of the possible interpretations of these ambiguous figures as discrete perceptual elements, the experienced percept thus becomes the perceptual interpretation which is perceptually grouped with the presented spatial contextual figure(s).

Outline of thesis

This thesis aims to gain a better understating of contextual effects on perceptual bias, bi-stability and binocular rivalry. The thesis will focus on the motion properties of contextual influence. In what follows I will report the results of three psychophysical studies. Chapter 2 investigates the influence of a moving object on the perceived motion trajectory of an ambiguous moving object. Chapter 3 investigates the effect of a rotating context figure on the perceived rotation direction of an ambiguous figure. Chapter 4 addresses the question of whether an identical contextual figure similarly influenced the perceived rotation direction of an ambiguous and a binocularly rivalrous figure.

Chapter 2

This chapter will explore whether the perceptual binding between a moving target and an identical secondary object moving with a similar trajectory is similar to that between a moving target and its cast-shadow.

Objects versus shadows as influences on perceived object motion

This chapter has been published as

Ouhnana, M., & Kingdom, F. A. A. (2016). Objects Versus Shadows as Influences on Perceived Object Motion. *I-Perception*, *7*(6), 2041669516677843. http://doi.org/10.1177/2041669516677843

Abstract

The motion trajectory of an object's cast shadow has been shown to alter the perceived trajectory of a casting object, an effect that holds even if the cast shadow appears unrealistic. This raises the question of whether a cast shadow per se is necessary for this influence, a question that has been studied only with stationary targets. We examined the relative influence of a shadow and a spherical object on the perceived motion trajectory of an identical spherical object, using a paradigm similar to Kersten, Mamassian, and Knill's ball-in-box animation. We recorded both depth and height estimates of the perceived end-point of the target trajectory as a function of various target and context trajectories. Both shadows and objects significantly influenced the perceived trajectory of the target, though the influence of the shadow was overall stronger. We conjecture that the influence of the object reveals the assumption that similar objects moving at the same speed and in similar directions are perceived to move within the same plane, a plane subject to a fronto-parallel bias.

Introduction

A single retinal projection may correspond to a multitude of perceptual interpretations, yet our visual experience is seldom ambiguous. Contextual cues in our visual environment tend to disambiguate an otherwise ambiguous retinal input. Shadows are such a cue; they have been shown to have a profound impact on the perceived position and motion trajectory of the casting object to which they are perceived to be attached (Kersten, Knill, Mamassian, & Bülthoff, 1996; Kersten, Mamassian, & Knill, 1994; Ni, Braunstein, & Andersen, 2004).

When one talks about shadows, one may refer to either attached or cast shadows: the former refers to a region of an object receiving weaker illumination from a light source, whereas the latter refers to a region of a surface or object that is occluded from the light source (see review by Mamassian, Knill, & Kersten, 1998). In this communication, we are interested in cast shadows, and how their influences compare to those of a secondary object on the perceived motion direction of a target object.

The first exploration of this question was undertaken by Ni et al. (2004) using a display comprised of two flat disks, one above the other, set against a grass-textured ground plane. The lower disc had various thicknesses and was textured to either resemble the upper disc or shaded dark to resemble a shadow. The discs were either stationary or moved horizontally back and forth along with the background to simulate head movements, but without motion parallax. The display therefore simulated a stationary object in both situations. The task was to estimate the perceived distance of the upper disc. Results showed that the upper disc was perceived to be significantly closer to the ground-contact position of the lower disk when the latter appeared as a shadow compared to when it was shaded to resemble the upper disc.

Cast shadows are tied to the objects that cast them, with the exact position of a shadow also determined by the position of the light source. Thus, when an object moves, so does its cast shadow. Hence, the influence of a shadow is likely to be greatest when both the casting object and the shadow are in motion. Kersten et al. (1994) demonstrated that the motion trajectory of a cast shadow significantly altered the perceived motion trajectory of the casting object—a spherical ball. When the shadow moved diagonally, the casting object appeared to recede and move along the ground, whereas when it moved horizontally, the casting object appeared to rise up in the frontal plane. Interestingly, the influence of the shadow was found to occur even when its contrast polarity was reversed. An interesting corollary to the flexibility with which shadows exert their influence is that incongruent shadows, that is, those whose positions and shapes do not accord with the laws of physical optics, are nevertheless perceived as shadows (Casati, 2008; Cavanagh, 2005; Mamassian, 2008), albeit with some measurable perceptual costs (Castiello, 2001). Such perceptual flexibility raises the possibility that Gestalt principles alone might be sufficient to explain the perceptual shifts observed in Kersten et al.'s (1994) moving ball experiments. The Gestalt principle that springs to mind is common-fate, which states that objects that appear to move together group together (Wertheimer, 1923). Common-fate may have a corollary: Objects that group together are perceived to move together. Such a grouping principle might even predict a stronger influence of a similar object on the target than a shadow. On the other hand, shadows may be a more powerful contextual cue due to the tight constraints under which they occur in the natural world.

The aim of this communication is to examine the relative influence of shadows and objects on the perceived motion trajectory of a primary object. For this purpose, we have employed a paradigm similar to Kersten et al.'s (1994) ball-in-box-animation, with observers making perceptual estimates of the motion trajectory of a dynamic target object in the presence of diverging trajectories of either cast shadows or a secondary bottom object.

General Methods

Subjects

Five observers participated in the experiment. All observers were naïve to the experimental aims and had normal or corrected-to-normal visual acuity. Written consent was obtained from all test participants, and all experimental protocols were approved by the McGill University Research Ethics Board.

Apparatus

The stimuli were created using open-source 3D computer graphics software Blender version 2.67 b. The stimuli were displayed using MATLAB version 8.1 with the Psychophysics Toolbox Version 3 extensions. The stimuli were presented on an Apple LED Cinema display connected to an Apple Mac Pro. The resolution of the display was set to 2,560 1,440 with a refresh rate of 60 Hz. Observers were seated 60 cm away from the display.

Stimuli

A display similar to the (Kersten et al., 1994) ball-in-box-animation was recreated using Blender. A spherical light source simulated lighting from above and placed in XYZ Blender space (0, 0, 20). The energy and distance of the light source were set to 6 and 100, respectively. The camera was placed in XYZ Blender space of (0, –34, 8) with a XYZ rotation of (80, 0, 0) using perspective view with a focal length of 35. Rendered animations had a resolution of 1,980 1,280 and each lasted 2 s. Illustrations of the stimuli are shown in Figures 2.1 to 2.3.

Target Stimulus

The target stimulus was created using Blender's UV sphere mesh, shaded gray with specularities set to 0.5. The UV sphere mesh size was set to 0.41 for all XYZ Blender scale units. See the top sphere in both left and right hand side of Figure 2.1 for an example of the target sphere. The target sphere was rendered following a specific fixed diagonal trajectory for each of the starting location conditions and was described later in the text.

Context Stimuli

Context stimuli were created in a similar manner to the target stimulus and positioned below the target. The UV sphere mesh size was set to 0.37 for all XYZ Blender scale units. The size difference between the top and bottom UV sphere was designed to compensate for the camera's perspective scaling, in other words to equate the perceptual size of both target and context stimuli. Two types of context figure were used, a cast shadow and a bottom- sphere identical to the target figure. When the context was a sphere, the shading was similar to the target stimulus. When the context was a cast shadow, the UV sphere was set to 100% transparency and set to cast a shadow. Figure 2.1 shows both context stimuli with the target. The context stimuli could follow four possible trajectories each starting at one of three possible locations, and these are described later.

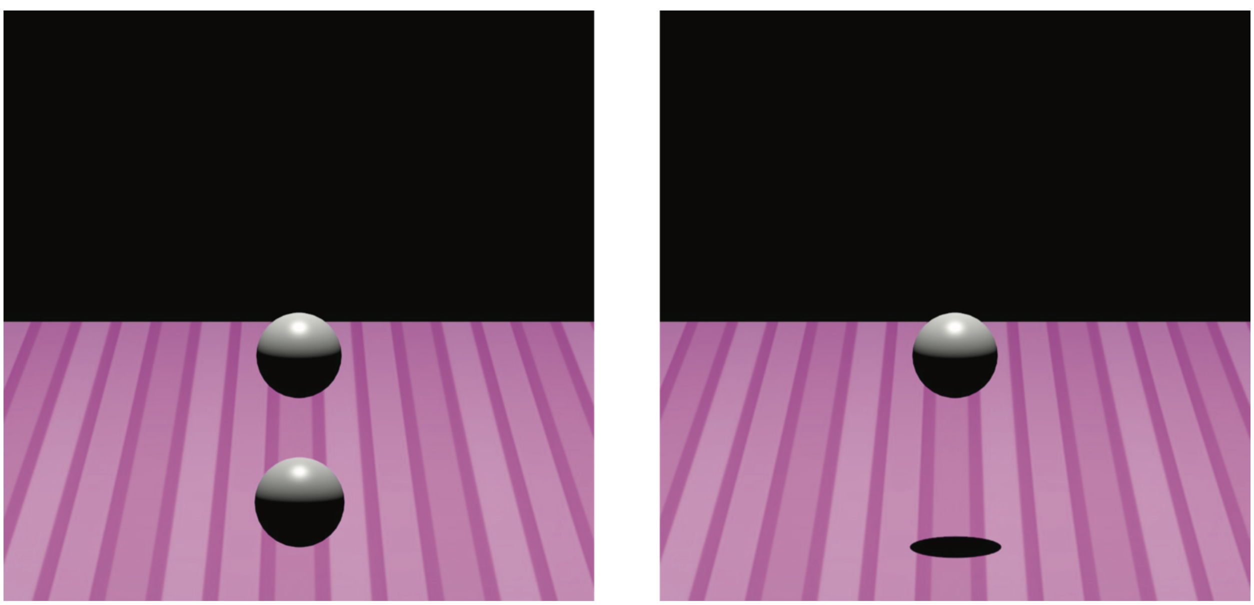


Figure 2.1. Target (top-sphere) and two context conditions: left, an identical secondary sphere; right, a cast shadow.

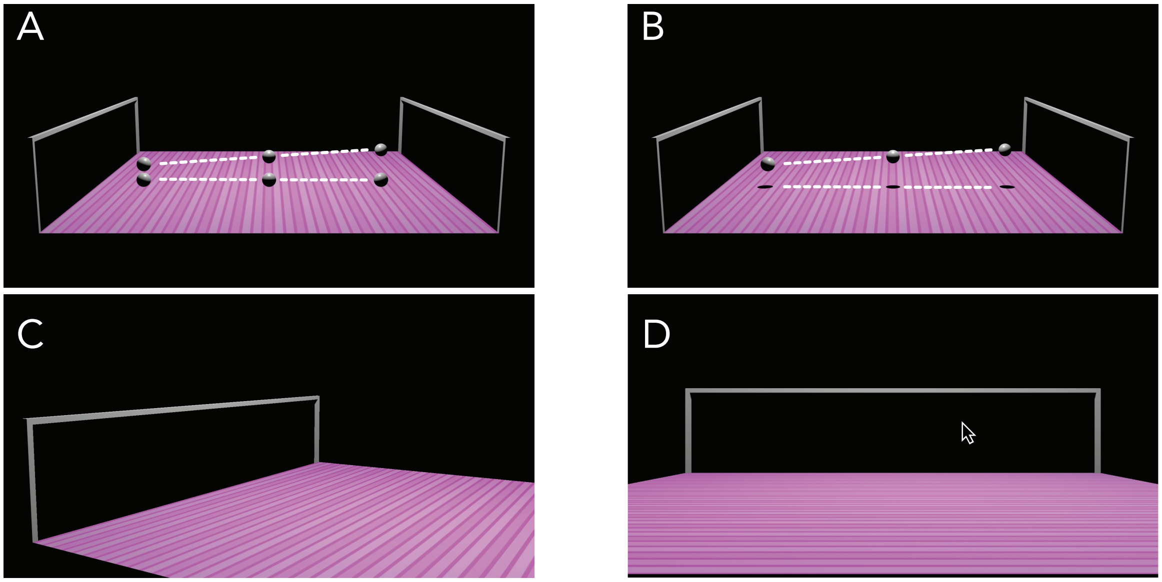


Figure 2.2. Procedure for measuring the perceived trajectories of the target, for bottom-sphere (A) and cast shadow (B) contexts. Following the animation, the target and context were removed and the display rotated (C) to present the skeleton goal posts in planar view (D), at which point a cursor appeared. The observer's task was to indicate via mouse-click the location where they thought the target would cross the goal, termed the 'end-point' of the perceived trajectory.

Ground Surface

A textured surface was rendered with an XYZ Blender scale of (10, 10, 01) and placed in the center of the 3D space. Rectangular skeleton "goal posts" were rendered on each side of the textured surface and represented the region within which observers could position their responses, as shown in Figure 2.2. The goal posts were rendered with an XYZ Blender scale of (0.1, 0.1, 2.1) with the crossbar defined by an XYZ Blender scale of (0.1, 10, 0.1) units.

Procedure

Observers were first presented with an animation of the target and context figure, or without the context during baseline trials, moving across the display. Then, the target, and context if applicable, disappeared, and the display was rotated by changing the camera's perspective to present the goal posts in frontal view; see Figure 2.2 for an example trial. Observers were instructed to indicate via a mouse-click the location where they thought the target would cross the goal.

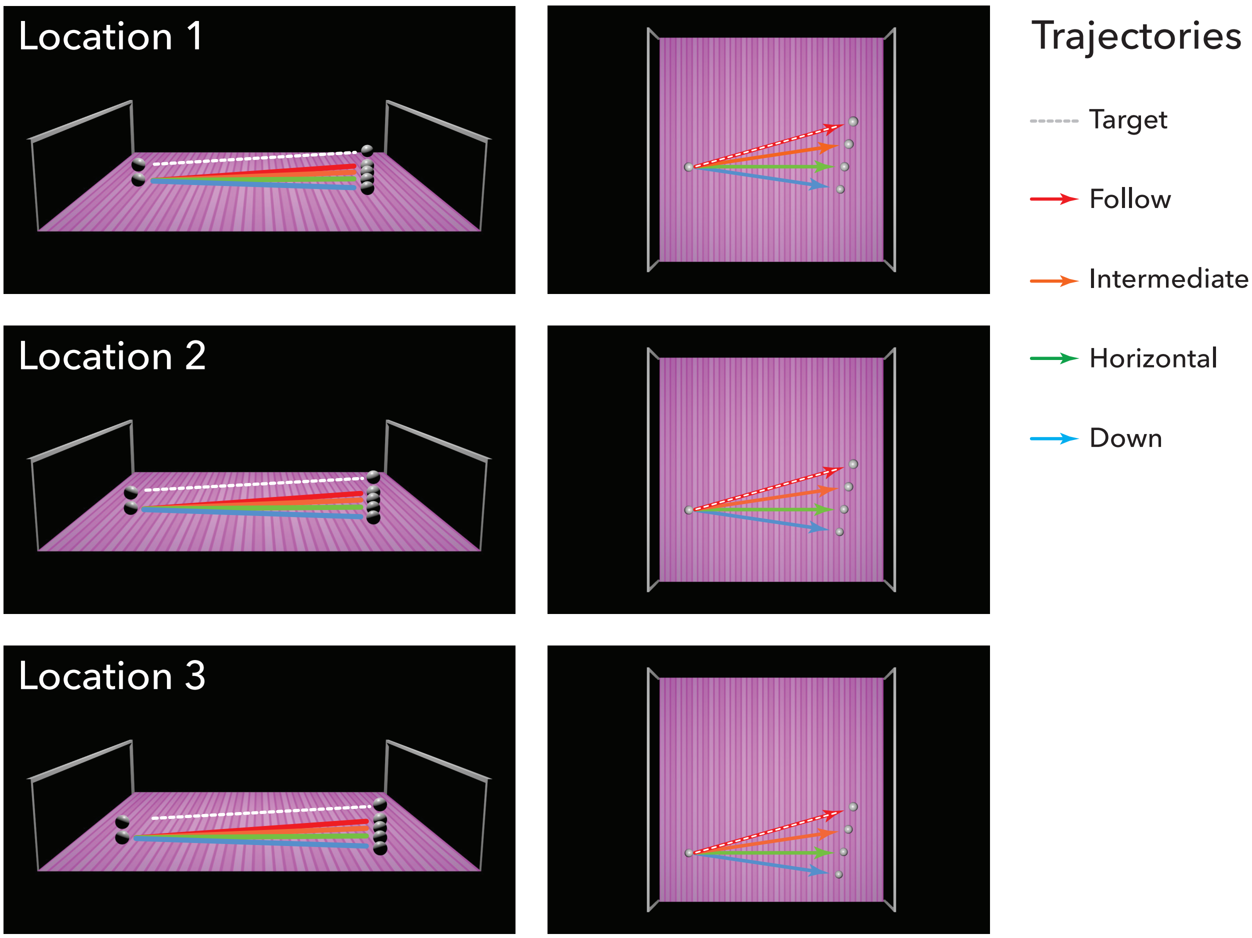


Figure 2.3. Starting locations and context trajectories. The figures on the left illustrate the three different starting locations of the target and context figures. The trajectory of the target sphere is shown by the dashed line and is fixed for each starting location. The four trajectories of the context figure (here the sphere) are the continuous lines, and their names are shown on the right. The figures on the right show a bird's eye view of the trajectories. Horizontally flipped versions of all trajectories were also employed but are not illustrated here.

Observers were presented with one of three possible target sphere start and end locations. The start locations are described by the following XYZ Blender positions. See Figure 2.3 for an illustration of each of the possible starting locations.

Location 1: starting location (–7.2, –1.5, 1.33), ending location (7.2, 2.5, 1.45)

Location 2: start location (–7.2, –3.5, 1.33), ending location (7.2, 0.5, 1.45)

Location 3: start location (–7.2, –5.5, 1.33), ending location (7.2, –1.5, 1.45)

Regardless of the start or end location of the target sphere, its motion followed a linear trajectory oriented at 4 from horizontal. Each target-sphere animation was accompanied by one of four context (secondary sphere or cast shadow) trajectories, as well as a no-context baseline trajectory. For all context trajectories, the XY Blender start locations were identical to those of the target-sphere, with Z Blender space set to a value 0.41, and the context's X Blender location adjusted to be perceptually below the target. The following context trajectories, illustrated in Figure 2.3, were tested:

Follow: The Y Blender position was the same as the top-sphere and followed the same linear trajectory oriented 4 from horizontal.

Intermediate: A value of 2 units was subtracted from the Y Blender ending position of the top-sphere. This motion trajectory followed a linear path orientated 2 from horizontal.

Horizontal: A value of 4 units was subtracted from the Y Blender ending position of the top- sphere. This motion trajectory followed a horizontal path, that is, 0 from horizontal.

Down: A value of 6 units was subtracted from the Y Blender ending position of the top-sphere. This motion trajectory followed a linear path oriented –2 from horizontal. Horizontally flipped versions of all conditions were also presented. All conditions, horizontally flipped and nonflipped, were interleaved and tested 10 times.

Results

To assess the influence of the context trajectories on the perceived trajectories of the target, end point estimates were separated into depths (X) and heights (Y). To normalize the data across experimental conditions, the observer's baseline depth and height estimate for each starting location and for both flipped and non-flipped animations was subtracted from the corresponding trajectory estimates. We term these measures shifts from baseline. The normalized depth and height estimates were then averaged for each condition. The shifts from baseline averaged across all subjects' trial data are summarized in Figure 2.4 for both cast shadow and secondary sphere context conditions and for each context trajectory.

To analyze the influence of context on the perceived target trajectories, two separate two- factor (Context Trajectory) within-subject ANOVAS (analyses of variance) were conducted, one on the depth estimates the other on the height estimates. There was a significant main effect of trajectory for both depth, $F(3, 12)$ 1⁄4 37.06, $p < .0001$, and height, $F(3, 12)$ 1⁄4 16.61, $p < .001$, estimates, showing that the contexts had a significant impact on the perceived trajectories of the target. There was no significant difference between the two contexts (shadow and bottom- sphere) for either depth, $F(1, 12)$ 1⁄4 0.65, $p < .46$, or height, $F(1, 12)$ 1⁄4 1.30, $p < .32$, estimates. However, this lack of difference between the two contexts obscures differences in the pattern of perceived target trajectories, which can be seen in Figure 2.4 and is expressed in the significant interaction between context and trajectory, for both depth, $F(3, 12)$1⁄432.38, $p<.0001$, and height, $F(3, 12)$ 1⁄4 13.94, $p < .001$, estimates. As the figure shows, the form of this interaction is that the depth and height estimates of the target followed more closely the changes in trajectory for the cast shadow compared with those of the secondary sphere contexts.

To investigate whether the shifts in target trajectory estimates occurred for both the cast shadow and secondary sphere contexts, a simple effects test was conducted between trajectories for each context condition. The analysis revealed significant effects of context trajectory for both context conditions (for the shadow condition, depth: F(3, 20)1⁄49.15, p<.01; height: F(3, 20)1⁄45.33, p<.05. For the secondary sphere condition, depth: F(3, 20) 1⁄4 113.7, p < .0001; height: F(3,20) 1⁄4 48.95, p < .0001.

To assess whether each change in context trajectory was accompanied by a change in target trajectory estimate, we conducted Tukey's Honestly Significant Difference (HSD) pairwise comparison tests between the various pairs of trajectories. For the cast shadow context, Tukey's HSD test revealed that all trajectories had a significant differential influence on target estimates. In other words as the trajectory of the cast shadow deviated away from the target, there was an accompanying significant shift in depth estimates: all p's < .01. This coherent relationship between cast shadow trajectory and target estimates was also found for height (p's < .01), with the exception of the follow with intermediate trajectory comparison (p 1⁄4 1), and the horizontal with down trajectory comparison (p 1⁄4 .23). In the case of the secondary sphere context condition, Tukey's HSD test revealed a significant shift in depth estimates between the horizontal and follow context trajectories (p < .05) as well as between the down and follow context trajectories (p<.05). None of the other secondary sphere trajectory comparisons were significantly different from one other, that is, produced similar shifts from baseline in both depth and height estimates.

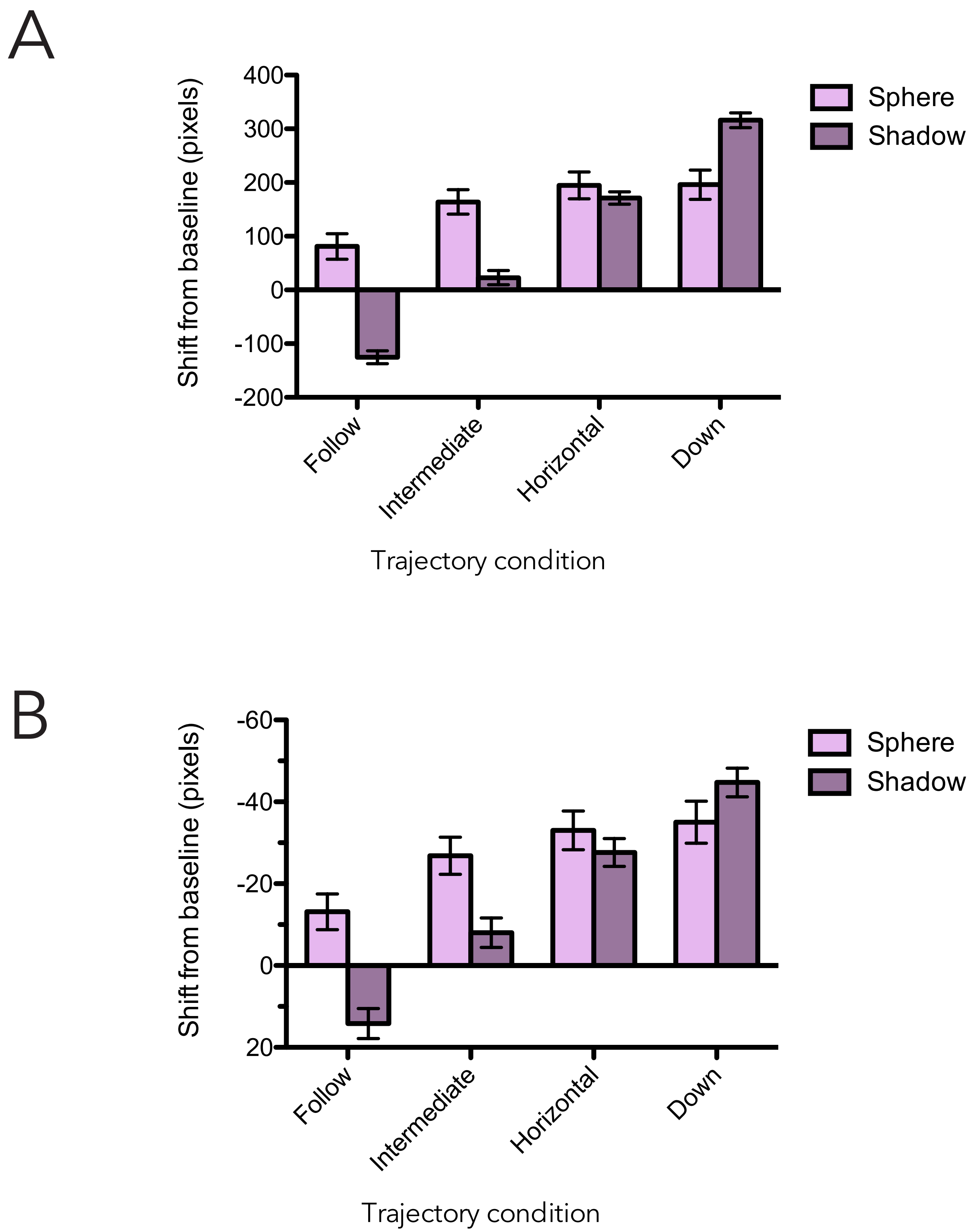

A
400
300
200
100
0
-100
-200
Shift from baseline (pixels)
Sphere
Shadow
Follow
Intermediate
Horizontal
Down
Trajectory condition
B
-60
-40
-20
0
20
Shift from baseline (pixels)
Sphere
Shadow
Follow
Intermediate
Horizontal
Down
Trajectory condition

Figure 2.4. Shift in end-point from baseline for the sphere and shadow contexts in terms of (a) depth and (b) height. The error bars correspond to 95% confidence intervals calculated across trials.

To further illustrate the relationship between target end-point estimates and context, the Euclidean distances between target trajectory end-point estimates for each adjacent pair of context trajectories were calculated and are summarized in Figure 2.5. The figure reveals consistent shifts in estimates for the cast shadow, while estimates for the secondary sphere display asymptote beginning at the second context trajectory.

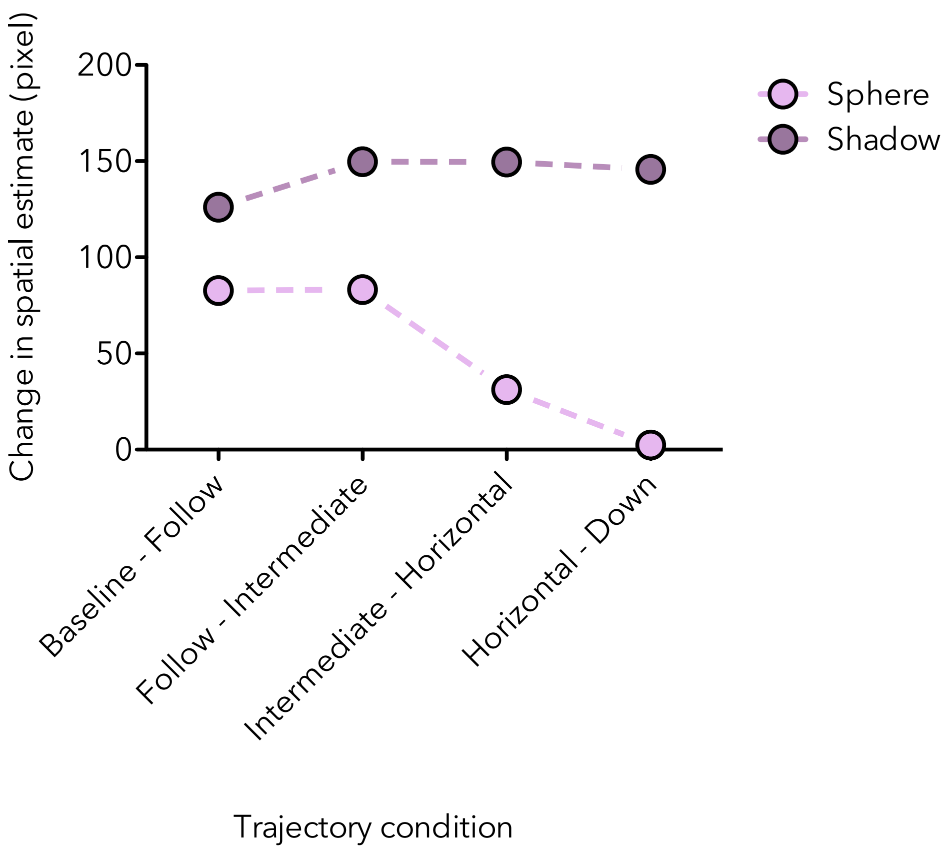


Figure 2.5. Euclidean distances between end-point estimates for adjacent pairs of context trajectories for both secondary sphere and cast shadow contexts.

Discussion

Our goal was to compare the influence of a secondary sphere and the cast shadow on the perceived motion trajectory of a target sphere. The main findings are as follows:

(1) The perceived motion trajectory of a moving target was found to be significantly influenced by the motion trajectory of both a cast shadow and secondary sphere context.

(2) The perceived target motion trajectory diverged along with the divergence of the cast shadow motion trajectory, but reached a plateau in the secondary sphere condition at greater motion trajectory divergences. Our data from the shadow conditions replicate the findings of Kersten et al. (1994) and

add the new finding that there is a similar influence of an identical object. The influence of the object however is overall less than that of the shadow, in keeping with the results of Ni et al. (2004), who compared shadow and object influences on the perceived location of a stationary object. The question however remains: Why does the secondary object have any influence at all, and why is the influence less than that of a shadow?

One idea suggested by an anonymous reviewer is motion repulsion. The angle between two superimposed dot patterns moving in different motion directions tends to be perceptually overestimated (Rauber & Treue, 1999), an example of a more general effect termed acute- angle expansion (e.g., Kitaoka & Ishihara, 2000). The largest overestimation occurs at a motion angle of about 7°, and diminishes in magnitude for both smaller and larger angles (Rauber & Treue, 1999). In our study, the trajectory angles ranged between 0° and 6°. Therefore, there may have been a perceived overestimation in the trajectory divergences in our displays. However, the task for the subject was not to estimate the angular difference in trajectory, but to estimate the height and depth of the perceived end-point of the target trajectory. The effect of motion repulsion would, if anything, tend to shift the end-point of the target away from the trajectory of the sphere, whereas our results showed the opposite.

Thus, while we cannot rule out an influence of motion repulsion on our results, it seems highly unlikely that it is their main cause. Another possibility is that the context sphere, which is identical to the target sphere, is perceived as its mirror-reflection, as if the ground plane in the display acted like a mirror. The physics of reflection is compatible with any of the trajectory combinations of perceived target motion direction and context motion direction used in the study, though normally of course the shading on the reflected sphere would be above the sphere, not below it as in our display. Mirror-reflection is therefore a tantalizing possibility.

The Gestalt principle of common fate states that objects that appear to move together group together. The target and secondary sphere share a degree of common fate in that they exhibit similar motion speeds, not too dissimilar motion directions, and a common vertical axis in their 2D projections. Grouping might lead to the interpretation that the two objects are perceived to lie on the same, flat, rigid surface. For the conditions in our experiment, any surface formed from such object grouping would be increasingly slanted away from the observer as the trajectory of the secondary sphere diverged from that of the target toward the observer. However, if there was a bias toward perceiving that surface as close to fronto- parallel, similar to the bias observed in the stereopsis domain (Deas & Wilcox, 2014; Fahle & Westheimer, 1988; Liu, David, & Ronen, 1999), then the target object would tend to appear to move upwards, as we found. This reduction in perceived surface slant would occur in conjunction with the trajectory of the secondary sphere, which would define the tilt of the surface.

As to why the cast shadow exerts a bigger influence than the secondary object, the answer probably lies in the fact that shadows are more constrained in the natural world than other objects in terms of their physical relationships with the objects that cast them. Kersten et al. (1994) suggest that the visual system assumes a stationary light source when dealing with shadows. Coupled with the fact that light is typically assumed to come from above (Morgenstern, Murray, & Harris, 2011; Ramachandran, 1988) and that shadows are assumed to lie on the ground plane, this reduces significantly the possible interpretations of the target-shadow trajectories in Kersten et al.'s (1994) and our moving ball experiments.

Conclusion

The perceived motion direction of an object such as a sphere is influenced by the motion direction of an identical object, not just a cast shadow. Future research will be needed to determine the underlying cause of this influence, exploring such possibilities as mirror- reflection and a bias towards motion surfaces being fronto-parallel. In conclusion, we have demonstrated systematic shifts of end point estimates to trajectory divergences of a secondary object.

Chapter 3

Following Chapter 2 this chapter investigates the properties that mediate the magnitude of contextual influence of an unambiguous object on an adjacent ambiguous target.

Perceptual-binding in a rotating Necker cube: The effect of context motion and position

This chapter has been published as

Ouhnana, M., & Kingdom, F. A. A. (2016). Perceptual-binding in a rotating Necker cube: The effect of context motion and position. *Vision Research*, *126*, 59–68. http://doi.org/10.1016/j.visres.2016.02.005

Abstract

Previous studies have shown that spatial context influences the perceptual interpretation of ambiguous figures such as the Necker cube; however, the properties that mediate the influences of an unambiguous spatial context have yet to be investigated. Here we consider the effect of the motion and position of an unambiguous rotating skeleton cube on the perceived motion direction of an ambiguous rotating Necker cube. We aimed to determine whether the motion of the two figures could be perceptually bound, and if it could, to determine the properties of the binding. We employed a novel procedure analogous to reverse correlation to establish the correlation between the rotation directions of the context and the perceived rotation directions of the target, across 32 s trial presentations. Our results showed that changes in the rotation direction of the context triggered above-chance changes in the perceived rotation direction of the target. However, the relative speeds of rotation of the context and target had little effect on the cor- relations.

Position on the other hand had a significant effect: correlations were higher when the context was below compared to when above the target. Our results reveal that change-synchrony not common fate is the factor mediating perceptual motion binding between the context and Necker cube. We also suggest that prior knowledge of friction forces could underlie the position dependency of the context and Necker-cube correlation.

Introduction

Ambiguous figures are a class of visual stimuli that possess more than a single perceptual interpretation. A classic example of such a figure is the Necker cube, illustrated in Figure 3.1. First described by Necker (1832), the Necker cube is a bi-stable figure that can be interpreted as an upward or downward facing skeleton cube. Each percept is relatively short-lived and alternates between the two interpretations. Such perceptual instability is argued to result from a dynamic winner-take-all process between competing perceptual interpretations, driven by either low-order, high-order, or as more recently argued, a combination of both processes – see Long and Toppino (2004) for a review.

In this communication we investigated the influence of a rotating context figure, in the form of an unambiguous skeleton cube, on the perceived motion direction of a nearby rotating ambiguous Necker cube. Previous studies (see below) have considered the effect of context on both stationary and rotating Necker cubes, but in the latter case using as the context an ambiguous Necker cube. Our aim is to firstly determine whether perceptual motion binding exists between an unambiguous context and an ambiguous Necker cube, and secondly, if it does, to determine the properties of the motion binding. A broader goal of this research is to further our understanding of motion binding in object perception in general.

Many experimental manipulations have been shown to alter the perceptual interpretation of a Necker cube. These manipulations include prior adaptation to unambiguous versions of the figure, presentation of spatial or temporal context figures, and learned associations between a particular figural interpretation and a particular cue. Adaptation effects in ambiguous figure perception have been reported in numerous studies; their influences are repulsive in nature, that is they increase the occurrence of the interpretation in the direction away from that of the adapting figure (Intaitė, Noreika,Šoliūnas,&Falter,2013; Long&Moran,2007; Long, Toppino, & Kostenbauder, 1983; Long, Toppino, & Mondin, 1992; Toppino & Long, 1987). Context effects on the other hand, both spatial and temporal, have an attractive effect; they increase the occurrence of the interpretation towards that of the unambiguous context figure (Dobbins & Grossmann, 2010; Goolkasian, 1987; Intaite˙ , Noreika, et al., 2013; Intaitė , Šoliūnas, Gurcˇiniene˙ , & Rukšėnas, 2013; Sundareswara & Schrater, 2008). In one study, after observers had learnt to associate a disambiguated Necker cube with either its position or its motion, or with a paired sound, the interpretation of a subsequently presented ambiguous Necker cube presented under the same conditions was found to be biased towards that of the disambiguated figure (Haijiang, Saunders, Stone, & Backus, 2006).

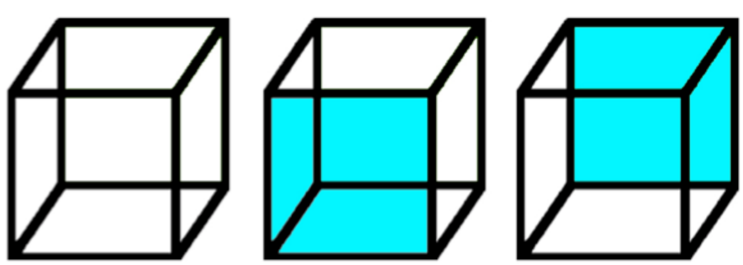


Figure 3.1. Example of the Necker cube (left). The bi-stable figure can be perceived as a downwards tilting cube (middle) or a cube that is tilted upwards (right).

Contextual effects have been demonstrated even when the context figures are themselves ambiguous. In an experiment by Dobbins and Grossmann (2010), observers reported the rotation direction of a rotating Necker cube presented in isolation or within a large array of other rotating Necker cubes. The axis of rotation was varied between vertical and horizontal for both the isolated cube and the cube array condition. Dobbins and Grossmann (2010) found a view-from-above bias in the reported rotation direction of the Necker cube for the vertical but not horizontal rotation axis condition. The contextual influence of the flanking Necker cubes was found to reduce the variability of observer reports, i.e. strengthen the view-from-above bias, compared to when the Necker cube was presented in isolation.

Contextual effects have also been shown to overcome inherent biases in ambiguous figures, as well as overcome the effects of incoherent influences. For example in one study using a static display, a Necker cube was presented within a surround of stacked cubes that appeared either lit and viewed from above, or lit and viewed from below (Sundareswara & Schrater, 2008). The resulting interpretation of the Necker cube was strongly tied to its surrounding context, overcoming the viewed-from-above bias. In another demonstration, Intaitė, Noreika, et al. (2013) used a display consisting of overlapping squares that appeared either facing upwards or downwards. The target was presented in one of three conditions: 1. following adaptation to an unambiguous version of itself; 2. simultaneously flanked vertically and horizontally by similar unambiguous versions of the target; 3. a combination of the two preceding conditions. The adaptation and context influences were manipulated so that their influences on the ambiguous figure were either consistent or inconsistent. Intaitė, Noreika, et al. (2013) reported that the effects were additive when the adapting and context figures were of opposite configurations, such that the interpretation-repulsive effect of the adaptor and the interpretation-attractive effect of the context were consistent. However, when the adaptor and context figures were of the same configuration, i.e. their influences were inconsistent, the percept was dominated by the influence of the context figure.

Visual perception seldom occurs in isolation, so the discovery of contextual influences on ambiguous figures should not be surprising. One can think of such contextual influences as examples of perceptual binding, as the influence is always to bias the interpretation towards that of the context figure. If contextual influences operated as external information to disambiguate an otherwise noisy visual input, then it might not be necessary for the context to be particularly similar to the target to exert an influence. On the other hand, perceptual binding might be highly selective, such that we might only expect the context to exert an influence when very similar to the ambiguous figure.

In this communication we have investigated for the first time two potentially important context properties, relative motion speed and relative spatial position, on the perceived motion direction of a rotating Necker cube. To measure the influence of the context we have employed a novel procedure that is analogous to reverse correlation. Previous studies of ambiguous figure perception have relied primarily on two measures: first-percept reports, i.e. the interpretation of the ambiguous figure when it is first presented, and average rates of perceived reversals (Adams & Haire, 1958; Adams & Haire, 1959; Intaitė, et al., 2013; Li & Kingdom, 1999; Long et al., 1983, 1992; Long & Batterman, 2012; Strüber & Stadler, 1999; Toppino & Long, 1987). These metrics, however, are agnostic to any changes in reversal rates over time. By considering how reversal rates change over time, the results of our study have helped to refine not only our understanding of the mechanisms involved in the perception of ambiguous figures, but also of contextual effects in the perception of motion in general.

General methods

Subjects

Five observers participated in the first experiment and 9 observers participated in the second experiment. The second experiment consisted of two variants; 5 observers participated in the first while 4 participated in the second. All observers were university students and participated voluntarily. One observer in the second experiment was excluded for failing to report any perceptual reversals of the rotating Necker cube. All observers were naïve to the experimental aims and all had normal or corrected-to-normal visual acuity. Written consent was obtained for all participants and all experimental protocols were approved by the McGill University Research Ethics Board.

Apparatus

The stimuli were created using open-source 3D computer graphics software Blender version 2.67b and were displayed using MATLAB version 8.1. The stimuli were presented on an Apple LED Cinema display connected to an Apple Mac Pro. The resolution of the display was set to 2560 1440 with a refresh rate of 60 Hz. Observers viewed the stimuli using a chin and forehead rest to guarantee a constant viewing distance of 60 cm. Routines from the Psychophysics Toolbox were used to present the stimuli.

Stimuli

The stimuli consisted of a pair of rotating skeleton cubes, a Necker cube and an unambiguous context. The figures were created separately and their frames of animation were rendered separately at 60 frames per second. A sample frame of the target Necker cube and context figure are shown in Figure 3.2. Each cube was centered at 2.86° of visual angle from a centrally presented fixation dot on a black background.

Stimuli – Necker cube

The Necker cube figure was created using Blender meshes with materials set to white and shading turned off. To remove perspective cues, the stimuli were rendered under orthographic projection. Blender XYZ rotational space corresponded to roll, pitch, and yaw, where yaw can be thought of as the figure's rotation along its vertical axis. The Necker cube was set to 20° in X rotational space, 0° for Y, with Z dictating its rotation. The rotational speed of the target Necker cube was set to a quarter revolution per second for all conditions.

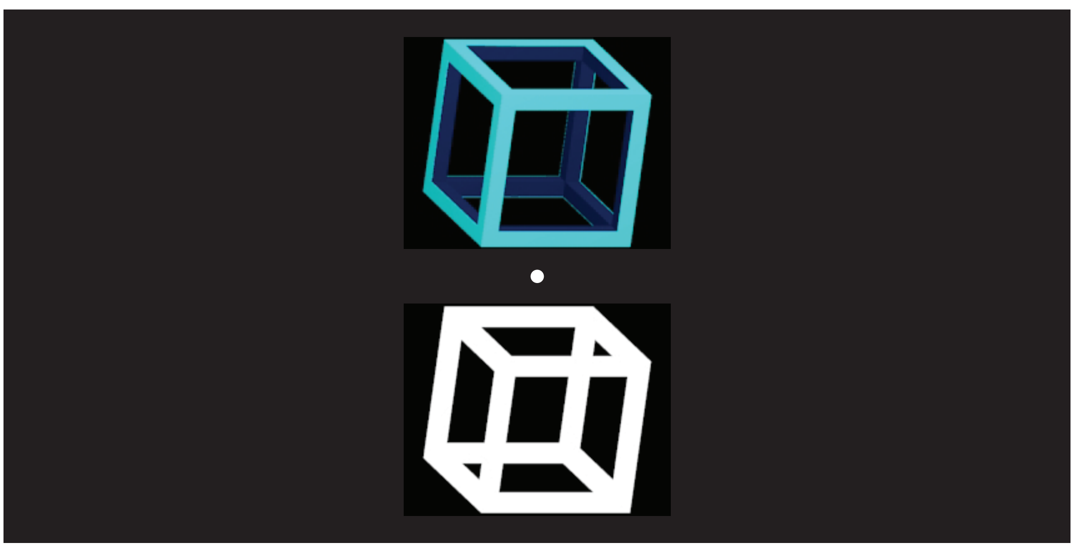


Figure 3.2. A sample frame of the context figure (top) and target Necker cube (bottom) rendered in orthographic projection. Shading, color variation and occlusion was used in the context figure to help disambiguate its motion direction.

Stimuli – context figure

The context figure was created using Blender meshes with colors set to teal on the outer surfaces and dark blue on the interior surfaces. The color difference between the external and internal surfaces added occlusion information that disambiguated the motion direction of the context figure. In addition, the shading option was turned on to simulate the effects of a light source placed above and in front of the figure. To confirm that the context figure was indeed unambiguous, four observers were asked to view the rotating context figure, presented at the various tested motion speeds, without any introduced rotation reversals. None reported any rotation reversals for any of the context motion speeds. As with the target Necker cube, Blender XYZ rotational space was set to 20° for X rotational space, 0° for Y, with Z dictating the figure's rotation. The rotation speed of the context figure was varied depending on the experimental condition.

Procedure

Experiment 1

Observers were instructed to maintain their gaze on the fixation dot while attending to the perceived motion direction of the target (Necker cube) during the entirety of the 32s trial. The context and target figure were arranged vertically above and below the fixation dot, the order of the arrangement depending on condition. Observers were instructed to report the motion direction of the frontal plane of the target figure via a key press throughout the trial. All participants underwent practice trials and were aware of the reversible nature of the Necker cube.

Three context rotation speeds and two context positions were tested. The three context rotation speeds were termed 'fast', 'same', and 'slow' in order to define the speed of the context relative to that of the target, and were respectively a half, a quarter, and an eighth of a revolution per second. At the start of each trial, the context figure's motion direction was randomly assigned to clockwise or counter-clockwise, and both the context and target figure's shared starting frame was also randomized. During each trial, the context figure's rotation direction was changed at intervals randomly-selected from 2, 4, or 8 s. For each of the three context speed conditions, two context positions were tested: context figure above and context figure below the target figure. All speed and position conditions were tested 10 times, resulting in approximately 5.5 min of data per condition.

Experiment 2

This experiment aimed to elucidate the effect of relative context position in more detail. The context and target figure were positioned either side of the fixation dot, but this time there were eight different target/context locations. Only one speed of context was tested, the same speed as the target in Experiment 1. Each of the figures were presented at 2.86° of eccentricity between fixation and figure center. The context figure was presented at specific orientations around the fixation dot: 0, 45, 90, 135, 180, 225, 270, and 315, while the target figure was presented opposite the context. Figure 3.3 shows the positions of the context and target for all orientation axes.

As in Experiment 1, observers were instructed to maintain their gaze on the fixation dot while attending to the perceived motion direction of the target cube during the entirety of the 32s trial. All orientation axes were tested 10 times.

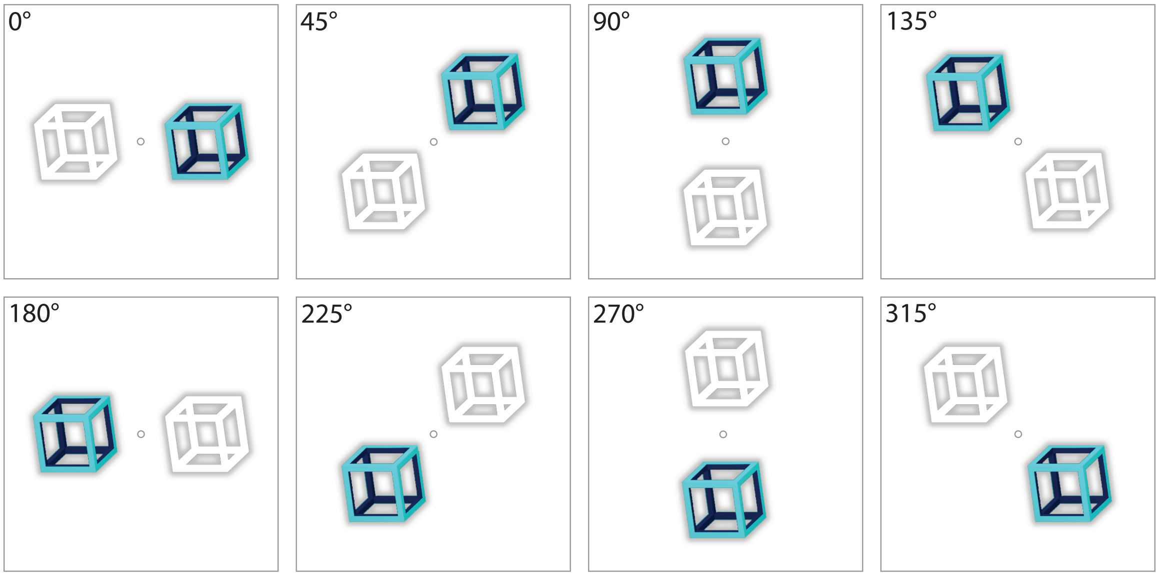
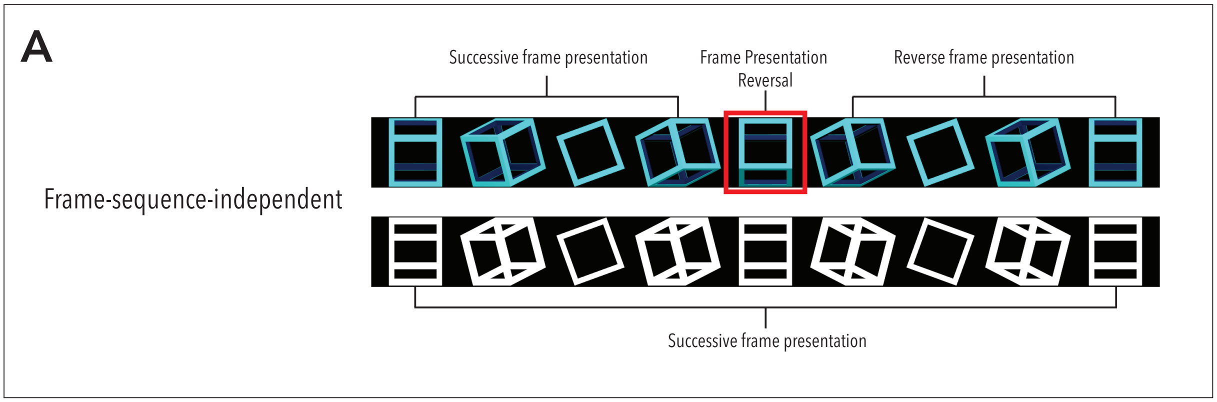


Figure 3.3. Target and context figure positions in Experiment 2.

Two variants of the experiment were tested, and these are illustrated in Figure 3.4. In Experiment 2a, as in Experiment 1, the target and context figure frames were independent: a change in direction of the context figure, achieved by reversing the frame sequence, was unaccompanied by a change in the frame sequence of the target. The result, as the figure shows, is that the coherence of the tilts of context and target were not consistently preserved. In Experiment 2b on the other hand, the target and context figure were frame-locked: a reversal in the frame sequence of the context was accompanied by a reversal in the frame sequence of the target. This consistently pre- served the coherence of the context and target tilts. All other procedures and instructions were otherwise the same as in Experiment 1. We refer to these two variants in Experiment 2 as the 'frame-sequence-independent' and 'frame-sequence-locked' conditions.



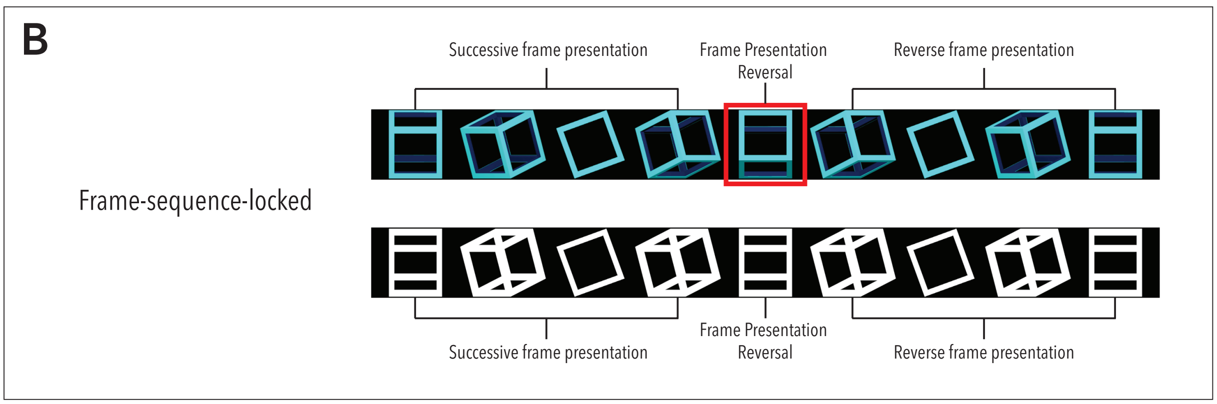


Figure 3.4. Sampled sequences of frames for context and target for a) frame-sequence-independent and b) frame-sequence-locked conditions. Each figure shows a sequence of frames for the context figure which reverses at the middle frame and below the corresponding target frame sequences. In a) there is no corresponding reversal of the target frame sequence at the middle frame, with the result that the tilts of the context and target loose coherence. In b) on the other hand there is an accompanying reversal which preserves the orientation coherence.

Results

Experiment 1: effect of relative speed and relative vertical position

First percept

A common metric used to characterize the influence of a context figure on the interpretation of an ambiguous figure is the concordance between the two at the start of each trial (Adams & Haire, 1958; Intaitė, et al., 2013; Strüber & Stadler, 1999). Figure 3.5 summarizes the proportion of times the target appeared to move in the same direction as the context at the start of each trial.

A one-sample t-test was conducted between each mean value and chance level (i.e. 0.5) for each of the speed and context position conditions. All mean values of first percept reports were significantly greater than chance for all conditions tested. For the context-above-target conditions, $t(98) = 1.731$, 1.731, and 2.364 for respectively the fast ($M=0.62$ $SEM=0.07$), same ($M=0.62$ $SEM = 0.07$), and slow speed ($M = 0.66$ $SEM = 0.07$) conditions, with $p < 0.05$ in all cases. For the context-below-target conditions, $t(98) = 12.969$, 4.731, and 4.731 for the fast ($M = 0.94$ $SEM = 0.24$), same ($M = 0.78$ $SEM = 0.42$), and slow speed ($M = 0.78$ $SEM = 0.42$) conditions respectively, with $p < 0.05$ in all cases.

A two factor (Speed Context Position) within-subject ANOVA showed a significant main effect of context position ($F1–49 = 12.25$, $p < 0.01$). The main effect of speed ($F2–98 = 12.25$, $p = 0.39$) and the interaction between speed and context position ($F2–98 = 12.25$, $p = 0.15$) were not significant. Subsequent pairwise comparison tests using Tukey’s HSD (Honestly Significant Difference) test indicated that the mean proportion of first percept measures was significantly higher when the context figure was presented below compared to above the target figure at the $p < 0.05$ level.

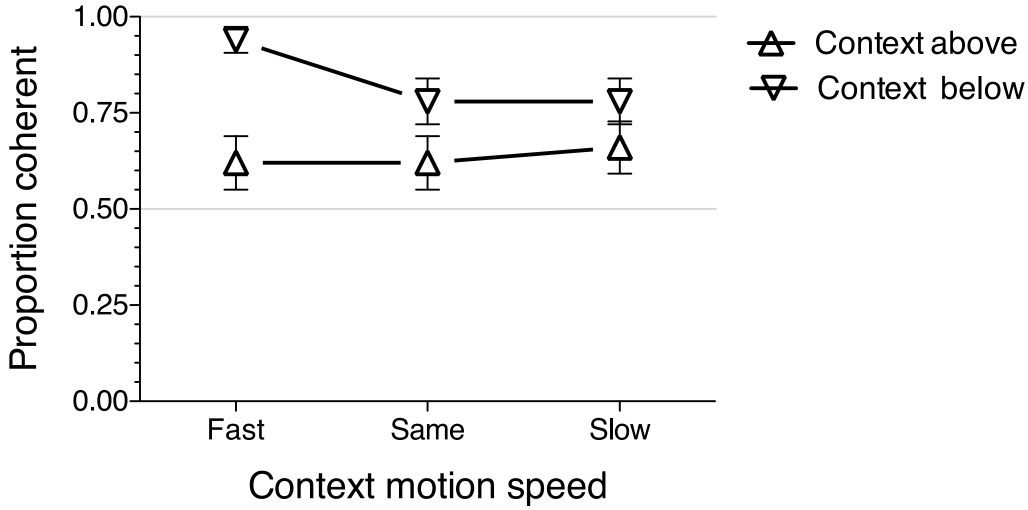


Figure 3.5. First percept measures for each speed and context position. The figure shows the proportion of times the target figure was reported as being coherent with the context figure at the start of each trial, averaged across observers. Error bars are standard errors of the mean.

Flip ratio

The number of perceived reversals is another common metric used to describe the perceptual instability of an ambiguous figure (Adams & Haire, 1959; Intaitė, Šoliūnas, et al., 2013; Li & Kingdom, 1999; Long et al., 1983, 1992). The ratio of perceived reversals of the target figure to the number of physical context figure reversals is summarized in Figure 3.6.

A two factor (Speed Context Position) within-subject ANOVA revealed once more a significant main effect of context position ($F1–49 = 24.13$, $p < 0.001$). No main effect was found for context speed ($F2–98 = 1.20$, $p = 0.3046$) nor was there a significant interaction between speed and context position ($F2–98 = 1.79$, $p = 0.1715$). Subsequent pairwise comparison tests, conducted using Tukey's HSD test, indicated that the ratio of perceived reversals was significantly higher when the context figure was presented above compared to below the target figure ($p < 0.0001$).

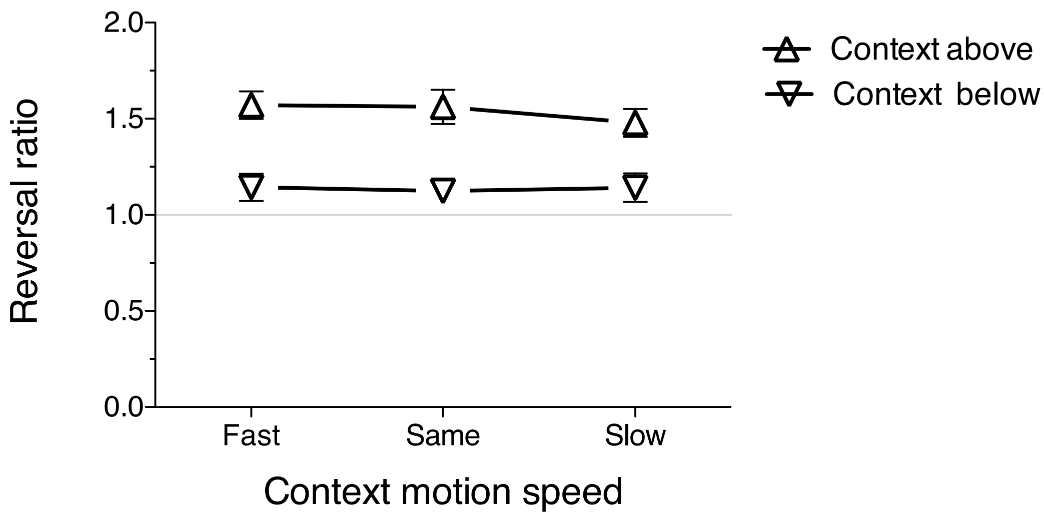


Figure 3.6. Ratios of the number of reported target figure rotation-direction changes to context rotation-direction reversals, for each speed and context position, averaged across observers. The value of 1 implies equal numbers of alternations. Error bars are standard errors of the mean.

Phi coefficients

To assess in more detail the influence over time of the context figure on the perception of the target, we used a method analogous to the method of reverse correlation used in both psychophysics and neurophysiology. We compared the timelines of the context figure with the observer responses within each trial. An example of such a timeline is shown in Figure 3.7. Each timeline indicates the rotation direction of the context and the reported rotation direction of the target figure. These timelines reveal the moments when the target appeared to change in direction as well as the periods of stability of both target and context.

We used as the measure of the correlation between the context and response timelines the Phi coefficient introduced by Karl Pearson, which is the Pearson product-moment correlation applied to binary data (Everitt, 1992). We shifted the start of the response timeline by an amount given in multiples of 100ms, and for each multiple calculated the Phi-coefficient. The analysis was conducted for 10 multiples, i.e. up to the maximum shift of 1 s. From the 10 Phi coefficients obtained for each trial, the maximum was taken and retained. Since Phi coefficient analyses require balanced data- sets, we truncated the context timeline by the amount of shift applied to the response timeline. The following values corresponded to the mean (M) and standard deviations (SD) of the shifts in seconds that produced the maximum Phi coefficient values for fast up and down, same up and down, and slow up and down: M=.61 SD=.40, M=.65 SD=.28, M=.57 SD=.37, M=.72 SD=.20, M=.52 SD=.43, and M=.61 SD=.37, respectively. Figure 3.8 presents individual data as well as summarizes the (maximum) Phi coefficients obtained using this method, calculated for each condition.

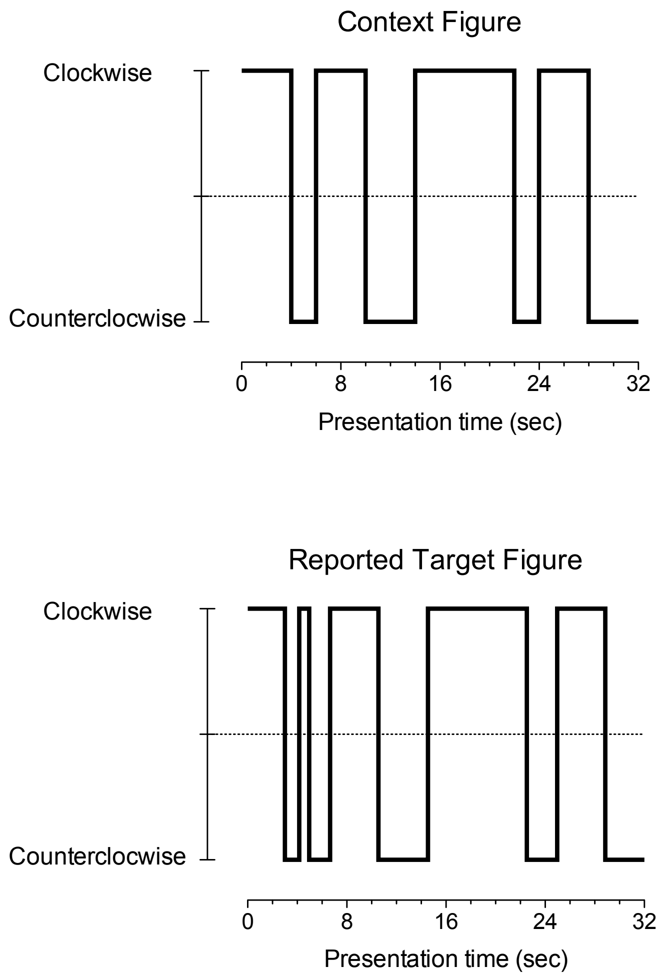


Figure 3.7. Example of context and target timelines taken from a sample trial. The square waves describe the context figure rotation directions (top) and the reported rotation directions of the Necker cube (below).

A two factor (Speed Context Position) within subject ANOVA was conducted on the Phi coefficients in order to assess the effects of the two factors across the time-line. A main effect of context position ($F_{1-49} = 64.58$, $p < 0.0001$) was found and this time also a significant main effect of speed ($F_{2-98} = 14.30$, $p < 0.0001$). The two factor ANOVA also yielded a significant interaction between speed and context position ($F_{2-98} = 3.73$, $p < 0.05$). Separate one- way ANOVAs were conducted between speeds at each context position revealed no significant effect of speed when the context figure was presented above the target ($F_{2-196} = 1.54$, $p = 0.2172$);

however, speed had a significant effect when the context figure was presented below ($F_{2-196} = 16.07$, $p < 0.0001$). Subsequent pair- wise comparison tests, conducted using Tukey's HSD test, indicated that the highest Phi coefficient value was found when the context figure was presented below and at the same speed as the target, as compared to the fast and slow speed conditions, and that the fast speed condition yielded a higher Phi coefficient than the slow speed condition, all ps < 0.01.

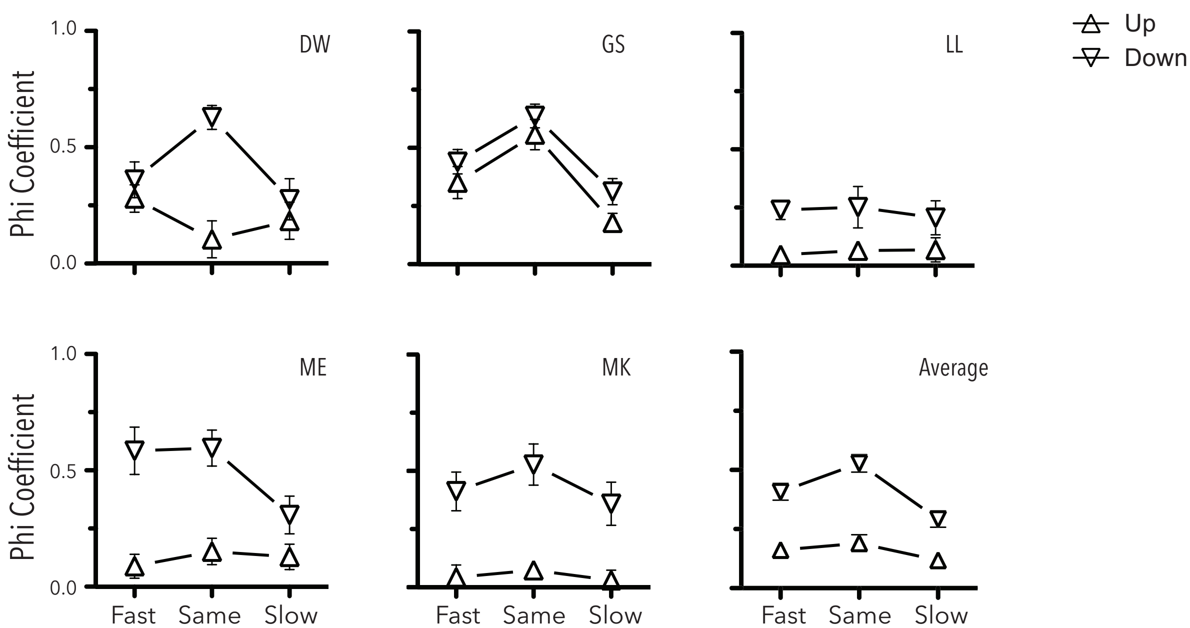


Figure 3.8. Individual Phi coefficient measures for each speed and context position are presented for 5 observers. The bottom right figure presents the average measures across observers. Error bars are standard errors of the mean.

Experiment 2: effect of orientation axis

Given the previous finding of a main effect of context position when context and target were arranged on a vertical axis, we decided to measure the context effect across a range of orientation axes. For this we measured the context effects for a series of context positions, defined in terms of an orientation axis. For example, the orientation axis of 90° means a vertical arrangement with the context above the target, whereas the orientation axis of 270° means a vertical arrangement with the context below the target. Orientation axes of 0°, 45°, 90°, 180°, 225°, and 315° were tested, as shown in Figure 3.3. Each orientation axis was compared to its corresponding opposite, for example 0 vs. 180, and we refer to this as an orientation axis pair. The orientation axis pairs that were compared were as follows: horizontal (0 vs. 180), right-oblique (45 vs. 225), vertical (90 vs. 270), and left-oblique (135 vs. 315).

We also compared context effects in frame-sequence- independent and frame-sequence-locked conditions (see Methods). In Experiment 1, frame-sequence-locking was not possible for speed conditions other than the one in which both context and target had the same speed, so frame-sequence-independence was employed. The effect of frame-sequence-locking was to pre- serve the coherence of the context and target tilts, as illustrated in Figure 3.4.

First percept – frame-sequence-independent

First percept reports were first analyzed to test whether they were different from chance (i.e., 0.5). One-sample t-tests indicated that only the 270° (M = 0.78 SEM = .43) orientation axis was significantly greater than chance: t (39) = 4.113 p < 0.001. For orientation axes 0 (M=0.48 SEM=.51), 45 (M=0.60 SEM=0.50), 90 (M=0.53 SEM=0.51), 135 (M=0.63 SEM=0.49), 180 (M=0.53 SEM = 0.51), 225 (M = 0.50 SEM = 0.51), 315 (M = 0.55 SEM = 0.51), the t (39) values were respectively 0.31, 1.27, 0.31, 1.61, 0.31, 0.0, 0.63, with all ps > 0.05.

A one factor within-subject ANOVA revealed a non-significant effect of orientation axis (F7–273 = 1.47, p = 0.17) on first percept reports. Figure 3.9A summarizes the proportion of times the target appeared to move in the same direction as the context at the start of each trial for each of the orientation axis.

First percept – frame-sequence-locked

First percept reports were analyzed once more to test whether they were greater than chance. One sample t-tests indicated that all the mean values of first percept reports were greater than chance except for the 135 ($M = 0.65$ $SEM = 0.48$) orientation axis $t(39) = 1.964$ $p = 0.057$. For orientation axes 0 ($M = 0.90$ $SEM=0.30$), 45 ($M=0.80$ $SEM=0.41$), 90 ($M=0.90$ $SEM=0.30$), 135 ($M = 0.65$ $SEM = 0.48$), 180 ($M = 0.88$ $SEM = 0.33$), 225 ($M=0.98$ $SEM=0.16$), 270 ($M=1.00$ $SEM=0.00$) and 315 ($M = 0.95$ $SEM = 0.22$), the $t(39)$ values were 8.33, 2.841, 4.68, 8.33, 7.08, 19, and 12.89 respectively, all $ps < 0.01$. For the 270° orientation axis, all participants reported a concordance of first percept reports between context and target figure 100% of the time.

A one factor within-subject ANOVA revealed a significant effect of positional axis ($F7–273 = 5.37$, $p < 0.0001$). Subsequent planned comparisons were conducted using independent groups t-tests with a Bonferroni-adjusted alpha level of .0125 (0.05/4) for each pair of spatial locations: horizontal, vertical, right-oblique, and left-oblique. Differences of first percept reports were found to be non-significant for all orientation axes tested, horizontal $t(39) = 0.36$ $p = .72$, right-oblique $t(39) = 2.55$ $p = .013$, vertical $t(39) = 1.46$ $p = .14$, except for the left-oblique orientation axis, where the proportion of first percept reports was significantly higher in the 315 than 135, $t(39) = 4.37$ $p < 0.0001$. Figure 3.9B summarizes the proportion of times the target appeared to move in the same direction as the context at the start of each trial.

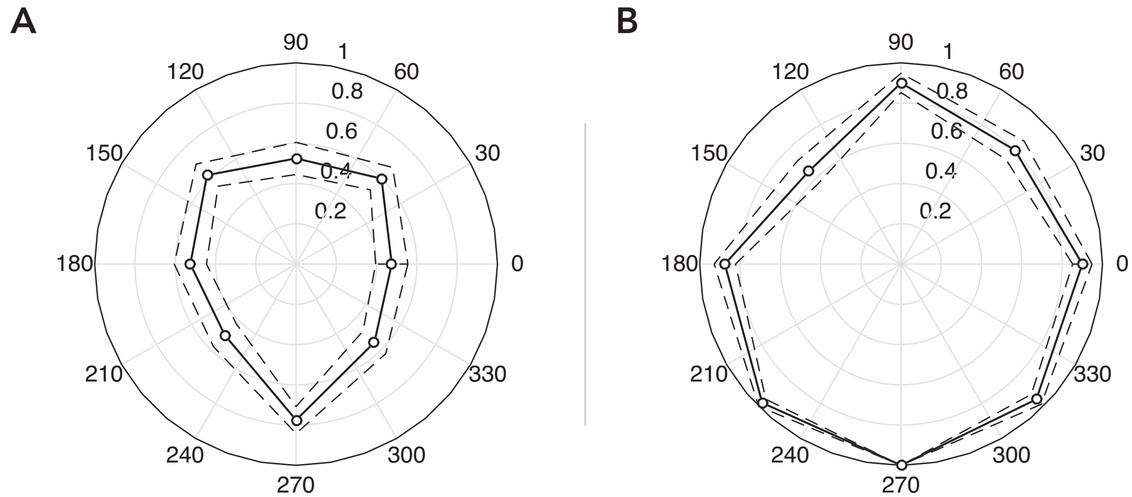


Figure 3.9. First percept measures for frame-sequence-independent A and frame-sequence-locked B for each of the orientation axes tested. Numbers around the circle are orientations, numbers within the circle give the proportion of times the target figure was reported as being coherent with the context figure at the start of each trial, averaged across observers. The value of 1 (outer edge of the plot) implies perfect consistency between first-percept reports of context and target. Error bars (dashed lines) are standard errors of the mean.

Flip ratio – frame-sequence-independent

The ratio of the number of perceived reversals of the target figure to the number of physical context figure reversals is summarized in Figure 3.10A. A one factor within-subject ANOVA revealed a significant effect of orientation axis ($F_{7-273} = 4.67$, $p < 0.0001$). Bonferroni post hoc contrasts revealed no significant response-to- context ratio differences for the horizontal, right-oblique, and left-oblique orientation axis pairs, $t(39) = 0.26, 2.07, 1.48$, respectively all $ps>0.05$. Significant response-to-context ratio differences were found for the vertical orientation axis pairs $t(39) = 3.27$ $p < .001$, where ratios were found to be smaller for 270° compared to 90° orientation axis and where 90° was closer to 1.

Flip ratio – frame-sequence-locked

A one factor within-subject ANOVA was performed and revealed a significant effect of orientation axis ($F_{7–273} = 2.73$, $p < 0.01$). Using Bonferroni post hoc contrasts, the differences in response-to-context ratios were found to be non-significant for all orientation axis pairs tested, t (39) = .70, 1.91, 2.27, and, 2.19 for horizontal, left-oblique, vertical, and right-oblique, respectively, all ps > 0.05. The ratio of perceived reversals of the target figure to reversals of the physical context figure reversals is summarized in Figure 3.10B.

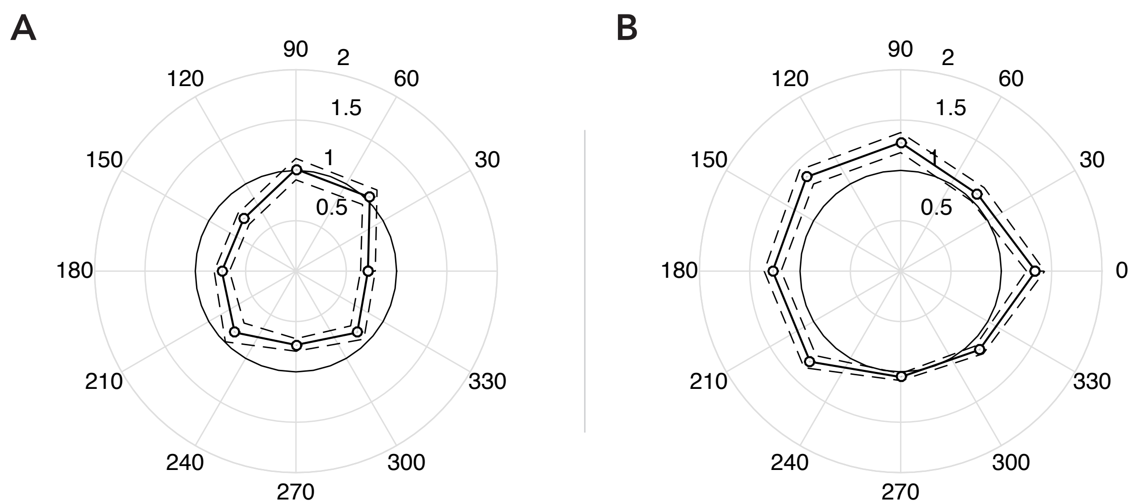


Figure 3.10. Ratios of the number of the reported target rotation-direction changes to context figure rotation-direction changes, for each orientation axis, averaged across observers. A frame-sequence-independent, B frame-sequence-locked conditions. The value of 1 (inner circle) implies equal numbers of alternations. Error bars (dashed lines) are standard errors of the mean.

Phi coefficient – frame-sequence-independent

As with the first experiment, we determined the delay between the reported motion direction of the target figure and the motion direction of the context figure that produced the maximum correlation between the two. The following values correspond to the mean (M) and standard deviations (SD) of the shifts, in seconds, that produced the maximum Phi coefficient values orientation axes 0, 45, 90, 135, 180, 225, 270, and 315: M=.64 SD=.46, M=.72 SD=.45, M=.81 SD=.38, M=.73 SD=.43, M=.66 SD=.46, M=.70 SD=.43, M=.86 SD=.28, and M=.64 SD=.47, respectively. The left hand plot of Figure 3.11A presents individual data, while the right hand plot Figure 3.11A summarizes the resulting maximum Phi coefficients. There was a main effect of orientation axis ($F_{7-273} = 3.76$; $p < 0.001$). Bonferroni post hoc contrasts revealed a significant difference in maximum Phi coefficient between the vertical orientation axis pair $t(39) = 3.31$ $p < 0.01$, while all other orientation axis pairs did not differ significantly; horizontal, left- oblique, and right-oblique: $t(39) =$ 1.07, .39, 1.28, respectively, all ps > 0.05.

Phi coefficient – frame-sequence-locked

The following values corresponded to the mean time shifts that produced the maximum Phi coefficient values for orientation axis 0, 45, 90, 135, 180, 225, 270, and 315: M=.53 SD=.33, M=.66 SD=.35, M=.68 SD=.36, M=.66 SD=.39, M=.61 SD=.29, M=.67 SD=.31, M=.63 SD=.15, and M=.68 SD=.23, respectively. A one factor within-subject ANOVA was performed and revealed a significant effect of orientation axis ($F_{7-273} = 20.97$, $p < 0.0001$). The left hand plot of Figure 3.11B presents individual data, while the right hand plot of Figure 3.11B summarizes the resulting maximum Phi coefficient. Subsequent pairwise Bonferroni post hoc contrasts revealed significant differences in Phi coefficient values the horizontal ($t(39) = 3.18$ $p < 0.01$), vertical ($t(39) = 9.80$ $p < 0.0001$) and left-oblique ($t(39) = 4.30$ $p < 0.0001$) orientation axis pairs, but not between the right-oblique orientation axis pair ($t(39) = 0.55$ $p = .58$).

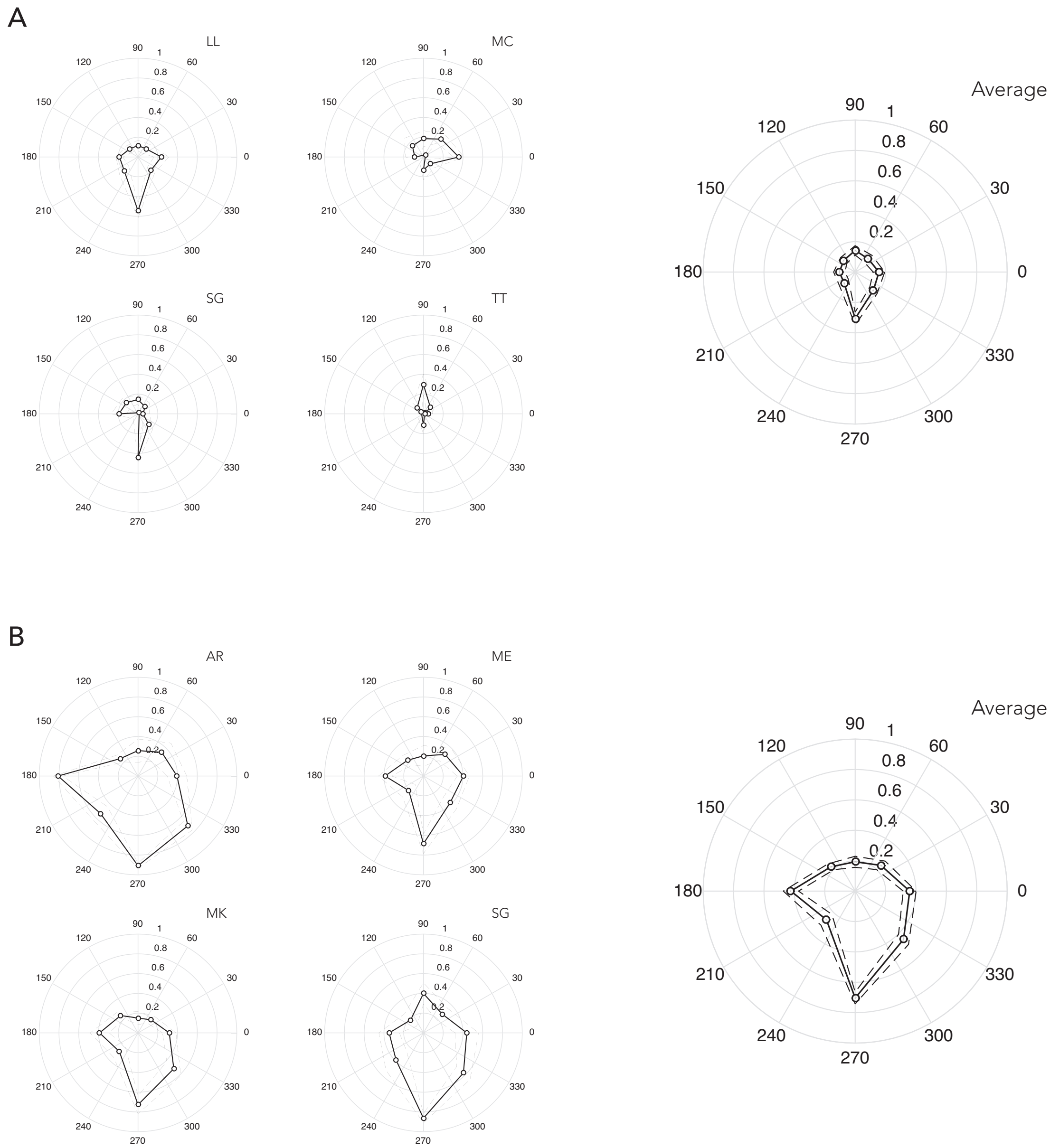


Figure 3.11. Phi coefficient measures for A frame-sequence-independent and B frame-sequence-locked conditions, averaged across observers, for each orientation axis. The left hand plots present individual Phi coefficient measures while the right hand plot gives the average measures across observers. Error bars (dashed lines) are standard errors of the mean.

Discussion

Our goal was to determine whether perceptual motion binding existed between an unambiguous context figure, in the form of a skeleton cube, and a Necker cube, and if it did, to determine the properties of motion binding. The main results are:

Perceptual motion binding: The rotation direction of the context had a significant influence on the perceived rotation direction of the target, revealing perceptual motion binding between the two figures.

Context speed: Regardless of relative position, the relative speed of context and target figure had little to no influence on the strength of perceptual binding. Context position: Regardless of relative motion speed, perceptual binding was greatest when the context was presented below the target figure.

Context speed/position interaction: An interaction was found between speed and context position, whereby perceptual binding was greatest when the speed of the context and target were the same and the context was presented below the target.

Context speed: common-fate versus change synchrony

A stationary Necker cube alternates in perceived spatial configuration, either appearing to face upwards or downwards. As a consequence, a rotating Necker cube spontaneously changes in its perceived rotation direction. Previous studies have showed that an ambiguous context biases the perceived spatial configuration of a Necker cube (Dobbins & Grossmann, 2010; Sundareswara & Schrater, 2008). We find that an unambiguous rotating context influences the perceived motion of a Necker cube. Intuitively we would expect that the influence of the context would be greatest when context and Necker cube were moving at the same speeds. However, this was not found. Instead, perceptual motion binding was found to be unaffected by the relative speed of context and target, at least when analyzed across context positions. How might this be explained?

The Gestalt law of common fate states that objects that move together, i.e. have a common motion direction and speed, are perceived to group together (Wertheimer, 1923). Grouping by common fate typically results in objects being perceived to form a rigid body in motion. With our stimuli, the possibility of perceiving a rigid body of context and target only existed when the target and context moved at the same speed. However, since we found no main effect of the relative speed of context and target on perceptual binding, it does not appear that common-fate is driving motion coherence between these figures. So what else could produce perceptual motion binding between context and target?

Change-synchrony refers to the idea that objects that are perceived to change together, along one dimension or another, are perceived as bound together (Alais, Blake, & Lee, 1998). The context stimuli in our experiment were generated such that their motion direction changed only when the face of the figure was fronto-parallel. We did this because Pastukhov, Vonau, and Braun (2012) showed that perceived reversals in a rotating stimulus, created from two perpendicular bands of bright dots, occurred pre- dominantly at rotation angles at which the face or edge of the figure was fronto-parallel. Therefore, our context was made to change motion direction at rotation angles that were optimal for producing perceived target reversals. This made possible synchronous, though not necessarily coherent, reversals between context and target, irrespective of relative motion speed. When analyzed across context positions, we indeed found that such synchronous motion direction changes occurred, that these changes were coherent, and that they occurred irrespective of relative motion speed between context and target. This finding supports the idea that change synchrony not common fate mediated perceptual binding in our stimuli.

Further support for change-synchrony as the critical factor was found when we changed the frame presentation sequence for both context and target from independent to locked. In the frame- sequence-independent condition, context reversals, achieved through a reversal of the context's animation sequence, did not alter the animation sequence of the target. As a result, the figures' tilt relationships could be coherent or incoherent, as illustrated in the top portion of Figure 3.4. In the frame-sequence-locked condition on the other hand, context reversals were always accompanied by reversal of the target's animation sequence. By locking the frames of animation of both context and target, the tilt- coherence between these figures was maintained during the entirety of the presentation, as illustrated in the bottom of Fig. 4 (it is important to bear in mind however that even with frame- locking, there is nothing inevitable about coherent motion rotation between context and target). We found that synchronous changes of both motion and tilt in the frame-sequence-locked condition produced greater perceptual motion binding than those observed in the frame-sequence-independent. This finding strengthens the idea that change-synchrony plays the critical role in the perceptual binding between our unambiguous context and ambiguous target figures.

Spatial position – an asymmetry of context influence

Context position was found to significantly affect the strength of the contextual influence, with the strongest influence observed when the context was presented below the target, and regardless of relative motion speed. Why might this be so? One possibility is that the visual system makes the assumption that the upper figure ''rests" on the lower figure and is bound to it by friction. There are good examples in the literature of the influence of in-built knowledge of friction physics on perception, most notably the study by Gilroy and Blake (2004). In Gilroy and Blake's study, observers were presented with side-by-side pairs of shape-from- motion-defined spheres, where one of the spheres was ambiguous in rotation direction. It was found that when the two spheres appeared to touch, they appeared to counter-rotate, as if bound by friction. When the velocity of the unambiguous sphere was inconsistent with friction or when the spheres were presented with a small gap separating them, the probability of perceived counter-rotation was significantly reduced. This second finding may seem inconsistent with the speed-irrelevant context-below advantage we find; however, our context and target were always vertically separated, not horizontally separated or abutting, and this may be the reason.

Further support for the friction explanation of the position effect comes from the interaction we observed between relative speed and context position: when the context was below the target the biggest influence was observed then the speeds of context and target were the same. Thus there is the possibility that when the position relationships between the figures favored a frictional interpretation, the advantage of common speed manifested itself.

Another possible reason for the context-below advantage is the contribution of the view-from-above bias demonstrated in previous Necker cube studies (Dobbins & Grossmann, 2010; Sundareswara & Schrater, 2008). These studies showed that the dominant configuration was consistent with the observer looking down on the cube. A view-from-above bias has also been shown to significantly influence the interpretation of rotating ambiguous stimuli such as the famous rotating ballerina by Japanese Flash designer Nobuyuki Kayahara. Troje and McAdam (2010) set the viewpoint of the ballerina to above and found that the silhouette became unambiguous. In our study when the context was above the figure there may have been a tendency to perceive the target as viewed from above (because it was below the fixation point), and this could have weakened the influence of the context. Future experiments will be needed to test between the friction and view- point explanations of the positional context effect found here.

Summary and conclusion

Using a novel method analogous to reverse correlation we have measured the influence of an unambiguous rotating skeleton cube on the perceived direction of motion of a nearby target Necker cube. Our measurements have revealed that changes in the context figure's motion direction trigger changes in the perceived motion direction of the target, irrespective of the relative speed of context and target. This suggests that change-synchrony not common-fate underpins perceptual binding between context and target. We also found that the influence of the context was greater when below rather than when above the Necker cube. We considered possible explanations for this finding, including the influence of prior knowledge of friction forces and the view-from-above bias.

Chapter 4

Given the results presented in Chapter 3, this chapter explores whether the contextual effects reported generalise to a different class of perceptually unstable figures, namely binocularly rivalrous figures. This was investigated by measuring the effect of a spatial context on the perceived interpretation of a bi-stable and a rivalrous target that were perceptually equated.

Common contextual influences in ambiguous and rivalrous figures

This chapter has been published as

Ouhnana, M., Jennings, B. J., & Kingdom, F. A. A. (2017). Common contextual influences in ambiguous and rivalrous figures. *PLoS ONE*, *12*(5), e0176842. http://doi.org/10.1371/journal.pone.0176842

Abstract

Images that resist binocular fusion undergo alternating periods of dominance and suppression, similarly to ambiguous figures whose percepts alternate between two interpretations. It has been well documented that the perceptual interpretations of both rivalrous and ambiguous figures are influenced by their spatiotemporal context. Here we consider whether an identical spatial context similarly influences the interpretation of a similar rivalrous and ambiguous figure. We developed a binocularly rivalrous stimulus whose perceptual experience mirrors that of a Necker cube. We employed a paradigm similar to that of Ouhnana and Kingdom (2016) to correlate the magnitude of influence of context between the rivalrous and ambiguous target. Our results showed that the magnitude of contextual influence is significantly correlated within observers between both binocularly rivalrous and ambiguous target figures. This points to a similar contextual-influence mechanism operating on a

common mechanism underlying the perceptual instability in both ambiguous and rivalrous figures.

Introduction

Our perception of the visual world, while in most situations stable and close to veridical, is sometimes ambiguous and unstable. The Necker cube is a well-known example of an ambiguous figure - a skeleton cube whose perceptual interpretation alternates between facing upwards or downwards (Necker, 1832). Another form of ambiguity is binocular rivalry, in which different stimuli presented to the two eyes alternate perceptually in dominance (Blake & Logothetis, 2002). Usually it is the percept of the two stimuli that alternates, but in some cases one perceives a piecemeal mosaic of each eye's input (Kovacs, Papathomas, Yang, & Feher, 1996).

The perceptual instability experienced in both ambiguous and rivalrous stimuli has been argued to arise through competition and inhibition between the neurons sensitive to each stimulus, via feedforward, feedback, or as recently suggested, a hybrid of these mechanisms see (Long & Toppino, 2004) for a review on ambiguous figures, and (Alais, 2012; Blake & Logothetis, 2002; Brascamp, Klink, & Levelt, 2015) for reviews on binocularly rivalrous figures.

Various manipulations alter the perceived interpretation of both ambiguous and rivalrous figures. These manipulations range from the effects of adaptation to temporal priming (Long, Toppino, & Mondin, 1992; Toppino & Long, 1987) and extend to manipulations such as learned associations (Haijiang, Saunders, Stone, & Backus, 2006). Long et al. (1992) showed that when an *unambiguous* version of an ambiguous figure is presented prior to an ambiguous target, the perceptual interpretation depends on the length of time the former figure is presented. When briefly presented, a priming effect causes the interpretation of the ambiguous target to shift *towards* that of the previously inspected unambiguous figure. However when the prior unambiguous figure is presented for an extended length of time, the interpretation of the figure is shifted *away* from that of the adapted figure. Haijiang et al. (2006) demonstrated that a particular interpretation of an ambiguous Necker cube figure could be cued by the motion or position of a disambiguated Necker cube, or even by a sound, that had been previously associated with that interpretation.

Not surprisingly, simultaneous contextual influences also alter the perceptual interpretation of both ambiguous and rivalrous figures (Intaitė, Noreika, Šoliūnas, & Falter, 2013; Kovacs et al., 1996; Ouhnana & Kingdom, 2016; Sobel & Blake, 2002), and are thus considered to be valuable tools for probing their underlying mechanisms (Klink, van Wezel, & van Ee, 2012). Intaitė et al. (Intaitė et al., 2013), using a display of overlapping skeleton squares which alternated between being perceived as stacked upwards or stacked downwards, showed that when the squares were flanked by unambiguous coherently stacked versions, observers' first percepts were biased towards that of the flanking context. The influence of spatial context on ambiguous figures has also been revealed through the introduction of backgrounds favoring a particular percept (Sundareswara & Schrater, 2008), or when the context itself is ambiguous (Dobbins & Grossmann, 2010). Using a Necker cube as stimulus, Dobbins and Grossmann (2010) reported that the variability of first percept reports, which favored the view-from-above bias, was reduced when the target Necker cube was presented in an array of other similarly ambiguous cubes than when presented in isolation.

For binocularly rivalrous figures the influence of context has been studied using a variety of paradigms. Kovacs et al. (1996) reported that when individual grids of red and green dots were presented dichoptically, the resulting percept alternated between the two colours. When half of the coloured dots were swapped across the two grids, the resulting percept was a grid defined by a single color. Sobel and Blake (2002) studied the influence of global coherence using a dichoptic display of four gratings, where one of these rivaled with a rotating radial checkerboard. The non-rivalrous flanker gratings drifted in motion directions that were either consistent or inconsistent with a coherent global percept. When the context motion was consistent with a global percept, the predominance of the grating percept in the rivalling pair was found to be significantly higher. The effect of global coherence in binoculary rivalrous displays has also been shown to occur using combinations of monocularly rivalrous figures. Vergeer et al. (2016) presented a shape-from-motion-defined cylinder with the leftward drifting dots to one eye and the rightward drifting dots to the other. Observers predominantly reported perceiving a cylinder rather than a convex or concave shell of horizontally drifting dots.

In this communication, we ask whether such simultaneous contextual influences operate similarly for both ambiguous and binocularly rivalrous stimuli, implying a common context-influence mechanism. Our study serves also to address the wider issue of whether the two forms of perceptual instability are mediated by a common underlying mechanism.

While it is tempting to draw parallels between the contextual influences observed in both rivalrous and ambiguous figures, it is important to note that the contexts described in the literature for both classes of figures are never the same. In the case of ambiguous figures, the contexts are always unambiguous in interpretation, whereas with binocularly rivalrous figures the contexts are invariably one component of the competing stimuli. Thus in order to investigate whether context has a similar influence on ambiguous and rivalrous figures one has to equate both the context figures as well as the perceptually unstable figures themselves.

When a rotating Necker cube is observed, the perceptual experience is of a skeleton cube rotating in a clockwise or counter-clockwise motion direction. A binocularly rivalrous version of this figure can be created by presenting clockwise and counter-clockwise rotating unambiguous skeleton cubes to each eye. The perceptual experience of this rivalrous skeleton cube is similar to that of its ambiguous Necker cube counterpart, that is, its rotation appears to fluctuate from clockwise to counter-clockwise and back. Importantly, although the visual experience of our rivalrous skeleton and ambiguous Necker cube are comparable, the physical properties of the figures differ in crucial ways: one is defined by a dichoptic pair of non-ambiguous figures, while the other is defined by a single binocular ambiguous figure.

In the current study we investigated whether the introduction of an identical spatial context similarly influenced the perceptual interpretation of rivalrous and ambiguous stimuli when both context, rivalrous and ambiguous figures are equated. A paradigm similar to Ouhnana and Kingdom (2016) was employed to compare the degree of coherence between the target and context figures. In their study, observers were instructed to report the motion direction of a rotating Necker cube in the context of an unambiguous rotating skeleton cube that changed motion direction randomly throughout the trials. Through a procedure akin to reverse correlation, a Phi coefficient was computed that described the extent to which the context influenced the perceived motion direction of the target. A similar procedure is used here.

General Methods

Subjects

Five observers participated in both experiments, two were authors (MO and BJ), while the others were naïve to the experimental aims. All had normal or corrected-to-normal visual acuity. Written consent was obtained for all participants and all experimental protocols were approved by McGill University Research Ethics Board.

Apparatus

The stimuli were created using Blender (version 2.67b), an open-source 3D computer graphics rendering environment, and displayed using Psychophysics Toolbox (Version 3.12) running under Matlab (MathWorks Inc., version 8.1). The stimuli were presented on a ViewSonic G225f Graphic Series CRT monitor connected to an Apple Mac Pro. The display was driven at 60 Hz with a resolution of 1024x768.

Participants viewed the display through a modified Wheatstone stereoscope, which employed four front-surfaced mirrors per eye, with an effective viewing distance along the light path of 55 cm. The stereoscope had a visible window (to each eye) of approximately 9.8 deg x 12.4 deg.

Stimuli – General

The stimuli consisted of rotating unambiguous skeleton cubes and rotating ambiguous Necker cubes. The difference between the skeleton and Necker cubes was that in the skeleton cubes the back parts of the skeleton were occluded by the front parts, rendering the cubes unambiguous. The skeleton cubes were used for the contexts and for the binocularly rivalrous targets, while the Necker cubes were used as the ambiguous targets. Depending on the experimental condition, the context figure was either presented as a single or a pair of figures, where the pair of figures were identical. The rivalrous skeleton cubes were always presented in pairs regardless of the experimental condition. Each of the rotating skeleton cubes was presented with an opposite rotation direction. In other words, if the cube was facing upwards in one eye and rotating clockwise, the other eye would be presented with a cube facing downwards with an anti-clockwise motion direction. The ambiguous pair was either presented as an identical pair, one to each eye, or as single figure presented to one eye. All cubes measured ~3 x 3 degrees of visual angle. All figures rotated at a quarter of a revolution per second and all animations were rendered at 60 frames per second. Orthographic projection was used for all stimulus conditions, i.e. there were no perspective cues.

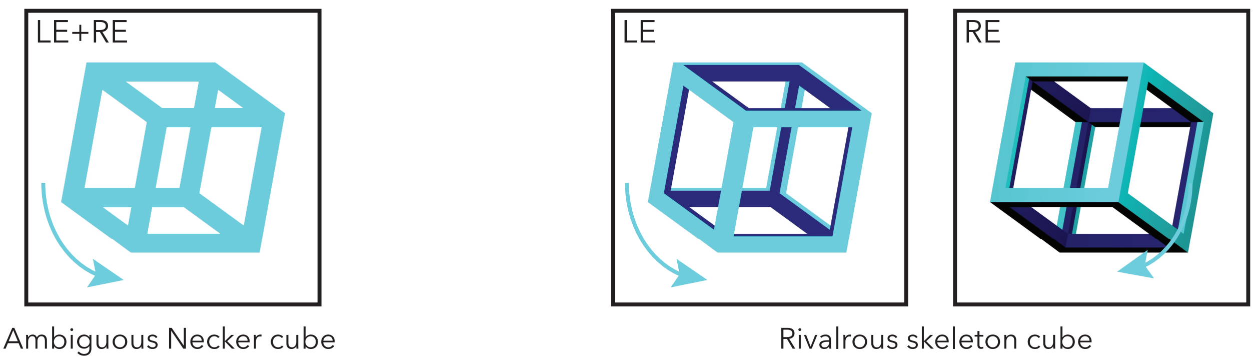


Figure 4.1. Example of the ambiguous Necker cube (left) and binocularly rivalrous skeleton cube (right pair). One of the skeleton cubes also served as the context figure. Orthographic projection was used to render ambiguous the motion direction of the Necker cube, while the addition of shading and color difference disambiguated the motion directions of skeleton cubes.

Stimuli - Ambiguous Necker cube

The ambiguous Necker cube, shown on the left of Figure 4.1, was created using Blender meshes with materials set to teal with shading turned off to maintain ambiguity. The Necker cube was presented either as a single figure or a pair of figures depending on the experimental condition. The figure's XYZ rotational space was set to [20,0,0], with Z dictating the cube's rotation. Blender XYZ rotational space corresponded to roll, pitch, and yaw, where yaw dictated the figure's rotation around its vertical axis, each trial starting at zero.

Stimuli - Context and binocularly rivalrous skeleton cubes

For the context and rivalrous skeleton cubes, as shown on the right of Figure 4.1, the figures were created using Blender meshes with materials set to teal for their outer surfaces and dark blue on their inner surfaces. Due to the orthographic projection used to render the stimuli, the external and internal surface color differences added occlusion information serving to disambiguate the figure's motion direction (clockwise or anti-clockwise). The figure's XYZ rotational space was set to [20,0,0], with Z dictating the cube's rotation.

Experimental conditions & Procedures

There were three eye presentation conditions for both types of target (ambiguous Necker cube and rivalrous skeleton cube), as illustrated in Figure 4.2. In addition the context was positioned either above or below fixation. The three eye presentation conditions were as follows: 1. A pair of identical context figures presented binocularly - see the leftmost column of Figure 4.2 highlighted in blue. 2. A single context figure presented monocularly: in the case of the rivalrous skeleton cube this was to the same eye as the test skeleton cube that matched the context's motion direction - see the top middle column of Figure 4.2 highlighted in green; in the case of the ambiguous Necker cube this was to the same eye as the monocularly presented ambiguous target - see bottom middle column of Figure 4.2 highlighted in green. 3. A single context figure presented monocularly: in the case of the rivalrous target this time to the opposite eye as the skeleton cube that matched the context's motion direction - see the top right of Figure 4.2 highlighted in red; in the case of the ambiguous Necker cube this time to the opposite eye as the monocularly presented ambiguous target - see the bottom of the rightmost column of Figure 4.2 highlighted in red.

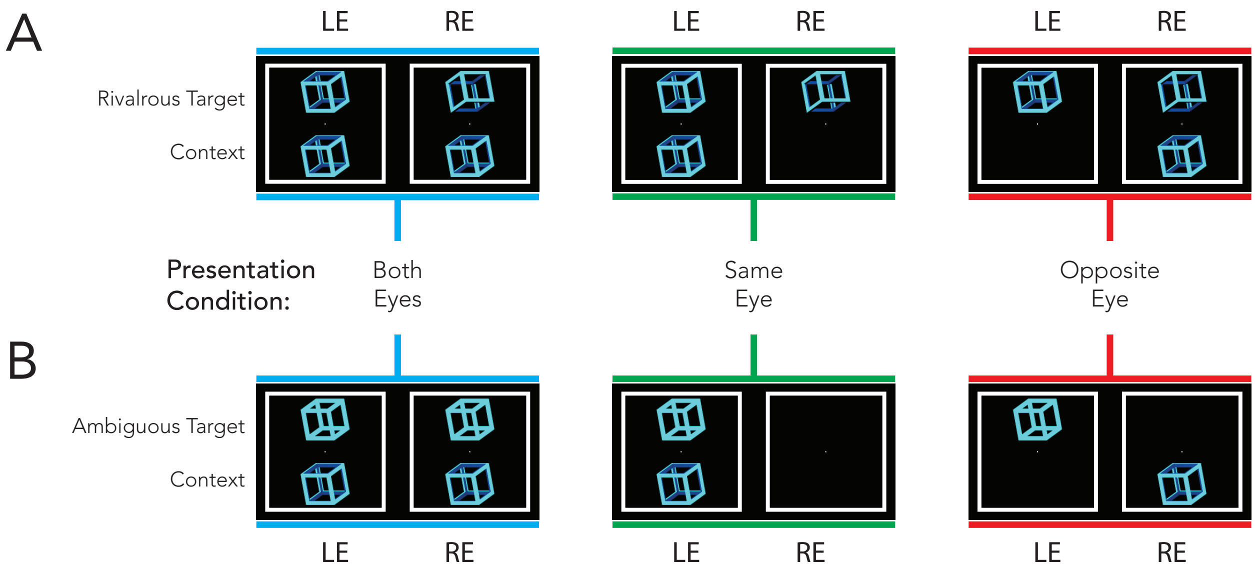


Figure 4.2. Sample frames illustrating each context and target condition for the rivalrous target (top), and ambiguous target (bottom). The context figure was presented either to both eyes (left hand side), the same eye (middle), or the opposite eye (right hand side) of the target. For each condition, the context was presented either above or below the fixation dot.

Observers were instructed to report via a key-press when the context and target were perceived to rotate in the same motion direction and by a different key press when they were perceived to rotate in opposite motion directions. The motion direction of the context figure was randomly selected at the start of every trial. During each of the 30s trials, there was a 0.5 probability that a change in context motion direction would occur when the context was aligned fronto-parallel. The context and target figure were either swapped across eyes (experiment 1) or remained in the same eye (experiment 2). All conditions were interleaved and 10 repetitions were performed.

Results

Analysis

For both Experiments 1 and 2, an analysis of the magnitude of contextual influence was first analyzed in terms of the concordance between the motion direction of the context figure and the perceived motion direction of the target across each trial. This was accomplished through a method analogous to reverse correlation, as in our previous study (Ouhnana & Kingdom, 2016). The magnitude of influence was calculated using the Phi coefficient, which is a measure of association between two binary variables. The two variables in our study were the *clockwise* and *counter-clockwise* motion, or perceived motion direction of respectively the context and target figure.

To extract the perceived target motion direction from the observer responses, the reported motion directions (same versus opposite) throughout each trial were compared to the clockwise versus counter-clockwise motion direction timeline of the context figure. An example context motion sequence (top), reported observer response (second from the top), and extracted perceived target motion sequence (three from the bottom) is illustrated in Figure 4.3.

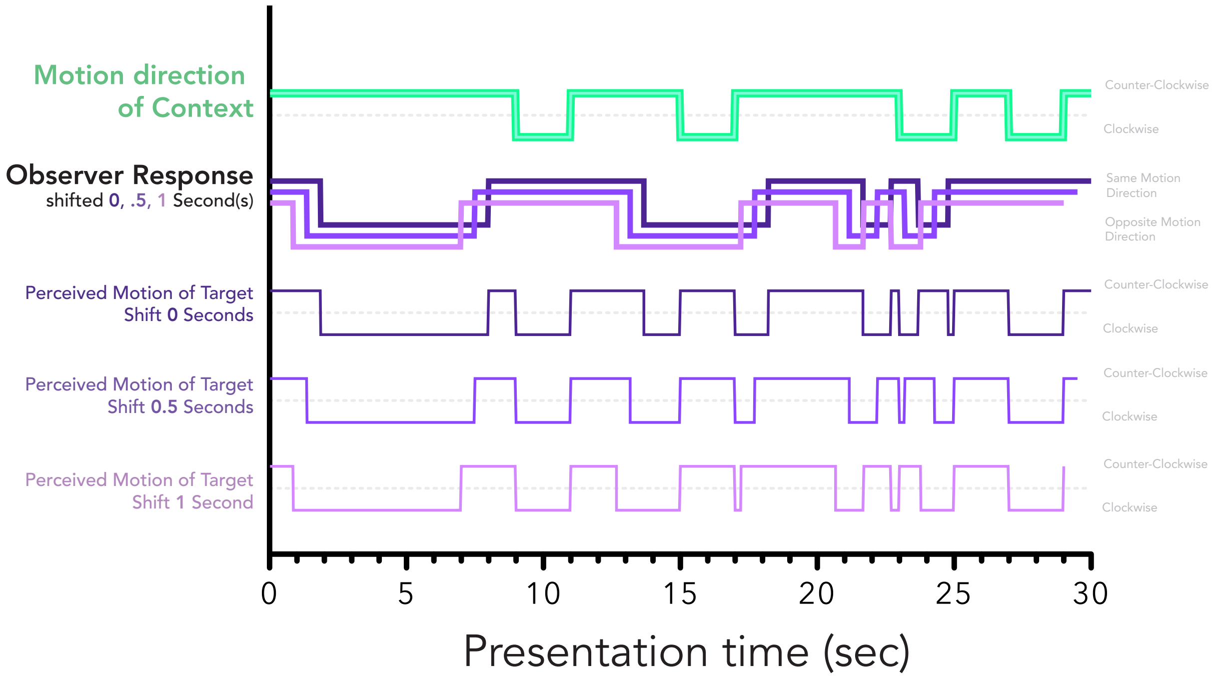


Figure 4.3. Example of context motion sequence (top), observer response (second), and extracted perceived motion of the target sequence (three from the bottom) taken from a sample trial for three example shifts of the timeline.

Prior to the Phi coefficient analysis, the observer response timeline was shifted by various amounts to account for any observer response delays, prior to extracting the perceived motion direction timeline of the target. The shifts of the response timeline were in multiples of approximately 15 milliseconds, or a single frame presentation, up to a maximum of 1 second, or 60 presentation frames. Since Phi coefficient analyses require balanced data sets, we truncated the context timeline by the amount of shift applied to the response timeline. Examples of shifts applied to observer responses and resulting extracted perceived target motion directions are illustrated in Figure 4.3. In the figure, shifts of 0s, 0.5s, and 1s are shown both in terms of the change to the observer response sequence and to the resulting extracted perceived motion direction of the target figure for each of the given shifts.

Following these shifts, the motion direction of the context was compared to the extracted perceived motion direction of the target, and the number of instances of the four possible pairings of responses and context entered into the following 2x2 table. From the possible Phi coefficients, i.e. for each of the possible shifts, the maximum coefficient was taken and retained.

	Target perceived clockwise	Target perceived counter-clockwise	total
Context clockwise	n_{11}	n_{10}	$n_{1\cdot}$
Context counter-clockwise	n_{01}	n_{00}	$n_{0\cdot}$
total	$n_{\cdot 1}$	$n_{\cdot 0}$	n

The Phi coefficient ϕ was calculated as follows:

$$\text{Eqn. 1} \quad \phi = \frac{n_{11}n_{00} - n_{10}n_{01}}{\sqrt{n_{1\cdot}n_{0\cdot}n_{\cdot 0}n_{\cdot 0}}}$$

From the equation one can see that a Phi coefficient of 1 indicates a perfect correspondence between the perceived target and the context motion direction. The Phi coefficient values were then averaged across conditions and then correlated between the ambiguous and rivalrous figures.

Experiment 1: With interocular swaps

When a rivalrous pair is swapped across eyes, the previously suppressed percept escapes inhibition resulting in a single perceptual change (Blake, Westendorf, & Overton, 1980). On the other hand when a monocular ambiguous figure is swapped across eyes there is a significant increase in the subsequent *rate* of perceived reversals in comparison to normal binocular presentation (Spitz & Lipman, 1962). The changes in interpretation observed in the rivalrous figure following an eye swap are thus not equivalent to the increase in reversal rates observed in an ambiguous figure:

the former is a reversal in interpretation that is temporally tied to the swap of the rivalrous pair, while the latter is a general increase in reversal rates throughout inspection of the ambiguous figure. We were therefore curious as to the effect of this manipulation with our stimulus protocol. Following a motion direction change of the context figure, both context and target were swapped between eyes, unaware to our observers.

The mean context-target Phi coefficient for each observer and target condition was calculated to assess the magnitude of contextual influence. The Phi coefficients for the ambiguous Necker and rivalrous skeleton cubes were then correlated using Pearson's correlation coefficient r, and are summarized in Table 4.1 and plotted in Figure 4.4, along with p values that determine whether the values of r are significantly different from zero. Contextual influences were found to be highly and significantly correlated between target types for all conditions.

These significant correlations are found regardless of the temporal shifts introduced to take into account any observer response delays. The mean ratio of shifted to non-shifted Phi coefficients across both the ambiguous and rivalrous conditions was 1±0.02 (mean±SEM); in other words the shifts had a minimal effect on the coefficients. Furthermore, the shifts had even less impact on the correlation of the magnitude of context influence between the rivalrous and ambiguous target: the mean ratio change was 1±0.005 (mean±SEM).

Table 1

Table 1 *Correlations r between the mean Phi coefficients of the ambiguous and rivalrous targets for Experiment 1*

Condition	Target Above		Target Below	
	r	*p*	*r*	*p*
Both eyes	0.86	<.05	0.91	<.05
Same eye	0.84	<.05	0.93	<.05
Opposite Eye	0.98	<.01	0.96	<.01

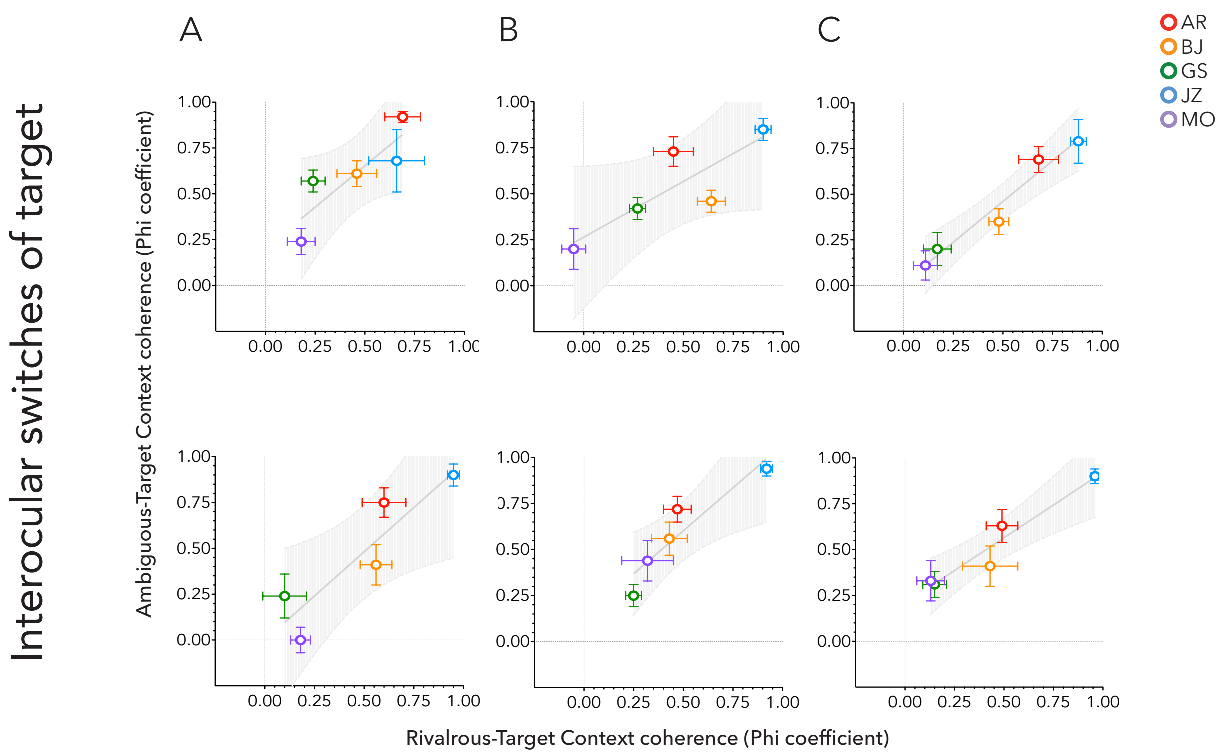


Figure 4.4. Figure 4.4. Mean Phi coefficients for context-target coherence averaged across trials. Each graph plots the bi-stable target coefficients against the rivalrous target coefficients. Different graphs are for different context condition. A, context presented to both eyes; B, context presented to the same eye as the matched target; C, context presented to the opposite eye as the matched target. The upper panel is for when the target figure was presented above the fixation cross, the lower panel when below it.

Experiment 2: effect of eye dominance

With rivalrous figures, the interpretation in the dominant eye tends to dominate (HANDA et al., 2004). Since the first experiment involved swapping the context and target across the eyes following a context motion direction change, it would seem prudent to investigate the effect of eye-dominance on the contextual influences studied here. We employed a similar methodology to that of Experiment 1 with the added differences that: 1. the location of the context and target was *not* swapped across eyes following a change in context motion direction; 2. the ambiguous target or the rivalrous skeleton cube that matched the context's motion direction was presented in separate conditions to the observer's dominant and non-dominant eye.

Eye dominance for each observer was measured using the Miles test (Miles, 1930). For the ambiguous Necker cube condition in which the context figure is presented to both eyes, the manipulation of eye-dominance is not possible. In this case the data is that from Experiment 1.

The correlations between the Phi coefficients for the ambiguous and rivalrous data are summarized in Table 4.2 and plotted in Figure 4.5 and Figure 4.6 for the target to dominant eye and target to non-dominant eye, respectively. Contextual influences were found to be highly correlated between target types for all except for when the target was presented to both eyes above fixation in the dominant eye condition where the correlation was highly correlated and approached significance, $p = 0.07$. As with Experiment 1 the addition of the delay had a negligible effect on the Phi coefficient values (ratio change 1.09 ±0.11 (mean±SEM)) or subsequent correlations between ambiguous and rivalrous figures (mean ratio change 1 ±0.003 (mean±SEM).

Table 2

Table 2 *Correlations r between the mean Phi coefficients of the ambiguous and rivalrous targets for Experiment 2*

	Target Above		Target Below	
Condition	*r*	*p*	*r*	*p*
		Target to dominant eye		
Both eyes	0.76	0.07	0.99	<.01
Same eye	0.88	<.05	0.94	<.01
Opposite Eye	0.92	<.05	0.88	<.05
		Target to non-dominant eye		
Both eyes	0.86	<.05	0.91	<.05
Same eye	0.84	<0.01	0.93	<.05
Opposite Eye	0.98	<0.01	0.96	<.01

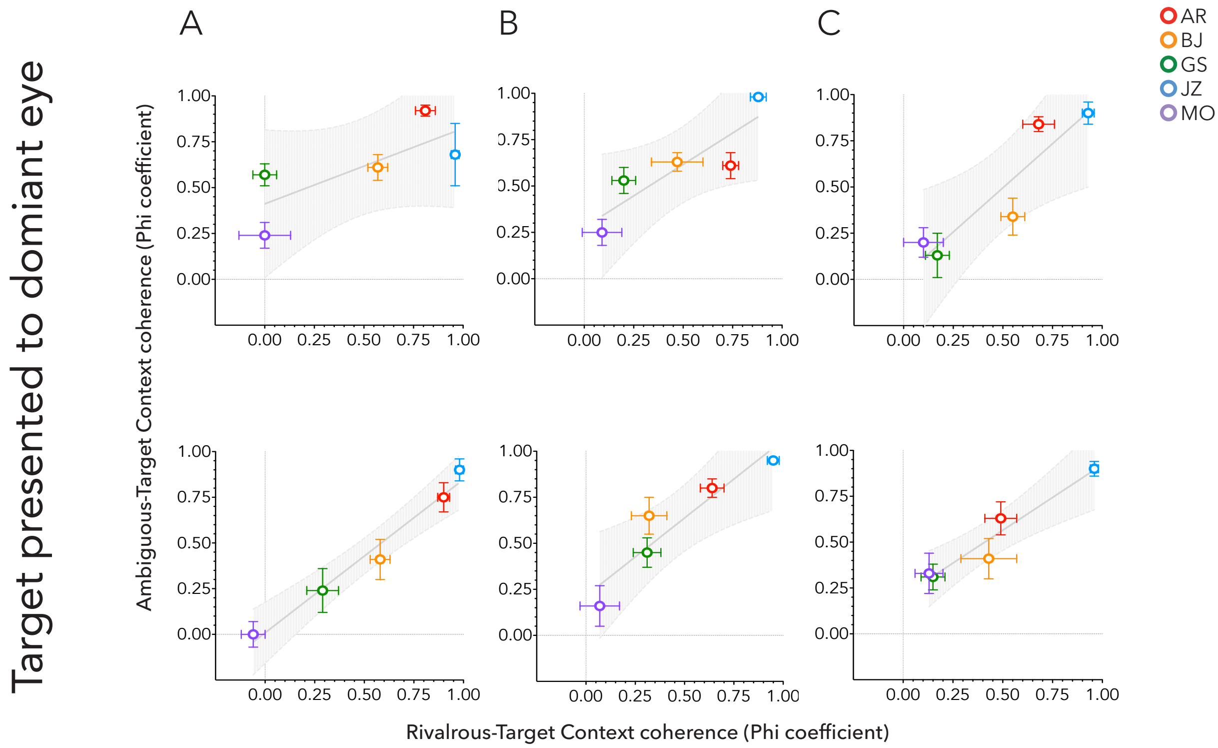


Figure 4.5. Mean Phi coefficients for the bi-stable target condition are plotted against those obtained for the rivalrous target condition by each context condition. The matched target was always presented to the observer's dominant eye. A, context presented to both eyes; B, context presented to the same eye as the matched target; C, context presented to the opposite eye as the matched target. The upper panel is for when the target was presented above the fixation cross, the lower panel when below it.

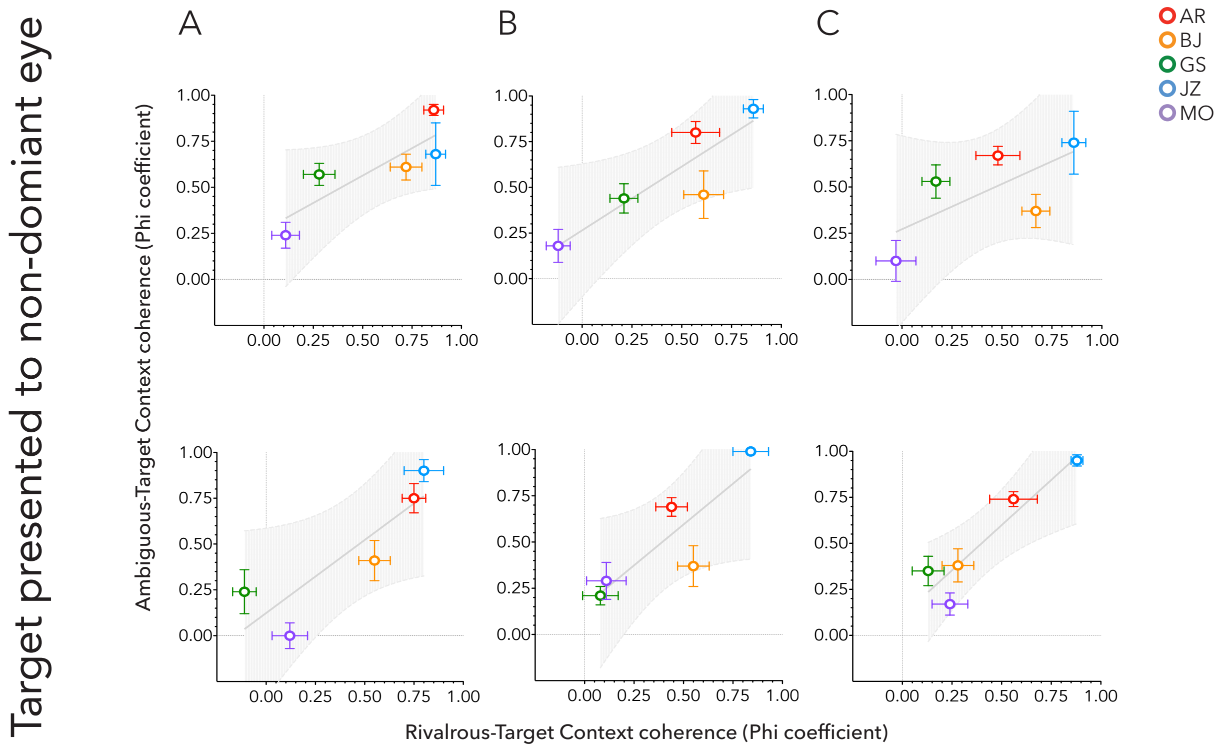


Figure 4.6. Mean Phi coefficients for the ambiguous target are plotted against those for the rivalrous target, for each context condition.

The matched target was always presented to the observer's non-dominant eye. A, context presented to both eyes; B, context presented to the same eye as the matched target; C, context presented to the opposite eye as the matched target. The upper panel is for when the target figure was presented above the fixation cross, the lower panel when below it.

Discussion

The results with the Necker cube confirm our previous report (Ouhnana & Kingdom, 2016) and extend it to a binocularly rivalrous figure. Our goal in this study however was to determine whether spatial contextual information similarly influenced the perceptual judgements of both types of stimuli, when designed to elicit a similar perceptual experience. The data show that the context influenced both classes of stimuli similarly: the magnitude of contextual influence between rivalrous and ambiguous target conditions was positively and significantly correlated across observers for nearly all experimental conditions tested. It has been argued that similarities between different types of perceptually unstable figure, such as binocular and monocular rivalry, reflect similar or even identical disambiguation mechanisms (van Boxtel, Knapen, Erkelens, & van Ee, 2008). The findings of the present study strongly support this idea and extend it to include ambiguous figures such as the Necker cube.

This might at first seem surprising. The perceptual experience of rivalrous and ambiguous figures differs markedly even when subjected to similar manipulations. For example, when a rivalrous figure is swapped across eyes, the suppressed interpretation assumes dominance (Blake et al., 1980). Swapping a Necker cube from one eye to the other on the other hand, which if it produced a switch in interpretation would result in reversal rates being more-or-less constant across time, instead results in an increase over time in reversal rates when compared to binocular viewing (Spitz & Lipman, 1962).

However, these differences between rivalrous and ambiguous figures are arguably bottom-up influences, whereas the present study focuses on what is presumably the top-down influence of context. We argue that the high correlations found in this study are due to the overriding influence of the context figures. There are good examples in the literature where contextual influences overcome competing manipulations such as in the study described by (Intaitė et al., 2013). They demonstrated that the influence of spatial context, which biases the interpretation of an ambiguous figure *towards* that of context, overcame any adaptation effects which on their own bias the interpretation of the target figure *away* from that of the adaptor figure.

Although the competing interpretations in perceptually unstable figures, whether produced by rivalry or ambiguity, are known to be influenced by contextual information, the contexts employed in previous studies for the two types of figure have been very different. In previous rivalry studies, the context perceptually combined with the rivalrous target to form a globally coherent percept, biasing the interpretation towards that of the global percept. A good example of this is the study by Sobel and Blake (2002) described in the Introduction. On the other hand in previous ambiguous figure studies, the spatial context was an independent figure, albeit one that also biased the target's percept towards that of the context (Intaitė et al., 2013); (Ouhnana & Kingdom, 2016)). In the present study, we demonstrated that an independent context, i.e. one that does not form a globally coherent percept with the target, influences the perceptual interpretation of a rivalrous figure similarly to that of its ambiguous counterpart.

Ouhnana & Kingdom (2016) argued that the Gestalt grouping principles of common-fate and, importantly, change synchrony underpinned the perceptual binding they observed between a simultaneously presented rotating context and Necker cube. The principle of common fate is that objects that *move* together are perceptually bound together, while the principle of change synchrony is that objects that *change* together are perceptually bound together (Alais, 2012; Wertheimer, 1923). In the current study, we found that these binding principles also extend to a rivalrous stimulus that perceptually resembled a Necker cube, and furthermore that the magnitude of perceptual binding was significantly correlated between the two types of figure.

One might think that the perceptual instability in the rivalrous figure was not due to competing monocular interpretations but instead due to ambiguity in a single, binocularly-fused cube, and that this was the reason for the similarities of contextual influence. To test this possibility, we created a demo in which each of the dichoptically presented cubes were defined by a different colour. The changes in perceived motion direction were always observed to be accompanied by changes in color. If the alternations were due to ambiguity in a single fused cube, then one would expect that the reversals would not be tied to the underlying cube colors. We therefore conclude that the changes in interpretation experienced in our rivalrous target are due to binocular rivalry between competing monocular percepts - not ambiguity within a single fused percept.

In our first experiment, both context and target figures were swapped between eyes following a motion direction change of the context figure. This manipulation was undertaken to maximize differences between the two types of targets, in order to assess any potential differential influence of the context. However, because we found no difference between the results of this experiment and the subsequent one in which we did not introduce an eye-swap, we conclude that this manipulation had no effect.

Our second experiment included an examination of the effects of eye dominance. In previous studies, when the rivalrous figures were presented to the same locations in both eyes, the perceptual interpretation was biased towards the input in the dominant eye (Handa et al., 2004). In our study we found significant correlations between the different types of figure regardless of the eye-dominance relationships between target and context.

Conclusion

In conclusion, our results argue that both our rivalrous and ambiguous rotating targets share a common contextual-influence mechanism. This finding has a corollary in that it also argues for a common underlying mechanism mediating the perceptual instability of ambiguous and rivalrous figures.

Chapter 5

Summary and conclusion

This thesis has attempted to determine the effects of spatial context on the motion properties of ambiguous, bi-stable and binocularly rivalrous stimuli. We aim to understand the mechanisms that underpin our perception of stimuli that have a range, or a limited number of categorically alternative perceptual interpretations. A number of new findings have emerged from the thesis, and these have led to some new ideas about the mechanisms of perceptual ambiguity. Below is a summary of the main findings and the new ideas:

1st study

1a. We found that the influence of a cast shadow - the 'context' - on the perceived motion trajectory of a casting object – the 'target' - could be partially replicated by a different type of context – a secondary object near-identical to the target.

1b. The above findings led to the idea that one form of the Gestalt grouping principle of common fate is that objects translating at the same speed and overall direction tend to be perceived as lying in a common plane biased towards appearing fronto-parallel.

2nd study

2a. We found that the perceived rotation direction of a Necker cube – the 'target' - was influenced by the rotation direction of a nearby skeleton cube – the 'context'. The magnitude of the influence was unaffected by the relative rotation speeds of the target and context, but was influenced by their relative positions, with the greater influence exerted when the target was above compared to below the context.

2b. The above findings led to the idea that change-synchrony not common fate underpinned the perceptual binding of the target to the context, and that prior knowledge of frictional forces exerted an additional influence.

3rd study

3a. We found a similar within-subject influence of a rotating skeleton cube – the 'context' - on two types of ambiguous figures – the 'targets': a rotating bi-stable Necker cube and a binocularly-rivalrous skeleton cube.
3b. The above results support the idea of a common mechanism mediating the ambiguity in bi-stable and binocularly-rivalrous figures.

In what follows, we discuss two general themes that emerge from the studies reported in this thesis: 1. concerning the assumptions about physics that appear to determine contextual influence on ambiguous figures and 2. the importance of top-down versus bottom-up influences of contextual influence.

Does prior knowledge about physics determine contextual influence?

When the visual system is faced with an isolated stimulus with more than one perceptual interpretation, whether the stimulus is a single figure or a dichoptic pair, the perceptual interpretations effortlessly alternate, with both interpretations experienced more-or-less an equal amount of time. However, this state of affairs can be significantly altered by various types of manipulation, of most interest here the introduction of spatial context. In all the cases studied in this thesis the effect of spatial context is always to draw the perceptual interpretation *towards* that of the unambiguous context figure, albeit to varying degrees. The question arises as to what extent can this influence be explained by prior assumptions about the physics of the external world.

In our first experiment, we first replicated the previous report by Kersten et al. (1996), namely that manipulating an object's cast shadow influences the perceived motion trajectory of a sphere. We then demonstrated that the effects could be partially replicated when the context, instead of being a cast shadow, was instead an identical secondary sphere, though importantly the overall magnitude of influence was less than that of the cast shadow. We argued that the different magnitudes of influence of the two contextual figures reflected differences in the assumptions made by vision concerning the physical properties of object motion in the natural visual world.

In nature cast shadows are physically linked to the objects that cast them. However, because of the physics of illumination and the 3D geometry of natural scenes, a given shadow can arise from more than one possible object, akin to the way that any two-dimensional retinal projection can arise from a number of three-dimensional worlds. Sufficient constraints nevertheless exist to limit the number of possible object-shadow relationships, with the result that shadows provide important information about object properties, for example their motion trajectories. These constraints include the assumptions that the light source is stationary, that illumination is directed, i.e. non-diffuse, and that light comes from above (Morgenstern, Murray, & Harris, 2011; Ramachandran, 1988).

What possible physical constraints underpin the influence of the secondary sphere? We posited a version of the Gestalt grouping principle of common fate, which states that when objects move together, they are perceptually bound together. We suggested that the perceptual binding between the target and secondary sphere resulted from the assumption that close-by objects moving in similar directions tend to move along a common planar surface that tends to be fronto-parallel. If we are right, it must result from the learned observation that many natural objects tend to move along common surfaces, such as the balls on a billiard table or the ripples on a pond. The fronto-parallel bias for multi-object motion is harder to explain, but it is worth noting that it is analogous to a similar surface bias in stereopsis, in which vertical bars with different disparities are perceived to lie more fronto-parallel when connected by horizontal lines, making them form a surface (Deas & Wilcox, 2014; Fahle & Westheimer, 1988; Liu, David, & Ronen, 1999).

Moving to our second study with rotating Necker cubes, prior knowledge of physics had already been shown to influence the perception of rotating figures. Gilroy and Blake (2004), using a pair of rotating dot spheres in which one sphere's direction of motion was ambiguous while the other was unambiguous, found that when the two spheres had parallel axes of rotation and abutted, the dominant perception was of counter rotation. However, when there was only a small gap between the spheres, there was no consistent relationship between their perceived directions of rotation. This, Gilroy and Blake argued, implied the influence of perceived frictional forces. In our second study, we found that an unambiguous rotating skeleton cube also influenced the perceived rotation direction of the ambiguous Necker cube, but the direction of influence was opposite to that found by Gilroy and Blake – in our study it was attractive whereas in Gilroy and Blake's zero gap condition it was repulsive. However, the condition that produced the biggest contextual influence in our study was when the contextual figure lay beneath the Necker cube, which given that the cubes rotated around a vertical axis meant that the two axes of rotation were coaxial, unlike in Gilroy and Blake's study in which they were parallel. With coaxial rotation axes, frictional forces would be expected to produce same rather than opposite directions of motion, as we found. This does not explain however why we observed same rotation direction influences when the context and Necker cube were positioned along the horizontal and hence had parallel axes of rotation. There must be a default mechanism which causes same axis of rotation based on common fate irrespective of the spatial relationship between the context and Necker cube, but one that can be augmented by perceived frictional forces.

Perhaps the most important finding of the Necker cube study however is that the amount of perceptual binding is very similar irrespective of the relative speeds of rotation of the two figures. This implies the operation of an additional grouping operation: change synchrony (Alais, Blake, & Lee, 1998). With change synchrony objects that *change* together along a shared dimension bind together. Motion change requires an external physical force, so change synchrony may be underpinned by the assumption that groups of similar objects are subject to the same forces, such as that driving the coherent changes in direction seen in shoals of fish or flocks of birds.

With regards to our third study that compared contextual influences on ambiguous and binocularly rivalrous figures, we did not manipulate the relative positions and relative speeds of the contexts and figures, so did not test whether perceived friction forces and change synchrony influenced the perception of the rivalrous figures. However, the strong within-subject correlations between the magnitude of contextual influence on ambiguous and rivalrous figures suggests that whatever underlying physical assumptions mediate the influence of context on bi-stable figures, apply to rivalrous figures too.

Top-down versus bottom-up processes

We have argued that prior knowledge about physical forces, in some cases mediated through Gestalt grouping principles, drives the magnitude of influence of spatial context on ambiguous or bi-stable figure perception. Furthermore, we have demonstrated that when bi-stable and binocularly-rivalrous target figures are designed in such a way as to produce a similar perceptual experience, the influence exerted by spatial context on the two types of figure is very similar, providing evidence towards a common underlying mechanism.

Prior knowledge and grouping principles arguably reflect the operation of top-down or feedback processes. As we noted in the Introduction, top-down as well as bottom-up processes mediate ambiguous figure perception (Blake & Logothetis, 2002), with recent hybrid theories stressing the importance of both top-down and bottom-up processes, for both bi-stable (Long & Toppino, 2004) and binocularly-rivalrous (Alais & Blake, 2005) figures. For all the studies reported in this thesis, it is reasonable to conclude that the context effects themselves were top-down and involved feedback from higher to lower visual centres. However, the perceptual reversals studied in the second and third studies were also presumably mediated by relatively low-level visual processes.

Although the bi-stable and binocularly-rivalrous figures in our third study were designed to produce a comparable perceptual experience, previous studies have shown that the two types of figure have different perceptual properties. For instance, swapping a binocularly-rivalrous figure interocularly releases the alternative inhibited interpretation, locking the perceived alternations to the physical swaps (Blake, Westendorf, & Overton, 1980). On the other hand interocular swapping of a bi-stable figure increases its reversal rate beyond that expected from the pattern of physical swaps (Spitz & Lipman, 1962). The fact that we found such strong within-subject correlations between the influence of context on both types of figure is perhaps then testament to the overriding influence of top-down over bottom-up processes in ambiguous and rivalrous figure perception, in line with previous findings by Intaitė, et al. (2013).

Conclusion

We have demonstrated that Gestalt grouping principles coupled with prior knowledge of physical forces can exert a strong contextual influence on the perception of ambiguous, bi-stable and binocularly-rivalrous figures. Furthermore, when a bi-stable and a binocularly-rivalrous figure are such as to elicit a similar perceptual experience, the magnitude of the influence of spatial context on the two types of figure is also similar, supporting the idea of a common underlying mechanism.

Our findings build on hybrid theories that stress the combined influence of bottom-up and top-down processes on ambiguous figure perception. The profound influence of context on ambiguous figure perception revealed in the studies reported in this thesis testify to the importance of top-down processing in ambiguous figure perception.

References

Adams, P. A., & Haire, M. (1958). Structural and conceptual factors in the perception of double-cube figures. The American journal of psychology, 548-556.

Adams, P. A., & Haire, M. (1959). The effect of orientation on the reversal of one cube inscribed in another. The American Journal of Psychology, 72(2), 296-299.

Alais, D. (2012). Binocular rivalry: competition and inhibition in visual perception. Wiley Interdisciplinary Reviews: Cognitive Science, 3(1), 87-103.

Alais, D., & Blake, R. (Eds.). (2005). Binocular rivalry. MIT press.

Alais, D., Blake, R., & Lee, S. H. (1998). Visual features that vary together over time group together over space. Nature neuroscience, 1(2), 160-164.

Blake, R., & Fox, R. (1974). Binocular rivalry suppression: Insensitive to spatial frequency and orientation change. Vision Research, 14(8), 687-692.

Blake, R., & Logothetis, N. K. (2002). Visual competition. Nature Reviews Neuroscience, 3(1), 13-21.

Blake, R., Westendorf, D. H., & Overton, R. (1980). What is suppressed during binocular rivalry?. Perception, 9(2), 223-231.

Brascamp, J. W., Klink, P. C., & Levelt, W. J. (2015). The 'laws' of binocular rivalry: 50 years of Levelt's propositions. Vision research, 109, 20-37.

Bruner, J. S., & Minturn, A. L. (1955). Perceptual identification and perceptual organization. The Journal of General Psychology, 53(1), 21-28.

Casati, R. (2008). The copycat solution to the shadow correspondence problem. Perception, 37(4), 495-503.

Castiello, U. (2001). Implicit processing of shadows. Vision research, 41(18), 2305-2309.

Cavanagh, P. (2005). The artist as neuroscientist. Nature, 434(7031), 301-307.

Deas, L. M., & Wilcox, L. M. (2014). Gestalt grouping via closure degrades suprathreshold depth percepts. Journal of Vision, 14(9), 14-14.

Dobbins, A. C., & Grossmann, J. K. (2010). Asymmetries in perception of 3D orientation. PloS one, 5(3), e9553.

Dutour, E. F. (1760). Discussion d'une question d'optique. De l'Imprimerie Royale.

Everitt, B. S. (1992). The analysis of contingency tables. CRC Press.

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

Gilroy, L. A., & Blake, R. (2004). Physics embedded in visual perception of three-dimensional shape from motion. Nature neuroscience, 7(9), 921-922.

Goolkasian, P. (1987). Ambiguous figures: Role of context and critical features. Journal of General Psychology, 114(3), 217.

Grünau, M. W., Wiggin, S., & Reed, M. (1984). The local character of perspective organization. Attention, Perception, & Psychophysics, 35(4), 319-324.

Haijiang, Q., Saunders, J. A., Stone, R. W., & Backus, B. T. (2006). Demonstration of cue recruitment: Change in visual appearance by means of Pavlovian conditioning. Proceedings of the National Academy of Sciences of the United States of America, 103(2), 483-488.
Handa, T., Mukuno, K., Uozato, H., Niida, T., Shoji, N., & Shimizu, K. (2004). Effects of dominant and nondominant eyes in binocular rivalry. Optometry & Vision Science, 81(5), 377-383.
Intaitė, M., Noreika, V., Šoliūnas, A., & Falter, C. M. (2013). Interaction of bottom-up and top-down processes in the perception of ambiguous figures. Vision research, 89, 24-31.
Intaitė, M., Šoliūnas, A., Gurčinienė, O., & Rukšėnas, O. (2013). Effect of bias on the perception of two simultaneously presented ambiguous figures. Psichologija, 47(47), 91-101.
Kersten, D., Knill, D. C., Mamassian, P., & Bülthoff, I. (1996). Illusory motion from shadows. Nature, 379(6560, 31-31.
Kersten, D., Mamassian, P., & Knill, D. C. (1994). Moving cast shadows and the perception of relative depth. Perception, 26, 171-192.
Kitaoka, A., & Ishihara, M. (2000). Three elemental illusions determine theZöllner illusion. Attention, Perception, & Psychophysics, 62(3), 569-575.
Klink, P. C., van Wezel, R. J. A., & van Ee, R. (2012). United we sense, divided we fail: context-driven perception of ambiguous visual stimuli. Philosophical Transactions of the Royal Society of London B: Biological Sciences, 367(1591), 932-941.
Köhler, W. (1940). Dynamics in psychology: Vital applications of gestalt psychology. Liveright, New York, NY.
Kornmeier, J., Hein, C. M., & Bach, M. (2009). Multistable perception: when bottom-up and top-down coincide. Brain and cognition, 69(1), 138-147.
Kovacs, I. L. O. N. A., Papathomas, T. V., Yang, M., & Fehér, Á. (1996). When the brain changes its mind: interocular grouping during binocular rivalry. Proceedings of the National Academy of Sciences, 93(26), 15508-15511.
Leopold, D. A., Wilke, M., Maier, A., & Logothetis, N. K. (2002). Stable perception of visually ambiguous patterns. Nature neuroscience, 5(6), 605-609.
Levelt, W. J. (1965). On binocular rivalry (Doctoral dissertation, Van Gorcum Assen).
Li, H. C. O., & Kingdom, F. A. (1999). Feature specific segmentation in perceived structure-from-motion. Vision research, 39(5), 881-886.
Liu, Z., Jacobs, D. W., & Basri, R. (1999). The role of convexity in perceptual completion: Beyond good continuation. Vision research, 39(25), 4244-4257.
Logothetis, N. K., Leopold, D. A., & Sheinberg, D. L. (1996). What is rivalling during binocular rivalry?. Nature, 380(6575), 621-624.

Long, G. M., & Batterman, J. M. (2012). Dissecting perceptual processes with a new tri-stable reversible figure. Perception, 41(10), 1163-1185.
Long, G. M., & Moran, C. J. (2007). How to Keep a Reversible Figure from Reversing: Teasing Out Top—Down and Bottom—Up Processes. Perception, 36(3), 431-445.
Long, G. M., & Toppino, T. C. (2004). Enduring interest in perceptual ambiguity: alternating views of reversible figures. Psychological bulletin, 130(5), 748.
Long, G. M., Toppino, T. C., & Kostenbauder, J. F. (1983). As the cube turns: Evidence for two processes in the perception of a dynamic reversible figure. Attention, Perception, & Psychophysics, 34(1), 29-38.
Long, G. M., Toppino, T. C., & Mondin, G. W. (1992). Prime time: Fatigue and set effects in the perception of reversible figures. Attention, Perception, & Psychophysics, 52(6), 609-616.
Mamassian, P. (2008). Ambiguities and conventions in the perception of visual art. Vision Research, 48(20), 2143-2153.
Mamassian, P., Knill, D. C., & Kersten, D. (1998). The perception of cast shadows. Trends in cognitive sciences, 2(8), 288-295.
Miles, W. R. (1930). Ocular dominance in human adults. The journal of general psychology, 3(3), 412-430.
Morgenstern, Y., Murray, R. F., & Harris, L. R. (2011). The human visual system's assumption that light comes from above is weak. Proceedings of the National Academy of Sciences, 108(30), 12551-12553.
Necker, L. A. (1832). LXI. Observations on some remarkable optical phænomena seen in Switzerland; and on an optical phænomenon which occurs on viewing a figure of a crystal or geometrical solid. Philosophical Magazine Series 3, 1(5), 329-337.
Ni, R., Braunstein, M. L., & Andersen, G. J. (2004). Perception of scene layout from optical contact, shadows, and motion. Perception, 33(11), 1305-1318.
Orbach, J., Ehrlich, D., & Heath, H. A. (1963). Reversibility of the Necker cube: I. An examination of the concept of "satiation of orientation". Perceptual and motor skills, 17(2), 439-458.
Ouhnana, M., & Kingdom, F. A. (2016). Perceptual-binding in a rotating Necker cube: The effect of context motion and position. Vision research, 126, 59-68.
Pastukhov, A., Vonau, V., & Braun, J. (2012). Believable change: Bistable reversals are governed by physical plausibility. Journal of vision, 12(1), 17-17.
Ramachandran, V. S. (1988). Perceiving shape from shading. Scientific American, 259(2), 76-83.
Rauber, H. J., & Treue, S. (1999). Revisiting motion repulsion: evidence for a general phenomenon?. Vision research, 39(19), 3187-3196.

Rubin, E. (1915). Synsoplevede figurer.
Schröder, H. (1858). Ueber eine optische Inversion bei Betrachtung verkehrter, durch optische Vorrichtung entworfener, physischer Bilder. Annalen der Physik, 181(10), 298-311.
Sobel, K. V., & Blake, R. (2002). How context influences predominance during binocular rivalry. Perception, 31(7), 813-824.
Spitz, H. H., & Lipman, R. S. (1962). Some factors affecting Necker cube reversal rate. Perceptual and Motor Skills, 15(3), 611-625.
Strüber, D., & Stadler, M. (1999). Differences in Top—Down Influences on the Reversal Rate of Different Categories of Reversible Figures. Perception, 28(10), 1185-1196.
Sundareswara, R., & Schrater, P. R. (2008). Perceptual multistability predicted by search model for Bayesian decisions. Journal of vision, 8(5), 12-12.
Tong, F., Nakayama, K., Vaughan, J. T., & Kanwisher, N. (1998). Binocular rivalry and visual awareness in human extrastriate cortex. Neuron, 21(4), 753-759.
Toppino, T. C., & Long, G. M. (1987). Selective adaptation with reversible figures: Don't change that channel. Attention, Perception, & Psychophysics, 42(1), 37-48.
van Boxtel, J. J., Knapen, T., Erkelens, C. J., & van Ee, R. (2008). Removal of monocular interactions equates rivalry behavior for monocular, binocular, and stimulus rivalries. Journal of Vision, 8(15), 13-13.
Vergeer, M., Moors, P., Wagemans, J., & van Ee, R. (2016). Visible and invisible stimulus parts integrate into global object representations as revealed by combining monocular and binocular rivalry. Journal of vision, 16(11), 14-14.
Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt. II. Psychologische forschung, 4(1), 301-350.
Wilson, H. R. (2007). Minimal physiological conditions for binocular rivalry and rivalry memory. Vision research, 47(21), 2741-2750.