The Resolution of Depth Ambiguity in the Kinetic Depth Effect

SHANE GREGORY BLACKBURN

Thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy.

2006

Department of Psychology,
University of Otago,
Dunedin,
New Zealand.

Supervisor: Dr Robert P. O’Shea
Acknowledgements

I would like to thank three people without whom I never would have completed my thesis. Robert O’Shea, my supervisor, was always patient, supportive, and enthusiastic. Holly Fentiman helped put the finishing touches to figures and text. Carolyn Smale never lost faith that I would complete, even when I had, and produced three beloved children while she waited and waited.
The projected two-dimensional motions of a rotating object are sufficient to enable us to perceive its three-dimensional structure—a phenomenon called the kinetic depth effect (KDE). Although the object’s local, three-dimensional (3D) shape tends to be seen correctly, its global depth organisation, along with its perceived direction of rotation, often spontaneous reverses. My thesis is that these aspects of the KDE can be resolved, disambiguated, by adding temporal or spatial context.

In Experiments 2 to 4, I showed that the KDE could be disambiguated by adding temporal context. When I added the depth cues of contrast and superimposition they cooperated or competed in different ways in different individuals to disambiguate the KDE. I also discovered that the frequency of reversals and their time courses were best predicted from temporal context: the length of observation and an observer’s previous viewing history.

In Experiments 5 to 10, I showed that the KDE could be affected by, and could affect, unambiguous spatial contextual stimuli. When I added unambiguous surround stimuli to an ambiguous KDE object, I observed rotational linkage and rotational contrast. Linkage occurs
when rotating objects share their axis of rotation. Contrast occurs when objects have different axes of rotation. These surround effects vary with speed, shape, rotation speed, and proximity. When I added an unambiguous surround stimulus that overlapped the ambiguous KDE object, I observed a phenomenon I call spreading concavity. In this, an unambiguous surround that by itself would be seen as a convex shape became concave when overlapping an ambiguous KDE cylinder.

The results of my experiments reveal the importance of temporal and spatial context for the KDE. Temporal context includes one’s history of viewing KDE stimuli and one’s history with different depth cues. These temporal contexts alter the frequency of reversals and can disambiguate the KDE. Spatial context has powerful effects on the KDE, disambiguating it, and, in the case of concavity spreading, supplanting alternative perceptual interpretation specified by depth cues present in the stimulus.

I discuss how context effects might be realized in terms of four hypotheses: Supplemented Physiology Hypothesis (SPH), Mechanical Constraints Hypothesis (MCH), Geometric Constraints Hypothesis (GCH), and Surface Reconstruction SFM Scheme (SRS).
### Contents

1. **CHAPTER 1: GENERAL INTRODUCTION** .................................................................1
   1.1. Motion-Based 3D Depth Perception and the Problem of Projective Ambiguity..............1
   1.2. Recovery of 3D Structure From Motion (SFM): The Kinetic Depth Effect (KDE) ........4
   1.3. Phenomenal Correlates of KDE Depth Organisation Ambiguity ............................6
   1.4. Thesis Rational: Resolution of Depth Organisation in the KDE ............................7
   1.5. Geometric Basis of the KDE: Projective Geometry ...........................................16
       1.5.1. Parallel projection ..................................................................................17
       1.5.2. Polar projection ....................................................................................19
   1.6. Neural Basis of the KDE: The Motion Pathway ..................................................21
       1.6.1. Lesion Studies .....................................................................................23
       1.6.2. Single-Cell Electrophysiology ..............................................................26
   1.7. Theoretical Basis of the KDE ...........................................................................31
       1.7.1. Helmholtz: Velocity and Relative Depth (motion parallax) .........................31
       1.7.2. Wallach and O’Connell: Length and Orientation Change .........................33
       1.7.3. Gibson: Perspective Transformation in the Optic Array ..............................33
       1.7.4. Johansson: Perceptual Vector Analysis ..................................................34
       1.7.5. Neurologically-Based Models of SFM ....................................................35
       1.7.6. Computational Approaches to the KDE ..................................................39
       1.7.7. Theoretical Basis of the KDE: Conclusions .............................................47
   1.8. Empirical Aspects of the KDE .........................................................................50
       1.8.1. Is Motion Necessary for the KDE? .........................................................50
       1.8.2. Is Rigid Motion Necessary for the KDE? ...............................................52
       1.8.3. What are the Minimal Conditions for the KDE? .......................................54
       1.8.4. Is Euclidean Structure Recovered in the KDE? ........................................55
   1.9. Depth Reversals in the KDE ...........................................................................55
   1.10. General Theories of Perceptual Reversal .........................................................57
       1.10.1. Satiation Theory ...................................................................................58
       1.10.2. Decision Theory ...................................................................................59
   1.11. Empirical Aspects of Perceptual Reversals .......................................................60
       1.11.1. Temporal Characteristics .......................................................................60
       1.11.2. Complexity ..........................................................................................62
       1.11.3. Rotation speed ......................................................................................63
       1.11.4. Cognitive factors ..................................................................................66
       1.11.5. Individual differences ............................................................................68

2. **CHAPTER 2: THE RELATIONSHIP OF DEPTH CUES TO THE DISAMBIGUATION OF THE KDE** .................................................................71
   2.1. Motion Perspective .........................................................................................71
       2.1.1. Full motion perspective: KDE .................................................................71
       2.1.2. Vertical versus horizontal motion perspective ...........................................73
       2.1.3. Horizontal motion perspective: KDE .......................................................74
       2.1.4. Horizontal motion perspective: Superimposed planar surfaces .................76
       2.1.5. Horizontal motion perspective: Velocity gradients ...................................78
       2.1.6. Horizontal motion perspective: Motion parallax ......................................82
       2.1.7. Why are Velocity Gradients Unambiguous when the Head or Monitor Moves? 84
       2.1.8. Relationship between motion parallax / velocity gradient displays and the KDE 89
   2.2. Dynamic Occlusion .........................................................................................93
   2.3. Binocular Disparity .........................................................................................96
   2.4. Luminance Contrast .......................................................................................99

3. **CHAPTER 3: INDUCTION EFFECTS IN THE KDE: SPATIAL CONTEXT ...101**
   3.1. Spatial Context ..............................................................................................101
   3.2. Perceptual Consequences of Spatial Context: Contrast and Assimilation 103
       3.2.1. 2D Motion Assimilation .........................................................................103
       3.2.2. 2D Motion Contrast ...............................................................................106
   3.3. Relationship Between Assimilation and Contrast .............................................107
   3.4. Spatial Context Effects in the KDE .................................................................109
   3.5. Effect of Spatial Context on Perceived KDE Shape .........................................110
   3.6. Effect of Spatial Context on KDE Reversals ...................................................112
6.7. EXPERIMENT 8: SIMULATED SPEED DIFFERENCE DUE TO SIMULTANEOUS MOTION CONTRAST AND FAILURE TO DISAMBIGUATE KDE ................................................................. 234

6.7.1. Experiment 8: Introduction ................................................................. 234
6.7.2. Experiment 8: Method ..................................................................... 235
6.7.3. Experiment 8: Results and Discussion .......................................... 236

6.8. EXPERIMENT 9: THE INFLUENCE OF SURROUND TYPE, CONFIGURATION, AND TRANSPARENCY ON ROTATIONAL LINKAGE AND ROTATIONAL CONTRAST ........................................ 238

6.8.1. Experiment 9: Introduction ................................................................. 238
6.8.2. Experiment 9: Method ..................................................................... 240
6.8.3. Experiment 9: Results ..................................................................... 243
6.8.4. Experiment 9: Discussion ................................................................ 245

7. CHAPTER 7. SPATIAL INDUCTION AND KDE DISAMBIGUATION 2: SPREADING CONCAVITY ............................................................................................................. 247

7.1. CHAPTER 7: INTRODUCTION .................................................................... 247
7.2. EXPERIMENT 10: SPATIAL DETERMINANTS OF CONCAVITY SPREADING ................................................................. 251

7.2.1. Experiment 10: Introduction .............................................................. 251
7.2.2. Experiment 10: Method ................................................................. 252
7.2.3. Experiment 10: Results ................................................................. 255
7.2.4. Experiment 10: Discussion .............................................................. 261

8. CHAPTER 8. GENERAL DISCUSSION ................................................................................................................................. 267

8.1. SUMMARY OF EXPERIMENTAL RESULTS .......................................................... 267
8.2. IMPLICATIONS FOR THEORY ................................................................. 270
8.2.1. Supplemented Physiology Hypothesis (SPH) ........................................... 271
8.2.2. Mechanical Constraints Hypothesis (MCH) .......................................... 272
8.2.3. Geometric Constraints Hypothesis (GCH) .......................................... 274
8.2.4. Surface Reconstruction SFM Scheme (SRS) ......................................... 274

8.3. SUGGESTIONS FOR FUTURE WORK ........................................................... 275
8.3.1. Future Experiment 1: Object Similarity and Spreading Concavity ............. 275
8.3.2. Future Experiment 2: Interaction of Cue-Specified and Context-Specified Depth Organisation ................................................................. 276
8.3.3. Future Experiment 3. Motion Parallax and KDE .................................... 277

8.4. CONCLUDING REMARKS ........................................................................... 279

9. REFERENCES ........................................................................................................ 280
Chapter 1: General Introduction

1.1 Motion-Based 3D Depth Perception and the problem of Projective Ambiguity

In everyday life we effortlessly perceive a three-dimensional (3D) world of solid objects and surfaces, despite our retinal images being inherently two-dimensional (2D). The projection of the distal world onto a flat surface introduces a theoretical problem, described by Berkeley (1709), for the perceptual reconstruction of the third dimension, namely projective ambiguity. Projective ambiguity refers to there being an infinite variety of 3D configurations projecting an identical 2D shape, and so the outline of an object on the retina does not uniquely specify its solid form. Yet as our experience testifies, the visual system generally does achieve a veridical 3D interpretation from the myriad possibilities.

Achieving a stable and depthful perceptual world would seem to be all the more remarkable considering that not only are retinal images two-dimensional but they also move. Typically the observer and parts of the environment move relative to one another, making the projected images of objects undergo complex, continuous deformations. Rather than posing a problem for vision, however, these movement-produced deformations provide a potent source of information about the
structure and motion of objects in the visual scene, because they represent perspective transformations in the optic array. Gibson (1950, 1966) has argued that these transformations, which he terms motion perspective, in themselves constitute important stimuli for perception and carry the structure of an object through those relationships between features which remain invariant under transformation. Two classes of perspective transformations resulting from observer and object movement are particularly informative about the 3D structure of the distal world: translational and rotational optic flow. As the observer moves through a cluttered environment, the projected images of objects change in ways systematically related to their 3D structure, the spatial layout of the environment, and to the movement of the observer. Since at least the time of Helmholtz (1878/1925) it has been known that the optical flow produced by positional changes provides the basis for motion parallax—the perception of depth from relative translatory retinal motions. The literature concerning motion parallax is reviewed in Section 2.1.6.

Of course, a change in object position is not the only consequence of observer movement. There are also progressive changes in the visual directions from which an object is seen. The equivalent transformation would also be produced by the physical rotation of an object before a stationary observer. Although actual object rotations
are relatively rare in natural settings (examples from nature might include leaves swirling in the wind, a boulder rolling down a hill), projected object rotations (or partial orientation changes) are commonplace as they accompany every movement of an object (or observer) except for those directly along the line of sight.

A rotational flow field potentially yields information about the 3D shape of an object because the distribution of projected relative 2D motions is directly related to the relative distances between feature points on the object. According to Gibson, the visual system “picks up” (Gibson, 1979) the rotational transformation directly from the deformation of the retinal image. This in turn uniquely specifies the 3D shape of the object thereby eliminating the problem of projective ambiguity. Although the introduction of perspective transformation might resolve projective ambiguity in the optic array (Gibson, 1979) Eriksson (1973) has shown analytically that, for the ambulant observer, perspective simply extends the ambiguity over space and time. A given retinal deformation might be produced by an infinite combination of 3D shapes and 3D movements. For example, a rotating cylinder would project the same retinal transformation as an elliptical solid with a greater curvature (i.e., relative depth) but lesser rotation speed. According to Eriksson (1973), the true 3D shape cannot be recovered unless the true rotation speed is also known. This point
illustrates the need for constraints (apriori assumptions about the structure of the world and/or auxiliary information about viewing conditions), such as the rigidity assumption, that allow the unambiguous recovery of 3D structure from motion. Just which constraints the visual system may potentially adopt to achieve this are described below (see Section 1.7.6).

In the face of such theoretical uncertainty, can the visual system actually use the 2D rotational flow field to reconstruct the 3D shape of objects under projection?

1.2 Recovery of 3D Structure From Motion (SFM): The Kinetic Depth Effect (KDE)

In a classic series of experiments, Wallach and O’Connell (1953) demonstrated that certain frontal plane motions (the 2D rotational flow field produced by a rotating object) did indeed give rise to the perception of 3D structure from motion—a phenomenon they referred to as the kinetic depth effect (KDE).

They adopted a shadow casting technique similar to that used by Miles (1931) in which various wire-frame and solid objects were placed on a turntable behind a translucent screen and backlit by a distant point light source. Observers sat on the other side of the screen
upon which they viewed the projected silhouette of the object. When
the object was stationary its silhouetted form appeared no more than a
flat, jumbled collection of lines. However once the object was rotated
on the turntable, the moving shadows on the projection screen were
reported spontaneously to organise producing a vivid impression of
the object’s 3D structure and motion in depth. The KDE clearly
illustrates that motion, in the absence of any other source of depth
information, is a powerful and sufficient indicator of depth.

The KDE shows that the visual system is able to use relative motion
to solve the problem of projective ambiguity and recover a stable 3D
shape, despite the theoretically ambiguous rotational flow field.
Observers did not report changes in either the perceived shape nor in
the perceived motion of the rotating objects. However, Wallach and
O’Connell (1953) noted a striking ambiguity that remained in the
KDE. Observers were unable to determine the object’s objective
rotation direction. Moreover, after prolonged observation, the object
frequently appeared to reverse in its direction of rotation. This
apparent reversal of rotation is the natural consequence of a reversal
of the apparent depth ordering of parts of the object (i.e., a change in
the perceived disposition of the object relative to the observer) such
that the most distant parts perceptually become the closest and vice-
versa.
1.3 Phenomenal Correlates of KDE Depth Organisation

Ambiguity

Illusory reversals of rotary motion in depth have been an object of scientific interest for more than a hundred years. Sinsteden (1860, cited in Miles, 1931) first observed that the rotating arms of a distant, silhouetted windmill (viewed at a slight angle) appeared to change direction repeatedly, with corresponding changes in the perceived orientation of the mill. Such reversals of rotary motion illustrate the perceptual consequence of depth-order ambiguity in the KDE.

The windmill illusion demonstrates the direct relationship between where an object’s parts are perceived to lie in depth relative to each other and how those parts appear to move through space.

The reason the windmill’s arms appear to rotate in one direction whenever the silhouetted mill apparently faces towards the observer, and in the opposite direction whenever the mill apparently faces away from the observer is from geometry. The patterns of retinal motion produced by rotating vanes are identical for both depth organizations, but are consistent with different directions of motion in the environment depending on the directions in which the mill apparently faces. Hochberg (1974) refers to the co-dependency between two
perceptual attributes of an object, such as its depth organisation and apparent motion in space, as perceptually coupled responses. Tracking the perceived rotation direction of such a stimulus therefore provides a means of indirectly determining its apparent perceptual organisation at any given moment.

1.4 Thesis Rationale: Resolution of Depth Organisation in the KDE

The central theme of this thesis concerns the determinants of perceptual organisation in the KDE. Why do depth reversals occur in the KDE? Which factors disambiguate depth-order ambiguity in the KDE and bias the perception of one particular 3D organisation over another? Which aspects of the stimulus contribute to the perceptual stability of the organisation?

The KDE stimulus represents an abstraction of a real-world situation (the change in projected object orientation due to observer or object movement)—a reduction to the minimum motion-based information possible to determine shape but not depth organisation. Yet, perceived changes in 3D shape organisation are rare under normal viewing conditions in the natural environment. Why then, should
depth reversals occur frequently with a KDE stimulus but not in the world?

KDE depth reversals are conventionally attributed to parallel projection (see Section 1.5.1), a technique usually applied to the KDE stimulus as a means of removing perspective-based information and thereby isolating relative motion as the sole source of 3D shape. The geometric consequence of parallel projection is that the veridical 3D organisation and its depth-inverted mirror organisation produce identical patterns of relative motions and are therefore theoretically indistinguishable. There exists no differential motion perspective in the stimulus to distinguish the near and far surfaces of a rotating object. In this widely accepted view, the absence of perspective results in equivocal stimulus information (depth-order ambiguity) to which the visual system produces an equivocal response (depth reversal).

Parallel projection cannot, however, provide a full account of depth reversals in the KDE. If this were the case then it would be expected that the addition of perspective to the KDE stimulus (through polar projection) would be a salient depth cue and effectively disambiguate depth organisation. Empirical investigations (reviewed in Section 2.1) have shown that perspective is a surprisingly weak kinetic depth-order cue. Unrealistically extreme perspective ratios (the equivalent of viewing a 5cm cube from 10cm) are required for depth
disambiguation. Shape recovery becomes unstable at such high levels of perspective; highly non-rigid 3D shape deformations are perceived as the object rotates (Braunstein, 1962). Moreover, the degree of perspective encountered in objects in most natural scenes is small enough to approximate parallel projection. Graham and Gillam (1970) have shown that objects that subtend a visual angle of less than 9 degrees produce negligible differential motion perspective. As perspective alone does not disambiguate the KDE, it follows that parallel projection alone does not explain KDE ambiguity.

Other depth-cues that are present under natural observation are removed from the KDE stimulus. Object features are deliberately made homogenous in appearance and so do not occlude one another when they overlap, and do not change in projected size, luminance contrast, or binocular disparity as they approach or recede from the observer.

Could other depth- cues (such as occlusion, luminance contrast, and size change) disambiguate the KDE? The addition of depth cues to the ambiguous KDE stimulus can determine which of the sources of depth information potentially available in the optic array are actually used by the visual system. If a cue-augmented KDE stimulus is perceived to rotate predominantly in the direction consistent with the cue-specified 3D organisation, then it may be assumed that that particular
cue is an effective source of depth information. The empirical literature examining the relationship between dynamic depth cues and KDE depth organisation is reviewed in Chapter 2.

The homogeneous appearance of object features, with the corresponding loss of occlusion information, is often unrecognised as a cause of KDE ambiguity. Element occlusion has been found to be a very effective means of KDE disambiguation (see Section 2.2). As occlusion is considered to provide only ordinal depth, it has proven difficult to vary systematically its strength in order to investigate parametrically its effect on KDE organisation. In Experiments 2, 3, and 4, a technique is introduced that produces a graded rather than ordinal element occlusion manipulation allowing a fuller characterisation of the occlusion cue. Previous investigations of luminance contrast have been confounded by the introduction of correlated element occlusion (due to the non-homogeneity of element luminance) that has made it difficult to determine the relative contributions of the two cues. Consequently the interaction between luminance contrast (the change in element contrast with observer-relative distance) and occlusion is also investigated in these experiments under conditions of cue-conflict and agreement. This provides an opportunity to examine how different (and often conflicting) cues combine to contribute to a unitary 3D organisation.
The paucity of differential depth information in the KDE due to the elimination of depth cues does not, however, fully explain the phenomenon of depth reversal. While one observer might perceive a clockwise rotation of a KDE stimulus and another observer a counterclockwise rotation, there is no a priori reason why either observer should experience a change in the appearance of an unchanged stimulus. That observers do not persist with the same arbitrary depth organisation suggests that depth reversals are mediated by observer-specific as well as stimulus-specific factors (although the absence of depth-order information in the KDE stimulus obviously facilitates the conditions under which reversals occur). The theoretical and empirical literature relating to the depth reversal phenomenon in general, and to reversals of rotary motion in particular, is reviewed later in this chapter. The conditions that promote KDE depth reversal are the subject of further investigation in Experiment 1. In this experiment I explore attributes of a depth-ambiguous KDE stimulus that are expected to influence its perceptual stability, based on the broader reversals literature. Characterising KDE stability with a representative baseline stimulus establishes the optimal conditions for later experiments.

Explicit depth cues are not the only potential source of 3D-organisation information missing in the ambiguous KDE stimulus.
There is also the elimination of an object’s surroundings, a factor that is rarely considered as an explanation for depth-order ambiguity. In a typical KDE experiment a single rotating object is presented against a featureless black background, as if suspended in a void. Objects in the natural environment are rarely observed in isolation, but rather within a rich spatial context provided by other objects and surfaces in a normally cluttered environment. When an observer moves (thereby producing the basis for the projective transformation underlying the KDE) not only is the object transformed but also its surroundings and its spatial relationship to those surroundings are transformed. I contend that the visual system makes use of the spatial context and the dynamic relationships between different objects in the scene, whenever this information is available, to determine the depth organisation of otherwise ambiguous KDE objects.

As in the natural environment, the projected motions of objects are directly determined by their 3D structures and the movement of the observer, it is theoretically possible to recover unambiguously the true depth organisation (and 3D structure and motion) of a given object if the parameters (speed, direction, and viewing distance) of the moving observer are known. Rogers and Rogers (1992) have shown, however, that these parameters are not recovered with sufficient accuracy by the proprioceptive senses in the absence of visual information. Instead this
information must be extracted from the projected transformations of objects (and the relationships between them) in the optic array (i.e., their spatial context).

Even if these parameters cannot accurately be determined from the optic array to allow precise quantitative recovery of 3D shape, there exist relationships between an object and its surroundings that provide potentially useful heuristics for determining its depth organisation. One example is the case of a laterally translating observer. In this instance the direction and speed of projected object translatory motion (i.e., motion parallax) varies as a combination of object distance, fixation distance, and observer movement. Consequently it provides ambiguous information regarding the true disposition of a single object (unless two of these parameters are known). However under these circumstances the direction of object rotation is always opposite to the direction of observer translation (and all objects rotate in the same direction). Therefore the depth organisation of an ambiguous object can be determined if the direction of rotation is known for surrounding objects (due perhaps to the existence of depth cues). Similarly, information regarding the direction of observer movement (perhaps through motion parallax of surrounding objects given knowledge of the fixation distance) could potentially establish the veridical depth organisation. If the observer made use of auxillary
visual information, theoretically correlated to observer movement, then spatial context should prove a potent means of disambiguating the KDE. Whether the depth organisation of an ambiguous object can be influenced by its surroundings is the focus of the remainder of the thesis. The literature concerning spatial context and KDE organisation are considered in detail in Chapter 3. Two new phenomena that demonstrate the role of spatial context in KDE disambiguation are investigated in Chapters 5 and 6.

The first phenomenon, *kinetic depth contrast*, occurs when an ambiguous transparent KDE cylinder is strongly biased to rotate in the opposite direction to the unambiguous motion of other objects (rotating opaque cylinders and translating textured sheets) in its immediate surroundings. This is a somewhat surprising finding because previous research, that had demonstrated KDE induction effects, had predominantly found rotational linkage rather than rotational contrast (reviewed in Section 3.7). In Experiments 5 and 6 the spatial and temporal parameters of kinetic depth contrast are examined in order to characterise the nature of this effect and test possible theoretical explanations. The possibility that rotational contrast is the result of 2D motion contrast induced by the object’s surrounding motion is examined and rejected in Experiments 7 and 8. Finally, in Experiment 9 I investigate the stimulus characteristics that
determine the nature of the induction effects produced by spatial context. In this experiment opposing effects of context – rotational contrast and rotational linkage are produced with the same stimuli by changing the configuration (and spatial relationships) between objects. Kinetic depth contrast clearly demonstrates that the presence of nearby unambiguous KDE objects can bias the perceived organisation of an otherwise ambiguous KDE object.

The second phenomenon, spreading concavity, leads to a quite different consequence of spatial context – here the combination, through spatial overlap, of two otherwise unambiguous objects (counter-rotating opaque cylinders) produces a new composite ambiguous object (a transparent cylinder). The perceived depth organisation of the composite object exerts a strong bias over the inducing objects rendering them ambiguous also. To do so requires that a very strong depth-order cue, namely self-occlusion (see Section 2.2) present in the inducers, be overridden through spatial context. Spreading concavity is a powerful demonstration that the spatial interactions between an object and its surroundings can provide a much more important organising principle than depth cues (in this case self-occlusion) localised in the KDE stimulus. In Experiment 10 the spatial parameters that determine the strength of concavity spreading are investigated.
1.5 **Geometric basis of the KDE: Projective geometry**

This section presents aspects of the projective geometry of object rotation that are relevant to this thesis (see Braunstein, 1976, for a thorough analysis of the projective geometry of 3D motion). Point coordinates in 3D space are defined with respect to three orthogonal axes – the Z (depth) axis represents the observer’s line of sight and the X and Y axes correspond to the horizontal and vertical dimensions, respectively, of the projection plane which is frontoparallel to the observer. The following discussion of parallel and polar projections considers the case of a random-dot cylinder rotating at a fixed angular velocity around its Y axis.

The “random-dot-cylinder”, or more properly, the dot motions generated by viewing the projection of a rotating cylinder covered in random dot texture, is typical of the type of stimuli used in the study of perceived 3D structure from motion.
1.5.1 Parallel projection

The vast majority of KDE studies have adopted parallel projection (sometimes referred to as orthographic projection) to remove all depth cues other than rotation from the stimulus display. Under parallel projection (depicted in Figure 1.1) rays projecting different visible features of the cylinder onto the projection plane run parallel to the line of sight.

With parallel projection the direction of rotation in depth is geometrically ambiguous. No information exists, in such a projection,
to distinguish whether a particular dot on the cylinder is rotating in front of or behind the axis of rotation. Dots located on the near and far faces of the cylinder move along horizontal paths in opposite directions, but in all other respects their projected trajectories are identical (projected dot velocities change as a sinusoidal function of their angular position, see Figure 1.1c). Thus a rightward moving point, for example, could equally correspond to a dot on the near face of a counter-clockwise-rotating cylinder (Figure 1.1a) or to a dot on the far face of a clockwise rotating cylinder (Figure 1.1b). This geometric ambiguity arises because under parallel projection, depth differences between features on the near and far faces of an object are vanishingly small relative to the distance between the object and the projection point.
1.5.2 Polar projection

Polar projection (depicted in Figure 1.2) represents the situation where object features are projected along rays that converge to a projection point located a finite distance from the object. Consequently the projection of a rotating cylinder is subject to perspective transformations, which produce asymmetrical trajectories for dots on the cylinder’s near and far faces. As discussed below, these asymmetries in projected motion paths provide information that
potentially disambiguates the direction of rotation in depth in the KDE.

Under polar projection dots follow elliptical motion paths (see Figure 1.2e) composed of both vertical and horizontal position changes. Braunstein (1976) has referred to these components as vertical perspective and horizontal perspective respectively.

The direction of vertical component motion depends on whether a dot approaches (in depth) or recedes from the observer (see Figure 1.2d). As dots rotate away from the observer their projected positions converge towards the centre of projection, such that a dot on the top rim of the cylinder will move downwards and a dot on the bottom rim will move upwards. During rotation away from the observer, dot projections diverge away from the projection centre. The curvature of the elliptical path is determined by (i) a dot’s distance from the observer along the Z axis (dots on the near face project a more curved path than those on the far face), and by (ii) a dot’s distance from the projection centre along the Y axis (larger distances produce greater curvatures).

Horizontal perspective refers to the asymmetry of dot velocities that occurs during different parts of the rotation cycle. Dots rotate more quickly around the near face of the cylinder than around the far face (see Figure 1.2c). Thus, under polar projection, the horizontal
trajectory of a dot is a combination of sinusoidal velocity changes associated with rotation in depth, and velocity changes inversely related to the dot’s distance from the observer. The latter component is essentially motion parallax. Having defined these key terms, we may now consider the neural basis of the KDE.

1.6 Neural Basis of the KDE: The Motion Pathway

The known physiological and anatomical structure of the visual system is generally believed to reflect a functional division of labour between two cortical pathways, extending from the primary visual cortex (area V1), which analyse different characteristics of visual input. The ventral pathway that projects to the inferior parietal lobe, involves visual areas such as V4 that contain predominantly orientation- and wavelength-selective neurons (Zeki, 1974). In distinction, the dorsal pathway that projects to the inferior temporal lobe, involves areas such as MT that exhibit directional- and disparity-selectivity (Zeki, 1974). Ungerleider and Mishkin (1982) have argued that the ventral, “what”, pathway analyses object properties like colour, shape, and size leading to object identification, while the dorsal, “where”, pathway analyses spatial relationships leading to object localisation.
Analysis of visual motion is thought to occur in specialised motion areas (MT, MSTd, MSTl) organised in a hierarchical fashion along the dorsal pathway (e.g., Newsome, Mikami, Wurtz, Wurtz, Dursteler, & Mikami, 1986). There is evidence to suggest that different aspects of motion information are extracted within each area of the motion pathway. For example, many MT neurons respond to the unitary pattern motion of a complex plaid whereas V1 neurons respond only to the component motions (Movshon, Adelson, Gizzi & Newsome, 1985). Also, MSTd neurons are sensitive to large-field optic flow components and may mediate visually-guided locomotion (e.g., Regan, 1991) while MSTl neurons respond preferentially to small moving stimuli and are involved in the control of pursuit eye-movements (Dürsteler & Wurtz, 1988).

Relatively little is known about the specific neural underpinnings of the KDE. Relevant data come from lesion studies (to be reviewed next), which show that damage to area MT selectively disrupts the perception of 3D structure from motion (SFM), and from single-cell studies (to be reviewed in Section 1.6.2), which reveal neural response properties that could mediate the analysis of depth from motion (and/or motion in depth).
1.6.1 Lesion Studies

1.6.1.1 Monkey Data.

Siegel and Andersen (1986) found that chemical lesions in area MT had a direct effect on the recovery of 3D SFM. In their experiment, monkeys were trained to detect the transition from an unstructured flow field (completely random dot motions) to a more structured flow field depicting a rotating random-dot cylinder among added dynamic noise. Their monkeys were able to detect 3D structure in the velocity field with performance equivalent to that of human observers (Siegel & Andersen, 1988) but were unable to perform the SFM task following lesions to area MT, even if the structured flow field contained no additional dynamic noise (a condition in which they had previously achieved more than 90% accuracy). Although the SFM deficit did not recover during the course of the experiment (which lasted 23 days), detection thresholds for shearing motion returned to preoperative levels after 4-5 days. This finding demonstrates that the SFM deficit was not due to a general deficit in motion processing.

1.6.1.2 Human data.

Focal lesions in human patients, resulting from strokes, have also been shown to be associated with specific deficits in SFM processing. Vaina (1989) found a double dissociation for perceptual tasks between
patients with unilateral lesions confined to either the right occipital-parietal (the ROP group) or right occipital-temporal (ROT group) areas.

ROP patients could readily identify 2D shapes defined by dots moving against a static textured background (2D form from motion). However, they could not discriminate which of two fields of moving dots had the faster velocity (2D speed discrimination), nor could they resolve random-dot stereograms. Of particular interest is the finding that these patients were completely unable to recover the 3D structure of a rotating random-dot cylinder. Instead the display was perceived as an incoherent collection of moving dots, and was often described as looking like ants crawling on the ground, or snow blown by the wind. ROT patients, on the other hand, showed no such deficit in the SFM task and spontaneously identified the 3D structure of the rotating cylinder. They also performed the 2D speed discrimination task with normal levels of accuracy. However, ROT patients were unable to identify 2D shape from motion or from random-dot stereograms, although they reported “seeing something” in the stereograms suggesting some degree of correspondence was achieved.

In line with Siegel and Andersen’s findings, this study also implicates area MT in the recovery of SFM because the human homologue of this area (see Zeki, 1990) was lesioned in the ROP
patients. From these data, however, it cannot be ascertained whether MT is the neural site for SFM recovery, or whether it provides the essential pre-processing of motion information for other visual areas (Andersen, 1990). Vaina (1989) suggests that the motion correspondence process and the extraction of velocity gradients through the integration of local motion measurements are likely to be achieved in MT. She ascribes the perceptual breakdown of the KDE to a loss of these functions in ROP patients. Their phenomenal reports of dot motions in random directions from KDE displays imply a failure of the correspondence process, which may be responsible for the loss of SFM. Livingstone and Hubel (1988) also report that incoherent random dot motion, rather than coherent 3D structure is perceived when displays of rotating RD objects are presented at equiluminance. The use of rotating wire-frame figures, where the motion correspondence problem is simplified, as KDE stimuli may help resolve this issue.

---

1 A technique which they consider functionally to lesion the “colour-blind” motion pathway (but see Michaels, 1992)
1.6.2 Single-Cell Electrophysiology

1.6.2.1 Rotation-in-Depth Neurons.

Sakata and colleagues (1985, 1986, 1994) investigated rotation-sensitive neurons in the parietal association cortex (area 7a) of the monkey. These neurons were found to respond exclusively to clockwise or counter-clockwise rotation (depending on their direction preference) of a stimulus, but not to its linear motion in any direction through their receptive field. Functionally, rotation-sensitive neurons appear to provide a plausible neural substrate for the perception of object rotation in space. Their response is largely independent of the shape of the stimulus and of its position within the receptive field. Sub-populations of rotation-sensitive neurons are selective for different rotation axis orientations (a preference for rotation-in-depth is more frequently represented than frontoparallel rotation), and their activity increases monotonically with angular velocity. These characteristics suggest that rotation-sensitive neurons may encode the speed, direction, and orientation of rotary motion in a KDE stimulus.

Further support for a link between perceptual experience and the activity of rotation-sensitive neurons is evident from their responses to a rotating Ames window (Sakata, Shibutani, Ito, Tsurugai, & Kusunoki, 1994). The preferred direction of rotation for a typical rotation-sensitive neuron was found to reverse during the half-cycle of...
stimulus rotation where the tall edge of the window moved behind the fixation point. This systematic reversal of direction selectivity is consistent with the oscillatory motion that is usually perceived by human observers of the Ames window. If the window was presented either binocularly, or monocularly at a near viewing distance, then the rotation-sensitive neuron maintained its usual directional preference (i.e., signalled rotation rather than oscillation)—a finding that is also consistent with human data.

Although rotation-sensitive neurons respond to the KDE stimuli, it is unclear exactly how they determine the direction of stimulus rotation, or whether they are involved in the recovery of 3D structure. Indirect evidence suggests that rotation-sensitive neurons may integrate different sources of depth information, such as binocular disparity, changing size, and perspective, directly to establish rotation direction. The activity of rotation-sensitive neurons is greater for binocular than monocular presentation, is inversely related to viewing distance, and is reliable (albeit weakly) for looming motion (Sakata, Shibutani, Ito, Tsurugai, & Kusunoki, 1994). Another possibility is that the depth organisation of a rotating object is computed (in parallel) elsewhere in the visual system, and the results of these computations modulate the directional preferences of rotation-sensitive neurons.
1.6.2.2 Centre-Surround Organisation of Direction-Selective Neurons.

Responses of many direction-selective (DS) neurons to moving stimuli are strongly influenced by simultaneous background motion in a larger region surrounding the classical receptive field (Allman, Miezin, & McGuinness, 1985; Frost, 1978; Gulyas, Orban, & Spileers, 1987; Saito et al., 1986; Tanaka et al., 1986; Tsai, 1990).

Such centre-surround interactions in direction-selective neurons may be important for the recovery of depth from motion. Gulyas, Spileers, and Orban (1990) have argued that centre-surround motion cells could function in a manner akin to disparity-selective cells to encode depth information. Buracas and Albright (1996), on the other hand, show that neurons with different centre-surround organisations could operate as filters that directly extract 3D shape descriptors. Both accounts will be elaborated below in Section 1.7.5. The nature of centre-surround interactions will now be described (see also Figure 1.3).
The modulation of direction-selective neurons by their silent surrounds can be either antagonistic or synergistic in nature (Born & Tootell, 1992). DS neurons with antagonistic surrounds typically show pronounced suppression of their normal classical receptive field response when centre and surround motions are in the same direction, and facilitation when the centre and surround motions are in opposite

Figure 1.3. Typical patterns of centre-surround modulation in the activity of direction-selective motion neurons (adapted from Tanaka, et al, 1986). (a) Responses to preferred stimulus (direction and speed) of the classical RF are measured when surrounded by a random-dot field moving in the opposite direction over a range of different speeds, generally form one of three velocity tuning curves: (b) V-shaped tuning, (c) Positive-slope inhibition, or (d) Negative slope inhibition.
directions. Synergistic-surrounds, which are less frequently reported, have the converse effect—classical receptive field responses are enhanced when centre and surround stimuli move in the same direction. Surrounds usually exhibit smooth directional-tuning, indicated, in antagonistic direction-selective neurons, by increasing release from suppression as the directions of centre and surround motions deviate (Allman et al., 1985). Furthermore, antagonistic surrounds often demonstrate one of three velocity-tuning functions: suppression either increases, or decreases with surround speed, or describes a V-shaped function where suppression is maximal when centre and surround motions have the same speed (Allman et al., 1985; Tanaka et al., 1986).

Several authors (e.g., Allman et al., 1985; Frost & Nakayama, 1983; Gulyas, Spileers, & Orban, 1990; Tanaka et al., 1986) have speculated on the functional significance of antagonistic-surround direction-selective cells for perceptual processes. Because these cells respond to differential motion between the classical receptive field and the surround, they are considered a suitable neural mechanism for extracting motion discontinuities and performing motion-based scene segmentation, for discriminating object motion from observer motion, for discounting uniform image flow arising from eye-movements, and in particular, for processing depth from motion. Synergistic-surround
cells have also been attributed a role in depth from motion processing (Buracus & Albright, 1996; Gulyas et al., 1990). In addition, neurons with this organisation may underlie grouping processes based on common fate, and function to reduce noise in the velocity field by integrating early motion measurements.

1.7 Theoretical Basis of the KDE

1.7.1 Helmholtz: Velocity and Relative Depth (motion parallax)

Helmholtz (1878/1925) gave an early account of how retinal motion could give rise to perceived depth. In an often cited statement Helmholtz observed that it is “as if seen through a good stereoscopic view” (p. 231). The relationship between velocity and distance he describes forms the basis for the well-known motion-parallax depth cue. Can the KDE also be explained by the assignment of individual element depths (relative to the fixation plane) based on their retinal velocities?

The experimental evidence (which is reviewed more extensively below in Section 2.1) regarding the role of 2D velocity in determining perceived depth is equivocal. In motion-parallax displays, reliable depth relationships are not perceived from the differential velocities of separate objects (or frontal surfaces) unless their projected motions are
yoked to head-movement (or some other indicator of the nature of observer movement is present). Although a depth separation may still be perceived, the apparent direction and magnitude of that separation is usually highly variable and unrelated to velocity. If, however, textured surfaces are oriented in such a way that their constituent elements project a gradient of velocities then reliable estimates of the direction and magnitude of surface slant are reported (Braunstein & Andersen, 1981; Gibson, Gibson, Smith, & Flock, 1959). This finding implies that the relationship between velocities is more important for depth perception than are individual velocities.

Another problem for a point velocity account is that both object motion and object depth contribute to the velocity of individual features. Consequently, veridical depths cannot be obtained from element velocities unless the rotation speed of the object is known (Ullman, 1979). If the KDE operated strictly on the basis of velocity, then two identical objects that rotated at different speeds would be perceived as having different depth extents (the faster rotating object would appear deeper). Observers do, in fact, seem to experience this phenomenon when two cylinders are superimposed (Ramachandran, Cobb, & Rogers-Ramachandran, 1988), and also tend to judge the same object to be more or less curved depending on its rotation speed (Todd & Norman, 1991). However Loomis and Eby (1988, 1989)
found no increase in reported depth with rotation speed, and also that the perceived depth of rotating cones was related to shear scaled by velocity but not directly to velocity. The problems in explaining the KDE solely from point velocities command other approaches discussed next.

1.7.2 Wallach and O’Connell: Length and Orientation Change

According to Wallach and O’Connell (1953), simultaneous changes in the projected lengths and orientations of contours undergoing rotary motion constitute the essential stimulus conditions for the KDE. Their hypothesis was extended to include random-element objects by positing imaginary lines that connected individual elements. Subsequent research revealed, however, that changes in both length and orientation represent neither necessary nor sufficient conditions for the KDE. Lines that change only in length may be reliably seen to rotate in depth (Braunstein, 1977; Johansson & Jansson, 1968; White & Mueser, 1960), while random changes in both the length and orientation of different line segments produce no perceived motion-in-depth.

1.7.3 Gibson: Perspective Transformation in the Optic Array

Gibson (e.g., 1950, 1957, 1966, 1979) emphasised the role of the mobile, active observer in the course of normal perception. As the
observer moves through the environment (and/or as parts of the environment move), the optic array undergoes continuous change called optical flow. It contains potentially useful information about the structure and layout of the environment, and also about the observer’s movement through it. Changes of the whole array indicate that the observer (or world) is moving, while partial changes in the optic array indicate the motion of detached objects. Those relationships within the optic array that do not change under transformation, known as invariants, reveal the persisting structure of the environment. Gibson argued that the visual system is particularly sensitive to perspective transformations (rotations and translations in space) and that these transformations are extracted directly from optic flow. Gibson considered perspective transformations, rather than individual spot motions, to be an important higher-order stimulus for perceived rigid motions in space. Although Gibson did not consider the KDE to be ecologically valid due to its absence of perspective transformation, his analysis of the optic flow field inspired later computational models of the SFM problem.

1.7.4 Johansson: Perceptual Vector Analysis

Johansson and his colleagues (Börjesson & von Hofsten, 1973; Johansson, 1950, 1964, 1970, 1973, 1986; Johansson & Börjesson, 1989; Eriksson, 1974) proposed that the visual system was
predisposed to interpret sets of frontal-plane motions as the movement of a rigid configuration in 3D space whenever those motions could be related through a perspective transformation. Vector analysis of the retinal motions is used to determine the relationship between sets of moving elements by extracting both a common motion vector and relative motion vectors. “Decoding principles” or heuristics which are “hard-wired into the visual computer”, are used to determine whether or not the common motion vector could have been produced by a perspective transformation. Residual relative motion vectors are treated as deformations of the object’s structure as it moves in space.

### 1.7.5 Neurally-Based Models of SFM

Several models have drawn upon the known physiology of the visual system to provide neurally plausible mechanisms for the recovery of SFM.

#### 1.7.5.1 Orban, Gulyas, and Spileers.

Orban, Gulyas, and Spileers (1987) and Gulyas et al. (1990) propose that depth from motion could be computed by centre-surround directionally selective neurons in a manner analogous to that hypothesised for binocular disparity detectors (e.g., Poggio & Poggio,
In their account, antagonistic-surround neurons (cf. far cells) respond to a background behind the fixation plane and an object in front (a situation where lateral observer movement would produce opposite directions of image motion), while synergistic-surround neurons (cf. near cells) respond if both the background and the object lie in front of fixation (producing the same direction of image motion). Speed differences between the object and background motions also determine the magnitude of their depth separation, and are encoded by the velocity tuning of the neuron’s surround. As the object motion for one direction-selective neuron may serve as the background motion for another neuron, a series of these cells could potentially recover the ordinal depths of visual features relative to the fixation plane. One problem with this scheme is that the direction of depth relative to the fixation point cannot be determined by centre-surround neurons unless the direction of observer movement is known (this would be true for disparity-detectors in the absence of eye-of-origin information).

1.7.5.2  *Buracus and Albright.*

Buracus and Albright (1996) ascribe a different functional role, in the SFM process, to the centre-surround organisation of motion cells. They have shown that antagonistic-surround neurons in area MT could act as differential motion filters, extracting 3D shape primitives
(surface orientation and curvature) directly from the optic flow field. In their model, the various surround speed-tuning functions observed in MT cells allow the recovery of different aspects of 3D shape. Neurons with V-shaped speed-tuning functions estimate surface orientation (i.e., tilt and slant) by performing a first-order spatial derivation of the velocity field. Neurons in which suppression either increases or decreases with surround-speed extract second-order spatiotemporal derivatives from which surface curvature is computed. Such monotonic speed-tuning allows the discrimination of concave and convex surfaces, as each class of neuron will respond only to one curvature sign, and the estimation of curvature magnitude, which corresponds to the velocity difference between centre and surround motions.

1.7.5.3  **Nawrot and Blake**

Nawrot and Blake (1991a) advanced a network model, composed of binocular units, both directionally and disparity selective, that mimicked human performance with KDE stimuli. Each unit, in the model, responds only to a particular direction of 2D motion in a particular depth plane—either behind, on, or in front of the fixation plane. The overall behaviour of the network is determined both by direct stimulation, and more importantly, by the interaction between
units via inhibitory and excitatory connections. Units responsive to a
given direction of motion in the fixation plane have excitatory
connections to same–direction near and far units, who, in turn, inhibit
their zero-disparity counterparts and each other. Units tuned for
opposite directions of motion, on any depth plane, are also mutually
inhibitory. These asymmetric interconnections, along with simulated
unit adaptation following sustained activity, produce activity patterns
in model simulations that resemble KDE phenomena.

When the model was stimulated with opposite directions of motion
in the fixation plane, activity spread rapidly to same-direction units
with opposite signs of disparity tuning. This network behaviour is
analogous to the perceived depth separation of superimposed, shearing
surfaces observed in the KDE (e.g., the front and back faces of a
rotating transparent cylinder). After prolonged stimulation these units
adapted and activity collapsed back to zero-disparity units, then spread
to the non-adapted same-direction units in the opposite depth plane
(i.e., a depth reversal). Simulations with opposing motions in different
depth planes produced network responses consistent with stereo
disambiguation and stereo adaptation of the KDE (Braunstein et al.,
1.7.6 Computational Approaches to the KDE

Ullman (1979) pioneered the computational analysis of how an object’s 3D shape could be determined from its projected motions. In his structure-from-motion (SFM) theorem, Ullman proved that a unique structure (or its reflection in depth) could be reconstructed from three distinct orthographic views of any four non–coplanar points, provided they formed a rigid configuration. This mathematical analysis demonstrated that a single a priori assumption about the distal world, namely object rigidity, was sufficiently powerful to constrain the 3D interpretation of retinal motion. Additional assumptions, like a fixed axis of rotation or a constant angular velocity, have been shown, for these special cases of motion, to reduce the theoretical minimum number of points and views required for the recovery of SFM (e.g., Hoffman & Bennett, 1986).

SFM models may generally be distinguished on the basis of the information from which they derive 3D structure. Feature-based models operate by analysing the positional changes, over different viewpoints, of identifiable features of a rotating object. These models require the prior solution of the correspondence problem—the matching of individual feature elements across each view. Velocity-based models, on the other hand, compute depth by measuring deformations in the retinal flow field of instantaneous velocities that
correlate with depth variations of environmental surfaces. Ullman’s incremental rigidity scheme (section 1.7.6.1) and Koenderink’s differential velocity scheme (section 1.7.6.3) are representative examples of feature-based and velocity-based models respectively, reviewed next.

1.7.6.1 Ullman: Incremental Rigidity Scheme

Ullman (1984) proposed the incremental rigidity scheme that recovers an internal 3D wire-frame model, representing an object’s estimated structure, from a temporally extended motion sequence. Initially the model is flat, unless static depth cues are present that determine a 3D structure. With each new view of the object, the model’s structure is transformed to minimise the change in 3D distances between points in the model and to account for the changes in the 2D positions of object features. This makes a robust model favouring, but not requiring, a rigid interpretation wherever possible.

Computer simulations (Grzywacz & Hildreth, 1987; Landy, 1987; Ullman, 1984) of the incremental rigidity scheme have demonstrated successful convergence to a single, accurate depth solution, even when the simulated object’s motion was not completely rigid. The convergence of the algorithm, and hence the recovery of estimated 3D shape, improved with additional views of the object, particularly when the angle of rotation between the views was large (>15 degrees).
Performance was much poorer when the views were separated by small rotation angles.

On occasion the estimated depth structure in the converged model became unstable and was temporarily lost. Following this, the model often reconverged to the mirror–reversed solution rather than returning to the correct 3D structure. Such behaviour in the incremental rigidity scheme is interesting because it represents the computational equivalent of a perceptual depth reversal.

Two further similarities exist between the performance of the incremental rigidity scheme and human observers. First, the KDE does not require fully rigid motion (see Section 1.8.2) although it should be noted that observers can tolerate much greater deviations from rigidity than can the incremental rigidity scheme. Second, the KDE builds up with prolonged observation. Although human observers can experience the KDE in less than 300 msec (Doner, Lappin & Perfetto, 1984; Husain, Treue, & Andersen, 1989), both the accuracy and the depth of the perceived 3D structure have been shown to increase systematically with extended viewing (Braunstein, Hoffman, Shapiro, Andersen, & Bennet, 1987; Braunstein, Hoffman, & Pollick, 1990; Eby, 1992; Hildreth, Grzywacz, Adelson, & Inada, 1990; Husain et al., 1989; Loomis & Eby, 1988, 1989; Petersik, 1979; Todd, Akerstrom, Reichel, & Hayes, 1988; Todd & Norman, 1991;
White & Mueser, 1960). Experiments conducted with oscillating rotation sequences established that it was the total extent of angular rotation rather than the observation time per se that was responsible for the temporal build-up of the KDE (Braunstein et al., 1987; Eby, 1992; Loomis & Eby, 1988, 1989). The maximum perceived depth and peak performance in 3D discrimination tasks was obtained with angular rotation extents in the range of 40 to 60 degrees (Eby, 1992; Hildreth et al., 1990; Husain et al., 1989).

An inconsistency between the empirical data and incremental rigidity scheme simulations is revealed by the manipulation of the rotation angle applied between successive views. Small inter-frame rotations produce optimal conditions for the KDE and perceived depth is lost with large rotation angles (see Section 1.8.1). The reverse relationship holds for the incremental rigidity scheme—the algorithm fails to converge properly with small, simulated rotation angles and performs best with large angular differences between views.

1.7.6.2 Hildreth et al. Surface Reconstruction

An elaborated version of the incremental rigidity scheme has been developed within a broader framework for computing 3D shape from motion by Hildreth, Ando, Andersen, and Treue (1995). In their surface reconstruction model, the estimated depths of individual moving features are computed by a SFM algorithm that provides a
depth map from which smooth surface functions are interpolated. The SFM algorithm alternately estimates those 3D velocities and depths that would simultaneously maximise the rigidity of the internal wire-frame model and account for the pattern of 2D image velocities. The surface reconstruction stage then fits the smoothest possible surface to the point-based skeletal structure obtained from the SFM stage using a surface interpolation algorithm described by Grimson (1981, 1983). This algorithm operates using simple local computations, and these estimates of the surface function are propagated over longer distances, unless explicit boundaries are present, to ensure that the surface representation is globally smooth.

Prior to the reconstruction stage, features are segregated into groups of points, belonging to a particular surface, on the basis of local differences in the direction and/or speed of their 2D motions. For example, features on opposite faces of a rotating transparent cylinder move in opposite directions and would be analysed separately by the interpolation algorithm, producing two independent surface maps. By maintaining separate surface maps, the scheme is able to recover the 3D structure of different surfaces that lie in a similar visual direction (as in the case of a transparent object or partially overlapping surfaces).
Features do not need to remain constantly in view for the recovery of 3D shape because an object’s representation is surface-based rather than point-based. When a feature disappears from view the surface structure persists, and when a feature appears it is assigned an initial depth value derived from the computed surface structure at its location. The interaction between the surface interpolation and SFM algorithms enables the model, like human observers (e.g., Treue, Andersen, Ando, & Hildreth, 1995), to recover 3D structure from limited-lifetime stimuli—a technique which eliminates texture gradients by repositioning each dot in a new random location on the 3D object after an arbitrary number of frames.

Simulations of the model produced further behaviour that resembled human performance. Ramachandran et al. (1988) found that the 3D structure perceived in a KDE stimulus was markedly affected by the 2D shape of its outline. For example, a random-dot cylinder viewed through a triangular mask was reported to look like a cone. The surface reconstruction scheme also recovered a 3D structure which was systematically distorted by its 2D outline, because the interpolation algorithm takes advantage of segmentation boundaries to “pin-down” estimated depths along the edges of smoothed surfaces. Furthermore, the finding that observers perceive continuous surfaces over featureless areas in KDE displays (Saidpour, Braunstein, &
Hoffman, 1992; Treue, et al., 1995) suggests that they too employ some form of surface interpolation.

1.7.6.3  

**Koenderink and van Doorn: Differential Velocity-based models**

Velocity-based SFM models (e.g., Koenderink & van Doorn, 1975, 1976, 1986; Longuet-Higgins & Prazdny, 1980; Waxman & Ullman, 1985) analyse the structure of instantaneous retinal flow to determine the 3D form and motion of the object(s) that structured the flow. These models recover the gradient and curvature of a surface in motion relative to the observer directly from its optic flow by taking the spatial derivatives of retinal velocity vectors. Koenderink and van Doorn (1975, 1976) suggested that the structure of the velocity field might be analysed through its decomposition into four independent components: translation, divergence (pure size change), curl (pure rotation), and deformation (pure shape change or shear). The latter three components, termed differential invariants, are useful because they independently describe different aspects of environmental events. From deformation alone it is possible to derive the local tilt and slant of a surface patch, and the derivative of deformation provides information about the surface curvature (Koenderink & van Doorn, 1986; Longuet-Higgins & Prazdny, 1980).

There is reason to doubt the practical viability of a depth analysis based on spatial derivatives of the velocity field (Adiv, 1989;
Longuet-Higgins, 1986; Nakayama, 1985; Verri & Poggio, 1989). These models are particularly sensitive to noise because any errors in the early measurement of velocities are amplified through differentiation, and also because the local velocity field is treated as being smooth – an assumption that is frequently violated in natural scenes. The further requirements for a large field of view, perspective projection, and densely textured surfaces, limits the applicability of velocity differentiation models to the KDE.

Koenderink and van Doorn (1986, 1991; Koenderink, 1986) advocated an alternative approach to computing SFM in which the instantaneous shearing motion between triplets of object features was used as a means of measuring deformation. In this motion-shear algorithm the essential 3D structure of a rotating and bending seven-vertex polyhedron is recovered, up to a relief transformation, from only two views (a stimulus which clearly violates the SFM theorem). These authors have argued cogently for value of partial depth solutions, such as the relief solution given by the motion-shear algorithm, leading to an affine rather than an Euclidean representation of structure. In their view, biological vision systems may take shortcuts and recover different levels of structure from motion depending on the visual information available and the task at hand.
1.7.7 Theoretical Basis of the KDE: Conclusions

How do SFM models account for depth reversal? Those based on SFM theorems under parallel projection all yield two potential 3D solutions related by a mirror reflection in depth. Even those models that operate on polar projection do not generally guarantee a unique solution. Polar projection (and high levels of perspective) creates computational problems that render the SFM problem largely intractable. Most models that employ polar projection therefore solve the SFM problem using local operations to approximate parallel projection. When considered within a local region, 3D motions produce highly similar patterns of 2D motions under both parallel and polar projections, hence both suffer the same reflective ambiguity. It is unlikely that the small, local, 2D velocity differences due to perspective could be registered by the visual system with sufficient accuracy to attain a unique depth solution given the inherently noisy measurements of 2D motion.

In other words, SFM models largely ignore the phenomenon of KDE depth reversal, considering only the unsigned depth structure.

Furthermore, under polar projection, a change in depth organisation necessarily leads to a change in 3D shape (and a possible loss of shape rigidity), so any theory of 3D shape recovery from motion must take depth reversals into account.
Several authors have examined the computational equivalent of depth reversal in their models and considered the reason for such a change in the 3D organisation of the recovered solution. Depth reversals in the incremental rigidity scheme correspond to errors in the convergence algorithm, where instead of settling in one local minimum, the algorithm moves to a different minimum during an iteration. Sperling and Dosher (1994) conceptualise the convergence of the SFM algorithm through physical analogy in which the solution space constitutes a saddle-shaped surface which contains wells (representing different local minima, or alternative solutions with opposite depth organisations) over which a marble (representing the convergence of the algorithm) rolls. Depending on chance factors (or stimulus characteristics such as depth cues that alter and mould the topology of the solution space) the marble will roll into one or the other well with equal probability. In the case of orthographic projection the wells are very shallow and the marble having rolled into one may have sufficient momentum to roll back out again and into the opposing well. This process would be the equivalent of a depth reversal. Landy (1987) suggested that such a failure to converge to a single minimum, in his implementation of the incremental rigidity scheme, was “clearly a function of errors made by the minimization
algorithm and … a more robust function minimizer need never suffer losses of structure” (p. 873).

An error-free convergence algorithm should not produce a depth reversal. Once a minimum is reached there is no reason for the algorithm to reconverge. The implication is that computational depth reversal is little more than an artefact of a faulty convergence process (rather than inherent to the process of 3D shape recovery) and that the change in organisation would not be predicted by the SFM theorem. As Sperling and Dosher (1994) have pointed out, reconvergence may be advantageous as it is always possible that the algorithm has settled into the wrong minimum, perhaps a local rather than a global minimum, and will yield an incorrect depth solution. For this reason, it is better for there to be some constant fluctuation in the convergence process, akin to shaking the marble during its traversal over the energy surface, to avoid settling on the wrong solution (a shallow rather than deep well). This explanation is consistent with the hypothesis-testing account of perceptual reversals (see Section 1.10.1).

Nawrot and Blake (1991b, 1993) specifically address the issue of depth reversals. In their model, depth reversals occur due to adaptation and inhibitory interactions between units signalling motion on opposite depth planes. This is strongly reminiscent of the satiation account (see Section 1.10.2) of depth reversals.
1.8 Empirical Aspects of the KDE

1.8.1 Is Motion Necessary for the KDE?

A surprising aspect of position-based SFM models is that they do not require motion per se. Different static views of a configuration will suffice to compute object structure, the temporal relationship between the views is unimportant. In contrast, not only is perceived motion required to produce the KDE but a number of studies (e.g., Dick, Ullman, & Sagi, 1987; Doner, Lappin, & Perfetto, 1984; Mather, 1989; Petersik, 1987; Todd et al., 1988) have shown that the KDE only occurs from motions within a fairly narrow spatiotemporal window.

Four lines of empirical evidence demonstrate an essential role of short-range motion (Braddick, 1974) for the KDE. First, the perceived depth in the KDE breaks down if the rotation angle between successive views is large (Dick et al., 1989; Ganis, Casco, & Roncato, 1993; Landy, Dosher, Sperling, & Perkins, 1991; Mather, 1989; Petersik, 1987). With larger rotation angles there is an increasing proportion of elements whose 2D displacements exceed the spatial limits of short-range motion. Second, prior selective adaptation of the short-range motion process decreases the inter-frame rotation angle at
which KDE collapse occurs (Petersik, 1991). Third, depth is eliminated in KDE displays by manipulations that specifically disrupt short-range motion such as dichoptic presentation (Mather, 1989), alternations of element contrast polarity (Dosher, Landy, & Sperling, 1989a, or interleaving of grey-fields between successive frames (Dosher et al., 1989). Fourth, the KDE is not obtained unless the timing parameters of the animation sequence are suitable for producing apparent motion (Doner et al., 1984; Petersik, 1979; Todd et al., 1988).

The above findings challenge the notion that the KDE is derived solely from positional information, and imply that feature velocities instead may constitute the primitives for the KDE. Grzywacz and Hildreth (1987) have offered an alternative interpretation. In their view, motion is not used directly in the depth recovery process, but rather serves to establish feature correspondences across different views. When the correspondence process fails, so too does the SFM process. Although a direct test of this hypothesis has not yet been performed, some indirect support is provided by qualitative reports. Under circumstances where the KDE breaks down, random patterns of element motions are usually perceived (Livingstone & Hubel, 1988; Vaina, 1989). These incoherent 2D motions would be the expected consequence of inappropriate correspondence matching.
Is Rigid Motion Necessary for the KDE?

Rigidity has long been considered to be an important organising principle of perceived 3D structure from motion (Gibson & Gibson, 1957; Johansson, 1964; Wallach & O’Connell, 1953), and almost all computational analyses of SFM assume some form of a rigidity constraint. If human observers do adopt an assumption of rigidity then they should be particularly good at discriminating rigid from non-rigid 3D motions.

Petersik (1987) found that observers could determine with 73% accuracy whether four points rotated rigidly from only three views, the minimum conditions of the SFM theorem. Braunstein et al. (1990) argued that earlier studies (cf. Sperling, Landy, Dosher, & Perkins, 1989) demonstrating reliable rigidity discrimination were confounded by the presence of 2D display artefacts. They rigorously eliminated 2D regularities related to simulated 3D rigidity from their displays and found that all observers could discriminate non-rigid motion with two views of five points, and that most observers could perform the task with only four points. These authors made the important qualifying point that the discrimination of rigid from non-rigid objects, on the basis of motion, requires no more than the detection of rigidity. Observers do not need to recover a rigid 3D structure to perform these
tasks. In fact, it is not possible to recover a unique, rigid Euclidean structure from only two orthographic views (Ullman, 1979).

Todd (1984) conducted another study qualifying the role of the rigidity assumption in the KDE. Although observers in this study discriminated rigid from non–rigid motion with high accuracy, their performance deteriorated sharply when the simulated objects rotated around a moving axis. Green (1961) also observed that tumbling motions appeared to be highly non-rigid. The movement of the rotation axis should not alter the rigidity of the stimulus, and so would not affect a SFM model that operated on a strict rigidity principle.

The rigidity assumption is challenged by other empirical data: the rigid motion of certain objects around a fixed axis can produce an overwhelming perception of a non-rigid 3D structure (Adelson, 1985; Braunstein & Andersen, 1984; Ishiguchi, 1988, 1990; Pomerantz, 1983), objects often appear to deform as they rotate under strong polar projection (Braunstein, 1962; Green, 1961; Petersik, 1980; Sperling et al, 1989), a depth reversal leads to a perceived 3D form that is grossly non-rigid (Mach 1872/1960; Schwartz & Sperling, 1983), and 3D structure can easily be perceived from non–rigid motions (e.g., Braunstein & Andersen, 1984; Cutting & Kozlowski, 1977; Jansson, 1977; Jansson & Johannson, 1973; Johansson, 1973b; Todd, 1984; von Fieandt & Gibson, 1959). Taken together these findings suggest
that the human visual system does not adopt a strict rigidity constraint to recover structure from motion.

1.8.3 What are the Minimal Conditions for the KDE?

According to the SFM theorem (Ullman, 1979), three views of four points represent the theoretical minimum conditions for recovery of 3D structure from motion. Somewhat paradoxically, human performance appears to exceed these limits. Compelling kinetic depth can be perceived from only three points in extended rotation sequences (Börjesson & von Hofsten, 1973; Hildreth et al., 1990; Lappin & Faqua, 1983) and from as few as two views of a rotating random-dot object (Braunstein et al., 1987; Doner et al., 1984; Lappin, Doner, & Kottas, 1980; Liter, Braunstein, & Hoffman, 1993; Todd et al., 1988; Todd & Bressan, 1990; Todd & Norman, 1991). The latter finding has been confirmed using limited dot-lifetime stimuli. The use of dot lifetimes lasting two frames still produces a strong KDE (Todd, 1985; Treue, Husain, & Andersen, 1991).

Todd (1985) has offered an explanation for this apparent out-performance of the ideal observer. He points out that while the SFM theorem states the requirements for computing a Euclidean structure, the visual system might recover a more qualitative depth representation in the KDE when provided only two views.
1.8.4 Is Euclidean Structure Recovered in the KDE?

In light of the evidence provided by Todd and his colleagues, it seems unlikely that a metric Euclidean representation is normally achieved in the KDE. Instead, the perceived 3D shape appears to be an affine structure. This claim is supported by the findings that (a) the perceived shape of a rotating object does not change when affine stretching transformations are added (Braunstein & Andersen, 1984; Todd, 1984); (b) increasing the motion sequence beyond two views does not improve objective judgements of 3D structure (Liter et al., 1994; Todd & Bressan, 1990; Todd & Norman, 1991); (c) observers perform better on affine-based tasks, such as the bisection in depth of rotating points (Lappin & Faqua, 1983) than on Euclidean-based tasks, such as 3D length comparisons between rotating lines oriented in different directions (Norman, Todd, Perotti, & Tittle, 1996; Todd & Bressan, 1990); and (d) that the perceived depth in the KDE is systematically overestimated (Braunstein, Liter, & Tittle, 1993; Tittle, Todd, Perotti, & Norman, 1995; Todd & Norman, 1991).

1.9 Depth Reversals in the KDE

Few theories of SFM, with the exception of Nawrot and Blake (1991a; see Section 1.7.5.3) have explained the occurrence of depth reversals in the KDE. Instead SFM theories have shown that the
orthographic projection of a KDE stimulus is ambiguous in that, theoretically, signed depth cannot be recovered without the additional effects of perspective (e.g., differential motion velocities between near and far surfaces) which are present in polar projection. The identification of ambiguity in a stimulus does not, however, account for the phenomenon of depth reversals. Logically, an observer could always perceive one particular depth organisation (and associated rotation direction) for an ambiguous KDE stimulus and never experience a depth reversal.

In addition to the lack of theoretical attention to KDE reversals, there has been little empirical work conducted on the basic phenomenon using KDE stimuli. The characteristics of depth reversals have been investigated with a variety of rotating figures—namely the Ames window, Necker cube, and Lissajous patterns, but these figures are not typical of KDE stimuli. Perspective information is potentially available in both the Ames window and the Necker cube due to their projected shapes. Consequently, static views of these stimuli may appear 3D whereas static views in the KDE appear 2D. The extensive literature on Ames windows is not directly applicable to KDE reversals since these studies are primarily concerned with the phenomenon of regular oscillation that is thought to occur as a result of the perspective present in the stimulus. Lissajous patterns also
differ from the KDE in that under rotation there are regular occasions
where the loops on the front and back faces of the stimulus coincide
exactly and the two surfaces become indistinguishable. In the KDE
the front and back faces of the stimulus (but not necessarily individual
elements) remain continuously in view.

Given the paucity of KDE studies directly concerned with the
phenomenon of depth reversal, it is necessary to consider the general
literature on perceptual reversals with other stimuli (in Section 1.10).
The degree to which KDE reversals resemble reversals with non-KDE
stimuli is considered in Experiment 1.

1.10 General Theories of Perceptual Reversal

Changes in fixation or eye-movements across an ambiguous figure
such as the faces-vase figure were originally thought to be the cause of
perceptual reversals (Tichener, 1906; cited by Pheiffer, Eure &
Hamilton, 1956). The eye-movement theory was later discounted
when reversals were obtained with afterimages (Mefferd et al., 1966)
and stabilised-images (Pritchard, Heron, & Hebb, 1960) of ambiguous
stimuli. Furthermore, Pheiffer et al. (1956) directly measured eye-
movements while observers reported apparent reversals of the
Schroeder staircase figure and found that eye-movements followed
from, rather than preceded, most reversals.
Following the eye-movement hypothesis, two opposing accounts of perceptual reversal—the satiation and the decisional theories, have emerged.

1.10.1 Satiation Theory

According to the satiation theory (Köhler & Wallach, 1944) alternative percepts of an ambiguous figure are represented by different underlying cortical organisations. The relative balance of activity between the competing cortical organisations determines which alternative is perceived. Mutual inhibition between these neural substrates ensures that only a single percept is seen at a given time. While a given cortical organisation is dominant (and the corresponding alternative of the ambiguous figure is seen) it is subject to neural fatigue. After prolonged viewing, fatigue reduces the level of activity in the dominant organisation to a point where the balance changes in favour of the fresher, alternative cortical organisation. At this time a perceptual reversal is experienced. The new cortical organisation undergoes the same fatigue process and is eventually replaced by the original cortical organisation, which has had time to recover, and the cycle of fatigue and recovery continues. Fatigue becomes more rapid with extended observation because the cortical organisations are assumed to undergo incomplete recovery after each reversal.
Over the last 50 years, the basic principles of the original satiation theory (i.e., mutual inhibition, fatigue, and recovery) have remained intact. Refinements to the theory have involved the hypothetical mechanisms underlying the satiation process. Satiation has variously been described as the non-neuronal electrotonic spreading and blockage of cortical potentials (Köhler & Wallach, 1944), as an auto-inhibitory process at the synaptic level (Howard, 1961), as analogous to electrical current leakage and resistance in a multi-vibrator flip-flop electronic circuit (Attneave, 1971), and as the strengthening and decay of synaptic patterns of activity in a neural network (Kawamoto & Anderson, 1985).

1.10.2 Decisional Theory

Several workers, notably Gregory and Rock, have disagreed with the view that reversals are simply the by-product of low-level neural processes during continuous passive stimulation. They characterise perception as an active process in which the observer subconsciously selects a particular perceptual interpretation of the visual stimulation from many possible alternatives through ongoing problem-solving (Rock, 1975) or hypothesis-testing (Gregory, 1970) activity. In the case of an ambiguous stimulus, the observer is presented with at least two equally plausible solutions to the perceptual puzzle. Consequently
perceptual reversals occur as the visual system vacillates between the alternatives.

Vickers (1972) formalised this process with a cyclic decision model based on an information-processing approach to perception. In this model, evidence favouring each alternative organisation is accumulated separately until a threshold is reached. At this point a decision is made and the alternative with the greatest accumulated support is perceived. The accumulators for both alternatives are then reset and the decision-making process begins afresh. Reversals are due to stochastic variations in the accumulation of evidence.

1.11 Empirical Aspects of Perceptual Reversals

A vast amount of data has been collected during a century of experiments with reversible stimuli. This brief review concentrates on the major variables that have been found to affect reversals in static and rotating ambiguous stimuli. The implications of the results for the two major theories of perceptual reversals are highlighted where appropriate.

1.11.1 Temporal Characteristics

Perceptual reversals have been found to follow a distinct time-course during extended observation of ambiguous stimuli. Typically there is a notable latency before the first reversal occurs. Following
this, the perceived organisation of the stimulus becomes increasingly unstable with continued viewing, and reversals increase in frequency as a negatively accelerated function of observation time. The reversal rate reaches an asymptotic frequency after rising for about 3 minutes (Brown, 1955; Cohen, 1959; Price, 1969). Towards the end of long (10-12 minutes) observation periods, the rate of reversals fluctuates widely and shows an overall decline (Bruner, Postman, & Mosteller, 1950; Price, 1969). A systematic increase in reversal frequency over time has been demonstrated with a variety of ambiguous stimuli (Necker cube: Babich & Standing, 1981; Cohen, 1959; Spitz & Lipman, 1962; rotating Necker cube: Howard, 1961; Long, Toppino, & Kostenbauder, 1983; Price, 1969; Maltese cross: McDougall, 1906).

The observed increase in reversal frequency with viewing time is predicted by the satiation theory. As a consequence of incomplete neural recovery the fatigue process occurs more rapidly, with successive reversals producing a positive feedback loop that speeds the fatigue-recovery cycle. The decisional theory requires the post-hoc addition of an associative learning component to explain the temporal character of reversals. Observers are thought to learn the alternative interpretations of ambiguous figures through perceptual experience. The negatively accelerated function relating reversal frequency to
observation time is considered evidence of a standard learning curve (Ammons, 1954).

Additional support for the perceptual learning hypothesis comes from the finding of practice effects with ambiguous figures. Long et al. (1983) found that when observers observed a Necker cube in testing sessions separated by a week, they reported more reversals in each successive session. The finding of long-term elevation of reversal rates is difficult to reconcile with the satiation theory.

1.11.2 Complexity

Stimulus complexity has been variously defined in terms of the number of lines in different ambiguous figures such as the Mach book, Schroeder staircase, Necker cube, and Honeycomb pattern (Washburn, Reagan, & Thurston, 1934; Donahue & Griffitts, 1931), the number of spikes in a Lissajous pattern (Philip & Fischelli, 1945), or as the amount of structural information (Leeuwenberg, 1971) favouring a concave interpretation of a Mach pyramid (Masulli & Riani, 1989). Evidence for the effect of stimulus complexity on the rate of perceptual reversal has been equivocal. Several authors (Washburn et al., 1934; Masulli & Riani, 1989) have reported that complex ambiguous figures undergo fewer reversals than do simple figures. Donahue and Griffiths (1931) obtained a slight reduction in reversal
frequency when the complexity of their ambiguous line drawings was increased but this effect was not statistically reliable. With Lissajous patterns, which are dynamic rather than static ambiguous stimuli, the tendency for reversal was found to increase with stimulus complexity (Philip & Fisichelli, 1945).

1.11.3 Rotation speed

Considering the large number of studies which have examined reversals in rotating stimuli, it is surprising that relatively few studies have actually tested the effects of motion on perceptual reversals. Those studies varying rotation speeds have reported conflicting effects.

Most of these studies have examined the effect of speed on stimuli viewed under polar projection. Polar-projected stimuli are theoretically unambiguous because differential motion parallax exists between near and far points on the stimulus that potentially specifies the veridical direction of rotation (and depth organisation). However stimuli that contain conflicting depth-order information, such as linear perspective in the case of trapezoidal planes or convexity cues in the case of the hollow face-mask, are often misperceived as oscillating rather than continuously rotating. Apparent oscillation arises from systematic depth reversals, where parts of the stimuli that are actually
most distant are perceived as being front-most during one half of the rotation cycle (Power & Day, 1973).

Increases in the rotation speed of polar-projected trapezoids from 3rpm to 30rpm have been found to reduce the frequency of depth reversals (i.e., oscillations) (Börjesson, 1971; Braden, 1978; Mulholland, 1956). A similar drop in reversals with an increase in the rotation speed of a hollow face-mask from 4rpm to 8rpm has also been obtained by Klopfer (1991). In all four studies, the perceived change from oscillation to continuous rotation was attributed to an increase in the effectiveness of differential motion parallax with speed.

This explanation predicts that reversal frequency will be unaffected by manipulations of rotation speed under parallel projection as here motion parallax remains ambiguous. In one condition of Börjesson’s (1971) experiment where rotating trapezoids were projected at a simulated distance approximating parallel projection, the reported oscillation rate remained constant despite a 10-fold increase in rotation speed. On the other hand, Braden (1978) presented rotating trapezoids and rectangles under true parallel projection and found a substantial drop in reversal rates with both stimuli following an increase in rotation speed. In fact, the magnitude of this effect was much higher in parallel than in polar projection conditions. As Braden (1978) noted, the inverse relationship between rotation speed and
reversal frequency cannot be entirely due to the effects of differential motion parallax in her experiment.

The opposite effect of rotation speed (i.e., reversal increases with speed) has also been reported. Mulholland (1956) observed that a polar-projected trapezoid appeared to oscillate more frequently at 90rpm than at 30rpm. Caelli (1979) examined the effect of different rotation speeds (from 30rpm to 180rpm) on the perceived rotation of simple KDE figures under polar projection. As the speed of rotation increased, reports of continuous rotation (i.e., no depth reversal) gave way to perceived oscillation (i.e., continuous reversals). An increase in reversal frequency with rotation speed (from 3rpm to 30rpm) has also been found with parallel projections of Lissajous patterns (Philip & Fisichelli, 1945).

To summarise, changes in rotation speed strongly affect the tendency for depth reversal, but the direction of this effect is inconsistent. Under polar projection the magnitude of rotation speed appears to determine whether reversals will decrease or increase with speed (i.e., a U-shaped relationship). The results with parallel projection are equivocal. Changes in rotation speed within a comparable range have been found to increase (Philip & Fisichelli, 1945), decrease (Braden, 1978), or have no effect on (Börjesson, 1971) reversal frequency.
Neither the decision nor satiation theory make specific predictions for the effect of rotation speed on perceptual stability. However there seems little reason to expect this stimulus manipulation to influence perceptual learning. On the other hand, fatigue of antagonistic neural mechanisms, as postulated by the satiation theory, is affected by similar stimulus properties in other domains, such as the speed of the inducing stimulus in the case of the motion aftereffect.

1.11.4 Cognitive factors

Support for the decisional account of perceptual reversals has arisen from a number of studies showing that reversals may be influenced by factors that cannot be reasonably attributed to the operation of a neural satiation process. Ulrich and Ammons (1959) found that highly experienced observers were able to reverse the perceptual organisation of physical objects such as tables, chairs, and even the dihedral corner formed between the walls of their laboratory. Furthermore, after some training naive observers could also invert these normally unambiguous objects.

Whereas it is generally accepted that observers cannot completely prevent or induce a perceptual reversal at will (Sadler & Mefferd, 1970), there is considerable evidence that the rate of reversals can be bought under conscious control. Pelton and Solley (1968) and also Bruner et al. (1950) instructed observers to either try and switch
perspectives of a Necker cube figure as quickly as possible, or to try and maintain the same organisation. More perceptual reversals were reported by observers under the switch instruction, and the usual increase in reversals with observation time was not found when observers were asked to avoid reversals. Other studies (Hochberg & Peterson, 1987; Liebert & Burk, 1985; Peterson & Hochberg, 1983; Phillipson & Harris, 1984; Washburn & Gillette, 1933) have shown that an observer’s preference for one organisation of an ambiguous stimulus over the other can be biased by instructions to hold that organisation. Due to the high probability of response bias and experimenter demand (Rosenthal, 1966) these findings should be accepted with some reservation. Under conditions of uncertainty with ambiguous figures, it may not be surprising that observers should report percepts that comply with the instructions they were asked to follow. More convincing support for the role of cognitive factors is provided by Girgus, Rock, and Egatz (1977) who found that naive observers often did not experience perceptual reversals until they were made aware of the possible alternative organisations of an ambiguous figure.

Attention is a further cognitive component that affects perceptual reversal. Observers in Reisberg and O’Shaughnessy’s (1984) experiment viewed a series of ambiguous figures while engaged in a
concurrent distracter task (mental arithmetic) on half of the trials. A significant slowing in the speed of perceptual reversals occurred when observers simultaneously counted backwards, compared to reversal rate obtained on control trials where observers simply observed the ambiguous figure.

1.11.5 Individual differences

Observers vary considerably in their tendency to experience perceptual reversals. Many studies have found large individual differences in reversal rates (Brown, 1955; Frederiksen & Guilford, 1934; Wieland & Mefferd, 1967) and also that these differences remain stable over time (Frederiksen & Guilford, 1934; Jackson, 1956; Howard, 1961; Spitz & Lipman, 1962).

In the past, the rate of perceptual reversals has been considered as a rather direct and objective test of cortical function. This assumption motivated several correlation studies that sought to relate individual reversal rates to some characteristic of the perceiver.

Frederiksen and Guilford (1934) examined McDougall’s (1906) hypothesis that reversal frequency was related to personality traits which corresponded to introversion and extraversion. Introverted individuals were expected to exhibit a greater tendency for perceptual reversal due to a relative paucity of a neurochemical known as
“Secretion X ... [a secretion that] operates by increasing the synaptic resistance in the cortex” (p. 470). No correlation was found between reversal rates and any of the 35 items on the introversion-extraversion scale that they used.

Mull, Arp and Carlin (1952) found a weak correlation between intelligence and the ability to control perspective reversals but this was not replicated by Jackson (1956) using a more reliable dependent measure. Age has also been implicated as another individual factor influencing reversal rate. Heath and Orbach (1963) reported that elderly observers experienced fewer reversals with a Necker cube than did young adults. This finding is difficult to interpret considering that reversals were only reported by less than 20% of the elderly sample, the remaining observers either did not see the cube as three-dimensional or did not experience any reversals.

Perceptual reversals have also been used as a tool in the investigation of mental disorders. Significant differences in the Necker cube reversal rate have been found between schizophrenic and clinically depressed patients (Eysenck, 1952; Hunt & Guildford, 1933). Phillipson and Harris (1984) pointed out the doubtful reliability of using schizophrenic observers in these studies that involve self-report measures, and sought to replicate these results using an indirect approach. They administered chlorpromazine, an antipsychotic drug
used to relieve schizophrenic symptoms, to normal observers. Observers in the chlorpromazine group experienced slightly fewer perceptual reversals than did observers in the saline control group.
2. **Chapter 2: The Relationship of Depth Cues to the Disambiguation of the KDE**

2.1. **Motion perspective**

Theories of the recovery of structure from motion attribute, either explicitly or implicitly, the ambiguity of KDE depth to the geometric ambiguity of parallel projection. If so, adding motion perspective via geometrically unambiguous polar projection should disambiguate KDE depth. Whether observers actually do use motion perspective to disambiguate the KDE, as predicted by SFM theories, is an important empirical question.

2.1.1. **Full motion perspective: KDE**

Braunstein (1977) varied the simulated projection distance (and hence the degree of motion perspective) of a rotating random-dot sphere. The sphere was projected at a distance of either three- or twelve-sphere radii from the observer. Veridical judgements of rotation direction were obtained in 80% (averaged over conditions) of the trials at the three radii distance, and were near chance levels for the twelve radii viewing distance. Petersik (1980) reported similar
levels of disambiguation through motion perspective. In this study, where a wide range of perspective strengths were tested, rotation judgment accuracy was found to decline as a monotonic function of simulated viewing distance ratios. Motion perspective becomes ineffective in the KDE at a simulated viewing distance between 12 and 15 radii.

Although high levels of perspective disambiguate the KDE and increase its apparent depth, there also is a corresponding loss of coherence accompanied by marked non-rigidity (Braunstein, 1962; Petersik, 1980; Dosher et al., 1989).

As described earlier (Section 1.5.2), motion perspective involves a combination of motion asymmetries related to relative depth, any of which could potentially disambiguate the KDE. Braunstein (1976) has suggested a distinction between those effects of perspective along the dimension perpendicular to the rotation axis (horizontal perspective) and those along the dimension of the rotation axis (vertical perspective). This terminology will be adopted here because every study described below involved rotation around a vertical axis or horizontal translation. For a further examination of the motion perspective cue I will review studies that have sought to isolate the

---

2 Of course, these labels are arbitrary and would be reversed if one were considering rotation around a horizontal axis.
contributions of vertical and horizontal perspective, and then consider a special case of horizontal perspective, namely motion parallax.

2.1.2. Vertical versus horizontal motion perspective

Vertical and horizontal perspective effects were varied independently in the projection of a rotating KDE sphere by Braunstein (1962, 1966). When full perspective was simulated, direction-of-rotation judgements were highly accurate (90% correct). An equivalent level of disambiguation was reported (91%) when only vertical perspective transformations were present, but performance fell to chance (55%) with horizontal perspective alone.

A similar dominance of vertical over horizontal perspective has also been found for rotating planar figures (Braunstein, 1968; Hershberger, Stewart, & Lauglin, 1976), motion parallax displays (Rogers & Rogers, 1992), and also with slant-in-depth judgements for static planes (Gillam, 1970).

Hershberger et al. (1976) placed the two cues in conflict in a KDE stimulus that emphasised horizontal perspective. A horizontal array of 13 short vertical lines, with a width to height ratio of 20:1, rotated around a vertical axis. Observers usually perceived rotation in the direction specified by vertical perspective unless an extreme level of horizontal perspective (which indicated the opposite rotation direction) was applied.
Although these studies indicate that horizontal perspective is a weak cue to depth order at best, evidence to the contrary is offered by work of Hershberger et al., and by studies involving translatory relative motion.

2.1.3. **Horizontal motion perspective: KDE**

Horizontal perspective transformations have been studied in considerable detail by Hershberger and his colleagues, and their research was motivated by the similarity between horizontal motion perspective and the motion parallax cue. Hershberger’s (1967) stimulus comprised a polar projection of a horizontal line of equally spaced dots rotating around a vertical axis. Such a stimulus isolates the effects of horizontal perspective since changes in dot positions occur only in the horizontal direction. Observers in this study were able to identify the veridical rotation direction with 90% accuracy. Simulated projection distance was found to have a strong influence on rotation judgements in a later study by Hershberger and Urban (1970a). Here, the correct rotation direction was perceived for 88% of the trial duration with a projection distance of 2.6 radii, and declined to 66% with a distance of 10.4 radii. This result is consistent with experiments that have simulated full motion perspective (Braunstein, 1977; Petersik, 1980).
Through a systematic dissection of horizontal motion perspective, Hershberger and associates (Carpenter, 1979; Carpenter & Dugan, 1983; Hershberger & Urban, 1970b; Hershberger & Carpenter, 1972; Hershberger, Carpenter, Starzec, & Laughlin, 1974; Hershberger & Starzec, 1974) identified four sources of information they claimed were sufficient to disambiguate rotation direction. The four cues were labelled as: direction (the direction of motion for dots at opposite ends of the line), order (the order in which different dots reversed their direction of motion), acceleration / displacement gradient (the ratio of dot acceleration to the distance from the rotation axis), and velocity (the differential velocity between dots behind and dots in-front of the rotation axis).

Observers were able to judge the simulated direction of rotation at above chance levels with displays that contained any one of these cues. Although a detailed discussion of these experiments lies beyond the scope of this review, one finding is relevant. Carpenter (1979) found that observers achieved similar levels of accuracy with displays that depicted only the velocity cue (the motion-parallax component of a polar-projected rotating object) and with those that simulated full horizontal perspective (91% and 95% respectively). This finding demonstrates that depth order can be determined on the basis of
velocity changes alone when all other components of motion perspective are removed.

2.1.4. Horizontal motion perspective: Superimposed planar surfaces

E. J. Gibson and colleagues (1959) examined the question of whether simulated motion parallax could provide unambiguous relative depth. Their stimuli were shadow projections of two overlapping transparent planes, textured with random elements, which translated in opposite directions. They varied the distance between the two planes (and thus their relative velocities). Observers could clearly distinguish two surfaces separated in depth, but the perceived depth ordering of the two surfaces was unstable and frequently reversed. Although elements on the physically near surface translated at the higher projected velocity, this did not determine the observer’s selection of the perceptually near surface. Subsequent research using computer-generated versions of these stimuli (Andersen, 1989; Mace & Shaw, 1974; McConkie & Farber, 1979) has confirmed the efficacy of relative motion for segregating surfaces in depth. However, with superimposed planar surfaces, depth ordering is usually ambiguous and largely independent of relative velocity.

McConkie and Farber (1979) presented evidence that observer-specific factors, rather than velocity differences, can affect the
interpretation of depth order. Their observers showed a strong and stable bias to perceive the surface moving in a particular direction (rightward, for instance) to be in front on every trial irrespective of its relative velocity. This directional bias was also found by Andersen (1989) and it dominated the responses of two of his five observers. The three observers who were indifferent to translation direction, often (80% of trials) perceived the depth arrangement of the surfaces that corresponded to relative velocities. This tendency increased with the velocity difference (i.e., the simulated depth) between the surfaces.

At least three factors may contribute to the depth ambiguity of superimposed surfaces containing a velocity-specified depth order. First, these stimuli represent two detached surfaces that could potentially move independently of each other. The pattern of relative motion in the stimulus could therefore be attributed to real movement of the objects rather than to their relative depth. If retinal motion is produced by object movement in a stationary environment there no longer exists any relationship between projected velocity and an object’s distance. Second, surface transparency may promote depth reversal. As the two surfaces were simultaneously visible when superimposed there was information to specify that the surfaces were transparent rather than opaque. Third, since the two planes were perpendicular to the line of sight there was no velocity gradient
present. Gibson et al. (1959) argue that such a gradient is necessary to determine correct depth order (direction of slant in particular).

2.1.5. *Horizontal motion perspective: Velocity gradients*

There has been some debate over whether velocity gradients can specify signed depth. Gibson et al. (1959) found that observers could correctly judge the direction of slant from the shadow projection of a moving surface that was tilted away from the observer. This stimulus provided a gradient of velocities that decreased toward the top of the display. Observers perceived this display as a receding surface in depth such that the more rapidly moving surface texture at the bottom of the display appeared to be lying nearer the observer. Flock (1964) applied the shadow-casting technique to present observers with a range of differently slanted moving surfaces. Slant judgements corresponded closely to the physical slant of the depicted surface, with mean regression coefficients ranging from 0.85 to 1.05.

One difficulty for the interpretation of these results is that the shadow-casting technique introduces a texture gradient that is confounded with the velocity gradient. Depth from velocity gradient displays could be ambiguous if supplementary sources of slant information, such as texture gradients, were eliminated. This issue was addressed by Braunstein (1968) who disentangled velocity from texture gradients using computer-generated displays to simulate
slanted surfaces. Braunstein (1968) found that observers judged the veridical direction of slant from velocity gradients alone, and that the magnitude of perceived slant was considerably underestimated. The texture gradient itself produced little effect when presented in isolation (either a stationary or uniform velocity display) but increased the degree of perceived slant when added to the velocity gradient.

The experiments outlined above were primarily concerned with the recovery of surface slant, and did not specifically investigate whether reversals in the direction of perceived slant occurred. Regarding this possibility, Faber and McConkie (1979) claimed that “although lateral motion parallax might provide a basis for perceived separation in depth, it should not be sufficient to determine the spatial ordering or sign of that separation” (p. 496). Their argument concerned the addition of a non-parallactic, uniform component to the proximal flow field, produced through eye rotation, which rendered ambiguous the relationship between projected velocity and direction in depth. From this, Faber and McConkie (1979) argued that depth order information in velocity gradients is fundamentally ambiguous unless the observer is able to partial out the non-parallactic contribution to the flow field.

To test this prediction Faber and McConkie (1979) developed velocity gradient displays that simulated a translating surface composed of two random-dot planes slanted in orthogonal directions.
These planes joined to form a dihedral angle that was either convex (pointing towards the observer) or concave (pointing away from the observer). The appearance of this surface resembled a textured sheet of cardboard bent along its horizontal midline. Observers were presented with velocity gradients that simulated either a convex or a concave surface that translated either left or right. The leftward convex and rightward concave velocity gradients were potentially related through the addition of a uniform velocity component as were the leftward concave and rightward convex gradients. Faber and McConkie (1979) found that observers systematically confused the depth organisation of these displays. Some observers consistently identified the correct direction of depth for leftward moving displays but perceived the reversed depth organisation for rightward moving displays. The remaining observers showed the opposite pattern of confusions.

Braunstein and Andersen (1981) used similar dihedral-angle velocity gradients but found, contrary to Faber and McConkie (1979), that observers could distinguish convex and concave surfaces with an accuracy that ranged from 73% to 94%, under optimal stimulus conditions. They identified several important stimulus and methodological differences to account for Faber and McConkie’s (1979) discrepant findings. Performance was found to improve when
eye-movements were permitted, and when the trial duration was increased from five to ten seconds. The latter effect may be attributed to their finding of an average latency of 3.2 seconds before the displays organised into rigid 3D surfaces. The maximum velocity of translating dots proved to be the primary determinant of depth-order judgement accuracy. Veridical directions of depth were reported in 69% of trials for velocity gradients with a maximum dot speed of 2.6 deg/sec (the speed used by Faber and McConkie), and accuracy increased to 81% with a four-fold increase in maximum dot speed. However, variation of the minimum dot speed had little effect on performance even though this would alter the degree of surface slant (and hence the magnitude of internal depth) simulated by the velocity gradient. This finding is somewhat surprising as with larger perspective ratios in the KDE, objects appear to have greater depth and their simulated direction of rotation is more accurately discerned. 

As mentioned earlier, one of the effects of increasing the perspective ratio in KDE displays is to increase the velocity difference between dots on the near and far faces of the rotating object. 

Braunstein and Tittle (1988) examined further the relationship between maximum to minimum velocity ratios, perceived depth, and perceived direction in depth. They found that although the perceived depth in velocity gradient displays increased with the ratio of
maximum to minimum velocity (and also increased with maximum velocity in displays with the same velocity ratio), the accuracy of depth-order judgements decreased. In fact, perfect performance was obtained with the lowest velocity ratio (1.12), which, if treated as a perspective ratio in the KDE, would not be sufficient to disambiguate the apparent direction of rotation.

In summary, velocity gradients can often produce the recovery of signed depths where the magnitude of perceived depth is in accordance with velocity of feature motion (i.e., the faster the 2D motion the nearer that feature appears). The degree of disambiguation achieved depends on the specific motion characteristics of the display (e.g., the velocity ratio).

2.1.6. Horizontal motion perspective: Motion parallax

Velocity gradient stimuli are an abstraction of the projected motions that result from motion parallax with head movement. These displays are presented on a stationary monitor to a stationary observer. Consequently this situation differs from true motion parallax where there exists relative movement between the observer and the environment.

Rogers and Graham (1979), in their seminal paper, established that motion parallax enabled the veridical perception of 3D structure and depth-order in the absence of any other depth cue. They created the
motion equivalent of random-dot stereograms, a stimulus developed by Julesz (1971) which has proved invaluable for the study of stereopsis from binocular disparity. In Rogers and Graham’s stimuli, parallax displacements of random dots on a simulated 3D surface were yoked to the lateral head movements of the observer. As the observer swayed the head from side-to-side, the dots moved across an oscilloscope display in the direction and velocity that would be produced from this same head movement if one were viewing a real surface.

These displays produced a strong impression of a stationary 3D surface where relative dot motions were perceived as variations in depth rather than element movements. When head-movement (and therefore the relative dot motion on the oscilloscope) ceased, the impression of three-dimensionality collapsed indicating that motion parallax was the only cue to depth present in the stimulus. Observers readily discriminated a variety of simulated corrugated surfaces (square-wave, sinusoidal, triangular, and sawtooth) and the perceived extent of depth in these surfaces varied with the magnitude of the relative dot velocities depicted. Similar results were obtained when the oscilloscope moved relative to a stationary observer, demonstrating that head movement *per se* was not necessary for the accurate perception of 3D structure from motion parallax.
Motion parallax also disambiguated the sign of depth in the depicted surfaces when relative motions were contingent on head- or monitor-movement. Observers identified, with perfect accuracy, whether a particular corrugation was a peak or a trough. Depth organisation was, however, ambiguous when the same pattern of dot motions was presented on a stationary monitor to a stationary observer.

Depth ambiguity is not surprising in these displays as (in contrast to velocity gradient stimuli and in common with KDE displays) depth-order information is carried by the direction, rather than the velocity, of dot motions. Relative dot velocities are uninformative regarding the sign of depth since peaks and troughs have symmetric velocity distributions, differing only in direction of motion. Dots nearer to the observer move in the direction opposite to head movement, while more distant dots move with the observer. What additional information, then, is present in the head- or monitor-movement parallax situation that allows the recovery of signed depth from a supposedly ambiguous velocity gradient?

2.1.7. *Why are Velocity Gradients Unambiguous when the Head or Monitor Moves?*

Braunstein and Tittle (1988) showed that one effect of head- or monitor-movement is to add a common motion vector that effectively
disambiguates the depth in an otherwise ambiguous velocity gradient. In their analysis they distinguished between two optic flow components contained in motion–parallax stimuli, (a) window-relative: the motions of elements relative to the frame of reference provided by the outline of the display monitor; and (b) observer-relative: the motions of elements relative to the observer’s head. When a moving observer, in Rogers and Graham’s experiment, viewed the square-wave grating, for example, a common motion was added to the observer-relative velocity gradient. Under these conditions, all elements move in the same direction but at different velocities, when considered as a motion with respect to the observer. Those bands (peaks) that translate in the direction opposite to head movement, in window-defined coordinates, now move faster relative to the observer than those bands (troughs) that translate with the head. In terms of the observer-relative flow field, a movement of either the head, or the monitor, restores the inverse relationship between velocity and depth.

To determine whether the recovery of signed depth was dependent on observer-relative or window-relative motion, Braunstein and Tittle (1988) used velocity gradients simulating a concave or convex dihedral surface that were displayed within a window which translated across a large monitor. The same observer-relative velocity field was presented in every condition, but the translation speed of the
surrounding window was manipulated to alter the distribution of velocities relative to the window. With this technique, the observer-relative and window-relative velocity gradient could be varied independently. Moving the window at a speed halfway between the minimum and maximum gradient velocities, produced a display with an ambiguous direction of depth when considered as window-relative motions. Either the opposite or the same directions of depth could be specified by the observer-relative and window-relative velocity gradient when the window moved slower than the minimum velocity, or faster than maximum velocity, respectively.

In every condition, observers predominantly judged the depth organisation in accordance with the observer-relative, rather than window-relative, velocity gradient. In other words, the translation speed of the window had little effect on whether a particular velocity gradient was perceived as being convex or concave.

Factors other than the observer-relative velocity gradient may have also contributed to the recovery of signed depth in Rogers and Graham’s (1979) study. Experiments in that study were conducted in a dimly lit room. Consequently when the observer (or oscilloscope) moved, dot motion was seen relative to visible surfaces and objects in the laboratory. In addition, the observer could see the perspective transformation of the oscilloscope display. From these considerations
Rogers and Rogers (1992) identified, and tested, several supplementary sources of information that might have disambiguated depth in Rogers and Graham’s original study. Any one of these sources was a potential indicator of the direction of observer (or display) movement relative to the display (or observer). Observers, in Rogers and Rogers (1992) experiment viewed a 3D random-dot sinewave grating through a tunnel that completely enclosed the oscilloscope display. The following five conditions were examined.

In condition one, visual cues (such as the optic flow of features external to the display, and the perspective transformation of the display itself) were eliminated through the use of the darkened viewing tunnel attached to the oscilloscope. As the observer moved the oscilloscope (and viewing tunnel) pivoted, rather than translated, thereby producing an ambiguous observer-relative velocity gradient. In this condition, non-visual (proprioceptive and vestibular) cues provided the only means of disambiguating the velocity gradient display. This was the only condition where the observer swayed.

The effects of motion perspective were examined in conditions 2, 3, and 4. In condition two, a full perspective transformation was produced by linking the velocity gradient to the movement of the oscilloscope that pivoted in front of the stationary observer. The vertical component of the perspective transformation (i.e., a
trapezoidal expansion and contraction) was simulated on a stationary display in condition three, and the vertical perspective was combined with a synchronised stretching and compression along the horizontal dimension of the velocity gradient in condition four. These three conditions are essentially KDE displays with polar projection.

In condition five, dot motions were yoked to the viewing tunnel that was pivoted back and forth while the observer and oscilloscope were stationary. In this condition the tunnel’s interior was covered with an irregular texture and was dimly lit. When the tunnel moved, a gradient of foreground optic flow was produced. As the texture elements in the tunnel were located in front of the fixation point, those bands of the sinewave grating that moved in the same direction as the foreground texture should have appeared nearest the observer.

Real perspective transformation of the display through rotation proved to be the most effective means of disambiguating relative depth from the theoretically ambiguous velocity gradient. Here, observers’ judgements of the convexity or concavity of a target corrugation agreed with the simulated depth order for an average of 90% of the 30-second trial duration. Although simulated vertical perspective transformation biased the perceived depth arrangement of the grating (64% accuracy), performance did not reliably exceed that found in the ambiguous control condition. The addition of a
component of horizontal perspective (width change) did not improve the accuracy of observers’ judgements (63% accuracy in this condition). Foreground flow and non-visual information also reliably disambiguated the sign of depth, giving mean accuracy rates of 75% and 73%, respectively.

2.1.8. Relationship between motion parallax / velocity gradient displays and the KDE

The majority of studies outlined above have shown that depth order can be recovered from translatory motion perspective displays that provide a gradient of velocities with considerable accuracy. Dots that move with a faster velocity are generally perceived as closer to the observer. On the other hand, Braunstein (1962) has shown in the case of the KDE that horizontal motion perspective, which introduces velocity differences between the front and back surfaces of a rotating object and should allow recovery of signed depth, is a rather ineffective determinant of perceived depth order. What is the reason for these contradictory findings?

One possibility is that the processes underlying depth recovery with motion-parallax/ velocity-gradient and KDE stimuli are fundamentally different, and as such, the results obtained with the former stimuli cannot be generalised to the KDE. This argument can be answered by two observations.
First, motion parallax and the KDE are related in the sense that both patterns of optical motions (i.e., translation and rotation) occur, under natural viewing conditions, when an observer sways the head. Here, objects are displaced (on the projection surface) relative to the observer and to each other, in accordance with their distance from the observer. This is, of course, the ecological basis for motion parallax. As Wallach and O’Connell (1953) originally noted, the same head movement would also produce a change in the orientation of objects relative to the observer (in the direction opposite to head movement) as they are seen from different directions. An identical change in orientation would be produced if an object actually rotated while the observer was stationary. This is the ecological basis for the KDE.

In motion parallax displays the normal relationship between rotational and translatory motion transformations is deliberately severed to determine the effectiveness of the latter cue. Consequently, simulated surfaces do not change their projected orientation with observer movement as they should. A real surface would need to rotate with the observer in order to maintain a frontal aspect in the manner depicted. Thus, motion parallax displays implicitly specify a surface rotation. Generally this rotation is not perceived (as is the case outside the laboratory), but on occasion it is. This point will be discussed shortly. Wallach (1987) has argued that motion transformations
accompanying observer movement go unnoticed because the visual system compensates for motions that are correlated with the movement of the observer rather than environment. The accuracy of the compensation process is defined through the “range of immobility” (p. 12) which refers to the minimum amount of actual object motion required for this object movement to be perceived by a moving observer. Wallach, Stanton, and Becker (1974) have shown that the immobility range for rotation is rather large. By way of example, their results indicate that if a given observer movement produced a relative orientation change for an object equivalent to a 20-degree rotation, the object would have to be physically rotated by eight degrees (40% of observer-produced rotation) before it was seen as rotating.

The second observation is phenomenological in nature. Motion-parallax and velocity-gradient displays are often perceived as 3D surfaces rotating around a vertical axis, under some conditions, even though only translatory motion is present in these stimuli (Braunstein & Andersen, 1981; Braunstein, Andersen, Rouse, & Tittle, 1986; Ono, Rivest & Ono, 1986; Rogers & Collett, 1989). For example, apparent rotation was reported by 27 of 36 observers in Braunstein et al.’s (1986) study.
Rogers and Collett (1989) examined the illusory rotation of surfaces defined through motion-parallax displays in some detail. Observers viewed the display binocularly which provided binocular disparities (or at least an absence of disparity) that indicated the true flatness of the stimulus. The perceived depth of the simulated 3D surface (a random-dot square-wave grating) was underestimated and occasional depth reversals were reported. Apparent rotation of the surface was also reported, especially when the surface was depth-reversed.

They argued that illusory rotation represents a compromise solution by the visual system that resolves a discrepancy between the extent of perceived depth and the objective motion in the display. When the perceived depth of a translating surface is non-veridical (i.e., underestimated, overestimated, or reversed), there exists relative motion in the stimulus that can no longer be attributed to depth variations. However, these residual motions can be accounted for through a simultaneous rigid rotation of the surface as it translates. Rogers and Collett (1989) pointed out that approximately the same pattern of relative motions (as for the simulated translating surface) would also be produced by one of three conditions: a surface with less depth that rotated against the direction of head movement, a surface with more depth that rotated slightly with the observer, or a reversed surface that rotated markedly with the observer.
Through manipulating the disparity of surface corrugations, Rogers and Collett (1989) showed that the nature of illusory rotation was determined by perceived depth. Rotation in the direction of head movement was reported when the disparity-specified depth was more than that simulated through motion parallax, and when disparity specified the reversed depth organisation. The opposite direction of rotation was perceived when disparities specified depth that was less than indicated by motion parallax. Clearly, there is an intimate relationship between depth from translation (motion parallax) and from rotation (KDE).

2.2. Dynamic Occlusion

In a well-lit environment the occlusion of one surface by another provides unambiguous information about the ordering of surfaces in depth. Surfaces nearer the observer block the light reflected from the more distant surfaces that they cover. Consequently a surface that is partially obscured by another will appear to be further from the observer. This occlusion information is absent in a silhouette or shadow projection of a 3D object, such as a tree seen at dusk. Although the overlap between two branches of a distant silhouetted tree is visible, there is no indication of which particular branch
obscures the other as both are black. This is also the case in the KDE when object features are homogenous in appearance.

Gibson, Kaplan, Reynolds and Wheeler (1969) identified another form of occlusion—the progressive covering or uncovering of surface texture that occurs when either the observer or objects move. As a nearby surface passes in front of another surface there is accretion and deletion of visible texture belonging to the more distant surface. For example, when a near surface moves from right to left, the texture of the occluded surface will be revealed (or accreted) along the right edge of the near surface and will be hidden (deleted) along the left edge. Visible contour interruptions are not required with this kinetic occlusion. The effectiveness of kinetic occlusion in determining the relative depth of planar surfaces has been shown in several studies (Kaplan, 1969; Royden, Baker, & Allman, 1988; Yonas, Craton, & Thompson, 1987; Ono, Rogers, Ohmi, & Ono, 1988).

---

3 One reviewer pointed out that relative depth information may still be present in this scenario through subtle binocular image differences at the branch intersections (Malik, Anderson, & Charowhas, 1999; van Ee, Anderson, & Farid, 2001). Here the absence of features in one eye’s view (due to occlusion) produces vertical image differences from which the visual system recovers proper depth ordering. For the sake of illustration I will treat the silhouetted tree as being at sufficient distance that the binocular differences are undetectable.
An experiment by Braunstein, Andersen and Riefer (1982) found two variations of occlusion information that disambiguated the depth organisation of an ambiguous KD sphere. The surface of the sphere was covered by irregular pentagon elements and viewed under orthographic projection. Edge-occlusion displays effectively simulated an opaque sphere. Pentagons disappeared from view while they rotated around the rear face of the sphere and reappeared when they rounded the front face. With element-occlusion a transparent sphere was simulated but pentagons on its front surface visibly obscured those pentagons on the rear surface which overlapped as the result of element motion. Pentagon size was also varied in order to increase the instances of element occlusions.

Edge-occlusion was a strong indicator of relative depth. Observers accurately identified the correct rotation direction of the opaque sphere on more than 90% of the trials. The efficacy of the edge-occlusion cue may indicate a tendency for the visual system to interpret opaque objects as being convex which is the most frequent depth organisation of objects encountered in the natural environment. To interpret the opaque sphere as depth-reversed requires an unusual percept. Here the sphere appears to be a concave dish with a surface texture that becomes transparent when it rotates around the front face. Despite the unlikely nature of this concave interpretation, an opaque
sphere was not completely unambiguous being occasionally perceived as depth reversed (Nawrot and Blake, 1991b, have also noted depth inversion of an opaque sphere, but only with experienced observers).

With element-occlusion displays, accuracy increased from chance levels to exceed 80% as the depicted size of the pentagons was increased. It was noted that these element-occlusion stimuli included static occlusion information on single frames. This prompted Andersen and Braunstein (1983) to test whether element-occlusion is effective when kinetic rather than explicit pentagon contours were used. To do this, the pentagon outlines were not actually drawn but were defined by dots randomly scattered throughout a pentagon-shaped area. Dots were deleted from a rear face pentagon when it was occluded by a pentagon on the front face. In any static view there was no indication of whether a given dot belonged to front or back face pentagon. These purely kinetic element-occlusion displays were comparably effective in disambiguating rotation in depth.

2.3. **Binocular Disparity**

Since Helmholtz’s (1878/1925) observation that the depth afforded from motion parallax was “just as ... looking at a good stereoscopic view” (p. 231) many researchers have ascribed a special relationship
between stereopsis and the KDE (e.g., Grosso, Sandini, & Tistarelli, 1989; Nasrabadi, Clifford, & Lui, 1989; Nawrot & Blake, 1989, 1991a; Richards, 1987; Rogers & Graham, 1982). Aside from their phenomenological similarities, KDE and stereopsis are comparable in terms of their geometries. With stereopsis, depth is determined from two simultaneous views of an object from slightly different orientations. With the KDE, depth is determined from two successive views. The interchangeability of these geometries can be demonstrated through the construction of stereoscopic versions of KDE stimuli. When two successive frames of a KDE sequence, which involves a small angular rotation of the 3D object, are viewed dichoptically, stereoscopic depth results. Similarly, two stereoscopic half-images can be used in alternation as an animation sequence that produces kinetic depth.

When considered separately, both stereopsis and KDE can provide information about the 3D structure of an object, yet each contains an inherent depth ambiguity. In stereopsis there is a scaling problem in which the same object will produce different binocular disparities when it is viewed at different distances. Unless the fixation distance is known, the object will appear distorted in depth. As elaborated above, the KDE suffers from a reflection ambiguity where two alternative depth organisations are equally likely under orthographic projection.
Richards (1987) has shown that combining stereopsis and kinetic depth will, in theory, resolve both these depth ambiguities and provide the veridical interpretation of 3D structure. On the basis of this theory, the addition of stereoscopic cues to the KDE should be particularly effective in disambiguating depth order.

Several studies (Dosher, Sperling, & Wurst, 1986; Braunstein, Andersen, Rouse, & Tittle, 1986; Nawrot & Blake, 1991b) have confirmed that stereopsis strongly biases the perceived depth organisation of the KDE. Dosher et al. (1986) found that direction of rotation judgements were determined by stereopsis when a maximum of 3 min arc of retinal disparity was added to a dynamic, perspective Necker cube. It could be argued that these experiments do not necessarily show that stereopsis disambiguates KDE. As was the case with contour-based element-occlusion, there exists static depth-order information on single frames of the KDE sequence. Observers may base their judgements on stereopsis alone and simply disregard kinetic depth.

This possibility may be rejected on the basis of two findings. First, some stereo-deficient individuals who fail on static measures of stereopsis, such as the RANDOT Circles test, nonetheless make effective use of dynamic disparity to disambiguate the rotation direction of a transparent KDE stimuli (Rouse, Tittle, & Braunstein,
1989; Braunstein et al., 1986). Second, static views of transparent random-dot objects are very difficult to fuse stereoscopically, involving a considerable latency before depth emerges (Akerstrom & Todd, 1988; Anderson, 1992). Pong, Kenner, and Otis (1990) reported that their stereo-normal observers could not see the 3D structure of a stereoscopic transparent random-dot cylinder until it was rotated. The direction of cylinder rotation was correctly identified on every trial. Thus, kinetic depth promoted stereopsis which in turn disambiguated the depth organisation of the KDE.

2.4. Luminance Contrast

Schwartz and Sperling (1983) investigated another source of potential depth-order information for the KDE. They varied the line luminance of a polar-projection rotating Necker cube such that lines on one face were brighter than those on the other face. Observers perceived bright lines as near and dim lines as far. They called this cue proximity luminance covariance (PLC). The depth organisation specified by PLC was perceived on more than 90% of all trials even when conflicting perspective required the perception of a deforming truncated-pyramid structure rather than a rigid cube. A continuous version of PLC was subjected to a fine-grained analysis by Dosher,
Sperling, and Wurst (1986) who found that it often remained an effective indicator of depth order despite conflicting stereoscopic cues.
Chapter 3: Induction Effects in the KDE: Spatial Context

As described in the previous chapter, adding depth cues to a KDE stimulus generally biases perceived depth organisation towards the cue-specified depth. Although this methodology has yielded valuable data concerning the nature and efficacy of dynamic depth cues, it has not directly addressed the issue of perceptual organisation in truly ambiguous stimuli. Cue-augmented KDE stimuli are, of course, no longer objectively ambiguous as explicit depth-order information is necessarily introduced into the stimulus. The degree to which these stimuli are perceptually disambiguated, therefore, provides an index of the effectiveness of the cue under consideration. But the question remains as to whether perceived depth organisation in the KDE can be biased without actually changing the test stimulus. This possibility may be realised through manipulating the spatial context within which an ambiguous KDE stimulus is presented.

3.1 Spatial Context

Spatial context refers to the presence of other objects or surfaces in regions of the visual scene surrounding the test object. An important
issue in perceptual research is raised when multiple objects (or surfaces) are present in a stimulus display: Is the appearance of a given object determined by its absolute proximal image characteristics (such as luminance, binocular disparity, retinal motion) alone, or in relation to its surroundings? Are KDE objects processed independently by the visual system, or do the characteristics of one object influence the perceived characteristics of another?

Considerable evidence exists to show that the nature of an object’s surroundings exerts a strong influence on how that object is perceived. Simultaneous brightness contrast, Belzold spreading, stereoscopic depth contrast, Mach bands, tilt illusion, Gelb effect, depth repulsion and attraction, border contrast, and induced motion are all textbook examples of spatial context effects.

In fact, the perceived character of an object may differ markedly when alone from when it is presented with other objects. The Land phenomenon (Land, 1959) provides a dramatic example. In this demonstration of colour constancy, a test patch is embedded in a larger Mondrian patchwork of different coloured rectangles. The test patch reflects predominantly medium wavelengths under white light illumination and consequently appears green. With an appropriate change in the spectral characteristics of the light source, the test patch is illuminated in such a way that it now reflects mostly long-
wavelengths. Under these lighting conditions, observers report that the patch still appears greenish (despite the fact that it reflects wavelengths that should make it appear red) provided it is surrounded by other patches similarly illuminated. If, however the surround is masked, thereby removing spatial context, the perceived colour of the patch is altered dramatically and now appears red.

3.2 Perceptual Consequences of Spatial Context: Contrast and Assimilation

Spatial context effects tend to manifest in one of two ways: spatial interactions may either exaggerate or attenuate the apparent differences between a stimulus and its surroundings along some perceptual dimension (such as colour, depth, or brightness). These modes of spatial induction are commonly referred to as contrast and assimilation (or capture) respectively. Spatial context effects are well-known in the domain of 2D motion perception that underlies the KDE. Any spatial interaction which influences the measurement of 2D motions may in turn influence the perceived organisation of the 3D structure recovered from such potentially misperceived motions.

3.2.1 2D Motion Assimilation

Spatial interactions are necessary for 2D motion computation because purely local motion measurements are generally ambiguous—
this is the well-known aperture problem (Marr & Ullman, 1981). Computational analyses have shown that a solution to the aperture problem can be obtained by integrating motion signals across space in order to constrain an interpretation of local motion signals that leads to globally coherent motion (Hildreth, 1984; Marshall, 1991; Grossberg & Rudd, 1992). In the case of the barberpole illusion (Wallach, 1948), unambiguous motion signals along the boundaries of an object strongly influence the direction of perceived motion within that object. The same translating diagonal grating appears to move upwards when viewed through a vertically elongated window and sideways when viewed through a horizontally elongated window. Here the shape of the window determines the nature of unambiguous motion in the display and consequently the spatial context in which local motion signals are interpreted.

Chang and Julesz (1984) provided further evidence to show that unambiguous motion in one region influences motion in another region through non-local interactions. They used a random-dot kinematogram that depicted alternating bands of unambiguous motion in one direction, and bands of ambiguous motion in which each dot was partnered with a potential match to both the left and right on successive frames (cf. Ullman’s, 1979, motion competition paradigm). Perceived motion in the ambiguous bands was strongly biased to
move in the same direction as the surrounding unambiguous bands, such that the whole display had a globally coherent motion.

A related spatial context effect is the phenomenon of motion capture. In a dynamic noise display, randomly moving dots that are surrounded by a moving inducing stimulus appear to adhere to it and drift coherently in the same direction. Motion capture has been demonstrated using moving outline shapes (MacKay, 1961), low spatial-frequency gratings (Ramachandran, Inada, & Kiama, 1986; Ramachandran & Cavanagh, 1987), and observerive surfaces (Ramachandran, 1985).

Ramachandran and Anstis (1983a, 1983b, 1985) demonstrated another aspect of motion assimilation using displays in which motion was ambiguous across the entire visual field. The basic stimulus was a bistable apparent motion sequence in which a pair of dots on diagonally opposite corners of an imaginary square were alternated with a pair of dots on the remaining two corners. When a collection of these figures were presented simultaneously, there was a global capture in which individual figures appeared to oscillate together as a group along the same motion-axis (either horizontal or vertical), and occasional direction reversals were synchronised across the field (cf. Attneave, 1971). The authors concluded from these results that:
“Whenever the brain applies a rule to resolve ambiguities in a given local region, there is a strong tendency to apply the same rule throughout the visual field and the system will go to tremendous lengths to achieve this” (p. 531).

3.2.2 2D Motion Contrast

Simultaneous motion contrast was first demonstrated by Loomis and Nakayama (1973) in a display which two large dots drifted at the same speed over a background velocity gradient of moving dots translating in the same direction. The perceived speed of each target dot was affected by its immediate surroundings such that the dot that drifted across the slower portion of the velocity gradient appeared to move substantially faster than the dot surrounded by faster moving texture.

Tynan and Sekuler (1975) found that simultaneous motion contrast, observed with a moving random-dot field embedded in a translating background, was strongly influenced by the direction of motion in the centre relative to that in the surround. When the centre and surround moved in opposite directions, dots in the central region appeared to move faster than their real velocity, and slower when the centre and surround moved in the same direction. Induced motion (Reinhardt-Rutland, 1988) was perceived when only the surround moved, where
stationary dots in the centre appeared to drift slowly in the opposite direction.

Several other motion contrast phenomenon inspired by luminance-based illusions, have been reported, including velocity analogues of: border contrast (Walker & Powell, 1974), the Craik-O’Brien illusion (Brigner, 1987), and White’s effect (Bressan, 1991).

3.3 Relationship between Assimilation and Contrast

Given the extensive literature on the effects of spatial context on perception, surprisingly little empirical attention has been paid to determining the relationship between assimilation and contrast. Although these seem to be distinct phenomenological effects, are assimilation and contrast two aspects of the same perceptual process, or separate processes? What are the stimulus parameters that produce one effect rather than the other?

Nawrot and Sekuler (1990) conducted one of the few studies in which this important issue has been addressed. They were able to induce opposing effects of spatial context, in a stimulus display similar to that used by Chang and Julesz (1984), by altering the widths of alternating bands containing either dynamic random noise (probe bands) or coherent motion in one direction (biasing bands). Motion
assimilation was observed when the bands were narrow whereas motion contrast resulted when the bands were wide.

To explain these data the authors proposed a modified version of Williams, Phillips, and Sekuler’s (1986) model of global motion processing. In this revised cooperative network model, the nature of spatial interactions, whether inhibitory or facilitatory, between local motion detectors depends on their directional-tuning and their proximity. Within a certain spatial extent, units tuned to the same direction facilitate one another while units with opposite direction tunings are antagonistic. Beyond this spatial limit, detectors responsive to the same direction of motion inhibit each other. In other words, Nawrot and Sekuler (1990) suggest that the “facilitatory and inhibitory influences of the network extend over different distances” (p. 1439).

They observed hysteresis in which the proportion of signal dots in probe bands required to null the direction of motion was much higher for ascending trials (where initially probe bands are pure noise and the perceived motion changes from capture to contrast) than for descending trials (where motion contrast gives way to capture as signal is reduced). Hysteresis is considered a hallmark of cooperative processing in a network with inhibitory and facilitatory interconnections (Williams et al., 1986).
3.4 Spatial Context Effects in the KDE

Relatively few investigators have examined the potential influence of spatial interactions on the recovery of 3D structure. On the whole, SFM models do not incorporate spatial interactions between objects as a factor in the determination of depth organisation. Instead models treat the total flow field as if produced by a single object. One exception is the model proposed by Hildreth et al. (1995) in which analysis in regions of superimposed opposite motions (typically produced by overlapping surfaces at different depths) is used to segregate objects prior to independent surface reconstruction.

Further empirical and theoretical attention to this issue is important for at least four reasons. First, strong spatial context effects pervade other perceptual domains and are often explicitly represented in theoretical accounts (e.g., ratio models of brightness and colour perception). Second, consideration of the geometry of KDE produced by movement of the observer or environment suggests that spatial context might provide useful constraints on recovered depth organisation. For example, in the optic array all objects rotate in the same direction when the observer translates laterally and so, under these conditions, the visual system may choose a depth organisation for each object that preserves a common rotation direction throughout
the scene. Third, spatial interactions modulate the activity of motion-selective neurons in areas that provide the likely neural substrate for the KDE (see section 1.6.2.2). Fourth, in the limited number of studies that have simultaneously presented multiple KDE objects substantial spatial context effects have been found. This empirical work can be divided along three aspects of the KDE phenomenon that spatial context has been shown to influence, namely: the recovered 3D structure, depth reversal, and depth organisation (direction of rotation).

3.5 Effect of Spatial Context on Perceived KDE Shape:

Ramachandran et al. (1988) demonstrated that the perceived structure of a KDE cylinder was altered when it was superimposed on another rotating random-dot cylinder. When the two cylinders were the same size but one cylinder rotated at half the speed of the other, the physically slower cylinder appeared to be flattened in depth and to lie inside the faster cylinder, and both cylinders were perceived to rotate at the same speed. A halving of the diameter of the faster cylinder produced the opposite effect. It was perceived as distended in depth, bulging at its centre to touch the outer surface of the larger, slower cylinder. As before, both cylinders apparently rotated at the same speed.
These two demonstrations show that spatial context, in this case the presence of another cylinder, can lead to a depth interpretation of the 2D velocity field that introduces substantial distortion in the recovered 3D shape. Here the visual system appears to apply a constraint in which all objects in the scene are perceived to rotate with the same angular velocity. This would be true of two stationary objects that shared a common axis (one inside the other) when viewed by a moving observer thereby producing the same change in orientation (or rotation) in both objects from the observer’s perspective. When the cylinders actually rotate at different speeds the 2D velocity field is not consistent with such an assumption and so the velocities are reinterpreted as being due to changes in the object’s 3D structure where features are treated as being closer (and therefore faster moving) than they actually are. In other words there is a trade-off between perceived rotation speed and perceived depth. Maintaining the assumption of a uniform rotation speed is so strong that non-rigid shape deformation occurs to explain the objective difference in angular velocity. When only a single cylinder of the pair was presented, its shape and rotation speed were perceived veridically. This suggests that the application of an angular velocity constraint was a direct consequence of the spatial context provided by multiple objects. According to most SFM models, neither rotation speed nor
the presence of other objects (assuming appropriate segregation) should influence perceived 3D shape.

3.6 Effect of Spatial Context on KDE Reversals

Howard (1961) reported that pairs of rotating wire-frame objects (a cube within a cylinder in one experiment, two adjacent skeletal cubes in another) tended to reverse together. This observation is consistent with demonstrations of synchronised reversals in other stimulus displays (e.g., bistable apparent motion: Ramachandran & Anstis, 1983b, 1985; adjacent Necker cube figures: Adams & Haire, 1959; pointing triangles: Attneave, 1971; Hemenway & Palmer, 1978; Palmer, 1983; shape from shading: Ramachandran et al., 1988). Further evidence for synchronised reversals in the KDE can be derived from the finding by Gillam (1976) that whenever multiple depth-ambiguous lines appeared to rotate in the same direction they most often appeared to form a planar surface. If the lines had reversed at different times then there would have been a rotational phase lag between the lines and the surface would have appeared jagged in depth.

Depth reversals between ambiguous KDE objects are usually not completely synchronised. If this were the case then KDE stimuli would always appear to have the same (or opposite) organisation as
one another throughout the observation period. Long and Toppino (1981) have argued that because rotating Necker cubes occasionally reverse at different times and rotate in opposite directions, there must be some independence between multiple ambiguous stimuli.

Although spatial context has some influence on the relative timing of depth reversals between ambiguous objects, it does not appear to affect their propensity for reversal. Long and Toppino (1981) showed that the reversal rate for a single rotating cube did not change when other cubes were present. Furthermore, Howard (1961) found that the average time before a rotating cube first reversed (the satiation period) was unaffected by the presence of other rotating objects.

Context appears to produce a general tendency for depth reversals to occur in different objects at the same time (a uniform change in organisation applied over the scene) but does not increase or decrease the perceptual stability of their organisation. Since the occurrence of reversal was monitored in both experiments cited above, but not the momentarily perceived direction of rotation, it is not possible to determine, from these results, how context affected the perceived depth organisation of one object relative to another—a pattern of rotational linkage or rotational contrast could equally lead to the same frequency and timing of reversals.
3.7 Effect of Context on KDE Organisation – 3D Rotational Linkage

3.7.1 Gillam et al: Fragmentation and Perceptual Grouping.

Gillam and her colleagues were the first to investigate systematically the determinants of perceived depth organisation of multiple ambiguous KDE stimuli. Their basic paradigm involved presenting oblique lines under parallel projection that rotated around a common axis. Direction of rotation was ambiguous. Observers were required to indicate continuously whether these lines appeared to rotate in opposite directions thereby providing an index for the degree of fragmentation within the configuration. If the rotation direction for each line was resolved independently, indicating no effect of spatial context on perceived depth organisation, then the lines would be expected to rotate in the same or opposite directions with equal probability. Thus, fragmentation periods less (or greater) than 50% of the trial duration would constitute evidence for rotational linkage (or contrast). In fact, with these stimuli, rotational contrast never occurred; in any experiment where non-independence was found, the perceptual consequence of spatial context was always rotational linkage.

The utility of this approach is that it provided a common framework within which the traditional Gestalt principles of perceptual
organisation (such as proximity, similarity, and closure) could be varied parametrically and their influence on grouping objectively measured (through the fragmentation index).

3.7.1.1 Similarity

Rotational linkage was found to be strongly influenced by the relative orientation of two rotating lines (Gillam, 1972). The tendency for the lines to rotate together decreased as a monotonic function of the difference in orientation between them. Parallel oblique lines almost always appeared to be linked (fragmentation period of about 6%) whereas convergent lines, with an orientation difference of 60 degrees, tended to rotate independently (fragmentation period of about 45%). Later, Gillam and McGrath (1979) showed that linkage was affected not only by the orientation of lines within a configuration but also by the retinal orientation of that configuration. Two configurations related by a 90 degree rotation in the picture plane produced different degrees of rotational linkage. Vertically-converging lines that rotated around a horizontal axis were found to fragment considerably less often (9%) than an equivalent configuration of horizontally-converging lines that rotated around a vertical axis (40%). The authors attributed this anisotropy to the greater salience of vertical-axis symmetry compared to horizontal-axis symmetry that has been reported with other stimuli (Julesz, 1971; Rock, 1973).
3.7.1.2 Proximity

Proximity was also found to be an important determinant of rotational linkage. Fragmentation time decreased monotonically as the vertical separation between two rotating lines was reduced (Gillam, 1972). At the smallest separation (30 min visual angle) convergent lines were rarely perceived to rotate in opposite directions, whereas the same lines rotated independently when separated by a one-degree gap. Gillam (1975) reported a related effect of proximity in an investigation of closure. Here rotational linkage between the top and bottom lines of a rotating trapezoidal frame was found to be inversely related to the size of a gap introduced into the longer vertical edge of the frame.

Gillam (1981) found that it is the relative line separation rather than the retinal separation that is the primary determinant of rotational linkage. Fragmentation time decreased as the lengths of two parallel, oblique rotating lines, with a constant separation, were increased—a manipulation that reduced the ratio of line separation to configuration size but did not change the retinal or distal separation between the lines. Furthermore, equivalent levels of linkage were found when the same configuration was viewed from different distances (a change in retinal separation), and with full-sized and half-sized versions of the same configuration (a change in both distal and retinal separation).
viewed from the same distance. Additional evidence for an effect of relative rather than absolute separation has also been obtained using different stimuli (Gillam, 1975; Gillam & Grant, 1984). Taken together these results suggest that separation constancy is achieved in which grouping strength is determined by ratios rather than taking distance into account to scale retinal separations. This is consistent with the adjacency principle (Gogel, 1954, 1974).

3.7.1.3 Configuration Effects: Collinearity, Numerosity, Closure and Framing

Considerable evidence exists to show that the tendency for rotational linkage often depends more on global aspects of the stimulus configuration than on basic stimulus properties such as orientation similarity and element proximity. In other words, the effects of relative orientation and proximity are determined by the relationships within a configuration rather than strictly by their absolute characteristics (determined with reference to the overall organisation of the configuration). Three studies (type of gap, multiple lines, and frame) demonstrate how the effects of relative orientation and proximity are determined by the nature of the configuration.

First, Gillam and Grant (1984) showed that the nature, rather than the size, of the gap separating elements was the more important determinant of linkage. Strong linkage between adjacent oblique
rotating lines (with the same orientation) was reported over large lateral separations that maintained the collinearity of the lines across the gap. There was a substantial increase in fragmentation when the same line segments were instead separated, along the vertical rotation axis, even with gap sizes that were smaller than those used with collinear lines. This effect was not strictly due to collinearity as strong linkage was also found between lines with a small but salient non-collinearity. Instead, as the results of Gillam (1992) demonstrate, the critical factor is more likely to be the degree of continuity (or relatability, cf. Kellman & Shipley, 1991) maintained between elements over a gap. Gillam (1992) found quite different levels of fragmentation with slightly different configurations of two oblique lines in which one line was separated from the other along a diagonal offset. Although both configurations had the same degree of separation and non-collinearity between the lines, the configuration with an offset in the direction opposite to the line orientation (a discontinuity in the direction of orientation change) produced fragmentation times that were more than twice those obtained for the configuration with an offset in the direction of line orientation.

Second, Gillam (1976) found that the tendency for rotational linkage increased with the number of lines present in a configuration, despite an increased potential for fragmentation given that any one of
the lines could be perceived as rotating independently. Linkage was strongest when the lines converged to a common vanishing point (i.e., the lines had the same perspective), particularly when the line endpoints were collinear, thereby providing implicit vertical edges to the configuration. Control experiments (Gillam, 1976, 1992) revealed that the effect of line numerosity was largely independent of a reduction in interline separation and orientation. The enhanced grouping effect appeared primarily due to the interpretation of the stimulus as a rotating planar surface promoted through the addition of lines.

Third, when convergent lines were enclosed by a rotating trapezoidal frame they became strongly linked (fragmentation period of 5%) to each other and to the motion of the frame (Gillam & Broughton, 1991). Capture by the frame was most effective when it shared the same perspective as the lines. The spatial context provided by the frame produced an interesting emergent property in that the lines appeared to oscillate with the trapezoidal frame rather than to rotate. Several further findings established that the effect of the frame depended on more global aspects of the configuration than on pairwise (local) linkages between the convergent lines and nearest contours of the frame. The reduction in fragmentation that was observed when the lines were enclosed by separate components of the frame (either the
vertical or horizontal contours alone) was insufficient to account for the strong linkage found with the complete frame. Moreover, a rotating wedge-shaped outline inserted between the convergent lines produced little reduction in fragmentation, even though it had the same local orientation and separation relationships to the lines as did the trapezoidal frame.

Gillam et al. interpreted their results as demonstrating “perceptual economy” whereby the visual system tends to group ambiguous stimulus features and treat them as belonging to a single object. In this way the visual system is able to apply the same perceptual solution in instances of organisational uncertainty. Although the application of a default organisation rule (and the resulting rotational linkage) might be expected when all objects are ambiguous due to parallel projection, the following studies demonstrate that rotational linkage is the norm even when explicit (and conflicting) depth cue information is present in the stimuli.

3.7.2 Landy et al.

Landy, Cohen, and Sperling (1984) presented observers with two counter-rotating Necker cubes, one above the other, which were observed under extreme polar projection (producing high levels of motion perspective). The two cubes most frequently appeared rotationally linked (or else rotated independently) even though this
required that one of the cubes remained perceptually depth reversed, overriding its perspective which signalled the opposite direction of rotation (and depth organisation). Perspective was not simply ignored by the visual system but rather reinterpreted such that the reversed cube appeared to change shape into a truncated pyramid, which grossly deformed as it rotated.

3.7.3 Eby et al.

Eby, Loomis, and Solomon (1989) have also shown that rotational linkage occurs even when unambiguous depth-order information signals that two objects are actually rotating in opposite directions. In their first experiment they tested the effects of rotation axis collinearity and perspective on perceptual linkage between two KDE spheres. One sphere was used as an inducing stimulus and its direction of rotation was disambiguated by a strong perspective cue (a projection ratio of 2.64 was used). The target sphere either rotated in the same, or opposite direction to that of the inducing sphere. Three levels of perspective were used for the target sphere thereby increasing the strength of information regarding the congruency of rotation (i.e., the direction of target sphere rotation relative to the inducing sphere). Axis collinearity was manipulated through presenting both spheres in either concentric or side-by-side configurations—a procedure that also varied the spatial separation
between the spheres. Observers accurately identified the target sphere’s direction of rotation when it was the same as the inducing sphere. When the two spheres counter-rotated, the target sphere’s rotation was often misperceived to be in the direction of the inducing sphere. This bias was greatest when the spheres shared the same axis of rotation, and when the target sphere’s perspective was reduced (making its depth organisation more ambiguous).

Spheres were replaced by two counter-rotating planar random-dot sheets that were stacked one above the other, in two further experiments. Two new factors, slant in depth and relative starting phase, were found to influence the degree of rotational linkage between the planes. The planes were slanted symmetrically, but in opposite directions, along the horizontal midline separating them. As the planes were tilted further away from the observer, the tendency for illusory co-rotation increased. The relative orientation of the two planes was also varied by introducing a starting phase difference between them of 0, 45, or 90 degrees. Rotational linkage was generally stronger when the difference in the orientation of the two planes was small. Changing the planes’ relative speeds or the relative alignment of their rotation axes (while maintaining a constant spatial separation) had no effect on the tendency to perceive them as linked.
As before, rotational linkage was greatest when the simulated perspective (signalling counter-rotation) was weak.

### 3.7.4 Rogers and Rogers

Rogers and Rogers (1992) observed the motion parallax analogue to rotational linkage (see Section 2.2.7). In their study observers viewed an ambiguous random-dot grating with sinusoidal depth corrugations. In the condition of most interest, the target stimulus was presented behind a textured ground plane that translated, with the movement of the display providing a gradient of translational surround motion. When the foreground flow was absent, the grating appeared to alternate between convex and concave depth corrugations, whereas when the flow was present, the grating was seen as unambiguously convex and appeared to rotate in the same direction as the foreground motion.

### 3.7.5 Rotational linkage: Summary

The experiments outlined above have shown that the visual system often applies the same solution to multiple rotating figures under conditions of perceptual uncertainty. While this may imply that global processing is the operative mode of the visual system, Long and Toppino (1981; Toppino & Long, 1987) have argued that individual figures do maintain some level of perceptual independence. This is an
important point. If true then the phenomenon of perceptual linkage is more than a default tendency of the visual system to apply a single rule to multiple objects that are thereafter treated as a single perceptual unit (as implied by the decisional/learning theory of perceptual reversal). Instead perceptual linkage may be due to long-range cooperative interactions between local processes. In this vein Gillam and Grant (1984) have offered a useful distinction between “aggregation” and “unit formation” that may clarify the consequences of perceptual linkage. Aggregation occurs when ambiguous elements that share similar characteristics (e.g., same orientation, shape, and size) are treated as class to which a common motion or depth may be applied, whereas unit formation occurs when all ambiguous elements are treated as forming parts of a single object thereby losing their perceptual identity and independence.

3.8 Effect of Context on KDE Organisation - 3D Rotational Contrast

The rotational linkage observed with KD stimuli implies a high degree of cooperative processing between multiple rotating objects. Is this assimilation effect the only type of spatial interaction that occurs with stimuli? As described above, there is considerable evidence from
other perceptual domains to expect that rotational contrast might also occur.

3.8.1 Kinetic Depth Contrast

Blackburn and O’Shea (1993) demonstrated 3D rotational contrast in which an ambiguous transparent KDE cylinder sandwiched between two adjacent sheets of rightward (for instance) moving dots appeared to rotate unambiguously in the opposite direction (i.e., the front face of the cylinder was perceived as rotating to the left). A similar phenomenon was independently reported by Sereno and Sereno (1991). They embedded an ambiguous KDE sphere in a horizontally translating conveyer-belt of random-dot texture. Again the transparent sphere appeared to rotate in the direction opposite to the translating dots in the surround. This effect was strong enough to override conflicting depth cues (luminance contrast, or binocular disparity) specifying the sphere’s true direction of rotation.
4 Chapter 4: Perceptual Reversals in the KDE

4.1 Experiment 1: Effect of stimulus complexity, form, rotation speed and observation time on KDE reversals.

4.1.1 Experiment 1: Introduction

In embarking on an investigation of the factors that resolve, or bias, the depth organisation of ambiguous KDE stimuli, a necessary first step is to explore the basic phenomenon of depth reversals in some detail. The primary purpose of this experiment is to examine the influence of a selection of stimulus parameters on depth reversal rates with the goal of producing an optimal test stimulus that undergoes depth reversals with suitable frequency.

If a test KDE stimulus reverses too frequently, then any systematic relationship between stimulus manipulations and their effects on KDE ambiguity could be masked. Only the strongest manipulations are likely to overcome this inherent KDE instability. On the other hand, a highly stable KDE stimulus that rarely reverses is also undesirable as it leaves little opportunity for performance differences between experiment conditions (due to ceiling or floor effects).
On the basis of previous research with static and dynamic reversible stimuli, rotation speed and stimulus complexity were manipulated in the present experiment. Conflicting effects of each of these factors on reversal rates have been reported in the literature. Depth reversals of the rotating Ames window (producing illusory oscillations) occur less frequently as rotation speeds are increased over the range of 3rpm to 30rpm (Braden, 1978; Börjesson, 1971; Mulholland, 1956), while increasing the rotation speed of Lissajous figures leads to a corresponding increase in depth reversals (Philip & Fisichelli, 1945). Increased stimulus complexity, which is usually measured by the number of lines in a stimulus, reduces the reversibility of ambiguous line drawings (Washburn & Gillette, 1933; Masulli & Riani, 1989) but increases the frequency of depth reversals in Lissajous figures (Philip & Fisichelli, 1945). These contradictory results warrant further testing of rotation speed and complexity as possible determinants of KDE reversal frequency.

One potential obstacle to generalising previous research on perceptual reversals to the KDE lies with the nature of the stimuli previously used. Although the reversibility of several different types of rotating figures (Necker cubes: Long et al, 1983; Lissajous figures: Philip & Fisichelli, 1945; bent wire-frame objects: Caelli, 1979) has been investigated, in all cases the KDE shapes tested were wire-frame
figures rather than the random-dot type KDE stimuli that have been adopted almost exclusively since Green’s (1961) original use of them.

It is conceivable that perceptual reversals of rotating wire-frame objects (which undergo changes in the size and orientation of line segments in addition to changes in the positions of line endpoints) will differ in frequency from those of random-dot KDE stimuli. For this reason, the type of KDE stimulus (random-dot versus wire-frame) was included as a third factor that might influence depth reversibility.

An additional aim of the study was to examine the temporal characteristics of KDE reversals. A consistent finding from a wide range of static and dynamic ambiguous figures (e.g., Babich & Standing, 1981) is that perceptual reversals do not occur entirely at random but rather follow a predictable increase in frequency over the course of the observation period. The typical negatively accelerated function relating reversals to viewing time has generated considerable theoretical interest and has variously been interpreted as evidence of either a cycle of neural fatigue and recovery (Köhler & Wallach, 1944) or a classic learning curve (Ammons, 1954).

This experiment sets out to map the effect of stimulus factors (rotation speed, shape, complexity, observation time) that are likely to influence the perceptual stability of ambiguous KDE stimuli. In doing
so I intend to provide the groundwork for producing a suitably malleable family of baseline stimuli for later experiments.

### 4.1.2 Experiment 1: Method

#### 4.1.2.1 Participants

One female and three males, whose ages ranged from 21 to 35 years, volunteered for the experiment. Two observers were experienced psychophysical observers and aware of the aims of the experiment. The other two observers were naive. All observers had normal or corrected-to-normal acuity.

#### 4.1.2.2 Stimuli and Apparatus

A Macintosh IIcx computer was used to generate and display the stimuli. Animated sequences of the KDE stimuli were created offline and saved to the hard-drive. During the experiment the stimulus animations were retrieved from the hard-drive and displayed on a 13-inch Apple colour monitor.

The ambiguous KDE stimuli were constructed by initially assigning a number of locations to random positions on the surface of an upright, virtual cylinder. Random-dot cylinder shapes were produced by drawing a filled circle (4 arcmin in diameter) at each of the assigned locations. Wire-frame shapes were created by connecting a random sequence of these locations with lines (2 arcmin wide)
through the volume of the virtual cylinder. Stimuli in the high-complexity conditions contained either 128 dots or 64 lines; stimuli in the low-complexity conditions contained 64 dots or 32 lines. Stereograms depicting representative stimuli are presented in Figure 4.1. These stereograms were produced from adjacent frames of the animation sequences.

The animated displays simulated a parallel projection of the stimulus shape rotating through $360^\circ$ around its vertical axis at an angular speed of 1, 2, or 4 degrees per frame. Each frame in the animation sequence was displayed (during the experiment) for either 16.7 or 33.4 msecs giving frame rates of 60 and 30 frames-per-second (fps) respectively. Smooth KDE motion was perceived with either frame rate.

The displays were viewed monocularly in a darkened room from a distance of 110cm. At this viewing distance, each pixel measured 1 minute of arc and the stimulus subtended a visual angle of 2x2 degrees.
4.1.2.3 Design

A total of 24 stimulus conditions were produced by the factorial combination of four variables: angular rotation (1, 2, 4 degrees/frame),

**Figure 4.1.** Stereograms of examples of the experimental stimuli. These stereograms can be free-fused by crossing the eyes. The stimulus conditions depicted are: (a) low-complexity random-dot cylinder, (b) high-complexity random-dot cylinder, (c) low-complexity wireframe, and (d) high-complexity wireframe. Note that when the high-complexity wireframe is rotated as a KDE stimulus, individual lines become clearly visible — (this is not readily apparent from the stereogram version).
frame rate (30, 60 fps), shape (random-dot cylinder, wire-frame), and complexity (low, high). Each condition was replicated four times over the course of eight 45-minute sessions. The design was completed in two parts with the 12 low-complexity conditions tested in the first 4 sessions, and the 12 high-complexity conditions tested in the last four sessions. This confounding of complexity with testing order is addressed below. The conditions within each session were presented in a randomised order.

4.1.2.4 Procedure

Prior to the experiment proper, an ambiguous random-dot KDE sphere was demonstrated to the naive observers. Both observers described the appropriate three-dimensional structure and, with continued viewing, spontaneously reported a depth reversal (change in rotation direction).

Observers were required to track the apparent direction of KDE rotation during each experimental trial. They were instructed to hold down the left-arrow key while the KDE stimulus appeared to rotate clockwise (front dots move leftwards), and to hold down the right-arrow key whenever the stimulus rotated counter-clockwise (front dots move rightward). If the observer was unsure of the rotation direction they were required to release the response keys.
Each trial was initiated when the observer pressed then released both response keys. The animated KDE display was then shown for 180 seconds during which the observer tracked its perceived rotation. At the conclusion of the trial there was an intertrial interval of at least 30 seconds while the next animation loaded from the hard-drive.

4.1.3 Experiment 1: Results

The frequency of depth reversals was calculated for each observer in each condition. A depth reversal was scored when the observer changed response from one direction to the opposite direction. Occasions where the observer responded with one direction key, released the key (signifying rotational uncertainty), then responded again with the same direction key were not considered reversals.

Descriptive data for individual observers are summarised in Table 4.1. Substantial individual differences are clearly evident in the total number of reversals experienced. For example, observer RT recorded more than seven times the number of reversals recorded by observer ROS. Reports of similarly large individual differences are commonplace in the ambiguous figure literature (e.g., Brown, 1955; Frederiksen & Guilford, 1934; Wieland & Mefferd, 1967). As a consequence of these initial differences between observers, separate analyses were conducted on the data of individual observers wherever possible.
Reversal frequency data were subjected to a repeated-measures analysis of variance (ANOVA) with rotation angle (1, 2, 4), frame-rate (30, 60), shape (random-dot cylinder versus wire-frame), and complexity (low versus high) as independent factors.

Table 4.1 Summary of Individual Reversal Characteristics. Standard errors are given in parentheses

<table>
<thead>
<tr>
<th>Observer</th>
<th>Total Reversals</th>
<th>Mean Reversals per Trial</th>
<th>Mean Time between Reversals in secs</th>
<th>Mean Time before First Reversal in secs</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROS</td>
<td>548</td>
<td>5.71 (0.55)</td>
<td>24.16 (0.94)</td>
<td>39.89 (4.22)</td>
</tr>
<tr>
<td>CS</td>
<td>770</td>
<td>8.02 (1.43)</td>
<td>17.13 (0.74)</td>
<td>41.38 (5.19)</td>
</tr>
<tr>
<td>SB</td>
<td>1414</td>
<td>14.73 (3.32)</td>
<td>9.06 (0.36)</td>
<td>45.30 (5.95)</td>
</tr>
<tr>
<td>RT</td>
<td>4073</td>
<td>42.53 (3.38)</td>
<td>4.07 (0.09)</td>
<td>4.61 (0.59)</td>
</tr>
</tbody>
</table>

4.1.3.1 Rotation Speed

A significant increase in the frequency of reversals with larger inter-frame angles was apparent for all observers, $F(2,6) = 15.07, 43.49, 67.96, 21.42$ for observers CS, ROS, RT, SB respectively; all $ps < .01$. There was a related main effect of frame-rate where reversals were more frequent in the 60 fps than in the 30 fps conditions. This difference was significant for three observers, $F(1,3)=157.38, 52.79,$
53.07 for observers ROS, RT, SB; all \( p < .01 \). As frame-rate and inter-
frame angle combined to determine the actual rotation speed of the
stimulus it is necessary to establish whether these main effects are
independent from, or the consequence of, rotation speed.

The interaction of the two factors is plotted in Figure 4.2 as a
function of rotation speed. It is evident that there is no independent
effect of frame-rate because those combinations of frame-rate and
inter-frame angle that produce the same rotation speed (i.e., 1 degrees
at 60fps versus 2 degrees at 30fps; and 2 degrees at 60fps versus 4
degrees at 30fps) also produced comparable reversal rates. Pairwise
contrasts between these combinations revealed no significant
differences for any observer. For conceptual clarity, the frame-rate and
inter-frame angle factors were collapsed (by averaging data across
conditions that gave the same rotation speed) into a single rotation
speed factor (5, 10, 20, 40 rpm) for the remaining analyses. Trend
analysis showed that reversal frequency increased as a significantly
linear function of rotation speed for all observers.
The linear increase in reversal frequency with rotation speed may be the consequence of the overall velocity of the stimulus (or individual dots/lines) in itself, or due to the corresponding increase in temporal frequency that is associated with rotation speed. A temporal frequency account would require two assumptions. First, that observers track recognisable landmark clusters of lines or dots throughout the rotation.

**Figure 4.2** Results of Experiment 1. Individual data. Effect of rotation speed on reversal frequency. Note that the absolute values on the ordinate axes vary between individual observers. Error bars are standard errors.

4.1.3.2 **Temporal Frequency or Rotation Speed?**

The linear increase in reversal frequency with rotation speed may be the consequence of the overall velocity of the stimulus (or individual dots/lines) in itself, or due to the corresponding increase in temporal frequency that is associated with rotation speed. A temporal frequency account would require two assumptions. First, that observers track recognisable landmark clusters of lines or dots throughout the rotation.
of the stimulus. Second, that depth reversals are more likely when the configuration of stimulus features reaches a particular orientation. These orientations where a reversal is likely to occur are points during the rotation cycle where the tracked configuration undergoes a substantial change in its motion. Possible orientations could be at the left (0°) or right (180°) edges of the stimulus where the configuration slows to its minimum velocity then reverses its direction of motion, 4 or at the front (90°) or back (270°) of the stimulus when the configuration accelerates then decelerates around its maximum velocity. Informal observation suggests that reversals tend to occur at the latter orientations. In either case the expected phase difference between the orientation of the stimulus at the beginning of a given depth organisation and its orientation when a reversal occurs would be 180° and the stimulus would be perceived as oscillating. With faster rotation speeds a configuration would reach a reversible orientation with greater frequency thereby increasing the likelihood of depth reversals.

In order to test the temporal-frequency explanation, the difference in orientation (or phase) between successive reversals was calculated

4 Power and Day (1973) have found that reversals with rotating planar figures (e.g., rectangles, trapezoids, ellipses) almost always occur when the stimulus is either approaching or leaving the frontoparallel plane.
and pooled across conditions (and sessions) for each observer. Frequency distributions are plotted for individual observers in Figure 4.3. If the effect of rotation speed is due to temporal frequency then a unimodal distribution peaking around 180° is expected. Alternatively, if reversals are independent of stimulus orientation then a uniform distribution is expected representing a pure effect of stimulus velocity.
Examination of individual frequency distributions reveals striking observer differences. At one extreme, RT’s reversals were clustered around 180°—a pattern clearly supportive of a temporal frequency
account, whereas ROS displayed the opposite pattern. His reversals were distributed uniformly over the full range of phase differences.

Cubic polynomials were fitted to the data to establish whether the reversal phase frequencies were randomly or systematically distributed. This analysis confirmed the strong organisation present in RT’s data with a highly significant cubic fit, $F(3,14)=66.73$, $p<.0001$, accounting for 93.46% of the variance. Some degree of organisation to SB’s data was also evident with the cubic fit capturing 42.35% of the variance, $F(3,14)=3.43$, $p<.05$. The remaining two observers appear to have experienced reversals in a more stochastic fashion. Cubic fits to their data did not attain significance and accounted for considerably less variance (33.51%, 14.19% for CS and ROS respectively).

4.1.3.3 Other Factors

Some individual differences were apparent for the effects of stimulus complexity and shape. Observers CS, RT, and SB recorded significantly more reversals under conditions of high complexity, $F(1,3)=17.98$, 17.67, 35.51 respectively; all $ps<.05$, whereas for ROS reversals were slightly more frequent in low complexity conditions, but this difference was not reliable, $F(1,3)=2.48$, $p>.05$. A main effect of stimulus shape was obtained for observers RT and SB, $F(1,3)=12.85$, 23.38 respectively; $ps<.05$, but the direction of these differences were in opposite directions. Random-dot cylinders
reversed more frequently than did wire-frame shapes for SB (a trend weakly present in the data of CS and ROS) whereas wire-frame stimuli reversed most frequently for RT. The difference in reversal frequencies between the two stimulus shapes became more pronounced at higher rotation speeds resulting in a rotation speed by shape interaction for RT, ROS, and SB, $F(3,9)=6.06, 6.58, 5.78; ps<.05$. A similar interaction between rotation speed and complexity (with an opposite effect of complexity for each observer) occurred for RT and ROS, $F(3,9)=13.01, 6.47; ps<.05$.

Two additional interactions were present in the data of RT alone. The effects of stimulus complexity were restricted to wire-frame shapes, $F(1,3)=22.63, p<.05$, and there was a marginal interaction between all three factors, $F(3,9)=4.29, p<.05$.

4.1.3.4 Temporal distribution of reversals within trials

Following from previous work it was expected that the frequency of depth reversals would increase as a function of observation time within each 180-second trial. To test this, each trial was divided into six 30-second bins. Every reversal was allocated to one of these bins on the basis of the elapsed trial time at which it occurred. The frequency data from all 12 conditions in a session were pooled to provide sufficient observations for each time period sampled. A scaling procedure was then adopted to minimise differences between
sessions in the total number of reversals reported. For each of the eight sessions (which were treated as replications) the number of reversals in each bin was expressed as a proportion of the total number of reversals in that session. If tendency for depth reversal was unaffected by observation time then the proportion of reversals in all bins should be close to 0.166 (i.e., 1/6).

Individual one-way repeated-measures ANOVAs were carried out for each observer and the data are plotted in Figure 4.4. Data from the two naive observers (CS, RT) indicated no effect of observation time; reversal frequencies remained constant throughout the trial. A significant effect of observation time was found for the experienced observers, $F(5,35)=4.22$, 6.17 for ROS and SB respectively; $p<.01$, with both observers recording a monotonic rise in reversal frequencies over the first 2 minutes. This trend continued over the final minute of the observation period for SB but dropped during this time for ROS.
Figure 4.4 Individual data for Experiment 1. Effect of observation time on reversal frequency. For each observer the number of reversals within successive 30-second interval bins is plotted as a proportion of the total number of reversals reported within each 3-minute trial. These observations are averaged over all experimental conditions. SB and ROS are experienced observers whereas RT and CS are naïve observers.
4.1.3.5 **Session effects**

The final analysis tests for a general increase in reversal frequency over successive testing sessions as has been reported with Necker cube figures (Long et al., 1983). Reversal frequencies were pooled, for each observer, across all conditions in a session. Individual session totals were then divided by the total number of reversals experienced by that observer for the entire experiment. A one-way ANOVA on these scaled reversal frequencies produced a marginal session effect, $F(7,21)=2.17, p<.07$, which is plotted in Figure 4.5. Trend analysis revealed a significant linear component in the data, $F(1,21)=11.86, p<.01$. Examination of the graph suggests that the relatively large number of reversals reported in the first session is at odds with the overall trend of increasing reversal rates with successive sessions. When this session is removed from the analysis a stronger session effect is obtained, $F(6,18)=3.13, p<.05$.

4.1.3.6 **Session and Complexity Confound**

The complexity effects described above require careful interpretation due to the confounding of this factor with test order. Low complexity stimuli were tested in the first four sessions and the high complexity conditions were tested in the last four sessions. The main effect of complexity effectively tests for differences in reversal rates averaged over the first half of the experiment with those
averaged over the second half. Given the trend of increasing reversal frequencies with session identified above, a difference between the two halves of the experiment is to be expected. An independent effect of complexity would be identified by a substantial increase in reversals between sessions four and five (the point at which there is a change from low to high density conditions) but a pairwise contrast between these two sessions failed to reveal any significant difference ($F < 1$). Instead the size of the difference is consistent with the general session trend. Thus it would seem appropriate to consider the complexity effect as another manifestation of the session effect.
The primary aim of this experiment was to establish effective parameters for producing ambiguous test stimuli in future experiments. To this end, one factor – rotation speed, was identified as being an important determinant of reversal frequency. All observers

![Figure 4.5](image-url)

**Figure 4.5.** Averaged group data from Experiment 1 showing a general increase in reversal frequency over successive testing sessions. The relative reversal frequencies, plotted here, are the raw total frequencies for a given session expressed as a proportion of the total reversals over all sessions for that observer. These individual reversal proportions so computed were then pooled over observers to give the group data plotted here. Note that low-complexity conditions were tested in sessions 1 to 4, while high-complexity conditions were tested in sessions 5 to 8.

### 4.1.4 Experiment 1: Discussion

The primary aim of this experiment was to establish effective parameters for producing ambiguous test stimuli in future experiments. To this end, one factor – rotation speed, was identified as being an important determinant of reversal frequency. All observers
experienced a strong linear increase in reversals with rotation speed. Between 78% (ROS) and 34% (SB) of the variance in individual reversal rates could be attributed to the rotation speed effect. This finding replicates (and extends to a new range of stimuli) previous reports of increased reversibility with rapidly rotating ambiguous figures (Caelli, 1979; Mulholland, 1956; Philip & Fisichelli, 1945).

Four possible interpretations of the rotation speed effect can be advanced. First, the temporal frequency account (described more fully in Section 4.1.3.2) attributes the rise in reversal rate to an increase in the frequency with which distinctive stimulus features rotate through particular orientations (180° apart) where depth reversal is most likely to occur. This account assumes that observers continuously track a particular cluster of dots or lines, which forms a recognisable configuration, and that reversals occur whenever the configuration undergoes a maximal change in its motion (e.g., at the front or back of the stimulus where the configuration accelerates then decelerates). Reversals, therefore, occur predominantly in the form of regular oscillations. Changes in temporal frequency with rotation speed appear to explain the performance of one observer, RT, who exhibited a strong tendency to experience oscillations. However there was little support for the temporal frequency hypothesis in the data of the remaining three observers. This implies that, at least for some
observers, there was an independent contribution of rotation speed per se to reversal tendency.

Second, the temporal frequency explanation can be modified by introducing the concept of slippage. Slippage is the assumption that immediately following a reversal, the observer loses track of the original configuration and then starts following a different configuration at some later time. When the new configuration rotates through 180° the KDE stimulus is most likely to undergo another depth reversal. The degree of slippage, or the difference in stimulus orientation between a depth reversal and tracking of a new configuration, essentially adds noise to the reversal phase distribution used in the preceding analysis. If the degree of slippage was large and normally distributed then the reversal phase frequencies would appear random and the observer would report few oscillations. As the stimulus is assumed to reverse whenever a tracked configuration rotates through some multiple of 180° (the reversible orientation assumption), the average reversal frequency remains determined by temporal frequency.

An in-depth investigation of the temporal frequency hypothesis lies beyond the aims of this thesis but the following approach is suggested. A critical assumption for the hypothesis is that observers track element configurations over time. Consequently the temporal frequency
account would predict an increase in reversal frequency with rotation speed with stimuli that promoted configuration tracking but not with stimuli eliminating the possibility of tracking. Limited point-lifetime KDE cylinders, in which individual dots are assigned to new locations on the cylinder after rotating through a random angle, provide ideal stimuli for the latter condition. With these stimuli, dot patterns are temporally uncorrelated and do not provide a stable configuration that could be tracked over time. A normal random-dot cylinder could be used in the former condition, with observers being instructed to track an explicitly identified (e.g., a group of adjacent dots highlighted with a different colour) configuration throughout the trial.

Third, the rotation speed effect may be the consequence of a breakdown in the recovery of 3D structure-from-motion (SFM) process at higher speeds. Caelli (1979) found that as the rotation speed of a simple wire-frame object was increased, there was an interesting change in the type of motion perceived. At speeds below 60rpm, observers predominantly reported 3D rotation; at speeds between 60 and 120 rpm there was a rapid transition to perceived oscillations. Above 120rpm the impression of 3D structure collapsed and instead sinusoidal planar motion was reported. Caelli (1979) considered these changes in perceived motion as evidence for the sequential loss of torsion (leading to oscillations) and curvature (leading to planar
motion) information in the 3D-structure recovery process. A similar sequence of perceived motions, with a rotating rectangular figure, has also been described by Mulholland (1956).

Although neither of these studies reported the actual number of depth reversals experienced with different rotation speeds, it is most likely that reversals increased with speed. An increase in the tendency for oscillations with rotation speed would necessarily increase the frequency of depth reversals. However, the effects of a KDE collapse to 2D motion are less certain. Temporary loss of 3D structure at regular intervals could increase the reversal rate because when the stimulus reorganises into depth there is a 50% probability it will assume the opposite depth organisation.

A breakdown in SFM cannot completely account for the results of the present experiment. Frequent oscillations were only evident in RT’s data, and the maximum rotation speed used (40rpm) was well below the minimum speed at which Caelli (1979) found notable oscillation frequencies. In fact, Caelli’s observers reported 3D rotation on almost every trial when a 30rpm speed was tested.

Fourth, the satiation theory of perceptual reversals offers an explanation of the rotation speed result. In this theory, reversals are attributed to a cycle of fatigue (adaptation) and recovery of neural populations underlying opposing depth interpretations of the rotating
stimulus. An increase in reversal frequency with rotation speed would be expected if the fatigue (or recovery) process occurs more rapidly at higher velocities. Fisichelli (1947) offered this explanation of their finding of a rotation speed effect with Lissajous figures. They hypothesised that neural units were fatigued each time a line of the figure passed through their receptive fields. With higher rotation speeds this occurs more frequently.

In contrast to the clear effect of rotation speed in the present experiment, the effects of stimulus shape and complexity on reversal frequency were less consistent. High complexity stimuli were generally subject to more reversals than were low-complexity stimuli, but this result was almost certainly a session effect. In any case it is unlikely that more reversals occur with low-complexity stimuli. The shape of the stimulus did not appear to be a strong determinant of reversal rates. Only two observers experienced a significant difference in reversals between random-dot cylinder and wire-frame stimuli, and these differences were in opposite directions.

On the basis of the general perceptual reversals literature it was expected that KDE reversals would increase in frequency with observation time. Evidence for such a trend was apparent for only two observers in the present experiment, and no overall trend was observed when all observers’ data were combined.
This discrepancy with previous studies lends support to the claim by Girgus et al. (1977) and Sadler and Mefferd (1970) that the widely cited effect of observation time is little more than an artefact of the averaging process with group data. When these authors examined individual data in their studies they observed that few observers experienced a change in reversal frequency over time. Because individual observers were found to begin reversing at different times throughout the trial, Girgus et al. (1977) suggested that “an increase in the number of observers who are reversing [as the trial progresses] gives the mistaken impression of an increase in rate of reversal for all observers. [W]hen the group data are averaged [they] suggest a group trend towards increased reversal rate, that is not apparent for any single observer taken separately” (p. 556). If a greater number of observers were included in the present experiment it is likely that an overall effect of observation time on reversal frequency would have been obtained.

In summary, KDE reversals were found to be predominantly affected by rotation speed and to a much lesser extent by the shape and complexity of the stimulus. The optimal stimulus for further experiments would appear to be a dense, random-dot cylinder with a rotation speed between 10rpm and 40rpm. In future experiments rotation speeds within this range were used. At lower rotation speeds,
a stimulus is unlikely to reverse sufficiently often to allow testing of different factors which may bias KDE depth organisation, especially when shorter trial durations are used. For example, with a rotation speed of 5rpm observer ROS reported only 2 reversals on average during the three-minute trial.
Chapter 5: The role of element occlusion and luminance contrast in the KDE

5.1 Chapter 5: Introduction

Artists and illustrators have long exploited the technique of brightening those elements in a depicted scene intended to lie closer to the observer in order to enhance the appearance of three-dimensionality. Although the depth cue of relative brightness received scientific attention some time ago, the efficacy of this cue in disambiguating the KDE was not formally investigated until quite recently (Schwartz & Sperling, 1983).

Schwartz and Sperling (1983) systematically varied the luminance of lines that formed a rotating wire-frame cube viewed in polar projection. The relationship between simulated distance and luminance they manipulated was labelled *proximity luminance covariance* (PLC).

In one condition (positive PLC) of their experiment, the luminance of each line increased as it rotated in depth closer towards the observer and decreased when the line rotated away from the observer. In the second condition (negative PLC) the relationship between line
luminance and depth was reversed such that lines objectively further from the observer were brighter and nearer lines were dimmer, whereas in the neutral PLC condition all lines were drawn with the same intermediate luminance. Observers showed a strong preference (on more than 90% of trials) for perceiving the depth interpretation of the rotating cube that corresponded to the PLC-defined organisation, which in the negative PLC condition required the non-veridical appearance of the cube as a deforming truncated-pyramid, despite conflicting perspective cues. Perspective alone (the neutral PLC condition) was found to have only a weak influence on the perceived depth organisation of the rotating cube.

Although the results of this experiment suggest that PLC is a strong cue to depth in the KDE, two issues need to be resolved before accepting this conclusion. The first, and most important issue relates to the confounding influence of element occlusion necessarily introduced by the PLC cue. In the traditional KDE stimulus the individual elements (lines or dots) which form the simulated 3D object are uniform in luminance and colour. Consequently, when two elements at different depths briefly overlap (a frequent occurrence with a dense stimulus) during the simulated object rotation there is no information to determine which of the elements occludes the other. With a PLC cube, however, the component lines are rendered with
different intensities and therefore explicit local superimposition cues arise whenever projected lines overlap during rotation. Braunstein, Andersen and Riefer (1982) have shown that element occlusion alone can effectively disambiguate the depth organisation of the KDE. Could KDE disambiguation by the PLC cue be due to the independent effect of depth-correlated luminance, the introduction of element occlusion, or some combination of the two cues?

The following three experiments attempt to tease out the effects of element occlusion and contrast by systematically varying the strength of one cue while pairing the other cue competitively or cooperatively. In this way it should be possible to isolate the contribution of both luminance contrast and occlusion cues to KDE disambiguation when non-homogenous features are used. Of further interest is how the visual system perceptually integrates multiple depth cues when present in a KDE stimulus.

The second issue surrounding the PLC cue is its theoretical basis. Dosher, Sperling, and Wurst (1986) argued that the PLC cue mimicked the reduction in light intensity of a self-luminous line with viewing distance implicating luminance as the operative stimulus variable. However, O’Shea, Blackburn, and Ono (1993) have challenged this explanation by demonstrating that the perceived proximity in depth of a light-grey square relative to that of a dark-grey
square reverses when they are placed on a white rather than on a black background. Changing the luminance of a background does not alter the luminance of the squares but it does alter their contrast. In three experiments O’Shea et al. confirmed that perceived depth is determined by the contrast rather than by the luminance of a stimulus, and suggested that the effect of contrast relates to the attenuation in contrast over distance that occurs in the natural environment as the result of atmospheric scattering of light. Consequently, an additional aim of these experiments was to determine whether the contrast or the luminance of elements determines perceived depth organisation in the KDE.

5.2 Experiment 2: Contrast with randomised occlusion is an ineffective KDE cue

5.2.1 Experiment 2: Introduction

One method of reducing the confounding influence of element occlusion, in order to assess contrast as a depth cue in the KDE, is to eliminate the correlation between the occlusion-specified depth and the actual simulated depth of each element, while maintaining the correlation between element luminance and depth. In the present experiment this was achieved simply by drawing the dots forming a KDE cylinder sequentially in a different random order on each frame
of the animation sequence. As a result of such a randomised drawing order, an objectively near dot would occlude *or* be occluded by a distant dot (if two such dots overlapped) with equal probability (see Figure 5.1). Consequently the global depth organisation of the KDE cylinder could not be determined on the basis of the spatially and temporally conflicting local occlusion cues present in the stimulus. Element luminance, however, specified a consistent global depth organisation as each dot was drawn with a luminance determined by its simulated depth throughout the animation sequence. Although this technique introduced a scintillating appearance to the KDE stimulus (similar to limited dot-lifetime stimuli), observers clearly perceived a coherent, rotating, 3D cylinder.

The aims of this experiment were to determine the role of differential luminance in disambiguating depth in the KDE when the influence of element occlusion is reduced, and also to examine whether luminance *per se* or luminance contrast is the operative depth cue.
Figure 5.1. Stereograms of demonstration stimuli from the black and white background conditions. These stimuli are intended for cross-eyed free-fusion. The sign of binocular disparity is in accordance with the contrast of the dots (i.e., high contrast dots have crossed disparities and low contrast dots uncrossed). The range of dot luminance portrayed correspond to a log F/R luminance ratio of 0.6. While low contrast dots are difficult to see in these static representations, they were clearly visible in the KDE displays used in the experiment. Similarly, the occurrence of dot overlap was more salient with rotating stimuli.
5.2.2 Experiment 2: Method

5.2.2.1 Participants

Two experienced observers participated along with two naïve observers. All four observers were male, their ages ranged from 22 to 35 years. Three observers completed one session (lasting approximately 50 minutes); observer SB completed two sessions. All observers had normal or corrected-to-normal acuity.

5.2.2.2 Apparatus and Stimuli

The KDE stimulus was constructed of 128 dots randomly positioned on the surface of a virtual cylinder. Animation sequences were created by rotating the cylinder through 180 degrees at an angular velocity of 3 degrees per frame. Each 60-frame sequence lasted for two seconds at the display speed of 30 frames per second and appeared as smooth rotation in depth.

The stimuli, which were generated using a Macintosh IIcx, were displayed on an Apple monochrome monitor and viewed through a light-tight chamber in a darkened room. The light-tight chamber was a viewing tunnel fit to the monitor and extended back to the observer’s view apparatus - a binocular cowl embedded into the closed end of chamber above a chinrest. Its interior was covered with black velvet to minimize stray light from the monitor.
At the 1.17m viewing distance the stimuli subtended a 3 degree visual angle both vertically and horizontally.

Each dot on the cylinder was a filled grey circle ten minutes arc in diameter. The luminance of each dot was determined, on a frame-by-frame basis, by its simulated distance from the observer using equation (5.1):

$$Dot_{lum} = Dot_{lum} + \frac{Dot_z}{Radius} \times \frac{L_{range}}{2} \quad (5.1)$$

where $L_{mid}$ corresponds to a mid-luminance value of 70cd/m$^2$, $Dot_z$ is the signed depth coordinate of the dot (negative when behind the centre of the cylinder, positive when in front), $Radius$ refers to the radius of the cylinder along the $z$ axis, and $L_{range}$ is a constant which determines the range of luminance differences between near and far dots.

The dots were drawn in a randomised order (without respect to simulated depth), as described above, in order to eliminate a systematic contribution of the element occlusion cue to the perceived depth organisation of the KDE cylinder.
5.2.2.3 Design

Two factors were varied in the experiment. First, the strength of the PLC cue was manipulated by using different luminance ranges given in Table 5.1. The terms ‘front’ and ‘rear’ refer to the luminance (or contrast) value given to a dot when at the nearest or furthest simulated distance from the observer respectively. In other words, columns (a) and (b) describe the minimum and maximum dot luminance values present in the KDE cylinder for a given cue level. $L_{range}$ is the luminance difference between these extremes, the F/R ratio contains the ratios of front-most dot and rear-most dot luminances, and the log F/R column gives the logarithmic value of these ratios. Columns (f)-(i) present the same front and rear dot luminances (as in columns a and b) expressed as contrast values when displayed against either a black or a white background. Contrast was calculated, for each dot luminance value, by substituting the appropriate dot and background luminances into the following equation: $\text{contrast} = \frac{(L_{\text{max}} - L_{\text{min}})}{(L_{\text{max}} + L_{\text{min}})}$.

Following Dosher et al.’s (1986) terminology, these ranges produce 14 different log Forward-to-Rear luminance ratios ($-0.81$, $-0.60$, $-0.44$, $-0.30$, $-0.18$, $-0.12$, $-0.06$, $0.06$, $0.12$, $0.18$, $0.30$, $0.44$, $0.60$, $0.81$) where negative PLC levels specify clockwise cylinder rotation and positive levels specify counter-clockwise rotation. Second, the
random-dot cylinder was displayed against either a white (140 cd/m²) or a black (1 cd/m²) background. See Figure 5.1 for illustrations.

The factorial combination of these factors produced 28 conditions. Each condition was presented ten times in a random order to give a total of 280 trials per session.

5.2.2.4 Procedure

The naive observers were briefed on the KDE. A model was used in order to describe the appropriate response when the cylinder appeared to rotate clockwise and the response for counter-clockwise rotation. Observers were instructed to view each two-second animation and then to report the apparent direction of cylinder rotation at the conclusion of the trial. They made their response by clicking the screen cursor on one of two on-screen icons that depicted either clockwise or counter-clockwise rotation. Observers initiated each trial by clicking the mouse button. Trials were separated by at least five seconds of darkness while the animation was loaded from the Macintosh’s hard-drive.

5.2.3 Experiment 2: Results

The effect of element luminance on perceived KDE depth organisation was masked by the presence of strong response biases in which observers continued to favour one particular direction of
rotation on successive trials irrespective of the direction specified by PLC. Both naive observers responded with the same direction (counter-clockwise for observer DOB and clockwise for observer DR) of apparent rotation on all 280 trials and consequently no further data were collected. This bias was also present for the two experienced observers and was most striking for the stimuli presented against a black background. ROS reported CW rotation on 89% of all trials; SB reported counter-clockwise rotation on 66%. Clearly these values depart from an average preference for clockwise (or counter-clockwise) of 50% expected if no response bias was operating. In Figure 5.2, the frequency of clockwise reports is plotted for observers ROS and SB over the range of luminance cue strengths for both black and white backgrounds.
The effect of dot luminance was most clearly evident when the KDE cylinder was presented against a white background. For both observers, the rotation direction of the darker dots biased the perceived cylinder rotation towards this direction and the bias generally increased with the strength of the ‘luminance’ cue. The observation that, on a white background, low luminance (but high contrast) dots appeared to lie on the front face of the KDE cylinder thereby determining its depth organisation, demonstrates that dot contrast rather than luminance was the effective depth cue.
Negative luminance ratios indicate CW rotation of the higher luminance cylinder dots (and CCW rotation of the lower luminance dots) while positive ratios correspond to CCW rotation of high luminance dots.

A similar but much attenuated and less reliable effect of dot contrast was evident with black background stimuli. The difference in the efficacy of contrast as a depth cue between these two background conditions was highlighted by ROS’s data. In conditions where contrast specified clockwise rotation (the direction opposite to his response bias) only eight clockwise judgements were made with a black background, 21 were made with a white background.

**Figure 5.2.** Individual psychometric functions showing the effects of luminance contrast on perceived KDE rotation in stimuli with randomised occlusion cues.
The most likely explanation for the asymmetry of luminance contrast cue effectiveness for the two background conditions relates to the difference in the luminance contrast range present in these conditions. From Table 5.1 it is apparent that the same range of dot luminances produces a much greater range of dot luminance contrasts when presented against a white background compared to a black background. Consequently the luminance contrast cue is objectively (and subjectively) stronger in the former condition.

5.2.4 Experiment 2. Discussion

Dot luminance contrast was found to be a fairly weak depth cue with briefly presented KDE cylinders containing inconsistent element occlusion cues. All observers exhibited strong response biases favouring one particular KDE organisation over the other. The use of luminance contrast depth cues, which objectively specified opposite directions of cylinder rotation during the course of the experiment, did not overcome the tendency of naive observers to perceive the same direction of rotation on every trial. With the experienced observers there was some evidence that the luminance contrast of dot elements did influence the perceived depth organisation of the KDE cylinder in a systematic fashion. However, this finding was only reliably obtained with the two experienced observers and only when a white background was used.
5.3 Experiment 3: Luminance contrast is an effective KDE cue when paired with consistent superimposition cues

5.3.1 Experiment 3: Introduction

The disruption of occlusion cues in a KDE stimulus containing elements with depth-correlated luminance contrast appears to weaken the degree to which the contrast cue disambiguates the KDE. This result suggests that the luminance contrast cue does not operate independently from the occlusion cues it introduces into the KDE stimulus. In the next two experiments the interaction between occlusion and luminance contrast cues is directly tested by comparing the effectiveness of each cue when paired cooperatively or competitively with the other cue. In the present experiment the strength of the luminance contrast cue was varied while the strength of the occlusion cue remained constant and in the final experiment the opposite relationship between the cues was tested.

5.3.2 Experiment 3: Method

5.3.2.1 Participants

The same two experienced observers and one of the naive observers from Experiment 2 participated. An additional naive observer was also
recruited. Three observers were male and one was females with an age range of 22 to 35 years. All observers had normal or corrected-to-normal acuity.

5.3.2.2 Stimuli

Animated KDE cylinders were generated in a manner similar to Experiment 2. The strength of the contrast cue was varied over six levels (log F/R ratios of –0.60, –0.30, –0.12, 0.12, 0.30, and 0.60) and, as before, KDE cylinders with these luminance ranges were presented against either a white or a black background.

The relationship between the occlusion and contrast cues was manipulated in two conditions. Dots were sorted on the basis of their simulated distance from the observer on each frame. In the cue-agreement condition, dots were drawn in order from the back to the front of the cylinder ensuring that higher contrast dots always occluded lower contrast dots whenever dots overlapped during rotation. In the cue-conflict condition, dots were drawn in the reverse depth order (from front to back) and consequently lower contrast dots always occluded higher contrast dots. Individual dots were given luminance values based on their simulated depth using equation (5.2) in both cue relationship conditions. Representative cue-conflict and cue-agreement stimuli are shown in Figure 5.3. All other stimulus, apparatus, and procedural details were identical to Experiment 2.
The twenty-four conditions that resulted from the factorial combination of contrast level, background luminance, and cue relationship were presented in a randomised order 20 times. The 480 trials were run over two sessions.

**Figure 5.3.** Stereograms of representative stimuli from Experiment 3. Dots are drawn with a luminance range corresponding to a log F/R luminance ratio of 0.6. Stereogram (a) demonstrates the cue-agreement condition where high contrast dots occlude low contrast dots, and stereogram (b) shows the cue-conflict condition where low contrast dots occlude high contrast dots.

### 5.3.2.3 Design

The twenty-four conditions that resulted from the factorial combination of contrast level, background luminance, and cue relationship were presented in a randomised order 20 times. The 480 trials were run over two sessions.
5.3.3  Experiment 3: Results

5.3.3.1  Cue-Agreement

Good–fitting psychometric functions relating the variations in the F/R luminance ratio to the frequency of counter-clockwise judgements were obtained from all observers when luminance contrast and occlusion specified the same direction of rotation. Individual data are plotted in Figure 5.4.

The tendency for perceived counter-clockwise rotation systematically increased as a function of the luminance ratio for KDE cylinders on a black background. The opposite relationship between dot luminance and perceived direction of rotation was obtained when a white background was used. In these conditions, cylinders with a negative F/R luminance ratio appeared to rotate counter-clockwise and those with a positive ratio appeared to rotate clockwise. As previously stated, the change from a black to a white background does not alter the dot luminances of the KDE cylinder but does reverse their contrast.

If dot luminance rather than contrast determined the apparent depth organisation of the cylinder then altering the luminance of the background should not affect the perceived direction of rotation. Clearly in this experiment, the depth organisation of the KDE
cylinders were determined by dot contrast thus confirming the findings of O’Shea et al. (1993) and those of Experiment 2.
Figure 5.4. Individual psychometric functions (observer initials in each panel) from Experiment 3. Data from white background conditions are shown in the left panels (unfilled symbols) and black backgrounds on the right (filled symbols). Cue-agreement is depicted with circles and cue-conflict by diamond symbols. In the stimuli positive luminance ratios indicate CCW rotation, as defined by PLC, while negative ratios correspond to CW rotation. The manipulation of background reverses this relationship when considered in terms of luminance contrast. For a white background it is negative F/R ratios that signal CCW rotation. Each data point represents 20 observations.
Individual data from the cue–agreement conditions were grouped and subjected to a two-way (contrast level and background luminance) ANOVA. A significant effect of contrast level, $F(5,15)=9.51, p<.001$, and an interaction between contrast and background luminance, $F(5,15)=29.05, p<.0001$ were obtained and there was no main effect of background, $F(1,3)<1$. The interaction between these two factors was most likely due to the greater effectiveness of contrast variations in the white background condition relative to the black background condition. Reflection on the individual functions in Figure 5.4 suggests that the asymmetry in contrast cue effectiveness between the two backgrounds is largely due to the weak effect of contrast with a black background in the data of ROS, although this pattern is also present for observers SB and JD.

Cumulative normal probability curves were fitted to the group data for both the white and black background conditions to illustrate the effect of the combined contrast-occlusion cues on perceived KDE rotation. Reliable fits of the data were obtained for both conditions ($r^2=0.91$ and 0.97 for white and black backgrounds respectively) and the slopes of the fitted functions were in opposite directions ($\beta= -4.12$ for white and $+1.59$ for black).
5.3.3.2 *Cue-Conflict*

Individual differences were apparent in the psychometric functions produced when the depth order specified by dot occlusion contradicted that specified by dot contrast. For observer SI, the perceived direction of cylinder rotation remained in accordance with the contrast cue but the reduced slopes of the cue-conflict functions relative to the cue-agreement functions show that the occlusion cue attenuated the effect of contrast.

![Image](image_url)

**Figure 5.5.** Grouped data (SB, ROS, JD) for experiment 3. Data from white background conditions are shown in the left panel (unfilled symbols) and from black background conditions on the right (filled symbols). Cue-agreement is depicted with circles and cue-conflict by diamond symbols. Error bars correspond to standard errors. Observer SI's data is not included because it showed the opposite cue weightings in the conflict function.
The other three observers produced a quite different pattern of results. For these observers the introduction of a cue conflict reversed the slopes of the psychometric functions relative to those obtained in the cue-agreement conditions. In other words when occlusion and contrast were pitted against each other, these observers tended to report a direction of cylinder rotation consistent with the occlusion-specified depth organisation thus overriding the contrast cue. The dominance of occlusion became more pronounced as contrast increased. Presumably this was because an increased luminance differential between overlapping elements makes the superimposition easier to see.

Data from the three occlusion-dominant observers were grouped and the averaged data from both background conditions were fitted with cumulative normal probability functions. These group data are plotted in Figure 5.5.

A very good fit to the data was achieved for the white background condition ($r^2=0.98$) but the fit for the black background function was rather poor ($r^2=0.40$). The slopes for the fitted psychometric functions were $+1.75$ and $-0.55$ for the white and the black background conditions respectively. These slope values were smaller in magnitude and opposite in sign to the corresponding slopes obtained for these three observers in the cue-agreement conditions ($\beta= -3.89$ and $+1.29$
for white and black backgrounds, both $r^2$s > 0.89). Taken together these results suggest that, for three observers, the introduction of conflicting occlusion cues reduced the impact of variations in the strength of the luminance contrast cue (shallower slopes of cue-conflict psychometric functions) and also biased perceived rotation in the direction specified by the occlusion cue rather than the contrast cue (reversed signs of cue-conflict function slopes).

### 5.3.4 Experiment 3: Discussion

At first blush the results of this experiment appear to support the view that element contrast was an effective cue to depth in the KDE. All observers perceived cylinder rotation in a direction consistent with the contrast cue provided that occlusion cues did not violate the contrast-specified depth-order of the cylinder (i.e., provided high contrast dots occluded low contrast dots). The fact that opposite directions of cylinder rotation were perceived for a given luminance ratio when the luminance of the background was altered (a manipulation which reverses dot contrast but leaves dot luminance unchanged) suggests that contrast rather than luminance determined the perceived KDE depth organisation.

In the cue-conflict condition, however, the majority of observers reported the direction of KDE rotation that corresponded to occlusion rather than contrast. On the other hand, observer SI continued to
report apparent rotation in the direction of the contrast cue (i.e., dependent on the background) despite conflicting occlusion cues.

Given that occlusion specified the same direction of rotation as contrast in the cue-agreement condition, it could be argued that the occlusion-dominant observers also based their judgements on the occlusion cue in this condition. The same pattern of results can be predicted by the depth organisation given by occlusion alone which raises the question of whether contrast had any effect for this group of observers.

Two aspects of the data for the occlusion-dominant observers suggest that although occlusion cues primarily determined KDE depth organisation, there was an additional (albeit weaker) contribution of luminance contrast cues. First, if luminance contrast did not influence the direction of perceived rotation then it would be expected that occlusion would be equally effective in both cue-agreement and cue-conflict conditions. The slopes of the fitted psychometric functions did, however, differ in magnitude between these two conditions. In the cue-conflict condition function, slopes were approximately halved relative to those from the equivalent cue-agreement conditions.

Second, occlusion is normally considered an all-or-none depth cue, the variation of the dot luminance ratios should have had little differential effect on the tendency to perceive rotation in the
occlusion-specified direction, resulting in a step function. It is clear that variations in the bias towards one direction of rotation or the other did follow from variations in the size of the luminance ratios used. This finding in itself is rather interesting. It might have been expected that a strong enough contrast cue would override any conflicting occlusion information and reverse the observer’s perceived organisation (rotation direction) preference. In fact the opposite occurred for three observers where the occlusion cue actually became systematically more effective as the strength of the luminance contrast cue increased. The greater the discrepancy between the two cues, the more observers tended to side with the depth organisation specified by the occlusion cue. As a result the occlusion-specified direction function was quantitative in nature even though the occlusion cue itself is ordinal. Increasing luminance contrast most likely rendered the occlusion cue more effective by increasing the visibility, or salience, of superimposition relationships whenever features overlapped. The depth ordering of (conflicting) element occlusions were therefore most apparent at higher contrast levels.

The findings of the present experiment are consistent with a weighted cue-combination model (Bruno & Cutting, 1988). Both luminance contrast and occlusion contributed to the overall perceived depth organisation of the KDE cylinder, but there were large
individual differences in the relative weighting of each cue. For one observer the contrast cue had a greater weighting than the occlusion cue, and for the other observers the opposite cue weightings were evident.

5.4 Experiment 4: The impact of element occlusion on perceived KDE organisation with cooperative and competitive luminance contrast cues

5.4.1 Experiment 4: Introduction

In order to test the complementary combination of cues, where the strength of the occlusion cue was varied while the contrast cue remained fixed, a new KDE stimulus was devised. Occlusion was introduced into the KDE cylinder by drawing dots that lay on the front-face with one luminance (either black or mid-grey) and dots on the back-face with a different luminance (mid-grey or black). As with the previous experiments this procedure distinguishes one dot from another on occasion when dots overlapped during rotation, thereby providing explicit occlusion information.

Of course assigning different luminances to near and far dots also introduces contrast cues to depth. The nature of the contrast cue used in this experiment differed from the previous manipulations in two
important respects—(i) luminance changed in a discrete step (from mid-grey to black, for example) rather than over a continuous range, as a dot rotated around the cylinder, and (ii) the same luminance range (i.e., mid-grey to black) was used in every condition. Consequently the strength of the contrast cue was held at a constant level.

An approach akin to a signal-noise paradigm was used to vary the effectiveness of the occlusion cue (which otherwise provides only ordinal depth information). In each condition a certain proportion of dots were tagged as occlusion-signal (OS) dots with the remaining dots treated as occlusion-neutral (ON). ON dots were drawn with the same luminance value throughout the cylinder rotation whereas OS dots changed from one luminance to the other as they rotated from the back face to the front face of the cylinder. As all cylinder dots were drawn in a depth-ordered sequence, OS dots on the front-face of the cylinder visibly occluded any back-face dots over which they moved. On the other hand, front-face ON dots did not provide explicit occlusion information; they were drawn with the same luminance as the back-face dots and therefore in the case of dot overlap, front-face dots were indistinguishable from those on the back-face leaving their depth-ordering ambiguous. The luminance contrast cue was either cooperatively or competitively paired with occlusion depending on the contrast assigned to the front-face OS dots.
Longer trial durations were adopted in this experiment to deal with the problem of the reduced variance observed in the preceding experiments. By increasing the observation period there is greater opportunity for depth reversal.

5.4.2 Experiment 4: Method

5.4.2.1 Participants

The same two experienced observers participated along with two new naive observers. Their ages ranged from 19 to 35 years. Two observers were female and two were male. All observers had normal or corrected-to-normal acuity.

5.4.2.2 Stimuli

KDE cylinders, similar to those of the previous two experiments, composed of 128 randomly positioned dots (10 min arc in diameter) were used as stimuli. Animated sequences were created by rotating the cylinder in a clockwise or counter-clockwise direction through a full 360 degrees at an angular velocity of two degrees per frame. The resulting 180-frame sequence was displayed at a rate of 60 frames-per-second (for 90 seconds) giving a cylinder rotation speed of 20rpm. The display subtended a visual angle of 3 degrees by 3 degrees. All other viewing conditions were identical to those of Experiment 3.
Occlusion strength was manipulated by assigning a proportion (0.2, 0.4, 0.6, 0.8, or 1.0) of cylinder dots as OS dots. These dots were drawn with one luminance (either grey or black depending on the contrast condition) when on the back-face of the cylinder and the other luminance (black or grey) when on the front-face. The remaining dots were drawn with the same luminance (that of the back-face OS dots) regardless of their location in depth. OS dots occluded any other dots that they overlapped during rotation.

Signed occlusion-signal values were used to signify the direction of rotation corresponding to the occlusion-specified depth organisation of the cylinder. Positive occlusion-signal values represent clockwise rotation and negative occlusion-signal values represent counterclockwise rotation.

Contrast was combined with occlusion in a manner that either agreed or conflicted with the depth organisation specified with occlusion. In the cue-agreement condition, all back-face dots and any front-face ON dots were grey (44 cd/m²), front-face OS dots were black (2 cd/m²). The KDE cylinder was presented against a white (140 cd/m²) background and so black dots had a higher contrast (and should appear nearer due to the contrast depth cue) than grey dots. In the cue-conflict condition, OS dots on the front-face were grey (i.e., low contrast) and all other dots were black. Therefore opposing depth-
orders were specified by contrast (high contrast dots should appear nearer than low contrast dots) and occlusion (near dots occlude far dots) cues whenever a grey OS dot overlapped a black dot. Stereograms of representative cue-agreement and cue-conflict stimuli with a 0.8 occlusion signal are presented in Figure 5.6.

![Figure 5.6](image-url)  
**Figure 5.6.** Stereograms depicting stimuli from Experiment 4 showing (a) cue-agreement and (b) cue-conflict cylinders with an 80% occlusion strength. Intended for viewing with crossed free-fusion.
5.4.2.3  Design

Twenty conditions resulted from the factorial combination of occlusion-signal (–1.0, –0.8, –0.6, –0.4, –0.2, 0.2, 0.4, 0.6, 0.8, 1.0) and cue-relationship (agreement, conflict). Each condition was presented twice in a randomised order over two 60-minute sessions.

5.4.2.4  Procedure

Animation sequences were presented for 90 seconds during which time the observer tracked the perceived direction of cylinder rotation. Observers depressed one response key whenever the cylinder appeared to rotate clockwise and a different response key for counterclockwise rotation. Successive trials were separated by at least 10 seconds after which time the observer could initiate the next trial by pressing then releasing both response keys.
5.4.3 Experiment 4: Results

The period of perceived counter-clockwise rotation was calculated for each trial by expressing the cumulative duration of counter-clockwise responses as a proportion of the trial where either direction of rotation was perceived.

**Figure 5.7.** Individual data from Experiment 4. The tendency towards perceived CCW rotation is plotted as a function of occlusion-signal strength. Negative values of occlusion-signal represent an occlusion depth organisation consistent with CCW rotation of the cylinder while positive values correspond to occlusion-specified CW rotation.
Figure 5.7 presents the counter-clockwise period data averaged over two sessions for individual observers as a function of occlusion-signal and cue-combination conditions. All observers showed a similar pattern of results, and their data were grouped and subjected to a two-way ANOVA. These group data are plotted in Figure 5.8.

The attempted manipulation of occlusion strength proved successful. Increases in occlusion-signal increased the tendency for observers to report rotation in the direction consistent with an occlusion-specified depth organisation of the cylinder—a finding supported by a significant main effect of occlusion-signal, $F(9, 27) = 45.59, p < .0001$.

Introduction of contrast cues that opposed occlusion cues reduced the degree to which occlusion-signal modulation biased the perceived cylinder rotation towards the occlusion-specified direction. This is apparent from the difference in psychometric function slopes in the cue-agreement and cue-conflict conditions and confirmed by a significant interaction between the occlusion-signal and cue-combination factors, $F(9, 27) = 2.93, p < .05$. There was no main effect of cue-combination, $F(1, 3) < 1$. 
Cumulative normal probability curves were fitted to the grouped data in order to quantify the impact of conflicting contrast cues. Good fits to data from both the cue-agreement ($r^2=0.94$) and cue-conflict ($r^2=0.97$) conditions were achieved, and the slopes of the fitted psychometric functions were $-1.04$ and $-0.60$ respectively. As negative slope values (i.e., consistent with the occlusion-specified direction of rotation) were obtained in both conditions it can be

**Figure 5.8.** Grouped data from Experiment 4. The tendency towards perceived CCW rotation is plotted as a function of occlusion-signal strength. Negative values of occlusion-signal represent an occlusion depth organisation consistent with CCW rotation of the cylinder while positive values correspond to occlusion-specified CW rotation.

Cumulative normal probability curves were fitted to the grouped data in order to quantify the impact of conflicting contrast cues. Good fits to data from both the cue-agreement ($r^2=0.94$) and cue-conflict ($r^2=0.97$) conditions were achieved, and the slopes of the fitted psychometric functions were $-1.04$ and $-0.60$ respectively. As negative slope values (i.e., consistent with the occlusion-specified direction of rotation) were obtained in both conditions it can be
concluded that occlusion cues dominated in determining the depth organisation of the cylinder. However, the finding that the cue-conflict function slope was approximately 58% of the cue-agreement function slope demonstrates that contrast cues also contributed, albeit to a lesser degree than occlusion cues, to perceived KDE organisation when occlusion and contrast were pitted against each other.

5.4.4 Experiment 4: Discussion

The results of the present experiment confirm, with a new technique, the effectiveness of element occlusion as a KDE depth cue as previously reported by Braunstein and colleagues (Braunstein et al., 1982; Braunstein et al., 1986). The signal-noise paradigm adopted here proved a successful means of systematically varying the strength of the occlusion cue, as is indicated by the quantitative nature of the psychometric function obtained. This finding reveals that local occlusion information present in individual features is globally integrated over the whole object to provide a unified depth organisation. It is only through such global integration that the signal to noise ratio has any meaning. To determine signal strength for a given condition, the visual system must compare superimposition information across all features in the object. If such global integration were not occurring then the manipulation of the occlusion signal should not produce any systematic variation in tendency towards a
given perceived rotation direction. Such a quantitative relationship between occlusion strength and rotation bias is not required apriori. As occlusion-neutral features do not specify any particular depth order it may have been the case that the presence of any signal (even a single OS dot) would be sufficient for disambiguating KDE depth organisation.

Data from the cue-conflict conditions confirm the findings of the previous experiment regarding the relative effectiveness of the two cues. When in conflict occlusion-specified depth ordering clearly dominated luminance contrast in determining the perceived KDE organisation. Rotational bias always tended in the direction consistent with occlusion in cue-conflict conditions as was also the case in the previous experiment (where the dominant direction of rotation reversed in cue conflict conditions). The role of luminance contrast can be illustrated by comparing responses for the cue-agreement and cue-conflict conditions when occlusion was 100%. At this occlusion signal strength all front-face dots were black (high contrast) in the cue-agreement condition whereas in the cue-conflict condition all front-face dots were grey (low contrast). Although observers saw rotation in the occlusion – direction in both these conditions, the strength of this rotational bias was reduced as a consequence of luminance-contrast conflict. Clearly the luminance contrast cue was
not simply ignored but rather modulated the effectiveness of the occlusion cue. This suggests that the two cues are perceptually combined with occlusion information given a much higher weighting than luminance contrast.

Even at 100% occlusion (complete cue-agreement) the period of perceived rotation in the cue–specified direction was less than unity (mean of 0.8) whereas the equivalent conditions in the previous experiment yielded near complete disambiguation. The reason for this difference is most likely methodological. In the present experiment the observers viewed the stimulus display for extended periods of time thereby increasing the possibility of a perceived depth reversal. Indeed the tracking method necessarily requires a certain degree of depth reversal to occur otherwise the same direction of rotation would be reported throughout the trial and the dependent measure (rotational bias) would be nominal, either 0 or 1 (continuous rotation in one direction or the other). Consequently the tracking procedure affords a much greater opportunity for depth reversal than does a 2AFC task of brief duration (as used in Experiment 3) and so provides a more conservative estimate of the level of rotational bias.
5.5 Chapter 5. General Discussion

The findings from the present experiments clarify the PLC cue. First, contrast rather than luminance is the basis for this depth cue. More importantly, although contrast does constitute an effective cue, it is the superimposition relationships (and the occlusion cue) which it renders visible that are much more important in determining perceived depth organisation. Without competitively pairing these two cues it may well have been erroneously concluded that PLC was a strong cue (given the cue-agreement findings).

To illustrate, consider the cue-agreement conditions of Experiment 3. Here increasing the level of the contrast cue systematically increased the tendency to see the contrast-specified direction of rotation. This would appear to be clear evidence for PLC. In fact it appears that contrast had a much more indirect role where it increased the salience of the occlusion cue thereby increasing the effectiveness of that cue. This is demonstrated by the conflict conditions of the same experiment. Here increasing the strength of the contrast cue actually decreased its effectiveness (in that more responses were made in the opposite occlusion-consistent direction). However contrast also contributed an independent effect (as confirmed in Experiments 3 and 4) on depth organisation.
These results also have theoretical implications demonstrating that any complete model of SFM must consider the impact of depth cues when heterogeneous (identifiable) features are used. Such models usually make the assumption of homogenous features but this is infrequently the case in real life (an exception would be silhouettes against the late evening sky). Making features heterogeneous in appearance has obvious consequences (as shown in these experiments) for depth organisation as it introduces the possibility of local superimposition information. The findings of Experiment 4 have further implications for SFM theory because it provides psychophysical evidence for the global integration of local feature information. It is only with such integration that a quantitative effect of varying occlusion signal strength could be expected.
Chapter 6. Spatial Induction and KDE

Disambiguation 1: Kinetic Depth Contrast

6.1 Chapter 6. Introduction

In the previous chapter I made the usual approach to disambiguating the KDE. I added explicit depth order information into the KDE stimulus then observed its impact on perceived depth. This approach, provides one means for investigating whether and how the visual system makes use of particular dynamic depth cues. However, adding such cues to KDE stimuli makes them no longer objectively ambiguous. The question remains as to whether perceived depth organisation in the KDE can be biased without changing the test stimulus.

In Chapter 3 I reviewed evidence to support the possibility that perceived KDE depth organisation being strongly influenced by the depth-order information provided by spatial context. Spatial context effects in the KDE were first established by Gillam (1972) who found that depth-ambiguous oblique lines, rotating around a common vertical axis, were more often perceived to be rotating in the same direction than rotating independently. This tendency towards grouping
was strengthened when the lines either had a similar orientation (Gillam & McGrath, 1979), were collinear (Gillam & Grant, 1984), were in close proximity (Gillam, 1972), or were surrounded by a rotating trapezoidal frame that promoted closure (Gillam & Broughton 1991). Eby, Loomis, and Solomon (1989) extended the work of Gillam et al. by exploring the spatial interactions between unambiguous and ambiguous KDE stimuli. In one experiment, two counter-rotating random-dot spheres were presented in two different spatial configurations. The depth organisation of the larger KDE sphere, the inducing stimulus, was rendered unambiguous by adding a high level of perspective information while the other sphere, the target stimulus, had only weak levels of perspective and remained perceptually ambiguous (when presented alone). Observers predominantly reported that the target sphere rotated in the same direction as the inducing sphere—a percept referred to by Eby et al. as rotational linkage. The arrangement of the two spheres was found to influence the degree of spatial interaction. Rotational linkage was strongest when the target sphere was enclosed by the inducing sphere, and greatly reduced when the two spheres were positioned side-by-side.

A quite different effect of spatial context on the KDE was described by Sereno and Sereno (1991). Their stimulus was an orthographic
random-dot sphere embedded in a large field of horizontally translating random-dot texture. Observers reported that for the majority of experimental trials, the ambiguous sphere appeared to rotate in the opposite direction to the surrounding dot-field. That is, the sphere tended to organise such that its perceived front face moved against the flow direction of the translating field, while those sphere dots that moved with the field formed the back face. In this perceived organisation the sphere appeared to be rolling against a translating background. Following Eby et al. (1989), this phenomenon is referred to as rotational contrast.

Rotational linkage (or grouping) is the predominant induction effect that has been reported in the literature, and the stimulus parameters that contribute to this phenomenon have been the subject of extensive investigation. Rotational contrast is a new finding and has consequently received little theoretical or empirical attention.

6.2 Experimental Aims

Why does rotational contrast occur? What are the essential stimulus parameters that determine the strength of rotational contrast? Why does spatial context produce opposing effects, and what determines whether rotational contrast or linkage will be seen in a particular stimulus?
I propose three accounts for rotational contrast (and rotational linkage). Experiments 5 and 6 examine stimulus factors involved in rotational contrast. In Experiments 7 and 8, I explore the possibility that rotational contrast results from the combined operation of simultaneous motion contrast and motion perspective. Finally, in Experiment 9, I test whether the nature of spatial context effects depends on how the stimuli are spatially arranged.

6.3 Candidate Explanations of Rotational Contrast

6.3.1 Supplemented Physiology Hypothesis (SPH)

Sereno and Sereno (1991) considered the rotational contrast effect to be a consequence of the neurophysiological organisation of the motion-processing pathway. In particular they implicated the action of silent-surround motion cells (see Section 1.6.2.2) but did not elaborate, in their abstract, on the exact mechanism relating this class of cells to the perceptual phenomenon they report.

Although silent-surround motion cells may provide the neural substrate of the rotational contrast effect, for a complete account of rotational contrast two issues need to be addressed: What is the perceptual consequence of inhibition/facilitation, and how does this neural process bias the perceived depth organisation of an ambiguous KDE stimulus?
A likely result of silent-surround inhibition/facilitation is to shift the peak population response of neurons sensitive to the moving elements of the two faces of the rotating cylinder. Consequently, the neural coding of element speed would be influenced by the presence of a moving surround. Responses to motion in the same direction as the surround will be inhibited causing elements on one face to appear perceptually slowed, while responses to motion opposite to the surround would be facilitated causing elements on the other face to appear to move faster than their true speed. In other words, the perceptual correlate of silent-surround inhibition / facilitation might be simultaneous motion contrast (Tynan & Sekuler, 1975).

A speed difference between the two faces of the cylinder (introduced through simultaneous motion contrast) may contribute to its perceived depth organisation by partially mimicking motion perspective. In this way the differences in registered rotation speed are interpreted, by the visual system, as the consequence of horizontal motion perspective arising from polar projection (see Section 1.5.2). Elements appear to move faster because they are nearer the observer. The depth organisation of a KDE stimulus has been shown to be biased by such horizontal motion perspective with the faster-moving face perceived as the front of the stimulus, although strong levels of perspective are required for reliable disambiguation (see Section 2.1).
6.3.2 Geometric Constraints Hypothesis (GCH)

An alternative explanation of rotational contrast is that it is consistent with the projected pattern of motions that would arise from a laterally moving observer viewing a real cylinder positioned in front of a textured background. The geometry of such a scene is that, for example, a rightward movement of an observer fixating the cylinder would produce a rightward displacement of the textured background (relative to the cylinder) along with a clockwise (front-face dots rotate leftward) rotation of the cylinder. In other words, on the basis of projective geometry, the cylinder should appear to rotate in the opposite direction to the translation of a more distant background.

Lateral observer movement produces two relevant transformations in the projected images of visible surfaces and objects. First, there is the well-known translation of objects relative to the fixated object that forms the basis for motion parallax. Objects in front of the point of fixation move in the opposite direction to head movement, while those objects behind fixation move in the same direction. The second transformation is the rotation of objects that occurs because as the observer moves, objects are progressively viewed in different orientations relative to the observer (the real-life basis for the KDE). All objects rotate in the direction opposite to observer movement.
irrespective of their distance from the observer (or fixation), while the extent of object rotation varies inversely with their distance.

The geometric hypothesis accounts for rotational contrast in the following way. The stimulus display is interpreted as an abstraction of the situation where the observer moves while fixating a cylinder that lies in front of a random-dot surface. The sheet is perceived as lying behind the cylinder due to accretion and deletion occlusion information at the sheet-cylinder borders in the display. Geometric constraints (outlined above) resolve the rotation direction of the ambiguous cylinder such that it is biased to rotate opposite to the random-dot sheet (which moves with the observer). In addition, the geometric hypothesis predicts that all 3D objects in the scene would rotate in the same direction thus accounting for rotational linkage.

6.3.3 Mechanical Constraints Hypothesis (MCH)

The final explanation proposed for rotational contrast is that observers apply a mechanistic model of the physical world to constrain the interpretation of the ambiguous stimulus display. If a real cylinder were held in place resting on a moving conveyer belt, or between two adjacent rollers, then it would physically rotate in the direction opposite to the conveyer belt (or rollers). In this account, observers adopt a form of naïve physics (McCloskey, 1983) or tacit understanding of natural physical laws, to explain perceptually the
motion and organisation of the ambiguous cylinder by reference to its physical relationship to its surroundings.

The MCH may also account for rotational linkage if the observer assumes that these stimulus displays represent objects that are physically connected to the same rotation axis (a reasonable assumption when the stimuli are collinear). In fact, in Gillam’s (1976) study, shadow projections were produced from wire figures attached to a stem (visible in the projection) mounted on a turntable. The existence of an explicit, shared rotation axis provides an optimal situation for application of the MCH, and observers in this study would be expected to perceive rotational linkage. If physical objects were affixed to a rotating axis they would objectively rotate in the same direction.

Although the MCH may appear to require high-level cognitive reasoning of which the observer is aware, and so could simply be tested by obtaining reports on any strategy the observer might consciously exercise, it is more likely that these mechanistic relationships have been learnt implicitly through experience and observation of the physical world and exist as perceptual heuristics (or constraints) of which the observer is not cognisant.

According to the MCH an observer will, in situations of stimulus ambiguity, favour the perceptual interpretation of the KDE that
corresponds most closely to that predicted by the perceived mechanical relationships between the ambiguous object and its surroundings. A physically impossible event is unlikely to be perceived. Rotation in the opposite direction to surround motion is perceived because this is what would occur in the physical analogue of the perceived scene.

Proffitt and Kaiser (1995) have provided empirical evidence to show that observers do indeed apply simple mechanistic heuristics in the perception of dynamical events. They distinguish between two classes of dynamic events: particle and extended-body motions. Particle motions refers to dynamic behaviours which depend only on the object’s perceived centre of mass whereas extended-body motions involve more complex physical interactions with other objects and surfaces and depend on other aspects of the object such as its shape and orientation. Generally observers perform well in tasks involving particle motion events but poorly in extended-body motion tasks. Consequently the validity of the MCH remains to be established as spatial context effects, such as rotational contrast, are most likely related to extended-body motion events.
6.4 Experiment 5: Surround speed, shape, and rotation speed, and their effects on KDC

6.4.1 Experiment 5: Introduction

The influence of spatial context on the KDE was examined using an ambiguous orthographic random-dot cylinder, which served as the test stimulus, that was sandwiched between two induction fields of moving random dot texture. Two aspects of the stimulus display were manipulated in this experiment.

First, the type of motion presented in the induction fields simulated either a flat sheet of horizontally translating dots (2D surround) or a pair of co-rotating, opaque \(^5\) random-dot cylinders (3D surround). By testing similar stimulus conditions to those used by Eby et al. (1989: 3D surround) and Sereno and Sereno (1991: 2D surround) in the same experiment, it is possible to assess whether the conflicting effects of spatial context found in these two studies were due to differences in type of inducing stimuli.

Rotational linkage between the ambiguous KDE cylinder and 3D surrounds might be expected as these stimuli share a common motion

\(^5\) Self–occlusion (producing opaque cylinders) rather than perspective (c.f., Eby et al. 1989) was used to disambiguate the 3D surrounds because previous research (Braunstein et al., 1986) has shown the former to be the more effective dynamic depth cue.
(3D rotation) and shape (cylinder), stimulus similarity being one factor that promotes perceptual grouping. Grouping may be less likely for 2D surrounds that share few characteristics with the KDE cylinder. Predictions from the SPH regarding the nature of spatial interactions with 3D surrounds are uncertain since the responsiveness of silent-surround motion cells to rotation-in-depth stimuli has not been reported. As the 3D surrounds contain the same average linear dot motion as the 2D surrounds, rotational contrast might also be expected with 3D surrounds.

Second, the average speed of motion of the surround and of the test cylinder was manipulated. Surround motion speed is likely to be an important factor in determining the strength of rotational contrast if this phenomenon is related to neural suppression via silent-surround motion cells. Electrophysiological studies (see Section 1.6.2.2) have shown that most of these cells possess speed-tuned silent-surrounds. Different speeds of surround motion produce different levels of suppression to an optimal motion stimulus simultaneously present in the classically defined receptive field. Consequently, the tendency for rotational contrast in the present experiment should be modulated by the speed of surround motion. Eby et al. (1989), however, found that the tendency for rotational linkage was unaffected by speed differences between their stimuli. A similar degree of rotational
linkage between counter-rotating random-dot planes was observed when both planes rotated at the same speed and when one plane rotated at twice the speed of the other.

6.4.2 Experiment 5: Method

6.4.2.1 Participants

Four male observers, aged between 22 and 38 years, participated in the experiment. Two were experienced psychophysical observers and aware of the experimental aims, the other two were naive. All observers had normal or corrected-to-normal acuity.

6.4.2.2 Stimuli and Apparatus

A Macintosh IIcx computer was used to generate and display the stimuli, and to record observer responses. The animation sequences were displayed on an Apple 13-inch colour monitor (66.7-Hz non-interlaced-frame rate, 72 pixels/inch) at a rate of 60 frames per second.

Animated sequences were created depicting the parallel projection of a random-dot, transparent cylinder rotating around its vertical axis. The transparent cylinder was flanked on either side by a rectangular strip (each half the width of the cylinder) of moving random-dot texture. These surrounds simulated either a rotating opaque cylinder (3D surround) or a sheet of uniformly translating dots (2D surround).
When dots in the 2D surround reached the edge of the surround they wrapped around to the opposite edge. The inducing cylinders were, therefore, half the diameter of the transparent cylinder.

Ninety-six small (2 arcmin) luminous dots were used to speckle the transparent cylinder; 48 similar dots were used to speckle each surround region, ensuring an equal average texture density between the centre and surrounds. The projected dimensions of the transparent cylinder were 90 x 90 pixels that corresponded to 3 deg x 3 deg visual angle at the 65cm viewing distance, and the surrounds were half this width (1.5 x 3 deg). Dots had a luminance of 96 cd/m² and were viewed against a black background with a luminance of 0.1 cd/m².

Some effort was made to reduce flatness cues (Eby, 1992). The stimuli were viewed monocularly (using the sighting dominant eye) through an artificial pupil (5 mm diameter) placed in the eyepiece of a light-tight chamber, situated in front of the computer monitor. The use of an artificial pupil reduced accommodation cues, hid the edges of the monitor, and restricted head movement. All testing was conducted in a darkened room.

6.4.2.3 Design

A repeated-measures factorial design was employed crossing the following independent variables: the speed of the test cylinder rotation (10 rpm or 20 rpm), the relative speed of the surround motion (0.25,
0.5, 1, 2, or 4 times the average speed of the test cylinder), and the type of surround motion (2D or 3D). The test cylinder speeds correspond to average 2D velocities of 1 deg/sec (10 rpm) and 2 deg/sec (20 rpm) respectively. Two further conditions, in which the test cylinder rotated at either 10rpm or 20rpm when sandwiched between two fields of stationary random dots (actually one frame from the 3D surround sequence), were added to establish a baseline measure for response variability with an ambiguous stimulus.

The direction of surround motion (leftward or rightward) and test cylinder rotation was randomised across trials with the constraint that the same direction did not occur on three consecutive trials. All 22 conditions were replicated four times over the course of four hour-long sessions. The order in which each condition was presented was randomised for each session.

6.4.2.4 Procedure

Observers were instructed to direct their attention to the test cylinder and were required to track its apparent rotation direction in the manner described in Experiment 1. Each trial lasted 90 seconds and successive trials were separated by a 30-second inter-trial interval.
6.4.3 Experiment 5: Results

Two measures were used to assess the influence of surround motion on the perceptual organisation of the ambiguous KDE test cylinder. The first measure, direction-of-rotation bias, was expressed as the amount of time in each trial where the ambiguous cylinder was judged as rotating in the direction opposite to the surround motion as a proportion of the total time where coherent rotation in either direction was reported. If the surround motion had no effect on the perceived organisation of the test cylinder then a bias around 0.5, signifying no preference for either rotation direction, would be expected. Biases greater or less than 0.5 would indicate a tendency for rotational contrast or rotational linkage respectively.

The second measure, reversal frequency, was included in order to examine whether changes in direction-of-rotation bias, due to surround motion, produced corresponding changes in the perceptual stability of the test cylinder. Reversal frequencies were calculated in the same way as in Experiment 1.

A general examination of the direction-of-rotation bias data revealed that rotational contrast (rather than linkage) was the predominant consequence of spatial interactions between the test cylinder and surrounds. Of a total of 320 experimental trials, a bias exceeding 0.5 was recorded on 253 (79%) trials.
Direction-of-rotation bias and reversal frequency data in each condition were averaged for individual observers across the four sessions, and the combined mean data for all observers were analysed in repeated-measures three-factor (surround type, relative surround speed, and cylinder speed) ANOVAs. Data from the two baseline conditions (stationary surrounds) were not included in the analysis, but are provided as reference points in the graphs.

Similar degrees of rotational contrast and reversal frequencies were obtained with both the 2D and 3D surrounds. There was no main effect of surround type in either the direction-of-rotation, $F < 1$, or reversal data, $F(1, 3) = 2.44, p > .05$, and this factor was not involved in any significant interaction.

6.4.3.1 Effect of speed on rotational contrast

A significant main effect of relative surround speed, $F(4, 12) = 6.87, p < .01$, was obtained with the direction-of-rotation bias data. The tendency for rotational contrast increased monotonically with the speed of the surround motion relative to that of the test cylinder. Reversal frequencies declined as a function of surround motion speed but this effect failed to attain significance, $F(4, 12) = 2.37, p > .05$.

A related main effect of cylinder speed was not present in the direction-of-rotation bias data ($F < 1$). Reversal frequency was influenced by the rotation speed of the test cylinder, a result consistent
with the findings of Experiment 1. Reversals were more frequent in the 20-rpm conditions (10.8 reversals, on average) than in the 10-rpm conditions (8.5 reversals, on average), $F(1, 3) = 12.89, p < .05$.

A significant interaction between cylinder speed and surround speed, $F(4, 12) = 4.36, p < .05$, was evident for the direction-of-rotation measure. This interaction is plotted in Figure 6.1 for individuals and 6.2(a) for average data.

The interaction arises because rotational contrast increases monotonically with surround speed for 10-rpm cylinders. Whereas it increases more rapidly with surround speeds but then saturates at the highest speeds for 20-rpm cylinders. There were no significant main effects or interactions.
It is possible that the observed differences were because different absolute surround speeds were tested for the two cylinder speed conditions. The issue of whether rotational contrast increases as a function of relative or absolute surround speed is addressed by replotting the data as a function of the absolute speed of surround motion in Figure 6.2(b).

Figure 6.1. Results of Experiment 5. Individual functions showing the effect of relative surround speed on rotational bias. Dashed line shows chance performance. 2D and 3D surround conditions are averaged.
Figure 6.2. Results of Experiment 5 averaged over all observers. Error bars are standard errors. The dashed line represents chance performance. (a) Effect of relative surround speed and test cylinder rotation speed on the perceived direction of cylinder rotation. The same pattern was obtained with both 2D and 3D surrounds and so the averaged data are plotted here. (b) Data replotted as a function of the absolute surround speed.
Here it is apparent that the two cylinder speed curves become more closely aligned implicating the role of absolute surround speeds. The alignment of the two functions is not perfect. There is a noticeable difference, at the 4 deg/sec surround condition, in the level of rotational contrast produced at the two cylinder speeds. A unplanned contrast confirmed this difference to be significant ($p < 0.05$).

### 6.4.3.2 Phenomenology

Observers reported, during debriefing, a number of perceptual changes in the display that accompanied rotational contrast. The test cylinder appeared to rotate at higher velocity and have a greater curvature when it was perceived as rotating in the opposite direction to the surround motion. Paradoxical movement was often experienced in which the cylinder appeared to drift away from the surround but did not change position. Finally, the perceived depth organisation of the cylinder influenced the apparent shape of the 2D surrounds. Under conditions of rotational contrast the edges of the random-dot sheets closest to the cylinder occasionally appeared to bend in depth away from the observer.
6.4.4 Experiment 5: Discussion

As with earlier work on rotational linkage, the perceived depth organisation of an ambiguous KDE object was found to be strongly biased by the presence of other moving stimuli in the visual scene. In distinction to these previous studies the predominant effect of spatial context in the present experiment was rotational contrast.

Conditions with 3D surrounds effectively produced the same degree of bias towards rotational contrast as obtained with 2D surrounds. This result is somewhat surprising given that Eby et al. (1989) observed a weak but significant bias for rotational linkage with a similar display. A further investigation of the necessary stimulus conditions required to produce linkage rather than contrast is described below in Experiment 9.

The finding that 3D surrounds failed to produce rotational linkage is evidence against the GCH that predicts that all 3D objects in the display should rotate in the same direction. Both the SPH and MCH can account for rotational contrast from either 2D or 3D surrounds.

Speed of surround motion emerged as an important determinant of rotational contrast with increases in surround speed generally producing stronger biasing effects. The effect of surround speed was qualified by an interaction with the test-cylinder rotation speed factor. This interaction suggests two aspects of the relationship between
motion speed and rotational contrast. First, the strength of the rotational contrast effect was determined by the absolute surround velocity rather than by the speed of surround motion relative to that of the test cylinder. Second, the rotation speed of the test cylinder may place an upper limit on the degree of surround bias possible. In the present experiment the maximum rotational contrast bias was obtained with the 10-rpm cylinder. The function for the 20-rpm cylinder appears to asymptote with a surround-motion speed of 2 deg/sec. Experiment 1 showed that the frequency of perceptual reversals in rotating cylinders increases with rotation speed. It may be that the influence of spatial context was not sufficient to override fully the higher level of inherent perceptual instability present in the 20-rpm cylinder.

An effect of surround speed on the strength of rotational contrast is consistent with the SPH as silent-surrounds are generally speed-tuned. Moreover Allman et al. (1985) reported that surround inhibition increased with increasing surround-motion velocity in the population of MT neurons tested. Precisely such a pattern of speed-tuning would be required to explain the results of the present experiment.

Other studies, however, have identified a variety of surround speed-tuning functions. Allman et al. (1985) found that the majority of silent-surround neurons in area V2 exhibited a V-shaped surround
speed-tuning with maximum inhibition when centre and surround motion had the same velocity. In addition, Tanaka et al. (1986) reported that MT cells could be classified as having either V-shaped, low-pass (inhibition increases with surround speed), or high-pass (inhibition decreases with surround speed) speed-tuning. Because a range of surround-tuning functions exists, almost any effect of surround speed on rotational contrast would resemble the speed-tuning of at least one class of silent-surround cell. In other words, the physiology of the SPH does not sufficiently constrain the psychophysics of rotational contrast.

6.5 Experiment 6: Effect of surround proximity on KDC

6.5.1 Experiment 6: Introduction

Previous research has shown that the strength of perceptual interactions between stimuli is often determined by their spatial proximity. A classic demonstration of the role of proximity in perceptual organisation comes from the work of Gestalt psychologists in the early 20th century. Wertheimer (1923) found that an array of elements was perceptually grouped into either rows or columns depending on the relative horizontal and vertical spacing between elements.
Gogel (1954, 1974, 1980) further investigated how spatial proximity affected stimulus interactions in a series of studies on perceived size, depth, whiteness and three-dimensional illusions. From the results of these studies he formulated an adjacency principle which stated that “the effectiveness of a cue system between objects or parts of objects in determining perceived object characteristics is inversely related to the relative separation of the objects or parts of objects” (Gogel, 1980, p.154).

The tendency for rotating objects to appear perceptually linked has also been found to be a function of their proximity. Rotating oblique lines were found to group more frequently when the gap separating them was systematically reduced (Gillam, 1972, 1975, 1976). Moreover, Gillam (1981) showed that the tendency for grouping depended on the size of the gap relative to the overall size of the configuration rather than to the retinal extent of the separation between lines.

In this experiment I examine whether the effects of spatial proximity reported in the perceptual grouping literature also extend to rotational contrast. Although proximity is a widely reported determinant of perceptual organisation an argument could be made for it to have no effect on rotational contrast.
The SPH predicts that the tendency for rotational contrast should diminish little over the range of separations between the inducing stimuli and test cylinder tested in this experiment, due to the large size of surrounds in silent-surround cells. Allman et al. (1985) measured the sizes of the effective inhibitory regions surrounding silent-surround cells in area MT of the owl monkey by masking part of the surround motion with annuli centred on the classic receptive field. They found that the smallest silent–surround had a diameter of 35° of visual angle with several surrounds covering the entire visual field. Similarly, Tanaka et al. (1986) reported a strong degree of surround suppression (65% of baseline suppression) for MT cells in the macaque monkey when a 20° x 20° region of surround motion was masked. A reduced, but measurable amount of surround suppression (23% of baseline) remained when the mask was extended to 40° x 40°. These studies suggest that some degree of surround suppression should be evident with separations between test and inducing stimuli of up to around 30° visual angle. On the basis of these physiological data, the SPH would predict that rotational contrast should be obtained with rather large separations between the surrounds and cylinder.

In this experiment I systematically vary the horizontal separation between an ambiguous KD cylinder and nearby fields of surround
motions (either 2D or 3D) to examine if, and how, the proximity of stimuli modulates any perceptual bias for the test cylinder.

6.5.2 Experiment 6: Method

6.5.2.1 Participants

One experienced and three naive observers participated. One of the naive observers had taken part in previous experiments. All were males and their ages ranged from 21 to 28 years. All observers had normal or corrected-to-normal acuity.

6.5.2.2 Stimuli

The stimuli comprised a subset of those from Experiment 5. The essential difference in the present experiment was that the inducing stimuli were horizontally offset from the test cylinder by varying amounts to introduce a gap between the stimuli. The test cylinder rotated at a speed of 10 rpm (average dot velocity of 1 deg/sec) in every condition. The artificial pupil used in Experiment 5 was widened into a narrow slit to ensure that stimuli with the greatest separation were entirely visible with central fixation. All other stimulus parameters, viewing conditions and procedural details were identical to Experiment 5.
6.5.2.3  Design

Three independent variables were factorially crossed in a repeated-measures design. The combination of relative surround speed (half or double), type of surround motion (2D or 3D), and surround separation (0°, 1.5°, 3°, 6°, or 12° visual angle) produced 20 experimental conditions. All 20 conditions were presented in a single session in a randomised order. Each observer completed four sessions, each of which lasted approximately 40 minutes. In two sessions the direction of surround motion was leftward for even conditions and rightward for odd conditions, and in the remaining sessions these directions of surround motion were reversed.

6.5.3  Experiment 6: Results

Direction-of-rotation bias measures were computed for each condition, and values obtained were averaged across sessions for each observer. The grouped data were analysed in a three-factor repeated-measures ANOVA. There was a significant main effect of surround separation, $F(4, 12) = 9.66, p < .001$, which described a decreasing tendency for perceived cylinder rotation opposite to surround motion as the separation between the cylinder and surrounds increased. This trend is evident in Figures 6.3 and 6.4 which plot the interaction between surround separation and surround speed for individual and group data respectively.
From these figures it can be seen that perceived rotation is biased opposite to surround motion providing surrounds are close to the cylinder. When the surrounds were separated from the cylinder by a gap of six or more degrees of visual angle, the perceived direction of rotation was ambiguous (with a bias close to 0.50). A further observation is that for small separations, faster surround motion produces a stronger bias towards rotational contrast than does slower surround motion. This is consistent with the finding of a surround speed effect in Experiment 5. It must be noted, however, that these differences between surround speed conditions were not statistically reliable.

When I analysed reversals data, there were no significant main effects or interactions.
Figure 6.3. Results of Experiment 6. Individual functions showing the effect of centre-surround separation on rotational bias. Dashed line shows chance performance. Cylinder rotation speed was 10rpm
The fact that a six-degree gap between the ambiguous cylinder and the surround eliminated any reliable tendency for rotational contrast is inconsistent with the SPH. Based on the physiological data at hand regarding the effective size of silent-surrounds it would be expected that surround suppression would extend over a much greater area. Of course, great caution should be exercised in extrapolating from the

**Figure 6.4.** Results of Experiment 6. Group data. Functions showing the effect of centre-surround separation on rotational bias. Dashed line shows chance performance. Cylinder rotation speed was 10rpm.

### 6.5.4 Experiment 6: Discussion

The fact that a six-degree gap between the ambiguous cylinder and the surround eliminated any reliable tendency for rotational contrast is inconsistent with the SPH. Based on the physiological data at hand regarding the effective size of silent-surrounds it would be expected that surround suppression would extend over a much greater area. Of course, great caution should be exercised in extrapolating from the
results of single-cell recordings. Neural processing of motion almost certainly involves the combined activity of a highly interactive network of silent-surround neurons. By analogy, one would be mistaken to calculate the limits of visual acuity on the basis of cone diameters. With this strong caveat in mind, the limited spatial extent of surround bias observed in this experiment raises some question of whether silent-surround neurons are directly involved in rotational contrast.

6.6 Experiment 7. Quantification of simultaneous motion contrast

6.6.1 Experiment 7: Introduction

This experiment provides a test of the assumptions underlying the SPH. Briefly restated, the SPH asserts that the presence of surround motion produces a perceived speed difference between the two faces of the cylinder (due to simultaneous motion contrast for which silent-surround cells are the neural substrate) and that this speed difference in turn introduces a motion-perspective-like cue in which the apparently faster moving face of the cylinder appears nearest. As an explanation of rotational contrast, three prerequisites of the SPH must be satisfied.
First, the face rotating in the opposite direction to the surround should appear to rotate faster than the other face. While simultaneous motion contrast has been demonstrated with 2D linear motion (Loomis & Nakayama, 1973; Tynan & Sekuler, 1975), it remains to be seen whether surround motion can similarly influence the perceived speed of 3D rotation. Second, the difference in perceived rotation speed between the cylinder faces must be of sufficient magnitude that if treated as an equivalent level of motion perspective, this level of perspective would be expected reliably to disambiguate the rotation direction of a KDE stimulus. Third, to account for the finding that rotational contrast increases with surround speed, the magnitude of simultaneous motion contrast (and hence the strength of the pseudo-perspective cue) should increase with surround speed.

6.6.2 Experiment 7: Method

6.6.2.1 Participants

The same four observers from Experiment 5 were used.

6.6.2.2 Stimuli

The test stimulus was an opaque rotating cylinder (essentially one face of the transparent cylinder used in the previous experiments) flanked, on either side, by the same translating random-dot sheets as
used previously in the 2D-surround condition. Dot density and visual angle parameters were the same as those of Experiment 5.

An opaque cylinder, which was identical to the test cylinder but was presented without any surround, was used as the variable stimulus. The rotation speed of the variable cylinder could be varied in real-time under the observer’s control. Moving a mouse in the direction of rotation increased the variable cylinder’s speed, moving the mouse in the opposite direction slowed the cylinder. With a 10-rpm test cylinder, the resolution of speed adjustments was 0.02 rpm for matched speeds slower than 10 rpm and 0.09 rpm for matched speeds faster than 10 rpm. Matching resolution was halved with a 20-rpm test cylinder.

6.6.2.3 Design

The design was the same as for Experiment 5 with the following exceptions: 2D surrounds were not used and two new factors: direction of cylinder movement relative to the surround and direction of surround motion, were introduced. On half the trials the test cylinder rotated in the same direction as the surround and in the opposite direction for the remaining trials.

The complete design consisted of the factorial combination of the following four factors: cylinder speed (10 rpm, 20 rpm), relative speed of the surround (x0, x1/4, x1/2, x1, x2, x4), relative direction of the
cylinder (same, opposite), and surround direction (left, right), which produced a total of 48 conditions. Speed matches were made four times (in separate blocks) for each condition. Order of presentation was randomised within each block subject to the constraint that the surround moved in opposite directions on successive trials. This constraint was adopted to avoid the build-up of simple motion after-effects in the surround. The direction in which the test cylinder rotated was randomised independently across trials.

6.6.2.4 Procedure

At the beginning of each trial, the initial rotation speed of the variable cylinder was randomly chosen. Whether the variable cylinder initially rotated faster or slower than the test cylinder was determined randomly. If slower, the rotation speed of the variable cylinder was at most 75% of the test cylinder speed. If faster, the rotation speed of the variable cylinder was at least 150% of the test cylinder speed.

The standard display was shown for five seconds and observers were instructed to fixate the test cylinder and to memorise its apparent speed of rotation. A dark interstimulus interval of 2.5 seconds was followed by presentation of the rotating variable cylinder. Observers adjusted the speed of the cylinder until they matched, to their satisfaction, the perceived speed of test cylinder. When a match was achieved the observer clicked the mouse button, the adjusted speed
setting was recorded, and the next trial began. Observers reported that they experienced little difficulty with the delayed matching paradigm. Testing was completed in one session that lasted approximately an hour.

6.6.3 Experiment 7: Results

Figures 6.5 and 6.6 plot the matched rotation speed as a function of cylinder speed, surround speed, and direction of cylinder rotation. Matched speeds were close to veridical in the baseline conditions where the surround was stationary (average settings of 9.98 rpm, and 18.99 rpm for the 10-rpm and 20-rpm test cylinders respectively) demonstrating that observers could perform the task adequately.

With moving surrounds, a difference in matched cylinder speeds as a function of direction of rotation is apparent. Cylinders that rotated opposite to the surround were perceived as rotating faster than those that rotated with the surround. The magnitude of the speed differences was, however, rather small. Speed differences can be conveniently expressed as the ratio of opposite-cylinder speed to same-cylinder speed. For the 10-rpm cylinders the maximum speed ratio (1.22) was obtained when the surround and cylinder moved at the same speed, the maximum ratio (1.10) for 20-rpm cylinders was obtained when the surround moved at half the speed of the cylinder.
The magnitude of speed differences did not increase monotonically with surround speed as predicted by the SPH. When the data are replotted in Figure 6.6(b) as speed ratios it is apparent that the effects of simultaneous motion contrast are better-described by an inverted U-shaped function of surround speed. Although the peaks of the 10-rpm and 20-rpm functions do not coincide when plotted as a function of relative surround speed, both maxima do occur for the same absolute surround speed (1 deg/sec).

Figure 6.5. Results of Experiment 7. Individual functions showing the mean matched speed of adjusted cylinder rotation as a function of surround speed.

The magnitude of speed differences did not increase monotonically with surround speed as predicted by the SPH. When the data are replotted in Figure 6.6(b) as speed ratios it is apparent that the effects of simultaneous motion contrast are better-described by an inverted U-shaped function of surround speed. Although the peaks of the 10-rpm and 20-rpm functions do not coincide when plotted as a function of relative surround speed, both maxima do occur for the same absolute surround speed (1 deg/sec).
Figure 6.6. Results of Experiment 7. Group Data. Mean matched speeds of adjusted cylinder rotation are shown as a function of surround speed. (a) Triangles are opposite-direction conditions, squares are same-direction conditions. (b) Data replotted as the ratio of matched speeds for the opposite-direction to same-direction cylinder conditions. Values greater than one indicate the opposite-direction face was judged to rotate faster than the same-direction face, less than one indicate the reverse. Unity (shown by the dashed line) represents no effect of surround motion direction. Note that the effect is visually exaggerated, for the sake of illustration, since the y-axis does not extent to zero.
The data were scaled to permit further analysis. Settings were expressed as a proportion of test cylinder speed (i.e., settings from the 20-rpm cylinder conditions were divided by 20 and settings from the 10-rpm conditions were divided by 10). Data from the baseline conditions were omitted from the analysis. A three-factor (surround speed, cylinder direction, cylinder speed) repeated-measures ANOVA was conducted on the scaled data.

Cylinder direction, $F(1,3)=14.60, \ p<.05$, and the interaction between cylinder direction and cylinder speed, $F(1,3)=13.49, \ p<.05$, were the only significant effects to emerge. These results can be easily summarised. Opposite-direction cylinders appeared to rotate faster than same-direction cylinders irrespective of surround speed, and this difference was more pronounced for the 10-rpm cylinders.

6.6.4 Experiment 7: Discussion

Simultaneous motion contrast appears to result from the addition of surround motion to a KDE stimulus. Rotation in the direction opposite to surround motion was perceived as faster than rotation in the same direction. This finding implies that a similar apparent speed difference between the two faces of a transparent cylinder is likely to exist in rotational contrast displays. Thus one assumption of the SPH is supported by the results of the present experiment. However, the small magnitude of the simultaneous motion contrast effect and its non-
monotonic relationship with surround speed are inconsistent with the SPH.

The maximum speed ratio obtained in any condition in this experiment was only 1.22 which raises the question of whether this speed difference would correspond to a sufficiently strong level of equivalent perspective to explain the degree of disambiguation obtained in Experiments 5 and 6. A simulated viewing distance of 20 cylinder radii (a perspective ratio of 1.22) would produce a similar difference in projected speed between two faces of a cylinder through motion perspective. Previous research has shown that very high levels of perspective are necessary for KDE disambiguation. For example, Braunstein (1977) found that a perspective ratio of 2.5 (a projection distance of 2 radii) was required to achieve 75% accuracy in determining the rotation direction of a KDE sphere.

There is another serious problem with the pseudo-perspective mechanism of the SPH. A speed difference between the near and far face of a rotating object is only one consequence of perspective (the compression component). Normally, perspective also involves vertical expansions and contractions in the motion path of rotating elements (the convergence component). When these two components of perspective are manipulated independently, the ubiquitous finding is that convergence strongly dominates compression and is largely
responsible for perceived depth organisation through perspective. Compression information alone does not effectively disambiguate depth order in the KDE (Braunstein, 1976).

Another related finding with motion-parallax displays of planar random-dot surfaces is that although differences in dot velocity give rise to perceived depth, the faster moving elements are not always perceived as nearer the observer; instead the depth ordering is often ambiguous (see Section 2.1).

The second finding from the present experiment that bears on the SPH is the observed relationship between surround speed and simultaneous motion contrast. Instead of increasing with surround speed, the speed-difference ratio followed an inverted U-shaped function. This result seems inconsistent with the findings of Experiment 5. Surround-speed conditions that produced different direction-of-rotation biases in that experiment yielded very similar speed-difference ratios in the present experiment (e.g., for the 20rpm cylinder, the x1/4 and x2 surround-speed conditions gave biases of 0.59 and 0.76 but ratios of 1.036 and 1.039 respectively).

In summary, while simultaneous motion contrast was demonstrated with KDE stimuli, the magnitude and nature of the effect is inconsistent with the SPH. Consequently it is highly unlikely that
perceived speed differences between faces of the KDE cylinder can account for rotational–contrast.

6.7 Experiment 8: Simulated speed difference due to simultaneous motion contrast and failure to disambiguate KDE

6.7.1 Experiment 8: Introduction

This experiment constitutes a final test of the simultaneous motion contrast component of the SPH. Explicitly simulating a speed difference between two cylinders should establish, in a more direct fashion, whether apparent speed differences can account for rotational contrast.

In order to do this a pair of counter-rotating opaque cylinders were superimposed to form the two faces of a composite, transparent cylinder. The rotation speed of each face could be independently controlled. With this stimulus, the speed of one face could be set to the matched speed (obtained from the previous experiment) for the same-direction cylinder with the other face set to the matched speed of opposite-direction cylinder. Consequently it should be possible to mimic the influence of surround motion, in terms of perceived speed differences, without the actual presence of any surround.
The prediction of the SPH is that if the front and rear faces of a cylinder are given matched speeds from Experiment 7, then its apparent direction of rotation should be biased in the direction of the faster moving face. The strength of the bias should also change as a function of the surround condition from which the matched speeds were obtained.

6.7.2 Experiment 8: Method

6.7.2.1 Participants

The same four observers from Experiments 5 and 7 once again participated.

6.7.2.2 Stimuli

A composite KDE cylinder was created by superimposing two opaque cylinders that rotated in opposite directions. The composite cylinder shared the same spatial parameters (size, average dot density) as the transparent test cylinder used in Experiment 5. No surround was present in this display.

6.7.2.3 Design

In Experiment 7, a pair of speed matches (one for the same-direction cylinder and one for the opposite-direction cylinder) was made in each surround-speed and cylinder-speed condition. Ten of these pairs (5 surround-speeds x 2 cylinder-speeds) were applied to
the composite cylinder in this experiment. From each pair, the same-direction setting was used as the rotation speed for one face of the composite cylinder and the opposite-direction setting was used for the other. The same-direction speed was assigned to the CCW rotating face for half the trials and to the CW face on the other half (vice-versa for the opposite-direction speed).

The resulting factorial combination of simulated cylinder speed (10 rpm, 20 rpm) and simulated surround speed (x1/4, x1/2, x1, x2, x4) produced ten conditions. Four replications of each condition were completed over two hour-long sessions.

6.7.2.4 Procedure

The present experiment adopted the same procedure as used in Experiment 5

6.7.3 Experiment 8: Results and Discussion

Observers agreed that the composite cylinder appeared to be a coherent transparent cylinder. Indeed they were surprised by details of its construction. Due to accretion and deletion of dots in the opaque cylinders the edges of the composite cylinder had a twinkling appearance (cf. limited dot-lifetime stimuli).
Rotational bias was expressed as the proportion of a trial where the cylinder was judged as rotating in the same direction as the faster moving face. These data are plotted in Figure 6.7 and were analysed in a 2-factor repeated-measures ANOVA. No significant main effects or interactions resulted. Examination of the graph confirms that simulated speed differences failed to produce rotational-biases exceeding chance levels. Clearly this experiment lends no support to the SPH.

![Figure 6.7](Image)

**Figure 6.7.** Results of Experiment 8. Group Data.
6.8 **Experiment 9: The influence of surround type, configuration, and transparency on rotational linkage and rotational contrast**

6.8.1 **Experiment 9: Introduction**

In this experiment I re-examine how different stimulus conditions can determine whether the consequence of spatial context is rotational linkage or rotational contrast. The results of Experiment 5 and 6 raise an important question—why was rotational contrast, rather than rotational linkage, obtained with 3D surrounds? A possibility is that the displays used in these experiments differed in two important respects (namely stimulus configuration and stimulus transparency) from those that have been previously shown to produce strong rotational linkage.

A vertical stimulus arrangement, where rotating objects are positioned above each other, is the most common stimulus configuration used in the experiments of Gillam (eg. 1972) and Eby et al. (1989). In fact, the horizontal configuration (where objects are positioned beside each other) has been tested only in one other study, Experiment 1 of Eby et al. (1989). In that experiment, rotational linkage was found between spheres arranged in a horizontal configuration, but the strength of linkage was markedly reduced
relative to another condition where the spheres were superimposed. It is not possible to determine whether the attenuation of rotational linkage was due to the configuration or to the increased separation between the two spheres as these two factors were confounded.

The use of a vertical configuration is likely to promote rotational linkage because with this arrangement all objects rotate around the same vertical axis whereas with a horizontal configuration objects rotate around independent axes. Gillam (1976) has shown that the collinearity of rotation axes is an important determinant of rotational linkage.

A second factor that may be important for rotational linkage is stimulus transparency. Eby et al. (1989) used perspective to disambiguate their inducing sphere whereas occlusion was used in Experiments 5 and 6. One consequence of this difference is that all stimuli in the former case were transparent while in the latter case the test cylinder was transparent but the inducing cylinders were opaque. Transparent inducing stimuli might promote rotational linkage either because they share a similar appearance to the test stimulus (i.e., grouping on the basis of similarity), or because both leftward and rightward motions are present in a transparent stimulus. The significance of a representation of opposing directions of motion in the inducing stimulus will be considered below.
In this experiment I explore likely stimulus factors that give rise to opposing perceptual resolutions (linkage vs. contrast), due to spatial context, by presenting the test stimulus in different relationships to its surroundings (horizontal vs. vertical arrangement) and also by varying the nature of those surrounds (2D motion fields, 3D opaque, stereoscopically unambiguous 3D transparent).

6.8.2 Experiment 9: Method

6.8.2.1 Participants

Five observers who were either staff or graduate students of the University of New South Wales participated. Three were experienced psychophysical observers. Two were completely naive, two had some awareness of the experimental rationale, and the remaining observer was the author. All had normal or corrected-to-normal acuity and good stereo vision.

6.8.2.2 Stimuli and Apparatus

Several stimulus changes were adopted in the present experiment. A random-dot sphere rather than a cylinder was used as the test stimulus. This change was to ensure that the boundaries of the test and inducing stimuli were effectively segmented, especially in the vertical configuration condition.
The use of transparent inducing stimuli required a new approach to KDE disambiguation. Binocular disparity was added to both the transparent and opaque inducing cylinders, by rotating the right-eye version of the cylinder through an additional angle of 6 degrees. The stimulus displays were produced as anaglyphs and were viewed through red-blue glasses. Identical random-dot sheets were presented to each eye (i.e., zero disparity) in 2D-surround conditions, and the test sphere was visible only to the right eye. When the displays were viewed with the red-blue glasses the sphere and the inducing stimuli appeared different in colour.

The size of the inducing stimuli was increased so that each subtended the same visual angle (3 deg x 3 deg) as the test sphere. A small gap (4 arcmin) between the nearest edges of the surrounds and sphere was introduced, and the total display subtended 9.13 deg x 3 deg of visual angle in the horizontal configuration (3 deg x 9.13 deg in the vertical configuration). Surrounds (when present) always rotated, or translated, at twice the average speed of the test sphere that rotated at 10 rpm (average dot speed of 1 deg/sec) in every condition. The test sphere, the transparent cylinders, and the opaque cylinders were

---

6 On average, only 128 dots would be visible for the opaque cylinders since the back face dots were not drawn.
speckled with 256 dots (each element had a diameter of 4 arcmin) and 128 dots were used for the random-dot sheet.

A Power Macintosh 6100/66 with a 14-inch Apple colour monitor was used to generate and present the displays. All testing was conducted in a darkened room, and observers used a chin-rest positioned 67 cm from the monitor.

6.8.2.3 Design

Two factors were manipulated—the type of surround used (2D random-dot sheets, opaque cylinders, transparent cylinders, or none) and the stimuli configuration (horizontal where the surrounds were located to the left and right of the sphere, or vertical where the surrounds were above and below the sphere). The factorial combination of these factors produced ten conditions, each of which was presented in a randomised block order for eight replications. The direction of surround motion was counterbalanced. Twenty trials were presented in each session and testing was completed over four sessions.
6.8.2.4  Procedure

Observers tracked the perceived rotation direction of the test sphere during each 60-second trial. Successive trials were separated by at least 15 seconds and were initiated by the observer.

6.8.3  Experiment 9: Results

Direction-of-rotation bias data for each observer were averaged over seven replications completed (the first block was treated as practice), and the mean biases were grouped across observers. A two-factor repeated-measures ANOVA yielded a significant main effect of surround type, $F(3, 12) = 13.29, p < .001$, a marginal effect of configuration, $F(1, 4) = 6.12, p < .07$, and a significant interaction, $F(3, 12) = 5.95, p < .01$. The interaction is plotted in Figure 6.8.

Random-dot sheets produced strong rotational contrast regardless of whether they were presented in a horizontal or a vertical configuration. Configuration had a substantial effect with opaque cylinders. Rotational contrast was obtained when the opaque cylinders were either side of the sphere (replicating the findings of Experiments 5 and 6) but the opposite effect of spatial context, namely rotational linkage, was observed when the opaque cylinders were located above and below the sphere. Rotational–linkage was produced with transparent cylinders in both configurations. The bias was substantially stronger with a vertical configuration.
Nine planned contrasts were used to test the significance of the configuration effect for each surround type, and also to establish in each surround condition whether the obtained rotational bias was significantly larger (or smaller) than the equivalent baseline condition. The increased tendency towards rotational linkage with vertical configurations was significant for opaque and transparent cylinders but not for 2D sheets. Significant rotational contrast was obtained for both 2D sheet configurations and the horizontal opaque cylinder configuration, whereas rotational contrast was reliable for the opaque vicinity.
and transparent cylinders vertically configured but not for transparent cylinders horizontally configured.

To summarise, three factors (configuration, transparency, and type of surround motion) appear to have a cumulative effect on the tendency for rotational linkage. For maximum rotational linkage, inducing stimuli should be transparent, rotating 3D objects, arranged in a vertical configuration on a common rotational axis. Rotational linkage occurred despite the test sphere and inducing cylinder differing in shape\(^7\), colour, and rotation speed.

6.8.4 Experiment 9: Discussion

The findings of the present experiment provides a critical test for each of the three proposed explanations of spatial context effects in the KDE.

As with experiments 5 and 6, rotational contrast was found when the ambiguous sphere was flanked by opaque cylinders. As the GCH predicts that all 3D objects should rotate in the same direction, both rotational linkage and contrast effects were observed, a result consistent with my earlier experiments and the literature. Spatial context effects appear robust over this change in the test stimulus.

\(^7\) As pointed out by a reviewer, changing the test stimulus to a sphere does introduce other factors, such as non-uniform motions at different latitudes. If anything making the inducing and test stimulus dissimilar might reduce the tendency for linkage.
regardless of their spatial arrangement, this reliable finding is evidence against the GCH.

Although rotational contrast was found with the opaque cylinders in the horizontal configuration, strong rotational linkage resulted from a vertical arrangement of the cylinders. The SPH cannot account for such an effect of stimulus configuration. Tanaka et al. (1986) presented motion to a limited part of the surround region of MT silent-surround cells. Surround motion in the regions above and below the classical receptive field (CRF) produced the same degree of inhibition as surround motion in regions to the left and right of the CRF. In other words, silent-surround regions are spatially isotropic. Based on the physiology of silent-surround cells, an equivalent level of rotational–contrast, rather than linkage, should have been obtained with vertically-stacked opaque cylinders.

Finally, the finding that opaque cylinders (in the horizontal configuration) produced rotational contrast but transparent cylinders did not (if anything there was a tendency towards rotational linkage) is inconsistent with the MCH. Both types of cylinders should make equally effective ‘rollers’ for the ambiguous sphere.
Chapter 7. Spatial Induction and KDE

Disambiguation 2: Spreading Concavity

7.1 Chapter 7: Introduction

Experiments from the previous chapter (and those reviewed in Chapter 3) have shown that under conditions of stimulus ambiguity the visual system can integrate additional sources of information over neighbouring regions of space to constrain the interpretation of depth from motion. Although it might be expected that the visual system should make use of auxiliary information when confronted with stimulus ambiguity, it is uncertain whether spatial context can also bias the depth organisation of unambiguous KDE objects. The purpose of this chapter is to explore another instance of spatial context which influences perceived depth organisation.

Context effects were tested with a stimulus created from the combination of two unambiguous KDE cylinders. Each cylinder’s depth organisation (and rotation direction) was disambiguated by eliminating dots from its back face. Self-occluded cylinders such as these appear to be both opaque and convex due to dynamic occlusion cues (Braunstein et al. 1982; Braunstein et al. 1986) and rarely depth-
invert. When these two cylinders are rotated in opposite directions and arranged so that they partially overlap, at least three different perceptual interpretations of the composite stimulus are possible (see Figure 7.1).
Figure 7.1. Three possible perceptual interpretations of overlapping counterrotating opaque random-dot cylinders. (A) Both cylinders appear convex and rotate through each other in the region of overlap. (B) The stimulus is perceptually segregated into three independent regions. (C) One of the cylinders appears convex while the other is perceived as concave (i.e., depth-reversed). The region of cylinder overlap appears transparent.
First, the two cylinders could simply be seen as opaque, convex surfaces rotating through each other (i.e., the simulated distal stimulus). In the region where the two cylinders overlap this interpretation presents a perceptual paradox. Two opaque surfaces cannot simultaneously occupy the same spatial location and both remain visible. One surface must obscure the other.

The following two alternative interpretations resolve this paradox by treating the region of overlap as a transparent KDE cylinder. This requires that the overlapping part of one of the self-occluding cylinders reverses its apparent curvature to form the concave (back) face of the transparent central cylinder.

Second, the central (overlapping) and end (non-overlapping) regions could be segregated and take on independent depth interpretations. In this case the central region would appear to be a transparent cylinder of ambiguous depth, bordered by two opaque, convex cylinders rotating in opposing directions.

Third, the central and end regions could combine into a single, long cylinder rotating in one direction. This would require one end to be convex and opaque, the centre to be transparent, and the other end to be concave. The concave end would appear as a pipe with its front face removed. Concavity presents a paradox because the rotation of
the dots specifying its surface require it to be continually created at one margin and destroyed at the other.

Remarkably, all observers reported \(^8\), at least for some of the time, just such a paradoxical display: one end appeared concave. Moreover, alternations in the depth organisation of the ambiguous, transparent, central region were accompanied by a reversal of the apparently convex and concave ends, along with a reversal of the direction of rotation of the entire cylinder. This suggests that the perceived depth organisation of the central area can bias the organisation of an unambiguous object to the extent of overriding otherwise effective depth cues. I refer to this phenomenon as concavity spreading.

7.2 Experiment 10: Spatial determinants of concavity spreading

7.2.1 Experiment 10: Introduction

The purpose of the present study was to quantify concavity spreading and to investigate the spatial parameters that promote it. By factorially manipulating the linear extent of the centre and ends, two questions can be addressed: First, what is the relative effectiveness of

---

\(^8\) The spreading concavity percept was observed by the author and by others through informal demonstrations. The naive subjects in this experiment, however, were not given a preview of the specific stimuli, but spontaneously experienced the effect as their results demonstrate and debriefing confirmed.
spatial context (in the centre) and dynamic occlusion (in the ends) for concavity spreading? Second, do the absolute or relative sizes of the centre and ends control concavity spreading?

7.2.2 Experiment 10: Method

7.2.2.1 Participants

Four graduate students volunteered for the experiment. There were two females and two males. Ages ranged from 22 to 27 years. Two were experienced psychophysical observers including the author. With the exception of SB, the observers were completely naive to the aims of the experiment. All had normal or corrected-to-normal acuity.

7.2.2.2 Stimuli and Apparatus

A Macintosh IIcx computer was used to generate and display the stimuli. An Apple 13-inch colour monitor was positioned 1.12 m from the observer’s eyes so that one pixel subtended a visual angle of 1 minute of arc. Animated sequences of the rotating KDE cylinders were presented at a rate of 60 frames-per-second ensuring smooth apparent motion. The displays were viewed in a darkened, light-tight room.

Each stimulus display was created by superimposing two rotating cylinders. Small dots (2 x 2 arcmin) were randomly distributed over the surface of two equal-sized virtual cylinders with an average dot
density of 85 dots per degree$^2$. Both cylinders were rendered opaque by hiding dots that were located on the back face (i.e., those dots with a negative z coordinate). The cylinders rotated in opposite directions (upwards or downwards) around their horizontal axes at a constant angular velocity of 60 deg / sec.

A small (8 x 8 arcmin) fixation cross was positioned in the centre of the display. The two cylinders were positioned side by side and arranged symmetrically around the central fixation cross. Overlap was introduced into the stimulus by offsetting the two cylinders in opposite (left or right) directions towards each other. A negative offset produced a stimulus where the two cylinders were separated by a gap while a positive offset produced partially superimposed cylinders. The size of the offset determined the extent of overlap (or separation). The size of the opaque region was changed by increasing the width of the two cylinders. The composite stimulus subtended a visual angle of 80 arcmin in height with an overall width that ranged between 20 and 320 arcmin.

7.2.2.3 Design

Two spatial aspects of the stimulus were varied independently. In any given display the extent of cylinder overlap was either -20, 0, 10, 20, 40, 80, or 160 arcmin and the total size of the opaque region was either 20, 40, 80, 160, 240, or 320 arcmin. The full factorial
combination of these factors yielded 42 possible stimulus displays. The maximum field size that could be smoothly animated at 60 frames-per-second, however, was 320 arcmin width for the total display. This eliminated six combinations that would have required greater widths, leaving 36 stimuli in the design.

Each stimulus condition was repeated four times yielding 144 trials. For half of the trials the opaque cylinder on the left of fixation rotated upwards and the right cylinder rotated downwards. These rotation directions were reversed for the remaining trials. The experimental conditions were presented in a fully randomised order.

7.2.2.4 Procedure

Prior to the experiment proper the difference between concavity and convexity was demonstrated using a curved model surface. Observers were also familiarised with KDE random-dot cylinders. During each trial observers were required to fixate on the central cross while paying attention to the apparent curvature of the opaque areas of the stimulus.

Two response keys were assigned to different sides of the computer keyboard. Observers were instructed to hold down the left key if the left-hand opaque region appeared to be concave. They were asked to keep the key depressed for as long as the region remained concave and to release it while the region appeared convex. They simultaneously
tracked the concavity/convexity of the right-hand opaque region in the same fashion using the right response key. Observers were given five practice trials.

Each of the experimental trials lasted 20 seconds and were started (following an intertrial interval of at least 15 seconds) when the observer pressed and then released both response keys. Half of the trials were completed in a single session that lasted approximately 45 minutes. The two sessions were separated by at least one day.

7.2.3 Experiment 10: Results

The cumulative time in which either opaque region appeared concave was converted to a proportion of trial duration. This measure represents the perceptual bias towards concavity. A concavity bias of 1.0 indicates that observers perceived one or the other opaque end as concave consistently throughout the trial while a bias of 0.0 indicates that both ends remained convex for the entire trial. The four replications of each stimulus condition were averaged for individual observers.

In those conditions where the two cylinders were separated by a 20’ gap, both cylinders appeared to be convex for 75% of the observation period (a concavity bias of 0.25). The convex interpretation of the cylinders was consistent with dynamic occlusion demonstrating that this depth cue was effective in non–overlapping stimuli.
Reports of perceived concavity dramatically increased when the cylinders overlapped, with an average concavity bias of 0.84. All three naive observers recorded a concavity bias of 1.0 on at least 57 trials (out of a possible 104) for those conditions where the cylinders were partially superimposed. Observer SB reported complete concavity bias on 9 trials.

7.2.3.1 Phenomenology

Concavity spreading did not always occur with overlapping cylinders and occasionally both cylinders were reported as being convex. In the experimental debriefing, observers were asked how the perceptual paradox created by these stimuli was resolved. They described two different percepts. Most frequently, the configuration appeared as a convex cylinder with a transparent end that was bordered by a smaller opaque cylinder that rotated independently (possibility B in Figure 7.1). The alternative percept, which was rarely experienced, was of two opaque cylinders rotating through each other with no appearance of transparency in the region of overlap (possibility A). This constitutes a failure of depth contrast and is also a violation of opacity since two opaque objects should not occupy the same position in space and both remain visible.
7.2.3.2 Effects of spatial manipulations

Figure 7.2 displays the full data set as a function of cylinder overlap with the total size of the opaque region as the curve parameter. From these data it can be observed that concavity bias systematically increased with the degree of cylinder overlap and was attenuated with larger opaque region sizes.

As the experiment design was not fully factorial it was necessary to analyse the effects of cylinder overlap and opaque region sizes using two subsets of the data. Subset 1 was produced by eliminating two levels of the opaque region size factor (240’ and 320’) because these conditions were not tested over all extents of overlap. For subset 2, three levels of the overlap factor (–20’, 80’, and 180’) along with the 320’ opaque region size conditions were removed. Between these two subsets, 33 of the 36 conditions were included. Both analyses gave the identical pattern of results and consequently only the former analysis is presented.

A 7 x 4 repeated-measures ANOVA confirmed a significant main effect of cylinder overlap, $F(6, 18) = 21.05, p < .0001$. The effect of opaque region size was not significant, $F(3, 9) = 1.83, p > .05$, but there was an interaction between the two factors, $F(18, 54) = 1.85, p < .05$. 
The interaction between the opaque and overlap factors suggests that the relative rather than the absolute sizes of these regions may determine the strength of concavity spreading. To explore this possibility, the data are replotted in Figure 7.3 as a function of a single

**Figure 7.2** Results of Experiment 10. Group data plotted as a function of cylinder overlap with the size of the opaque region (in min arc) as the curve parameter. Negative overlap values correspond to a separation between the cylinders. Each data point represents the mean of 16 observations (four observers x four replications). Error bar indicate standard errors.

7.2.3.3  **Relativity of spatial parameters**

The interaction between the opaque and overlap factors suggests that the relative rather than the absolute sizes of these regions may determine the strength of concavity spreading. To explore this possibility, the data are replotted in Figure 7.3 as a function of a single
independent measure—the proportion of the total stimulus area in which the cylinders overlapped. The proportional measure of overlap was calculated using the following equation:

\[
p_{\text{Overlap}} = \frac{a + 1}{a + b + 1}
\]

where \(a\) is the extent of cylinder overlap, and \(b\) is the size of the opaque region. The baseline conditions in which the cylinders were separated by a 20’ gap were eliminated from the analysis.

---

\(^9\) I added 1 into the expression to allow inclusion of the 0' overlap conditions. In these conditions where the two cylinders touched but did not overlap, concavity spreading was evident and appeared to be modulated by the size of the opaque region in the same fashion as in the conditions where the cylinders did overlap. Due to the relatively sparse dot density in the stimuli there were no sharp edges delineating the cylinders. Consequently in the 0' overlap conditions, the cylinders often appeared to overlap slightly.
Examination of Figure 7.3 shows that by expressing the strength of concavity bias in terms of the relative extent of cylinder overlap, the data can be organised along a single dimension. A logarithmic function produced a good fit to these data ($r^2 = 0.82$) compared to a poorer fit ($r^2 = 0.50$) obtained when the absolute extent of overlap was used instead as the independent variable.

Figure 7.3. Results of Experiment 10 replotted as a function of proportional cylinder overlap. The proportion of overlap corresponds to the size of the overlapping region as a proportion of the total stimulus size. Also shown is the logarithmic fit to these data.

Examination of Figure 7.3 shows that by expressing the strength of concavity bias in terms of the relative extent of cylinder overlap, the data can be organised along a single dimension. A logarithmic function produced a good fit to these data ($r^2 = 0.82$) compared to a poorer fit ($r^2 = 0.50$) obtained when the absolute extent of overlap was used instead as the independent variable.
7.2.3.4 **Region size versus eccentricity**

Any manipulation of overlap or opaque region sizes necessarily changes the overall width of the composite stimulus, thus introducing a confound between the spatial manipulations and the total size of the stimulus (or the eccentricity of the outer edges of the opaque regions). To rule out overall stimulus size / eccentricity as a factor in concavity spreading, the fifteen instances where the same overall stimuli widths are produced by two different overlap: opaque size combinations were compared. Pairwise comparisons between these conditions revealed significant differences ($p < .01$) in 12 of 15 cases. Such differences demonstrate effects of the two spatial manipulations that are independent of overall stimulus size / eccentricity.

7.2.4 **Experiment 10: Discussion**

Concavity spreading demonstrates the potency of spatial context in determining an object’s perceived three-dimensional organisation. Partially superimposing unambiguous KDE surfaces has two paradoxical effects. First, in the region of overlap, a coherent, transparent cylinder with ambiguous depth is perceived. Second, the apparent depth organisation of the composite cylinder controls the perceived curvature of the opaque ends thereby introducing ambiguity into these otherwise unambiguous KDE stimuli.
An opaque KDE object, presented in isolation, rarely reverses its apparent depth (see Section 2.2) – self-occlusion is a strong depth cue. However, when two opaque cylinders overlapped by even a small amount, as in the present experiment, there was a very striking tendency for one cylinder to invert its depth and appear concave. In the baseline condition where the cylinders were separated by a small gap, one of the cylinders appeared concave for about 25% of the observation period. When the cylinders overlapped, however, concavity spread over one or the other cylinder for about 80% of the observation period on average.

Two spatial manipulations that regulated concavity spreading were investigated. The first, and stronger, is a directly proportional influence of cylinder proximity and overlap. The second is an inversely proportional influence of the opaque region size. While it remains to be learned whether there are other spatial manipulations that have not been studied that might affect concavity spreading, it should be noted that the above two influences account for 66% of the variance of these data.

One issue that this experiment raises is the question of whether cylinder overlap was necessary for concavity spreading. When the cylinders were touching but not overlapping there was a notable bias towards concavity. Two interpretations of this finding can be
advanced. It may have been due to spatial mislocalisation such that the
two cylinders appeared to be slightly overlapping. There were no
sharp edges delineating the cylinders due to the relatively sparse dot
density in the stimuli. The other possibility is that the cylinders were
close enough to produce the same spatial interactions as in the truly
overlapping stimuli. In other words cylinder proximity, rather than
overlap, was the essential prerequisite for concavity spreading. The
significant reduction of concavity spreading with opaque region size
that occurred in the adjacent cylinders condition supports the latter
possibility. It is unlikely that any tendency for mislocalisation would
be systematically affected by changing the size of the opaque regions.
Rather this attenuation with larger opaque cylinders is the likely result
of an increase in the salience of self-occlusion information present in
the opaque regions.

I propose that the combined operation of two different spatial
interactions produces concavity spreading which is illustrated as a
simple process model in Figure 7.4. The first of these interactions,
depth contrast (or repulsion), occurs most strongly in the region of
cylinder overlap. When curved surfaces rotating in opposite directions
occupy the same location (or are in close proximity) the visual system
perceptually separates them in depth. As a consequence of this depth
repulsion, the overlapping portion of one cylinder undergoes a depth
reversal and appears concave. Several authors (Andersen, 1989; Mace & Shaw, 1974; McConkie & Farber, 1979) have investigated similar depth repulsion effects with superimposed random-dot fields that move in different directions. Their observers reported that the surfaces appeared to be separated in depth with an ambiguous depth ordering. Preliminary evidence for neural mechanisms in the primate striate cortex capable of segregating transparent moving surfaces has been reported by Erickson, Snowden, Andersen, and Treue (1989). Competitive processes segregate the surfaces to opposing depths while cooperative processes favour the perceptual solution of a coherent cylinder over independent surfaces.

The second interaction proposed to produce concavity spreading is perceptual linkage (Eby et al., 1989). Perceptual linkage between the overlapping and non-overlapping parts of a cylinder ensures that each part of the cylinder maintains the same depth organisation. Thus, the depth-reversed organisation in the overlapping region of one cylinder (due to depth repulsion) is extended along to its non-overlapping region. Consequently the opaque region of this cylinder also appears to be concave.

Opposing concavity spreading in these stimuli are dynamic occlusion cues specifying convexity. In this account, whether one or the other opaque ends of the composite cylinder is perceived as
concave depends on the balance between concavity spreading and convexity bias. Altering the relative sizes of the overlapping and non-overlapping regions has the effect of shifting this balance\textsuperscript{10}.

Figure 7.4. Process model for spreading concavity.

\textsuperscript{10} Of course, there are also dynamic changes in KDE organisation (i.e., when one end alternates between concavity and convexity within an observation period). My simple model does not address these changes.
Concavity spreading increases with the size of the transparent region, due to an increased tendency for depth repulsion between superimposed surfaces with opposing directions of motion. On the other hand, increasing the size of the opaque regions attenuates concavity spreading due to a corresponding rise in the salience of competing dynamic occlusion cues, an increased number of moving dots requiring perceptual capture, an increase in the distance (between the transparent region and the opaque outer edges of the composite cylinder) over which depth contrast must spread, or some combination of these factors.

In conclusion, the present experiment demonstrates a new phenomenon of the KDE, concavity spreading. When two KDE stimuli that are usually perceived as convex are brought close to each other or superimposed, one appears concave. Concavity spreading may be accounted for in terms of depth repulsion and perceptual linkage.
The experimental results show that ambiguous depth organisation in the KDE is resolved not only by reinstating depth cues into the stimulus, a form of temporal context, but also by spatial context. Spatial context yields perceptual contrast and capture despite contradictory depth cues. That is, objects in the scene are not processed independently. Instead, the interpretation of one object or surface biases the resolution of another. This finding, I propose, is important in the explanation of why depth reversals are common in the KDE but rarely perceived in the natural world. It is the removal of an object from the context of its surroundings undergoing perspective transformation that contributes to its unstable motion-based depth organisation. Although the use of depth cues has been shown (here and in other work) to disambiguate the KDE, it is clear that any theory of the perceptual stability of 3D shape recovery must also account for spatial interactions between separated objects.

8.1 Summary of Experimental Results
The interplay of luminance contrast and superimposition cues, investigated in Chapter 5, demonstrates that the depth cues present within the stimulus interact to facilitate disambiguation in the KDE. I consider this as an example of context that is local to the stimulus because the signal:noise paradigm employed in Experiment 4 demonstrates spatial integration. That is, the graded disambiguation that results from manipulations of superimposition cue frequency (which provides only ordinal depth) would not be expected unless local depth-order measurements are combined globally. Furthermore, the information afforded by each of the cues is interpreted in the context of the other. In the case of cue conflict, a stronger luminance cue enhances rather than diminishes the efficacy of conflicting superimposition signals. Rotation is seen in the direction of occlusion-specified depth more frequently when opposing luminance contrast is greater. The effect of context here is due to interactions within the KDE object.

The resolution of KDE depth organisation is strongly biased by the presence of other objects in its immediate surroundings. The experiments of Chapter 6 examined two manifestations of this spatial context effect – rotational contrast and rotational linkage. Surrounds with 2D linear motion (flat sheets) induce rotational contrast that is strengthened with faster surrounds speeds and smaller separations.
Rotational contrast is unaffected by whether 2D surrounds are positioned parallel to or perpendicular to the KDE rotation axis. Surrounds with 3D rotational motion (cylinders) also induce rotational contrast but only when positioned parallel to the ambiguous KDE stimulus. But when the 3D surrounds are perpendicular, sharing the same rotation axis, rotational linkage occurs. Rotational linkage is stronger when the 3D inducers are transparent. This anisotropic spatial context effect, which depends on the interaction of object-based (2D versus 3D surrounds) and scene-based (configuration) properties, is an important result for any theoretical account to explain.

The phenomenon of spreading concavity, examined in Chapter 7, demonstrates that the interpretation of dynamic depth cues is critically dependent on the context in which they are presented – an observation consistent with the broader perceptual constancy literature. Considered in isolation, a self-occluding KDE stimulus provides a very stable convex percept, but when placed alongside another in counter-rotation, it reliably inverts in depth. Accretion/deletion cues now give rise to a scintillating appearance along the stimulus edges – a percept unnoticed in the convex organisation – suggesting a non-depth based interpretation of the occlusion cue forced by spatial context.
When counter-rotating cylinders overlap there is a striking emergent phenomenon – the formation of an ambiguous transparent cylinder whose momentary depth organisation determines the apparent direction of curvature of the opaque ends. One opaque cylinder end or the other frequently appear concave (a depth inversion rarely experienced in an opaque object) – a perceptual resolution that provides a coherent solution for the configuration. The bias for concavity increases with greater overlap and decreases with greater non-overlap. Remarkably, the strong influence of spatial context introduces ambiguity into the whole configuration where none previously exists in its constituent parts.

8.2 Implications for theory

Faced with the challenge of recovering stable depth from ambiguous depth cues and object motion, the visual system clearly takes the contextual factors into account to arrive at a perceptually coherent solution. Why should an object’s organisation be resolved with reference to its surroundings? I have proposed four theoretical explanations of context effects: (i) the Supplemented Physiology Hypothesis: context effects are a by-product of the physiological
organisation of the motion system; (ii) the Mechanical Constraints Hypothesis: context effects reflect the use of heuristics based on mechanical physics; (iii) the Geometric Constraints Hypothesis: context effects result from the application of constraints based on the geometric relationships that exist between objects under common motion produced by a moving observer; and (iv) the Surface Reconstruction SFM Scheme: context effects are due to segmentation and interpolation processes used in computational algorithms that recover structure from motion. The extent to which these accounts fit the experimental data will be summarised below.

8.2.1 Supplemented Physiology Hypothesis (SPH)

Spatial context effects are not adequately explained by simple 2D motion contrast. The perceived speed differences between faces of the KDE stimulus due to surround motion, found in Experiment 7, were insufficient to give effective differential motion perspective. The SPH predicted the graded effects of surround speed and proximity on rotational contrast observed, but could not account for rotational linkage particularly with transparent inducers. The spreading concavity effect is also not easily explained by the SPH. Simple motion contrast in the region of cylinder overlap should cancel and have little (if any) effect on their opaque surrounds. There is also no
reason to expect that induced speed differences applied to the
surrounds would favour a concave over a convex solution – the
velocity gradients are equivalent.

8.2.2 Mechanical Constraints Hypothesis (MCH)

The MCH can explain the anisotropic context effect. A cylinder
would rotate against a moving sheet on which it rested, regardless of
the sheet’s position. Likewise, a cylinder would roll against other
cylinders on which it lay (the parallel configuration) but would rotate
in the same direction if the cylinders shared a common axis (the
perpendicular configuration). Spreading concavity can also be
accounted by a mechanical bias that minimises torsion (conservation
of energy). Otherwise, if the counter-rotating cylinders were treated as
a single cylinder and both ends maintained their occlusion, then this
would require non-rigid twisting of the composite object at its centre.

Several findings are, however, inconsistent with a hypothesis based
on the MCH. First, a small gap between objects should establish their
physical independence and invalidate any assumption of mechanical
constraint. Yet both rotational contrast and spreading concavity effects
were observed (although weakened) when an obvious separation was
present. Second, rotational contrast increased with the magnitude of
the difference in speed between the object and its surround – a
stimulus situation that again should reinforce mechanical independence. Third, changing the extents of overlapping and non-overlapping regions in the spreading concavity stimulus should not change the mechanics of the scene but it strongly influences the perceptual effect. Fourth, and most problematic, the strength of context effects differs for transparent and opaque surrounds. Yet this manipulation does nothing to change the physical relationship between objects.

It could be argued that perceived physical characteristics of the surrounds, in particular their solidity, might affect the mechanical influence they afford. The opaque 3D surround might well be interpreted as solid roller whereas the transparent 3D surround might appear less substantial (like a swarm of bees) and therefore exert less force. A loss of structural cohesion (and perceived solidity) in random-dot KDE structure is sometimes observed in the transitional period between alternative perceptual organisations. However the transparent 3D surrounds used in Experiment 9 were disambiguated through the addition of binocular disparity. Therefore both stereoscopic and structure-from-motion information in the surround stimuli indicated that they were solid (albeit transparent) 3D surfaces.

---

11 I am grateful to an anonymous reviewer for raising this point.
8.2.3 Geometric Constraints Hypothesis (GCH)

The GCH accounts for rotational contrast with 2D surrounds, rotational linkage with 3D surrounds in a perpendicular configuration, and the occurrence of concavity spreading. However, the contrast observed with adjacent 3D surrounds, as shown in Chapter 6, violates the rotational linkage that GCH predicts. This might be expected as the displays did not faithfully represent the veridical geometry of observer-produced motion perspective. The absence of translatory movement of the surrounds (particularly with greater separations) provides evidence that the objects are moving independently thereby decreasing the likelihood that geometry-based heuristics are applied. Furthermore, the GCH does not account for the effects of inducer transparency nor the increase in spreading concavity with overlap. A conclusive test of the GCH will require stimulus displays that abstract the real-world geometry more faithfully.

8.2.4 Surface Reconstruction SFM Scheme (SRS)

The SRS surface-interpolation process (employing smoothness and rigidity constraints) predicts rotational contrast and linkage observed with opaque 3D surrounds. It also explains spreading concavity (the percept of concavity supports the more rigid interpretation of the
configuration) but not the influence of cylinder overlap. According to the SRS, rotational contrast should have been: enhanced when adjacent 3D surrounds were transparent; eliminated when the 3D surrounds rotated at different speeds to the cylinder (a violation of rigidity); and not produced by 2D surrounds. None of these results were borne out by the experimental findings.

8.3 Suggestions for Future Work

Several experiments can be suggested to extend the findings presented in my thesis.

8.3.1 Future Experiment 1: Object Similarity and Spreading Concavity

The component cylinders used in the concavity spreading experiment were matched in terms of their diameter, rotational velocity, shape, and colour – only the direction of their rotation differed. This similarity provides the possibility of a composite configuration compatible with the component parts. To what extent do spatial context effects depend on object similarity? This question could be answered by systematically introducing differences in the
rotation speed (a manipulation that would violate the rigidity of a composite solution), diameter (a situation where a transparent cylinder in the region of overlap would not be locally smooth), and colour (a parameter that does not affect the geometry, mechanics, or rigidity of the configuration but does perceptually distinguish the surfaces) of the opaque cylinders.

8.3.2 Future Experiment 2: Interaction of Cue-Specified and Context-Specified Depth Organisation

With the concavity spreading stimuli the only available depth cue, occlusion, was violated in the region of cylinder overlap. If other depth cues (for example: binocular disparity, relative element size) are added to the opaque cylinders that support their convex organisation, will spatial context remain sufficiently compelling to produce perceived concavity? Or, can additional depth cues force a stable percept of opaque cylinders rotating through each other? This approach directly pits context-specified depth organisation against that specified by depth-cues. The strength of conflicting depth cues could be varied parametrically, in a nulling paradigm, as another approach to quantify concavity spreading.

The use of coloured cylinders would introduce superimposition cues that could be manipulated through the approach used in Chapter
5. Consider the overlap between a red and a blue cylinder. If red dots, in the region of overlap, always obscured blue dots could a reliably stable percept of a concave blue cylinder (the percept supported by superimposition) be achieved? Such a result would directly demonstrate that the perceptual resolution in transparent cylinder is applied, through context, to the opaque ends.

8.3.3 Future Experiment 3. Motion Parallax and KDE

If spatial context effects are in some part due to the geometric transformations due to observer movement, then context effects would be expected to be pronounced context is better simulated with motion parallax displays. To explore this possibility I constructed the following demonstration stimulus.

An orthographic, transparent KD stimulus was presented rotating around its vertical axis. Two random dot sheets translated horizontally in opposite directions. One sheet started from a location approximately three cylinder diameters to the right of ambiguous cylinder. It was placed such that its lower edge was slightly above the top edge of the cylinder bounds. When in motion it first moved to the left until it reached a position three diameters to the left of the test cylinder. At this point it reversed direction to complete the cycle. The other sheet started diagonally opposite (bottom left). The cylinder rotated through a partial revolution at rotation speed that produced an
average dot speed that matched the sheet translation, and reversed its
direction in phase with the sheets. In such a display the relative depth
ordering of all three objects is ambiguous.

The oscillation of the transparent KD cylinder was unnoticed when
presented alone. Subjective reversals did not share a phase
relationship with objective direction changes. The cylinder appeared
to rotate coherently (a percept that would require reversal) and the
direction of rotation was ambiguous. When the translating sheets were
observed there was a clear separation in depth but the depth ordering
was ambiguous. When the sheets and cylinder were combined in the
same display however, the stimuli constrained each other. Oscilations
in the cylinder were now synchronised with objective direction
changes – a percept compatible with the motion-parallax surrounds.
One depth interpretation (ordering of surfaces) stabilised the percept of
the other (rotation direction of the cylinder). Reversals in depth were
yoked - a change in the rotation direction of the cylinder was
accompanied by a reversal in the depth ordering of the motion
parallax components (and vice-versa). Adding objective depth order
through disparity cues to either stimulus (KDE cylinder, motion
parallax sheets) dictated the perceived depth ordering of the other in
the predicted direction.
A more formal, parametric investigation of these observations is warranted.

8.4 Concluding remarks

In conclusion, I have shown that internal and external context do help disambiguate the KDE. I have proposed it is the removal of an object from the context of its surroundings undergoing perspective transformation that contributes to its unstable motion-based depth organisation. This theory explains more of the critical findings than any other.
References.


References


References


References


References


References


