DOCTOR OF PHILOSOPHY

The modulation of inhibition of return by object internal structure : implications for theories of object-based attention and object shape representation.

Reppa, Irene

Award date:
2002

Awarding institution:
Bangor University

The Modulation of Inhibition of Return by Object Internal Structure: Implications for Theories of Object-Based Attention and Object Shape Representation

Irene S. Reppa B.A., M.Sc.

This thesis is submitted in part fulfilment of the degree of Doctor of Philosophy,
completed at the Centre for Cognitive Neuroscience,
School of Psychology,
University of Wales, Bangor.

March 2002.

Acknowledgements

The research presented in this thesis was supported by a three-year studentship to myself from funds available from Unilever Plc to the School of Psychology at the University of Wales, Bangor.

I would like to express my thanks and gratitude to a few people whose contribution to this work, direct and indirect, has been invaluable. First and foremost, I am indebted to Dr. Charles Leek, not only for the rescue from the 'Cyclops' but also for doing an excellent job as my Ph.D. supervisor. My parents, Rebecca and Sotiris, I thank for the bare essentials, my nature and nurture. I am grateful to my dear friend Panagiotis for suggesting this 'journey to Ithaca' to me and Julian Phillips for the comradeship. Noel Dempsey and his family I thank for the music.

Table of Contents

Abstract....................1

General Introduction....................3

1. Chapter One - Location-based and Object-based Models of Attention....................7

1.1 The focus of attentional selection....................8

1.2 Attention to locations....................9

1.2.1. Neuropsychological evidence for location-based attention....................13

1.3 Attention to objects....................15

1.3.1 Divided attention tasks....................15

1.3.2 Selective attention and spatial cueing tasks....................18

1.3.3 Neuropsychological evidence for object-based attention....................24

1.3.3a. Object-based neglect....................24

1.3.3b. Object-based extinction....................26

1.3.3c. Balint syndrome....................26

1.4 Summary....................28

2. Chapter Two - Object-shape representation and attentional selection....................30

2.1 The computational framework....................31

2.2 Shape primitives....................32

2.1.1 Higher-order parts representations....................35

2.1.2 Edge-based representations....................37

2.3 Frames of Reference....................38

2.4 Levels of Organisation....................40

2.5 Shape representations available to object-based selection mechanisms....................41

2.6 Theoretical Hypotheses....................42

2.6.1 The Globally Structured Representation hypothesis....................42

2.6.2 The Locally structured Representation hypothesis....................42

2.7 Evidence supporting the *global structure* hypothesis....................................44

2.8 Evidence supporting the *local structure* hypothesis....................................45

2.9 Aim of the present thesis....................................49

2.10 Plan of the investigation....................................49

2.10.1 The 'Inhibition of Return' paradigm....................................50

2.10.2 Object-based 'Inhibition of Return'....................................53

2.11 Empirical Predictions....................................57

3. Chapter 3 - Modulation of object-based inhibition of return (IOR) by the internal structure of 2D forms **61**

3.1 Experiment 1

3.1.1 Introduction 62

3.1.2 Method 62

3.1.3 Results and Discussion 67

3.2. Experiment 2

3.2.1 Introduction 71

3.2.2 Method 73

3.2.3 Results and Discussion 77

3.2.4 Discussion 82

4. Chapter 4 - Investigations into the nature of object structure modulating object-based IOR....................................**85**

4.1 Introduction 86

4.1.1 Surface-based representations in object recognition 88

4.1.2 Evidence for volumetric components in object recognition 93

4.2. Experiment 3

4.2.1 Introduction 90

4.2.2 Method 101

4.2.3 Results and Discussion 105

4.2.4 Discussion 108

4.2.5 Summary....................................114

4.3. Experiment 4

4.3.1 Introduction 116

4.3.1.. Predictions 118

4.3.2 Method 119

4.3.3 Results and Discussion 125

4.3.3.1 Object Present condition..125

4.3.3.2 Object Absent condition...129

4.3.3.3 Comparison between Object and No-Object 133

4.3.4 Discussion 135

4.4. Experiment 5 - Modulation of object-based IOR by volumetric component representations

4.4.1 Introduction 137

4.4.2 Method 139

4.4.2.1 Data analysis 144

4.4.3 Results and Discussion 145

4.4.4 Discussion 150

4.5 Experiment 6

4.5.1 Introduction 153

4.5.2 Method 154

4.5.3 Results 158

4.5.4 Discussion 163

5. Chapter 5 – General Discussion and Conclusions **168**

5.1 Introduction 169

5.2 Object-internal structure modulates object-based IOR 170

5.3. Nature of object-shape representations in object-based attention............172

5.4. Generalisation of attention across the whole object................................178

5.5. Comparison between the present findings and previous evidence on part-based modulation of attention...179

5.5.1. IOR versus facilitation......180

5.5.2. Implicit versus explicit measure of selective attention......180

5.5.3. Pattern of modulation of object-based IOR......180

5.6. The effect of display orientation on object-based IOR modulation......184

5.7. Epilogue......188

6. References......190

Abstract

Investigations into visual attention have led to the identification of location- and object-based mechanisms of attentional selection. This thesis is specifically concerned with object-based attention and aims to distinguish between two different hypotheses regarding the representations available to the *inhibitory* mechanisms of object-based selection. According to the *global structure* hypothesis, inhibitory mechanisms of selection operate over representations that *do not* have access to object-internal structural properties, such as surface boundaries. According to the *local structure* hypothesis, inhibitory selection mechanisms operate over representations that *do* make explicit object-internal structure. The second issue addressed in this thesis concerns the *nature* of object-shape representations that object-based inhibitory mechanisms operate over. Two candidate shape primitives, surfaces and volumetric components, are examined in order to ascertain whether they modulate object-based attention.

These issues are addressed using the inhibition of return (IOR) paradigm (Posner & Cohen, 1984), which has previously been used to examine object-based attention (e.g.Tipper, Driver & Weaver, 1991). The two aforementioned hypotheses make contrasting predictions about the modulation of object-based IOR by object-internal structure. The *global structure* hypothesis predicts that object-based IOR should not be modulated by the object's internal structural properties, irrespective of changes in the object's internal structure. In contrast, the *local structure* hypothesis predicts that object internal structure will modulate the magnitude of object-based IOR.

The results raise a number of interesting issues. First, they show that object-based IOR is modulated by internal discontinuities in object structure. Second, object-based IOR operates over representations that make explicit surface properties of volumetric forms. Third, the effect is attenuated when cues and targets appear on the same surface of an object, relative to when the cue and target are separated by an internal structural discontinuity.

These findings are consistent with the *local structure* hypothesis for object-based attentional selection and provide new evidence (a) to suggest that inhibitory mechanisms of selection can operate over shape representations that make explicit information about object internal structure, (b) about the surface-based nature of these mental representations, (c) to posit new constraints on hypotheses about the distribution of facilitation and inhibition in object-based attention.

General Introduction

Our visual system receives large amounts of data that form the basis for our visual experience. On the basis of this visual experience we are able to create a meaningful internal model or else a representation of our external world. At a very early stage in the visual process, our perceptual machinery selects information which is either salient, or relevant to our current intentions or goals (e.g. Moran & Desimone, 1985). Selection of salient or currently relevant information is the principal function of attention. In the contemporary, information-processing theory of attention, selection can occur over multiple mental descriptions or representations of a single stimulus. One of the earliest demonstrations of this comes from a study by Cherry (1953). Cherry had his subjects listen to two different messages, each played to either ear through headphones. When he asked the subjects to ignore one of the messages and concentrate on the other, the subjects could subsequently report little information relating to the ignored message. For example, subjects could not recall any changes in the content of the message or the language in which it was delivered, although they could recall whether the voice changed from male to female or vice versa. Research following Cherry's observations formed the early basis for the information processing approach to psychology, by promoting the role of multiple mental descriptions or representations of extrernal events and the idea that attentional selection can operate over *some* of those representations (Pashler, 1998).

Contemporary investigations into the cognitive mechanisms mediating attentional selection of visual information have concentrated upon two general issues. The first concerns the *way* attention operates; that is whether attention acts to enhance relevant or supress irrelevant visual information. The second issue concerns the *focus* of attentional selection.

By default, the very nature of selection[1] requires the enhancing of processing a relevant over irrelevant information. This process may also be reversed, whereupon selection would involve active suppression of the irrelevant stimulus for the benefit of dedicated processing of the relevant stimulus (e.g. Tipper, 1985). Both facilitation and inhibition of visual information have been observed under carefully controlled laboratory conditions[2]. The review of some of these investigations will form the focus of Chapter One of this thesis.

The second issue relating to attention concerns the nature of the representations over which attentional selection mechanisms operate. Some models of attention have posited that attention operates like a 'spotlight' or a 'zoom lens', exclusively selecting information that falls within the beam of the metaphoric 'spotlight' (e.g. Eriksen & Hoffman, 1972; Posner, 1980). In contrast, more recent accounts of attention posit that it selects from discrete objects in the visual scene, independently of, or inter-dependently with spatial factors (e.g. Duncan, 1984; Vecera & Farah, 1994). This distinction has led to the proposal of space-based and object-based models of attentional selection.

---

[1] The term selection is rather ambiguous. Alan Allport, in a review of 25 years of research into attention, points out the great variety of meanings associated with this term; selection may be taken to mean 'selection for facilitation' 'selection for shutting out any unwanted information', 'selection to make the relevant information available for further cognitive processing', 'selection of information to enter a limited-capacity store and so forth (Allport, 1993, p. 186). Such ambiguity has fueled most current debates about the *locus* (early vs. late selection) and the *focus* (space vs. object) of selection.

[2] In this thesis facilitation and inhibition will be referred to as *mechanisms* of selection. LaBerge (1995) makes the distinction between the expression of attention, which is the facilitation of relevant information, and the mechanisms for the expression of attention, which may be facilitatory or inhibitory (c.f. Umilta, 2001).

The present thesis is specifically concerned with the links between objects and visual attention. In spite of the fact that ample experimental evidence suggests that objects and attention are closely linked, most investigations are constrained by the ambiguous notion of *what constitutes an object*, not only in the field of visual attention but also in the field of object recognition. For example, in the attention literature, an object can be viewed as a single object, i.e. a car; a group of objects linked by Gestalt grouping factors, such as motion, colour, and so on; or a part of an object, i.e. the head on a body.

At the same time, object recognition literature purports that recognition of an object may proceed on the basis of representations, each making explicit, different kinds of information about the object. Much of the debate in this literature focuses on the *nature* of primitive units (building blocks) at each stage of an object's shape representation. For example, some theories posit that objects are represented as an arrangement of volumetric forms, including their spatial relations, this representation being matched against similar representations in memory (Marr & Nishihara, 1978; Biederman, 1987). Other theories propose that surfaces are a more reliable set of primitives for the representation of three-dimensional object shapes (e.g. Gibson, 1979; Marr, 1982). Finally, there are theories that propose that object-shape representation is dependent on low-level elements of the image itself, i.e. contours, edges, vertices, colour or texture homogeneity, creating a two-dimensional representation, which is subsequently matched against the existing object representation in memory (Lowe, 1985; Ullman, 1989).

The present investigation is concerned with the nature of object representations available to inhibitory selection mechanisms. The foundation of this work draws upon, and attempts to integrate, previous empirical research and theoretical proposals in studies of object-based attention, object-shape representation and object recognition.

Chapter One

1. Location-Based and Object-Based Models of Attention

1.1 The focus of attentional selection

What determines the kinds of perceptual information that is selected for further processing? The representational loci of selection in visual attention (that is, the mental representations available for further processing) is a long-lived debate balanced on behavioural, neurophysiological and neuropsychological evidence. Until recently, spatial location was the primary candidate as the medium for attentional selection. According to this *location-based* (also sometimes termed *space-based*) view, attention may be oriented towards locations in the environment (e.g. Eriksen & Eriksen, 1974; Eriksen & St James, 1986; Posner, 1980). On the other hand, central to most *object-based* [3] theories of attention, is the proposal that attention may select from mental representations of objects, independently of their location in the environment (e.g. Baylis & Driver, 1992; Driver & Baylis, 1989; Duncan, 1984; Kramer & Jacobson, 1991; Lavie & Driver, 1996; Tipper, Driver & Weaver, 1991; Vecera, Behrmann & McGoldrick, 2000; Vecera & Farah, 1994). Both location- and object-based theories of attention gain supporting evidence from studies with normal and neuropsychologically impaired individuals.

---

[3] Some theorists (e.g. Vecera & Farah, 1994) differentiate between the term object-based and object-centred attention. The term object-centred attention usually refers to attention to object representations that do not change as a result of transformation of position of the object in the visual field – also termed spatially invariant (Marr, 1982). In contrast, the term object-based attention often refers to attention to objects formed on the basis of gestalt grouping, such as proximity, connectedness, etc. Object-based attention in this sense does not exclude the influence of spatial location in selection. Since this distinction is of no theoretical relevance here, the term object-based attention will be used to denote attention to objects in either sense.

1.2. Attention to locations

The principal claim of location–based theories of attention is that attention selects from topographical representations within a specific region in space. One of the most renowned ways of conceptualising attention is the *spotlight* metaphor (Eriksen & Eriksen, 1974), according to which attention selects perceptual information for further processing *only* if this information falls within the 'beam' of the spotlight. Location-based accounts of attention differ in the way they describe the 'spotlight' and its attributes, i.e. what is the size of the spotlight's 'beam' and what factors determine the next focus of the spotlight; what happens when it shifts from one location to another; is all of the information under the spotlight processed the same way, and so forth (see Cave & Bichot, 1999). Despite of their differences in accounting for above issues, most location-based models agree on the topographical representation of the visual scene.

Most well-documented evidence on location-based attentional selection comes from the effects of location information on spatial cueing and distracter interference tasks (e.g. Posner, 1980; Eriksen & Hoffman, 1972; Eriksen & Eriksen, 1974), on feature integration tasks (e.g. Treisman & Gelade, 1980) and on divided attention tasks (e.g. Hoffman & Nelson, 1981).

The *spatial cueing* paradigm, first used by Posner and his colleagues (Posner, Nissen & Ogden, 1978, cited in Cave & Bichot, 1999; Posner,1980) involves participants making simple key press responses to a target appearing on either side of a fixation point at the centre of the computer screen. When the target appeared at the *same location* as a cue that was presented at various intervals before the target, participants responded faster as opposed to when the target appeared in a different location from the cue. This finding suggested that attention was oriented to the *location* of the cue, and when the target appeared at the same location its processing was facilitated (hence faster reaction times).

Following initial demonstrations of the effect that the location of the cue had on target detection times, other investigators have provided more examples of the way location information affects attentional selection. Hughes and Zimba (1985 cited in Cave & Bichot, 1999) found that targets were responded to faster when they appeared within the same hemifield (up, down, left or right) as the cue, irrespective of the cue-target distance. On the other hand, Downing and Pinker (1985) showed that reaction times increased as the distance between the cue and target increased. However, they also showed that this increase in reaction times was greater when the cue and target were within different hemifields (as in Hughes & Zimba, 1985).

One confounding factor in most spatial cueing studies is that cues and targets appear within squares in various positions around a central fixation point. This is problematic for two reasons: first, because it encourages attention to focus to the specific location of the square, and second, it is not clear whether attention is oriented to the *location* where the cue appeared or to the *object* within which the cued appeared, or even both (Cave & Bichot, 1999). However, recent investigations (e.g. Tipper, Driver & Weaver, 1991), reviewed in the next chapter, have found ways to disentangle these two stimulus properties, namely its location and its identity or form.

The role of location in selection is also demonstrated in tasks where the location of distracter items relative to that of the target determines naming as well as reaction times for that target. C.W. Eriksen and his colleagues (e.g. Eriksen & Hoffman, 1972; Eriksen & Eriksen, 1974) have shown that distracter items produce more interference (longer response times) when they appear near the target (up to 1° of visual angle) than when they appear further away from the target (as far as 5° of visual angle). Eriksen and Eriksen (1974) presented participants with a horizontal array of letters and asked them to identify the one in the middle of the array. The letters belonged to either of two groups of letters, i.e. H and K, or S and C, each group requiring a left or right response. Interference, yielded by slowing of reaction times, occurred when the letters adjacent to the target belong to the set of letters, requiring a different response from the target (e.g. SSS*H*SSS). No competition, however, occurred when the adjacent letters belong to the same set of letters and required the same response (e.g. KKK*H*KKK). This finding led Eriksen and Eriksen (1974) to suggest that interference from distracters surrounding the target occurs at the level of response selection. Importantly, however, and with respect to the role of location in attentional selection, the interference disappeared when the distracters were moved further apart from the target. Subsequent studies have shown that this interference re-appears, when the 'far' distracters and the target are grouped to form a perceptual whole (e.g. Driver & Baylis, 1989).

Visual search studies provide strong support for the important role of location information in selection. In the basic visual search paradigm participants view displays consisting of a *target* item among a number of *distracter* items. The target is only present in half of the trials, while in the other half of the trials there are only distracters in the display. The task is to respond as *fast* and *accurately* as possible to whether the target was present in the display or not. Pertinent to the importance of location when searching for a target, are the results from a variation of the basic visual search paradigm, where participants look for a conjunction target that combines two visual features, i.e. blue circle. Results suggest that, when a task requires integration of visual information from different *visual maps* (implying that visual information in a display is also registered in terms of its position in the spatial layout), such colour and shape, location is the indispensable link between the two maps (Vecera & Farah, 1994). Nissen (1985) argued that attentional selection by location is indeed a unique and necessary aspect of visual processing under conditions of 'cross-referencing' (integration of features from different dimensions, i.e. colour and shape). In two experiments Nissen (1985) showed that, when participants had to use location as a cue to identify the colour of a subsequent target and vice versa, accuracy in identification of either colour or shape was not significantly different. However, when the task was to use a colour cue in order to report the location and shape of a subsequent target, correct responses for the 'shape' attribute were dependent upon correct 'location' responses. It was, thus, inferred that in situations where cross-talk between two separate visual processing maps (colour and shape) is needed, location provides this unique link by allowing access of the target in a particular location in several visual processing maps (Nissen, 1985; Treisman & Gelade, 1980).

Tsal and Lavie (1988) have also found evidence supporting the role of spatial location in visual selection. They used displays consisting of nine coloured letters (each group of three letters coloured in red, green and brown) presented along a circular array, where adjacent letters were never of the same colour. The task was to report a single letter in a certain colour (i.e. a green T) and then to report any other item in the display they could. Tsal and Lavie found that participants were more successful in reporting letters that were adjacent to the first letter they reported (but see van de Heijden, Kurvink, de Lange, de Leeuw & van der Geest (1996) for possible limitations in Tsal and Lavie's (1988) findings). Tsal and Lavie's (1996) results, taken together with other studies (e.g. Cave & Pashler, 1995) suggest that attention to a spatial location is a mandatory process, that occurs irrespective of instructions to select a stimulus on the basis of a different feature (i.e. colour) or time limits that do not allow participants to make eye movements.

In summary, the evidence reviewed here from the spatial cueing, distracter interference and visual search paradigms strongly suggests that internal representations of location-based information play an important role in the allocation of attentional resources across the visual field.

1.2.1. Neuropsychological evidence for location-based attention

Evidence from neuropsychological research also supports the role of spatial location in the allocation of attention. Frequently, cases of neglect have been cited in support of a location-based account of visual attention. Neglect patients often fail to detect and report information about visual, auditory or tactile stimuli presented on the side contralateral to the lesion (see Bisiach & Vallar, 1988 for review). The condition has repeatedly been dissociated from a perceptual or a motor deficit, like hemiplegia or hemianopia in some patients, in the light of neglect cases without either of these deficits (Halligan, Marshall & Wade, 1990). Instead, neglect has, in some cases, been associated with inability to disengage attention from stimuli or events occurring in the ipsilesional side of space (as opposed to inability of moving and re-engaging attention to new stimuli, Posner, Walker, Friedrich & Rafal., 1984). It is generally agreed that the deficit underlying these cases of neglect is related to disruption of some element of selective attention (Bisiach, 1993). In some cases patients show impairment in disengaging attention from stimuli or events in the ipsilesional visual field, in order to re-engage attention to events in the contralesional visual field (e.g. Rafal, 1998; Posner & Peterson, 1990).

The important role of location information in selection has been demonstrated in visual search tasks that require some form of *binding* of different visual features. The notion of binding (Treisman & Gelade, 1980; Treisman, 1988) refers to the way different features of objects, such as colour, shape, location, are correctly combined to produce a coherent visual object. Treisman and colleagues (e.g. Friedman-Hill, Robertson & Treisman 1995; Robertson, Treisman, Friedman-Hill & Grabowecky, 1997) have proposed that location information may play a significant role in binding visual features of an object together (also see Treisman, 1999). For example, Friedman-Hill et al. (1995), investigated a Balint patient's (R.M.) performance on a task requiring conjunction of two visual features across two dimensions (shape and colour). The patient viewed displays of two letters (selected from T, X and O) coloured red, yellow or blue, and was asked to report which letter he saw first. His errors revealed a striking pattern of *illusory conjunctions*, reporting a letter in the colour of another letter in almost 40% of the trials. The high rate of illusory conjunction errors made by the patient, is likely to be explained in terms of impaired spatial representation, which acts to bind features to the relevant object. The results from Friedman-Hill et al.'s (1995) study are consistent with Treisman's idea of the importance of location in feature integration (e.g. Treisman & Gelade, 1980).

1.3 Attention to objects

Recent accounts of attention propose that selection may not necessarily operate solely over spatially defined representations of the visual scene. These accounts propose that selection on the basis of objects is achieved through representations of the object's shape *independently* of where it appears in the visual field. Although empirically differentiating between these two properties of an object, namely its location and its shape description, is not always easy, (since an object's shape occupies particular locations, and being at particular locations defines an object's shape) researchers have succeeded, using carefully designed tasks, to separate an object's identity from its location. A group of theories predict that other grouping factors apart from (e.g. Duncan, 1984; Baylis & Driver, 1993) or alongside spatial location (e.g. Vecera & Farah, 1994; Lamy & Tsal 2000), influence the allocation of attention.

1.3.1 Divided attention tasks

The first to show that one can selectively attend to one of several superimposed objects in the visual field was Neisser (1979) and Neisser and Becklen (1975). He devised a task, where participants viewed two superimposed films and were asked to selectively attend to either one of them, the task being to report information from the attended film. Participants were not only successful in reporting information correctly from the attended film but also remained unaware of unexpected events in the unattended film.

More recently, Duncan (1984) has put forward one of the clearest demonstrations of object-based attention. Duncan presented subjects with displays consisting of two static overlapping objects (Figure 1). The two objects were a box and a line, each of which varied across two dimensions: the box could be long or short and have a gap on its right or on its left side; and the line could be dashed or dotted and be tilted clockwise or counterclockwise. It is important to note that the spatial distance between the two dimensions of one object was the same as the spatial distance between the objects in either dimension. When subjects had to report two attributes of the same object their performance was better (more accurate) than when they had to report one attribute from each object. Duncan found that the objectness rather than the location of the stimuli determined performance on the task. However, various attempts have been made to investigate whether Duncan's results could be attributed to the activation of spatial locations occupied by the object rather than the representation of the object's identity (e.g. Vecera & Farah, 1994). These attempts bear mainly on the issue of whether attention is allocated on spatially invariant[4] object representations or on spatially variant, grouped location-based representations (that is representations of objects in the visual field that are determined by spatial grouping factors).

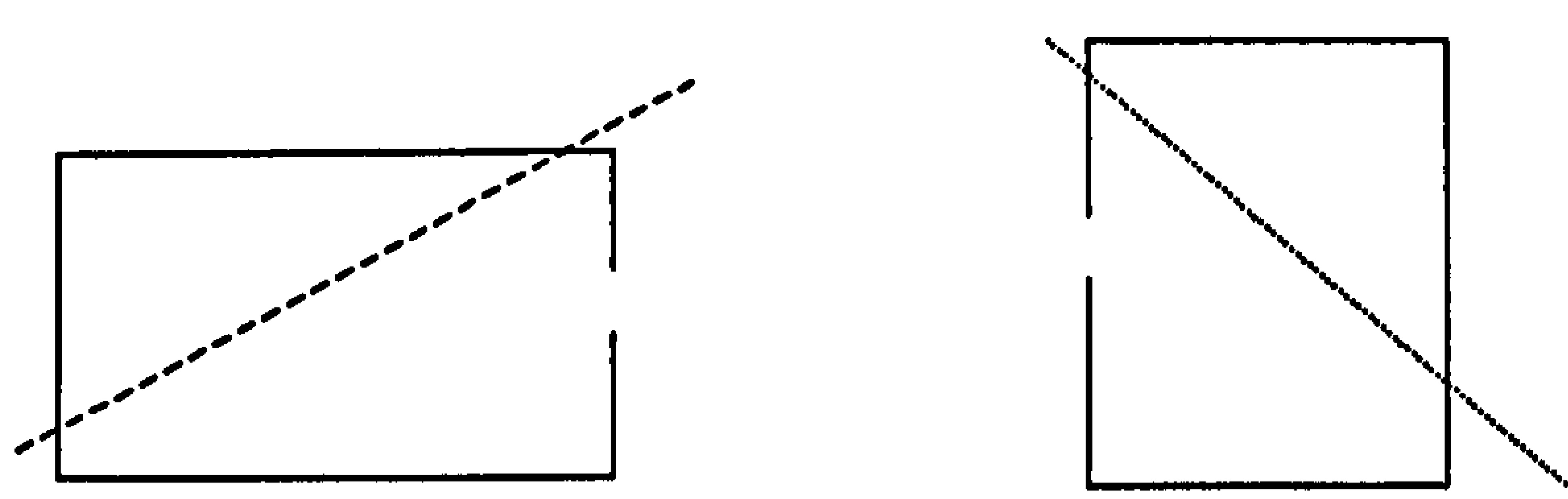


Figure 1. Examples of stimuli used by Duncan (1984) to demonstrate object-based attentional effects.

[4] The term *spatially invariant* is used throughout this thesis to mean the representation of an object's shape regardless of where it appears in the visual field (i.e. as in Marr's visual processing framework, where each object in the 3D model is represented by a certain set of generalised cylinders with specific spatial relations to each other. This 3D model representation remains constant throughout changes in the object's location or orientation).

Vecera and Farah (1994) replicated Duncan's findings using stimuli similar to those used by Duncan (1984), and further examined the underpinnings of the object-based effects found in their experiments. They found that, by instructing participants to ignore the objects and, instead, respond to the onset of targets appearing within the object the object-based effects disappeared. They proposed that selection during this target detection task was made on the basis of spatial representations coded by location, rather than on the basis of the object's spatially invariant representation. To account for these results Vecera and Farah (1994) proposed a *grouped location* account for attentional selection, according to which selection is based upon representation of an object's contours occupying particular spatial locations. This, however, presupposes that the behavioural goal is localisation rather than identification.

Vecera and Farah's (1994, Experiments 3 and 4) raise an important issue with regard to type of task under which object-based effects are found. They propose that only when the task requires identification of an object's properties (thus using the object's shape representation), do object based attentional effects come about. In contrast, when the task does not require responses to object related properties, but instead detection of an onset stimulus, the object effects disappear. Along a similar line of argument, Lavie and Driver (1996) have recently found that object-based limits on divided attention can operate across distances of up to 8 degrees of visual angle. They also found that when participants were instructed to expect targets in a narrow region of the display, the object-based effect disappeared. Thus, visual attentional selection can be location - or object-based depending on the demands of the task.

Baylis and Driver (1993) used a version of Rubin's (1915; c.f. Baylis & Driver, 1993) ambiguous faces-vase figures, where two possible interpretations of the image are possible (Figure 2). This type of display allowed Baylis and Driver (1993) to equate the spatial frequency of the two perceived objects. They used displays that could be perceived either as a central black figure against white background, or as two separate white figures on black background. Observers were asked to compare the height of two angles within the displays. Accuracy was best when the two angled points were perceived (by instruction) as belonging to a single figure, as opposed to when they were perceived to belong to two separate figures. Their findings can be coupled with Duncan's as evidence for object-based selection and against claims of spatial-frequency explanations for these (Duncan's) effects.

Figure 2: Examples of the ambiguous figure-ground displays used by Driver and Baylis (1993). Colour instructions manipulated whether the two central edges were perceived to belong to two separate figures at the extremes of the displays or one figure in the centre of the display. (Adapted from Driver & Baylis, 1998).

1.3.2. Selective attention and spatial cueing tasks

Evidence for object-based selection also comes from studies on selective attention and spatial pre-cueing paradigms. For example, Kramer and Jacobson (1991), using a variation of the flanker task, found that distracters that shared common properties with targets, such as contour or colour, produced larger interference (longer latencies) than distracters that belonged to different objects from the targets. Using the same paradigm Baylis and Driver (1992) found that grouping by good continuation and colour may also facilitate performance.

Despite some claims that object-based effects are revealed in tasks where objects and their properties are task-relevant, a substantial body of research has shown that this may not necessarily be the case (e.g. Tipper, Driver & Weaver, 1991; Egly, Driver & Rafal, 1994; Yantis & Moore, 1995 cited in Watson & Kramer, 1999; Avrahami, 1999). Egly et al. (1994) used displays of two rectangles positioned in parallel with each other and placed vertically or horizontally on either side of a fixation cross (Figure 3). They cued either end of either rectangle by the brightening of the end contours of rectangles. The target subsequently appeared in any of the four ends. The participants' task was to detect the onset of the target by pressing a single key. The displays allowed Egly et al. (1994) to measure facilitatory effects in the cued location as well as in the cued object, thus allowing both attributes of the cue, its location in space, and its location in the object, to be observed. Reaction times to validly cued targets (where the cue and target appeared at the same end) were faster than RTs to invalidly cued targets (where the cue and target appeared at different ends). The comparison of interest, however, in Egly et al.'s experiment was between invalid targets in the rectangle where the cue appeared and the invalid targets in the uncued rectangle. Although both types of targets were an equal distance apart from the cue, Egly et al. found that participants were faster to respond to invalidly cued targets when the target was in the cued rectangle. This was a clear demonstration of attention selecting from mental representations of objects independently of their location in space.

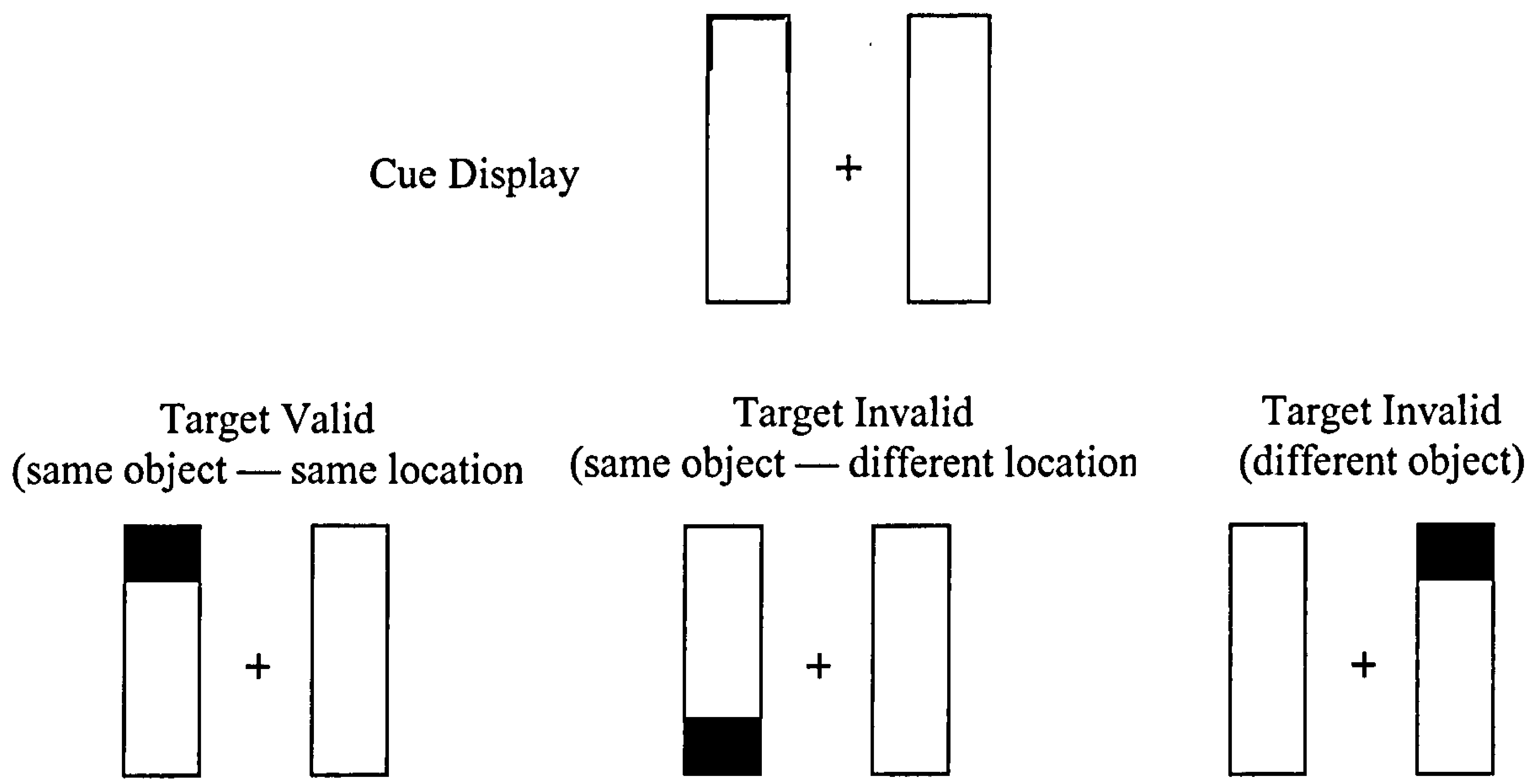


Figure 3: Examples of the displays used by Egly, Driver & Rafal (1994) to demonstrate the same-object advantage in a cueing task. (Adapted from Egly, Driver & Rafal, 1994).

Egly et al. s claim of object-based attention has, nevertheless, been challenged by Vecera (1994) as incorrectly describing what in fact is attention to grouped locations in space. When Vecera (1994) manipulated the distance between the rectangles, he found reduction in the between-object cost in trials where the two objects were closer together. This finding was in accordance to the Vecera and Farah s proposal of attentional selection on the basis *grouped array* representations. This proposal posits that spatial attention may conform to an object s shape by selecting the spatial locations occupied by the object. Furthermore, this proposal predicts that if attentional selection takes place from representations of the object, irrespective of where it appears in the visual field (or where it appears in relation to another visual stimulus), then distance would affect neither the magnitude of the same-object benefit, nor the magnitude of the between-object cost. Thus, Vecera s (1994) finding that the between-object cost reduced with distance, suggests that attention may be directed to a representation of the locations occupied by the object rather than to a spatially invariant representation of the object.

Nevertheless, Egly et al.'s (1994) task has provided a flexible paradigm within which to investigate attentional selection both of locations and objects. For example, Avrahami (1999) used stimulus displays similar to those used by Egly et al (1994) to demonstrate the effect of perceptual objects on the spread of attention. She used ribbon-like stimuli and compared RTs for targets appearing within the cued location or the uncued location of the cued ribbon with RTs to targets appearing within a distance-matched location on the uncued ribbon. She found that RTs to targets within the same ribbon were significantly faster than targets in the uncued ribbon, but only at long SOA. Avrahami (1999) posited that the presence of objects encourages tracing of the objects' contours and long cue-target stimulus-onset asynchronies (SOAs) allow tracing to be completed before attention is oriented to the object, resulting in a same-object advantage (the idea that line tracing takes longer when the line is curved has been previously shown by Joliceour, Ullman & Mackey, 1986, cited in Avrahami, 1999).

Lamy and Tsal (2000) employed a variation of the Egly et al (1994) stimulus displays, in a task where the location of the object and its features, such as colour could be dissociated. They achieved this dissociation by 'moving' the object between the presentation of the cue and the presentation of the target (in their Experiment 2). This way the features occupying the 'target' display need not only be the same features that occupied the 'cue' display. They found that attention may be oriented to the object as well as to the cued location in space. More importantly they found that only when the object features were relevant to the task in hand – that is participants were asked to attend to a certain shape as the target was very likely to appear in that shape- did the participants showed faster reaction times for cued features. On the other hand, the cued object location (location within the cued object) was always attended to, whether or not location is task relevant. Their results reinforce the idea that different types of representations are used depending on task demands.

Similar object-based effects (with static displays) were revealed using the inhibition of return (IOR) effect by Jordan and Tipper (1999). They used displays consisting of two rectangles (similar to those used by Egly et al., 1994) positioned tilted either +45 or-45 degrees relative to the vertical axis. Participants were required to respond to the onset of a target (white square) which could appear following the brief presentation of a peripheral cue, within either end of either rectangle. The results replicated the IOR effect for previously cued locations. Of particular interest were again the invalid trials, where targets appeared in the uncued location of the cued rectangle. Jordan and Tipper found that responses were slower to targets presented at uncued locations on the previously cued rectangle, than to targets appearing at equidistant uncued locations in the uncued rectangle. They claimed that it was the object that was inhibited as opposed to a single location in space. They results taken together with Egly's et al.'s (1994) finding that facilitation spreads across an object, suggest that attentional selection, both in the form of facilitating and in the form of inhibiting performance, takes place from representations relating to the object, rather than to spatial location alone.

Evidence in support of the idea that selection may take place from segmented regions of space (such as surfaces or features grouped by some visual attribute like colour or motion) rather than from arbitrary, object-free regions of space, also comes from cueing studies using three-dimensional displays. He and Nakayama (1995) used stereoscopically separated displays of groups of squares to examine the way attention 'spreads' across surfaces. They asked participants to detect a target, which was the odd square, in the middle depth plane, whilst ignoring items on the other two depth planes. When the items in the three depth planes were tilted so that they appeared to be lying along the plane's surface (Figure 4, Panel a) target detection was facilitated. However, when the items were tilted so they appeared to horizontally bisect the depth planes (Figure 4, Panel b), then target detection latencies increased. In the first case, items appeared to lie across a single surface, whilst in the second case items were not perceived as belonging to a single surface, thus hindering response times, as no depth plane can be selected for search. This finding indicates that attention spreads across individual surfaces.

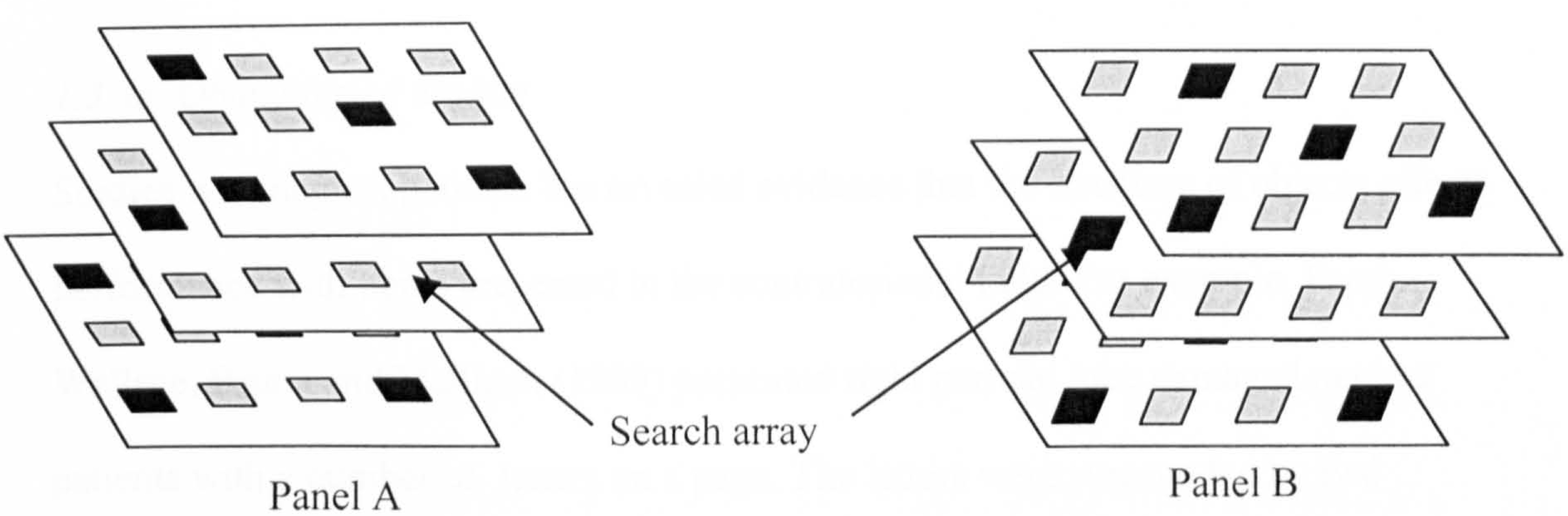


Figure 4: (Adopted from He and Nakayama ,1995). Participants were asked to search the middle plane for the target (a single black square). The items in Panel A lie along the surface that corresponds to that plane. In contrast, items in Panel B lie along a surface that vertically bisects the middle plane. Target detection times are slower for Panel B displays, where attention cannot select a single surface for search.

1.3.3. Neuropsychological evidence for object-based attention

Evidence from neuropsychology has also been accumulated to support the idea that selection may also be based on object representations. As mentioned earlier, disorders like unilateral neglect and extinction have often been associated with inability of the patient to attend to stimuli presented on side of space opposite the side of lesion (e.g. Rafal, 1998). This contralesional nature of neglect and extinction has repeatedly been explained, in some cases, in terms of impaired spatial attention, as patients ignore all stimuli (visual, auditory or tactile) within the affected hemifield. Some recent evidence from case studies of neglect and extinction, however, has shown that these disorders may also be object-based, in the sense that patients ignore the halves of objects on the contralesional field relative to the salient vertical axis of the object itself irrespective of where the object appears in the visual field (e.g. Driver, Baylis, Goodrich & Rafal, 1994; Driver & Halligan, 1991; Ward, Goodrich & Driver, 1994).

1.3.3a. Object-based neglect

Studies with neglect patients has revealed evidence that the structure of objects affects performance with items presented in the contralesional field. For example, Farah, Wallace, Brunn and Madigan (1989) presented right parietal lobe damaged neglect patients with a number of letters on a page. The letters were surrounded by two elliptic circles (like parentheses) that were placed either horizontally above and below the letters or vertically on the left and on the right of the letters. More letters were read in the contralesional side of the page when the letters were surrounded by horizontal eclipses, that when they were surrounded by the vertical ones. These results pointed to the possibility of modulation of unilateral neglect by grouping elements of the contralesional field, so they can be perceived as a single object (the horizontal eclipses served to connect the letters into a single object).

Driver and Halligan (1991) presented their patient with two nonsense figures – placed one above the other - and asked them to report whether they are same or different. When the two shapes were upright the patient missed any differences on the left side of the object. However, the patient still missed any differences on the left side of the object even when the two shapes were tilted 45 degrees, so that the top left of the object fell into the patient's right field. The patient's neglect had apparently 'moved' with the object's axis of elongation. Space-based accounts of attention alone cannot explain this failure to attend to features appearing on the unimpaired (ipsilesional) side of space. In contrast, this finding may be explained by object-based accounts of attention that predict that attention selects from objects irrespective of their location in the visual field. Therefore, failure to attend to visual information may be determined by object-based information , such as the object's axis of elongation.

Behrmann and Tipper (1994) provided further support for object-based frames of reference in neglect. In a reaction time experiment they presented their patients with displays consisting of a green and a red circle connected with a line, forming a barbell. Patients with right parietal lesions failed to see any circle in the left side. When the stimulus (barbell) was rotated 180 degrees, the patients still neglected the circle that they had fail to see in the left side of space which was now on the right side of space. Thus, the neglected side moved with the object as it rotated. Tipper and Behrmann (1996) also showed that this finding was only observed when the two circles were connected with a bar. When the two circles were presented as separate objects, neglect remained with the impaired (left) visual field. Taken together the above studies (Driver & Halligan, 1991; Behrmann & Tipper, 1994; Tipper & Behrmann, 1996) indicate that, depending on stimulus presentation conditions (a single object as opposed to two separate objects) attention may operate on location as well as object-based representations.

1.3.3b. Object-based extinction

The study of *extinction* has also provided evidence against a strict spatial bias against contralesional stimuli following parietal brain damage. Patients with extinction show impaired ability to detect stimuli or events on the contralesional or ipsilesional side of space, when another stimulus is simultaneously presented on either visual field. This inability can be ameliorated when the two stimuli are connected to form a single stimulus (e.g. Driver, Mattingley, Rorden & Davis, 1997, Ward, Goodrich & Driver, 1994). For example, Ward and colleagues have shown that extinction can be reduced by placing a familiar object in the contralesional field instead of a nonsense drawing, suggesting that familiar objects can survive the competition with a simultaneously presented object (Ward & Goodrich, 1994) or by grouping two stimuli by means of good continuation, hence the two stimuli form one single object (Ward, Goodrich & Driver, 1994).

1.3.3c. Balint syndrome

Studies with Balint syndrome patients provide some of the strongest evidence for accounts of object-based attentional selection. Balint syndrome is typically caused by (often bilateral) posterior parietal lesions and its main manifestation is the patient's inability to see more than one object at a time (also known as *simultagnosia*), or make any comparisons between two objects, irrespective of the size of the object (Rafal, 1997).

The object-based nature of attentional impairment in Balint syndrome was demonstrated early in the century by Holmes and Horax (1919; c.f. Scholl, 2001). Their Balint patients were unable to judge whether two straight lines were of equal length or not, whilst the same patients had no difficulty identifying a shape as a rectangle or a trapezoid (a decision requiring intact perception of the lines), where the lines would make up a single object.

Luria (1959) showed that when the patients were presented with the classic 'Star of David' figure (a star made up from two differently coloured triangles), they could recognise one triangle at a time but not the star. Furthermore, Luria (1959) showed Balint patients can report two objects in a display at a time when these objects are joined in some way to form a single object. He showed that when the patients are presented with two adjacent circles, they only report the presence of one. However, when a line is drawn to connect the two circles, patients report the presence of a bar-bell (or sometimes a pair of spectacles).

More recently, Humphreys and Riddoch (1993) have extended Luria's findings with two Balint patients. They presented their patients with displays of thirty-two circles that were in the same colour (red or green) or half were red and half were green. The displays also contained straight black lines that were either randomly placed amongst the circles or joined pairs of circles to form 'barbells'. The patients were simply asked to report whether the circles in the displays were all the same colour or not. Critically, in the two-colour displays, the patients were better at reporting whether the displays consisted of circles of a single or two different colours, when a line joined two differently coloured circles, to form a single object (a bar-bell). In contrast, when the lines connected two same-coloured circles, patients had greater difficulty reporting the presence of two different colours in the two-colour displays.

1.4 Summary

The primary focus in this chapter was the distinction between location-based and object-based accounts for attention. This issue has also been characterised in terms of the amount of visual processing before attentional selection takes place (or *pre-attentive* processing; Driver & Baylis, 1998). Each account gains supporting evidence from paradigms and effects in cognitive psychology and neuropsychology. There now exists enough evidence to suggest that selection may take place on the grounds of location-based representations, i.e. where something is, and on the grounds of spatially invariant object-based representations. Which kind of mental representation is used seems to depend on the experimental paradigm, the task instructions that define the participants intentions, as well as on the available visual information in the displays. Furthermore, some attempts have been made to integrate location-based and object-based accounts of attentional selection. One such attempt is Logan's (1997) CODE theory of attention, where selection operates over spatial representations of the visual field, but changes in the spatial distribution are determined by object representations.

One issue raised from the evidence on object-based attention concerns the specific kinds of shape representations that mediate object-based selection. That is, although it is now clear that attention may operate over representations of objects, there is little empirical evidence on the nature of these object-based representations from which attention may select. To complicate matters more, there is little consensus on what qualifies as an object (e.g. Scholl, Pylyshyn & Feldman, 2001; Avrahami, 1999 for more detailed discussions). Thus, whilst in some investigations an object is defined on the basis of Gestalt principles of grouping, such as proximity, closure, good continuation (i.e. Duncan, 1984; Tipper & Behrmann, 1996; Kramer & Jacobson, 1991), other investigations define an object on the basis of figure-ground segmentation (i.e. Baylis & Driver, 1993, reviewed later in the thesis). The present research aims to investigate the kinds of object-based representations selected by attention, on the basis of different proposals about the kinds of representations assumed to mediate object-shape perception. Ultimately this thesis attempts to draw together theoretical proposals from object-based attention and object shape perception. In the following chapter proposals from theories of object-shape representation are briefly outlined.

Chapter Two

2. Object-Shape Representations and Attentional Selection

2.1 The computational framework

One of the primary purposes of visual processing is to represent objects for the purpose of identification and action. This purpose is achieved by creating a mental representation of an external visual object in such a way that it can be compared with that object's representation stored in long-term memory or be used to guide action, such reaching and grasping. The visual system needs to extract certain clues from the two-dimensional retinal image and compute a mental representation of the perceived object by reconstructing the 3D scene (Marr, 1982). This view of reconstruction of the external world formed the basis for a general framework of visual perception, according to which the visual system achieves the task of perceiving[5] an object by proceeding from representations related to two-dimensional image structure to representations related to the three-dimensional object structure. This thesis adopts the computational framework derived from the work of Marr & Nishihara (1978), according to which perception of a visual object proceeds in stages of representation, each stage using different types of primitive elements (see Figure 5).

Theories of shape representation make different predictions about three issues with regard to the nature of shape description at each level of representation. First, there is the issue of *shape primitives* made explicit at each level of shape representation. Second, there is the issue of the *frames of reference* applied to these primitives. Third, is the issue of the *organisation* of shape primitives at each level of representation.

[5] The term perceiving (or generally perception) is used in a variety of ways and contexts. For example, sometimes the term simply refers to processing of sensory input; whilst sometimes refers to deriving meaning and significance from the sensory input. In the latter sense, perception is seen in the representational sense, aiming to facilitate recognition (Goodale, 1995).

2.2 Shape primitives

At each stage of representation different primitives are made explicit. In the first stage of object shape representation, simple, low-level features, such as local edges, line terminations, luminance boundaries, are detected in the two-dimensional retinal image. In the next stage of visual processing of an image surface properties are detected in order to compute an intermediate, two-dimensional representation of surfaces. Such properties include the orientation, slant and distance of the visible surface from the viewer. Finally, a higher-level representation of the object is computed to facilitate recognition of objects in the environment under distorted conditions such as partial occlusion, overlap between objects, distance or unusual views (Figure 5).

Although there is general agreement that these stages occur during recognition, different theories of object recognition propose for different representations as the primary ones to inform the recognition process. For example, *structural description* theories predict that objects are represented as an arrangement of viewpoint-invariant volumetric components or higher-order parts *and* their spatial relations, and it is their representation is matched against similar representations in memory (Marr & Nishihara, 1978; Biederman, 1987). On the other hand, there is evidence that surfaces are a necessary set of primitives for the representation of 3D object. For instance, in some models of shape representation shape can be recovered on the basis of surface properties alone, such as colour or shading (e.g. Koenderinck, 1990, cited from Palmer, 1999).

In contrast, *image*-based, viewpoint-specific models propose that object-shape representation is dependent on representations of low level visual features, i.e.

contours, edges, vertices or colour. These features are immediately derivable from the retinal image and form a two-dimensional, viewpoint-specific representation, which is subsequently matched against the existing object shape representation in long-term memory (e.g. Lowe, 1985; Ullman, 1989; Edelman & Bulthoff, 1992; Tarr & Bulthoff, 1995).

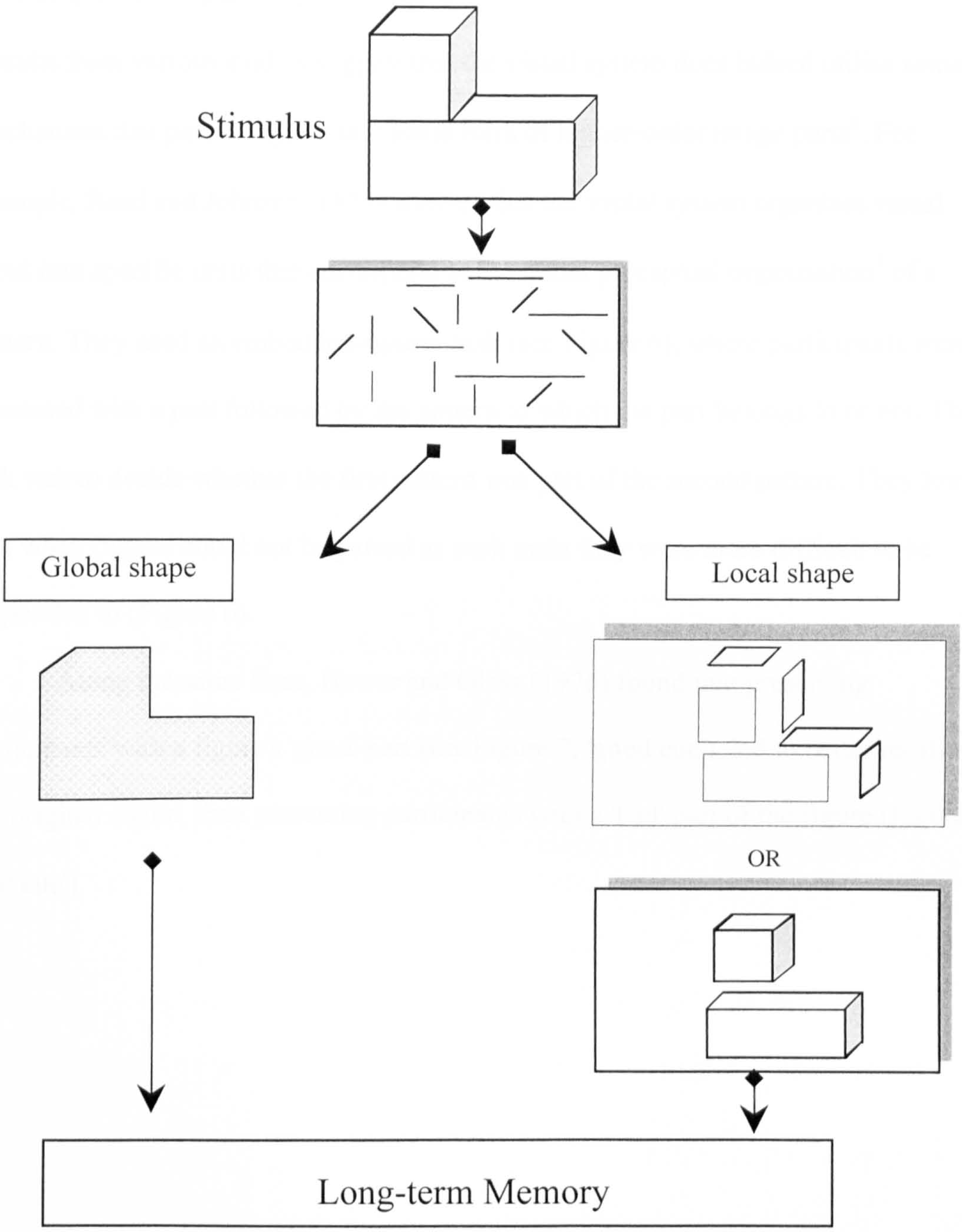


Figure 5: An illustration of the representational stages in perception loosely based on Marr's (1982) general theoretical framework. An edge-based (or primal sketch) representation makes explicit information about the 2D image, i.e. edges, boundaries, line terminations and their geometrical organisation. An intermediate-level representation is computed making explicit properties of visible surfaces, i.e. their orientation, distance from the viewer; or properties of the object's volumetric components. A 3D model representation can also be computed to describe the object's global shape, without making explicit the object's internal structure. (Also see Figure 8).

2.1.1 Higher-order parts representations

Results from various studies suggest that the visual system does indeed utilise some mechanism that parses objects into some form of higher-order image parts[6]. For example, Reed and Johnsen (1975) showed that the visual system organises visual input into specific units that correspond to the initial perceptual organisation[7] of a pattern. They used an embedded-figures task (see Figure 6), where participants were presented with a part followed by the pattern to which the part belongs to or not. The task was to decide whether the first pattern was part of the second pattern. They found that when figures could not be parsed as such units they were more difficult to be responded to (Figure 6).

Along the same lines, Bower and Glass (1976) found that presenting participants with a figure's 'good' part (i.e. Figure 7, 'good cue') led to better recall of the original figure, than presenting participants with a 'bad' part of the figure (Figure 7, 'bad cue').

[6] The term 'higher-order parts', as described in this section does not differentiate between volumetric components or surfaces. The emphasis is placed on the idea that a representation of the object is created on the basis of the objects' component parts, rather than the lines or vertices of the object.

[7] Reed and Johnsen (1975) do not specify the factors that influence the initial perceptual organisation of a pattern. As an example of initial perceptual organisation, they suggest that if the Star of David pattern is initially organised into two triangles, then observers should be able to recognise the triangle as part of the 'Star of David', but perhaps not the rhombus.

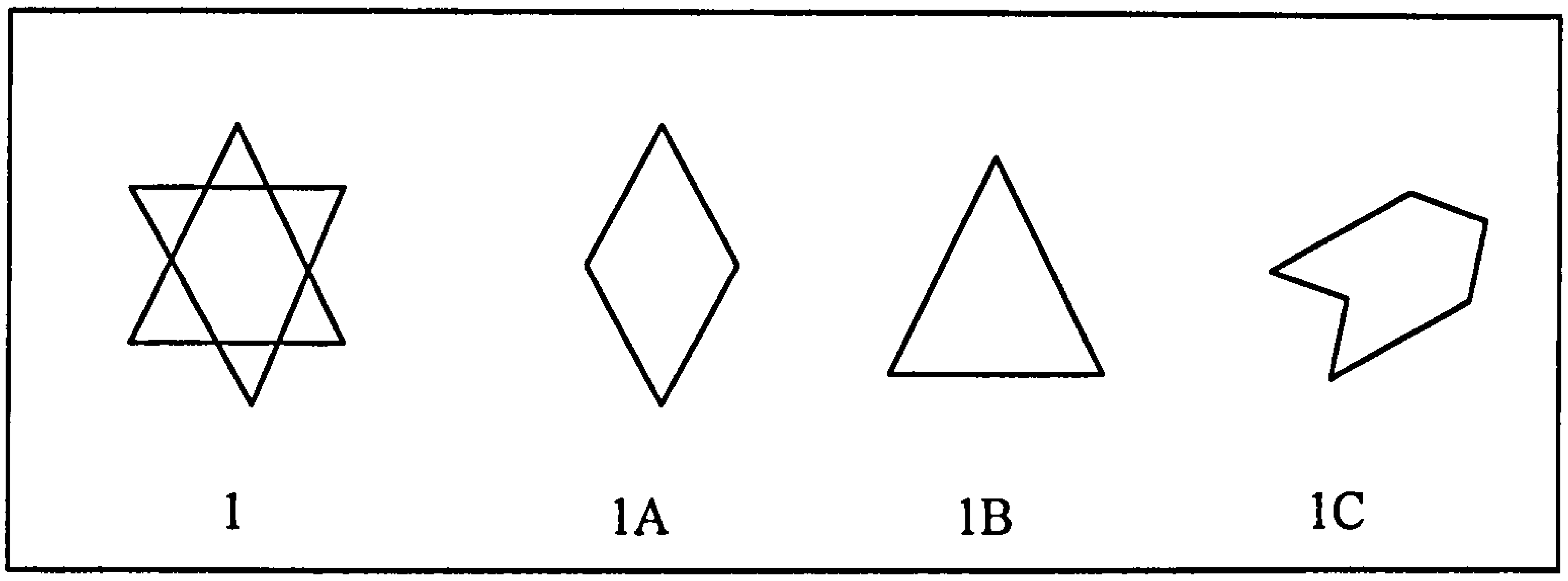


Figure 6. Some of the stimuli used in Experiment 1 by Reed & Johnsen (1975). Participants were asked to decide whether patterns 1A, 1B or 1C were part of pattern 1. Responses were faster and more accurate when deciding on pattern 1B as opposed to patterns 1A or 1C. Adopted from Reed and Johnsen (1977, p.570)

Bower and Glass proposed that a *good* cue is "one or more salient subparts as dictated by our intuitions and parsing rules" (Bower & Glass, 1976, p.459). On the contrary, in the case of *bad* cues, the "parsing rules assign to them in structural units that do not correspond to the units in the representation of the original pattern" (as above).

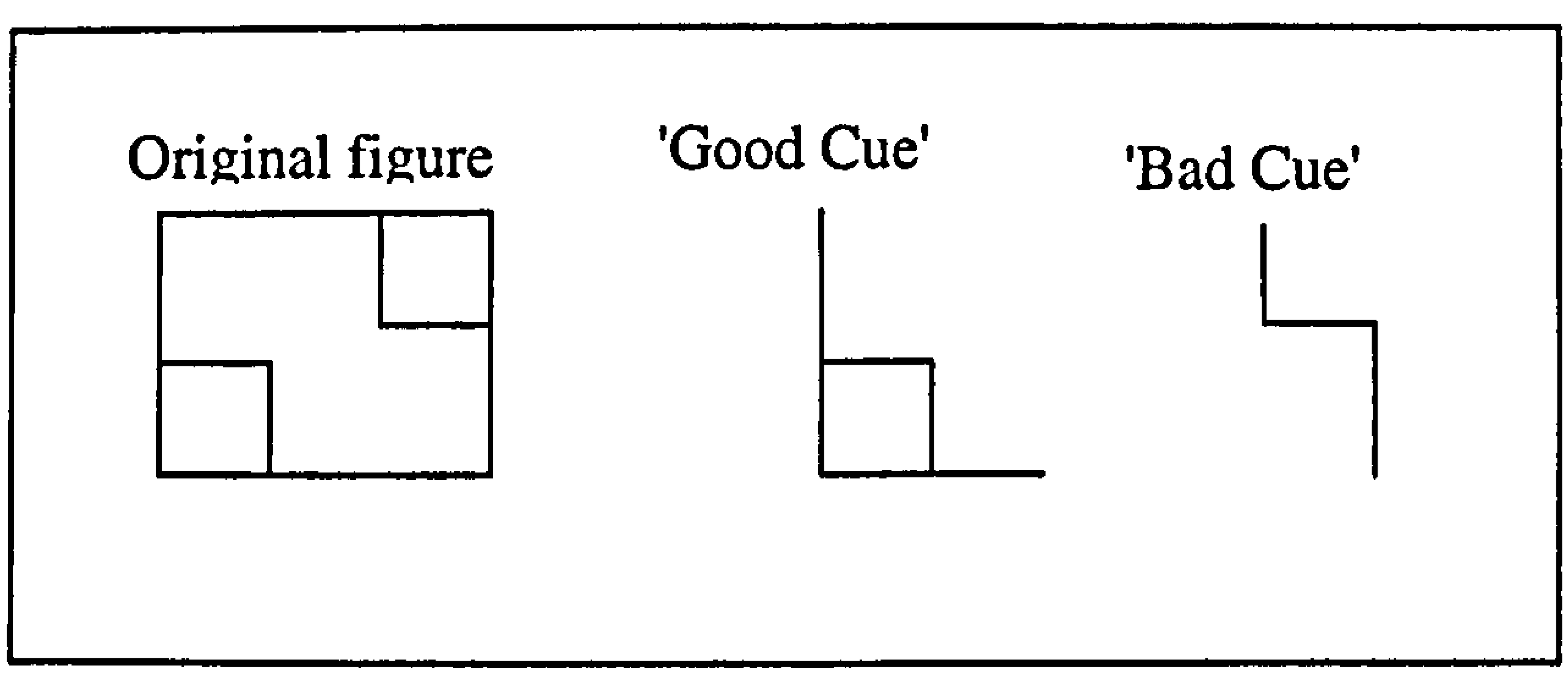


Figure 7. An illustration of the stimuli used by Bower & Glass (1976). The *good* cue is comprised of one or more salient parts of the original figure. In contrast, *bad* cues "are misleading in the sense that the parsing rules assign to them structural units that do not correspond in any way to the representation of the original pattern" (Bower & Glass, 1976, p. 459). (Drawing adapted from Bower & Glass, 1976).

Bower and Glass's study along with other previous studies (e.g. Fodor, Bever & Garrett, 1974; cited in Bower & Glass, 1976; Palmer, 1975), were some of the first to draw attention (in the everyday meaning of the word) to the role of structural units in a visual pattern *and* their hierarchical relation for its representation in memory. As if in anticipation of Biederman's Recognition-by-Components (RBC) theory of object recognition (Biederman, 1987), Bower and Glass suggested that such kind of research "would form the basis of a *grammar* of visual shapes to complement the linguist's grammar of language" (Bower & Glass, 1976, p.465).

2.1.2 Edge-based representations

Despite the apparent evidence for the importance of a figure's higher-order components for the mental representation of an object's shape, image-based accounts of object shape representation posit that object shape can be represented by the 2D features on an image, without any intermediate representations of higher-order parts (e.g. Lowe, 1985; Kosslyn, Flynn, Amsterdam & Wang, 1990; Tarr & Pinker, 1989).

This group of theories propose a different interpretation to the results obtained by proponents of structural description theories. That is, they posit that, if the object's image is indeed organised into higher-order parts, this may not necessarily imply that these parts play a role in object representation. Priming effects (in the sense of facilitation of performance, when the picture of an item facilitates naming or recognition of a subsequently presented item that is either identical or of the same category as the first item) from objects' parts (e.g. Biederman & Cooper, 1991) may occur after recognition has taken place, which does not necessarily indicate that objects are first parsed into parts as a prerequisite for the object's representation (Backer Cave & Kosslyn, 1993).

Backer-Cave and Kosslyn (1993, Exp. 1 & 2) challenged previous conclusions about the importance of component part-based representations in recognition. They used line-drawings of everyday objects, such as spectacles, scissors or telephones, and parsed them following various segmentation rules (i.e. 'breaking' the objects into their natural or un-natural parts [i.e. consisting of the bottom left corner and top end of the leg of a chair] and maintaining or disrupting their spatial configuration). They found that disrupting the spatial relations among object parts resulted in longer naming times than disrupting object parts themselves. Their results were predicted by the account proposed by Lowe (1985), according to which disruption of an object's spatial configuration, that is spatial relations between its component parts, would disrupt the object's shape, obscuring its representation, and thus resulting in long naming latencies. They concluded that a representation of the outline shape of the stimulus is encoded first followed by representation of its constituent parts.

2.3. Frames of Reference

Another issue in shape representation literature concerns the *frames of reference* that the visual system applies to these primitives, when representing the object. That is, in order to represent external objects, the visual system may apply some kind of co-ordinate system within which features of this object, i.e. edges, surfaces or volumetric components, are represented relative to the viewer's body, head or retina (called body-centred, head-centred and retinocentric frames of reference, respectively). This class of reference frames is known as *viewer-centred reference frames* and are applied to visible surfaces extending to 3D space. Alternatively, the object may be represented in relation to the gravitational axis, or in relation to other object in the environment. These are termed *environmental reference frames*. Both viewer-centred and environmental frames of reference are *external* to the object and the object's coding in terms of these reference frames results in that object's representation in long-term memory at one particular orientation.

In contrast, an object and its individual components can be assigned to an *object-centred reference frame,* a co-ordinate system whose reference axis is an *intrinsic* property of the object itself, for example its axis of elongation or some distinctive feature. This co-ordinate system defines the spatial relationship between the objects' edges, surfaces or volumetric components, and may be used independently of other frames of reference, such the gravitational axis or the viewer's main body axis, when representing the object. The frame of reference used in the representation of an object depends largely shape primitives assumed to represent the object (e.g. Marr, 1982).

An important distinction that should be made here is that between object-centred reference frames and object-based attention; in other words attention may select objects on the basis of perceptual grouping (such as a collection of edges or other simple visual features grouped by common colour, orientation, continuation) thus still applying a viewer-centred reference frame; or attention may operate over object shape representations, totally independently of their spatial locations (i.e. in the case of overlapping objects), thus applying an object-centred reference frame, where any of the object's intrinsic axes is applied to the representation[8]. This distinction is important as it speaks to the distinction made earlier (see footnote 1) between object-*based* and object-*centred* attention (e.g. Vecera & Farah, 1994; Goldsmith, 1998). Despite the fact that in this thesis these two terms are not differentiated (they are both called object-*based* attention), the distinction between object-based and object-centred attention may prove to be important when theorising about object shape representation, since the latter assumes attention to a three-dimensionally defined object shape.

2.4 Levels of Organisation

The third issue in theories of shape representation is the *organisation* of information imposed by the representation. Each stage of shape representation imposes different organisation on the object's shape. For example, at the earlier stage of low-level feature representation all elements (edges, discontinuities, line terminations etc.) have the same organisational status. However, at later stages of representation a hierarchical organisation of primitives becomes possible. Such hierarchical organisation allows the description of object shape at different levels of specificity (depending on the scope of the representation). For instance, a human body can be described by a single *global* volumetric primitive but it can also be decomposed into smaller shape primitives that also make the local structure of the object explicit, i.e. from the arm to the forearm and

[8] The issue of reference frames in object shape representation is very broad to cover here comprehensively. It is important, however, to point out its relevance to theoretical issues regarding the kind of primitive shapes used to represent object shape. For instance, whether objects are represented as 'wholes' or by their parts may interact (in the sense of either one influencing the other) with the frame of reference applied to such representations (also see McCloskey, 2001).

the hand (Figure 8). The advantage of such hierarchical organisation is that local spatial relations can be represented separately from global ones (Marr, 1982).

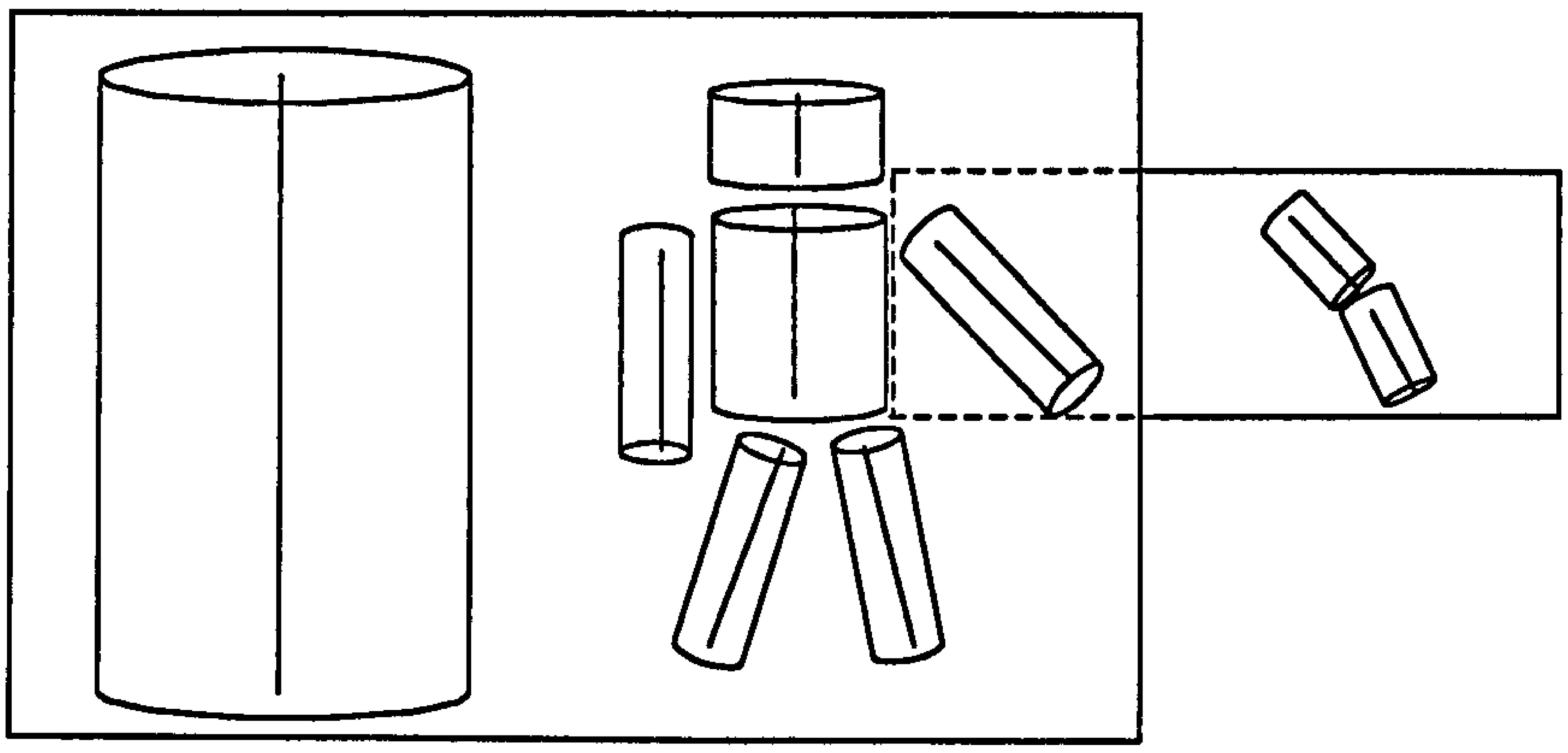


Figure 8: (Adopted from Marr, 1982). An illustration of the organisation of shape information in the 3D-model representation. The human body shape can be represented by a single volumetric primitive that makes explicit the object's global shape, size, orientation. Alternatively, the human body can also be described at a more local level of specificity, as an arrangement of the major components along their own axis of elongation.

2.5 Shape representations available to object-based selection mechanisms

Proposals about shape representations in visual perception motivate a number of hypotheses about the kinds of representations that may be available for object-based selection. Object shape may be represented on the basis of the object's global structural properties, such as its size and relative volume; or it may be represented on the basis of local shape descriptions, such as the individual components (surfaces or volumetric components) of the object. The present investigation focuses on two contrasting hypotheses regarding the nature of object shape representations that inhibitory attentional mechanisms operate over.

2.6 Theoretical Hypotheses

2.6.1 The Globally Structured Representation hypothesis

According to this hypothesis, object-based selection operates *solely* over globally structured object-shape representations, that is from representations that *do not* make explicit any information about internal shape properties. This hypothesis (hereafter referred to as the *global structure* hypothesis) is illustrated in the bottom panel of Figure 9.

2.6.2 The Locally Structured Representation hypothesis

A contrasting hypothesis predicts that object-based selection operates over internally structured representations that *do* make explicit internal shape properties *as well as* the object's global shape. This hypothesis allows for simultaneous availability of local *and* global structure information, as global structure can emerge from local structure representations. An illustration of this hypothesis (hereafter referred to as the *local structure* hypothesis) is shown in the top panel in Figure 9.

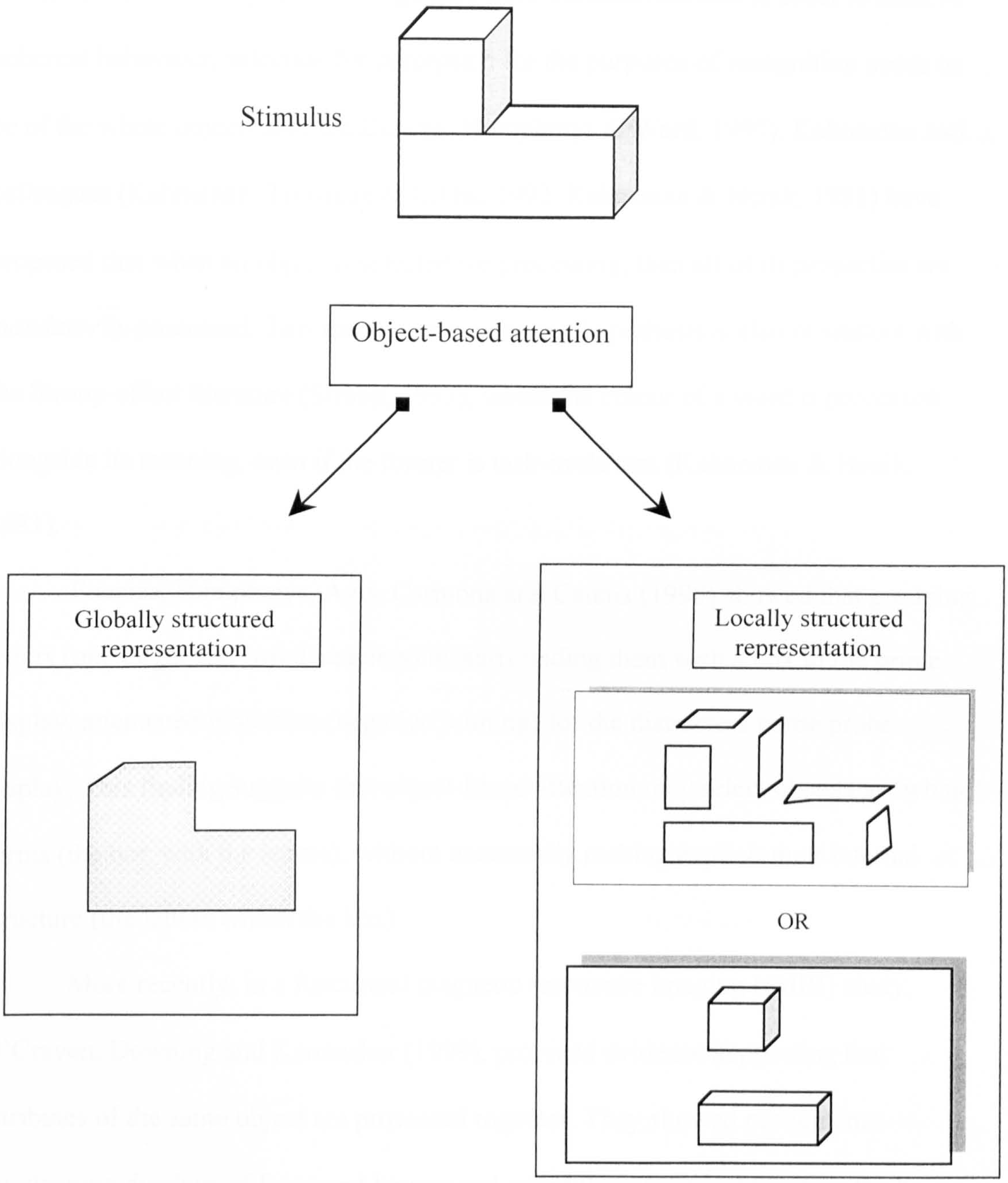


Figure 9: Schematic illustration of the two theoretical hypotheses contrasted in this thesis.

2.7. Evidence supporting the *global structure* hypothesis

The *global structure* hypothesis originates from the assertion that in order to achieve coherent behaviour, selection for perception for the purposes of recognition needs to be of the whole object (also see Duncan, Humphreys & Ward, 1997). Kahneman and colleagues (Kahneman, Treisman & Gibbs, 1992; Kahneman & Henik, 1981) have proposed that when an object is selected for processing, then all of its properties are *mandatorily* processed. This mandatory processing hypothesis is also consistent with the Stroop-effect literature (Stroop, 1935), where the colour of a word is processed alongside its meaning, even if the former is task-irrelevant (Kahneman & Henik, 1981).

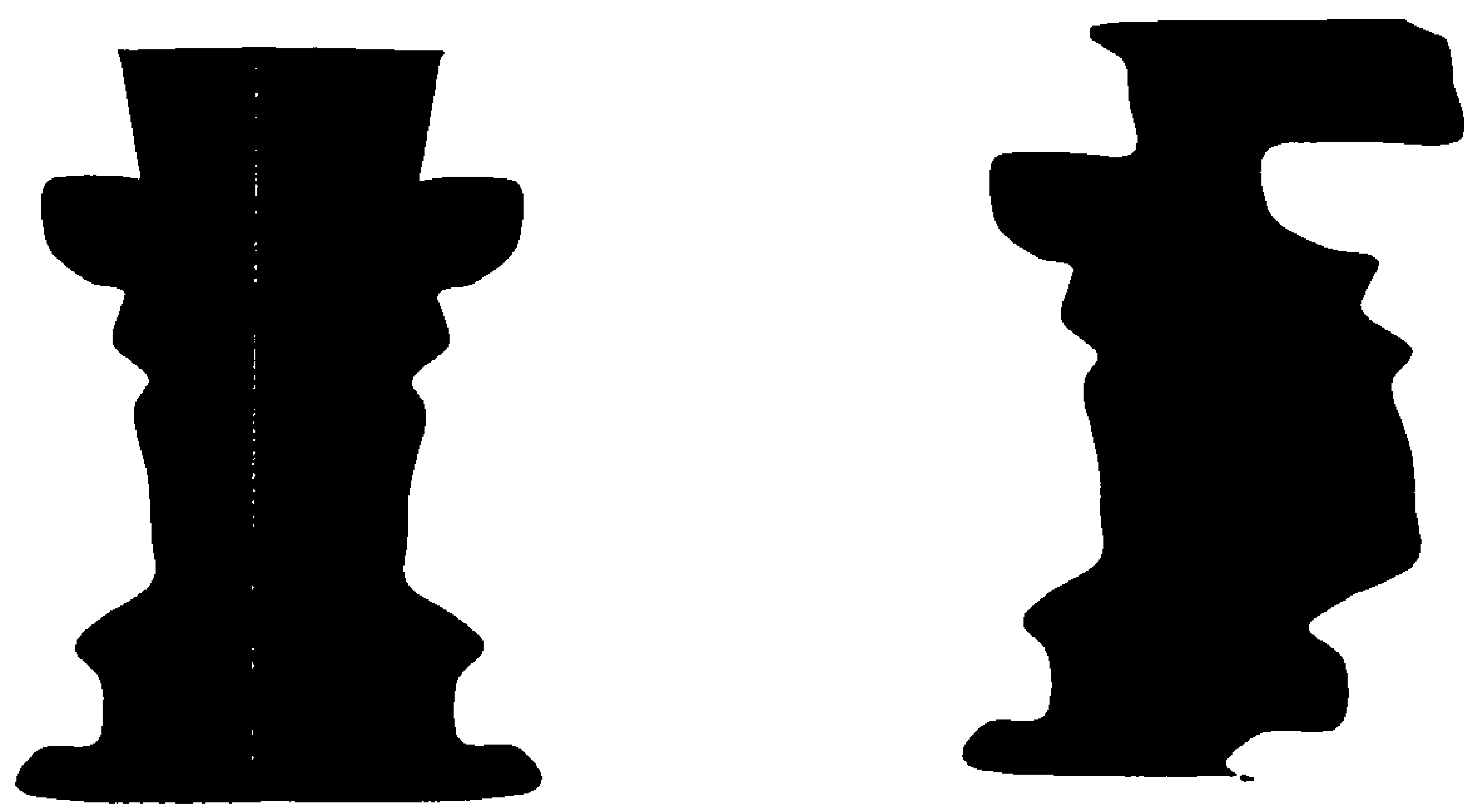
Fuentes, Humphreys, Agis, Carmona and Catena (1998) showed that grouping letters (one target and two distracters) by surrounding them with a box in the prime display, attenuated inhibition (negative priming) for the distracters in the probe display. This finding suggests that object-based attention may select objects as 'whole' forms (the box with the letters), without necessarily making explicit their internal structure (the letters within the box).

More recently, in a functional magnetic resonance imaging (fMRI) study, O'Craven, Downing and Kanwisher (1999), provided evidence supporting that attributes of the same object are processed together. They showed participants overlapping displays of faces and houses and asked them to attend to either the house, the face, or the direction of motion. The task was to report whether the house, the face or the direction of motion was repeated in two consecutive trials. Results showed that attending to one attribute (i.e. motion) resulted in more activation in the fusiform face area (FFA) when the face was moving, than when the houses were moving (and the faces remained stationary). They suggested that "even when the task requires only that the subjects select a given visual attribute, both attributes of the attended object are automatically selected" (O' Craven et al, 1999, p.586).

2.8 Evidence supporting the *local structure* hypothesis

In support of the alternative, *local structure*, hypothesis, according to which object-based attention selects from locally structured representations, there is a series of investigations that have been exploring the idea that object-based attention can also be 'sensitive' to representations of the parts that compose the object, and their spatial relations. For example, work with neglect patients by Humphreys and Riddoch (1994) has provided evidence for attentional modulation by between-object as well as within-object representations. Their findings indicate that attention can selectively ignore (neglect) representations of the internal structure of an object (e.g. neglect for the right side of a plate) as well as representations of spatial relations between objects (e.g. left/right visual field neglect). However, their findings did not relate to the *nature* of these within-object representations that are selectively attended to by object-based attention.

The issue of the specific types of shape descriptions selected by object-based attention was also addressed in a series of divided attention experiments by Driver and Baylis (1995). Using figure-ground displays, where the task was to judge symmetry and repetition of contours. They found that making symmetry judgements about two edges in displays like those in Figure 10, Panel A, was easier than the symmetry judgements in figures like the one in Panel B. Their findings bear important implications for accounts of part-based attention for at least two reasons. First, the pattern of accuracy in symmetry judgements indicates that participants did not focus their attention to the edges per se (as they were instructed to) but took into account the whole object defined by those edges. Second, the finding that accuracy was better when both edges to be compared were convex (Panel A), is in agreement with Hoffman and Richards (1984) claims that an object can be represented in terms of its parts, defined by regions separated at points of concavity (Figure 10).



(A) (B)

Figure 10: Examples of stimuli adapted from Driver and Baylis (1995). Participants found it easier to compare the contours in (A) than in (B).

In a different line of research, object shape representations mediating attention are assumed to specify object shape at different levels of specificity, that is from global to local feature structure (e.g. Navon, 1977, Palmer, 1977). Global-local research has demonstrated that each hierarchical element of an object can be attended within a hierarchically organised stimulus (i.e. a larger letter composed of smaller letters) and influence behaviour. Despite the important theoretical implications of findings in global-local literature (attention can be directed to different levels of representation of a stimulus), these do not speak to the specific nature of the part-based representations, as the smaller letters comprising the large letter cannot be considered structural parts of the letter's *shape*. To illustrate, most observers would agree that the shape of the letter E corresponds to the structural description of four distinct parts - one vertical line horizontally crossed at equally spaced intervals by three small lines of equal length (this point is also raised in Vecera, Behrmann & Filapek, 2001, p309).

The nature of object shape representations available to attention was also examined more closely in four experiments by Watson and Kramer (1999). Their experiments are based on the principle of *uniform connectedness*[9] (Palmer & Rock, 1994, cited in Palmer, 1999), where areas with uniform visual properties, such as colour, shading, texture, are organised into a single perceptual unit. They showed participants objects (wrenches) that were either uniformly connected or not. They found that depending on the task demands they would find a same-object attentional benefit both for parts of the wrenches and for the whole wrenches.

Watson and Kramer (1999, Experiments 3 and 4) also found that, when they presented participants with wrenches consisting of either well-defined or poorly defined parts, object-based attention (same-object benefit) effect was modulated by the goodness of the parts. Poorly defined parts (defined by the absence of concave discontinuities) showed a larger object-based attention effect than their well defined (defined by the presence of clear concave discontinuities) counterparts. That is, the absence of clear concavities encouraged participants to locate attention to the object as

[9] The principle of uniform connectedness is an organisational principle, which describes the tendency to perceive region of uniform visual properties- such as luminance or lightness, colour, texture, motion and possibly other properties-as the initial unit of perceptual organisation (Palmer and Rock, 1994, cited from Palmer 1999).

a whole. On the other hand, good concavities (that is points along an edge contour where the contour bend sharply towards the interior of the region (e.g. Hoffman & Richards, 1984) encouraged participants to locate attention to the individual ends of the objects. They concluded that in the poorly defined parts objects attention was not restricted to the parts making it easier for attention to 'spread' across the entire object. Watson and Kramer's (1999) findings constitute important evidence for the nature of object representations available for selection. In their words "...object-based attentional selection [...] can occur from at least three different representational levels: single-UC representations, grouped representations, and parsed representations" (p. 41).

More recently, Vecera, Behrmann and McGoldrick (2001) showed that attention can select from locally structured shape representations. In a divided attention task they asked participants to report features belonging to one or two different component parts of the same object. They found that reporting features from the same part resulted in better accuracy than reporting features from two different parts. This result indicated that facilitation may be constrained by object internal attributes. In a later study Vecera, Behrmann and Filapek (2001) replicated the finding that selection is more efficient when it is restricted to a single part than when it involves two parts. However, one limitation of the Vecera et al. (2000) study is that it relies on an explicit measure of the distribution of attention across object structure, and arguably, the task (to report features belonging to specific shape parts) may bias selection towards relevant object features[10]. Thus, it remains unclear whether object-

[10] Vecera et al (2000) attempted to address this issue by using displays of two multi-part objects, where not only did they replicate the part-based effect in their first experiment, but they also showed an object-

based selection ordinarily operates over internally structured shape representations. In addition, the Vecera et al (2000) and Vecera et al. (2001) data only speak to facilitatory attentional modulation. An outstanding issue is whether facilitation and inhibition operate on the same kinds of shape descriptions.

2.9 Aim of the present thesis

The aim of this thesis is to distinguish between two contrasting hypotheses outlined above in their predictions about the level of specificity of object shape structure made explicit in the representations selected by inhibitory attentional mechanisms. The experiments presented in this thesis make use of the inhibition of return (IOR) paradigm, as an implicit measure of object-based selection.

2.10 Plan of the investigation

This thesis utilises the inhibition of return paradigm as an implicit measure of selection. The rationale is based on findings that inhibition-of-return (Gibson & Egeth, 1994; Jordan & Tipper, 1999; Tipper et al., 1991), as well as facilitation of attention (Egly et al., 1994; Moore, Yantis & Vaughan, 1998), can spread along the surface of a cued object. Jordan and Tipper (1999) examined object-based selection using the IOR paradigm. Using displays of two rectangles (Egly et al, 1994) participants were required to respond to the onset of a target (white square) which could appear following the brief presentation of a peripheral cue, at a location within one of the two objects in the stimulus displays. The results showed that responses were slower to

---

based effect. The task, however, still required participants to explicitly direct attention to particular parts of objects in the stimulus displays.

targets presented at uncued locations on the surface of a previously cued object, than to targets appearing at equidistant locations in uncued objects – demonstrating object-based inhibition.

2.10.1 The 'Inhibition of Return' paradigm

Orienting attention to a peripheral location through endogenous or exogenous cueing (using a central arrow pointing towards a peripheral location or activating a peripheral location with an event respectively) results in facilitation of detection of a subsequent target that appears in the cued location, when the stimulus onset asynchrony (SOA) between the cue and the target is short (approximately 100-150 msec). Such facilitation is interpreted as being the result of alignment of visual attention to the cued location in space. At longer cue-target SOA, however, response times to the target are slower in the cued trials compared with uncued trials (Posner & Cohen, 1984; Maylor & Hockey, 1985). This effect is referred to as *inhibition of return* (IOR) by Posner and Cohen (1984) and has since been used to refer to attentional bias against re-orienting attention to a previously attended location. Such a mechanism is thought to facilitate efficient visual search, by 'tagging' already attended locations in a visual scene (Posner & Cohen, 1984; Klein, 1988).

The experimental procedure that elicits such facilitatory and inhibitory effects is the spatial cueing procedure and is shown in Figure 11. Three boxes are presented on the computer screen and participants are instructed to fixate in the central box. Following a short delay, one of the peripheral boxes is illuminated for 100 msec. This event constitutes the *cue*. Next the central box is illuminated summoning attention back to the centre of the display. Finally, a target is presented either in the peripheral box that was illuminated (*cued target*) or in the other box (*uncued target*). Reaction times are measured as a function of target status (or cueing condition), that is cued or uncued. When the target is presented in the cued peripheral box, response times are increased compared with response times when the target is presented in the uncued peripheral box. This pattern is observed when stimulus onset asynchrony between the cue and target is greater than 300 milliseconds. Inhibition of return is inferred from slower reaction times (RTs) for targets appearing to a previously cued location compared with reaction times to targets appearing within uncued locations (see Figure 12).

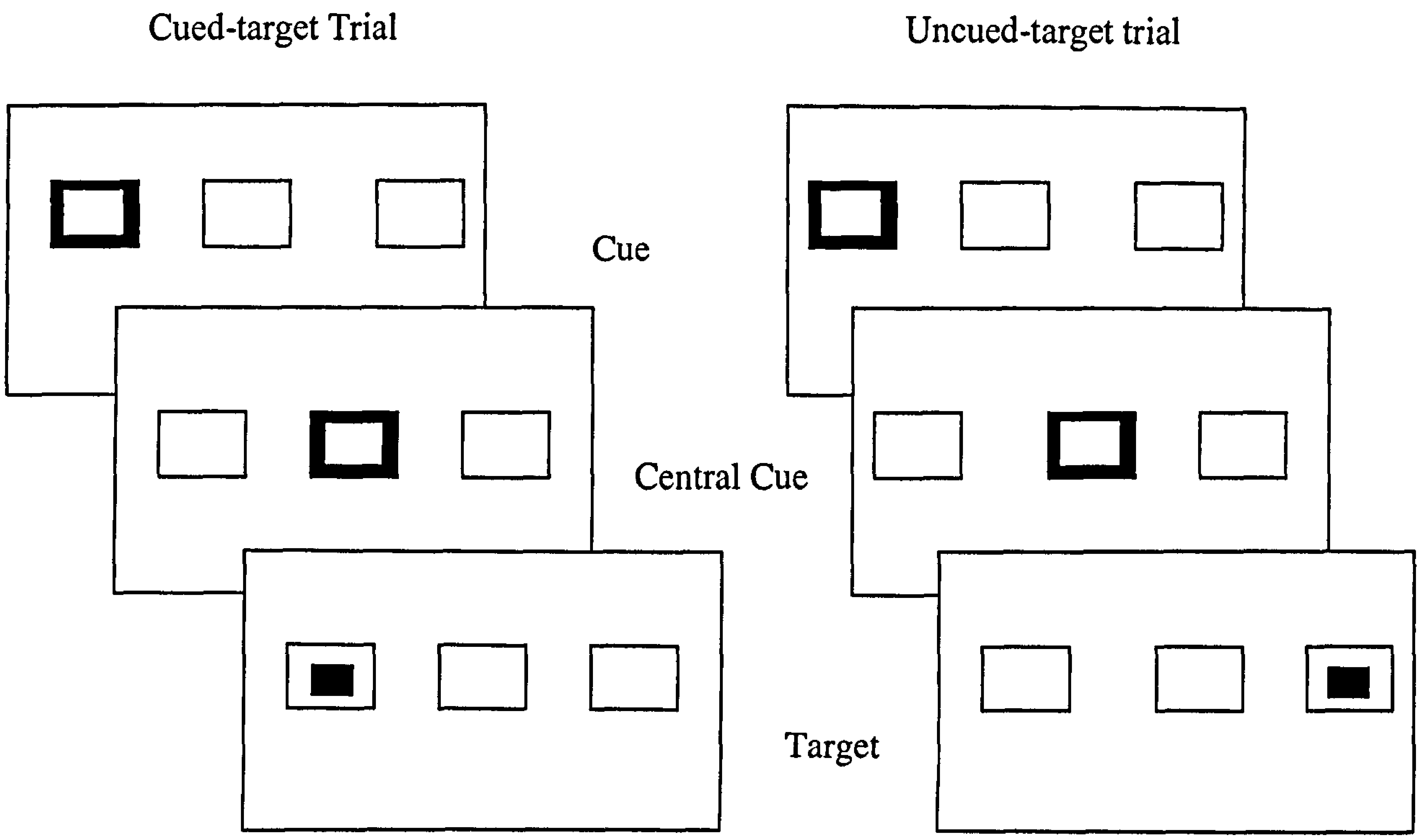


Figure 11. An example of a cued and an uncued trial using the exogenous spatial cueing procedure introduced by Posner and Cohen (1984). Facilitatory and inhibition of return effects are observed using this and other similar procedures.

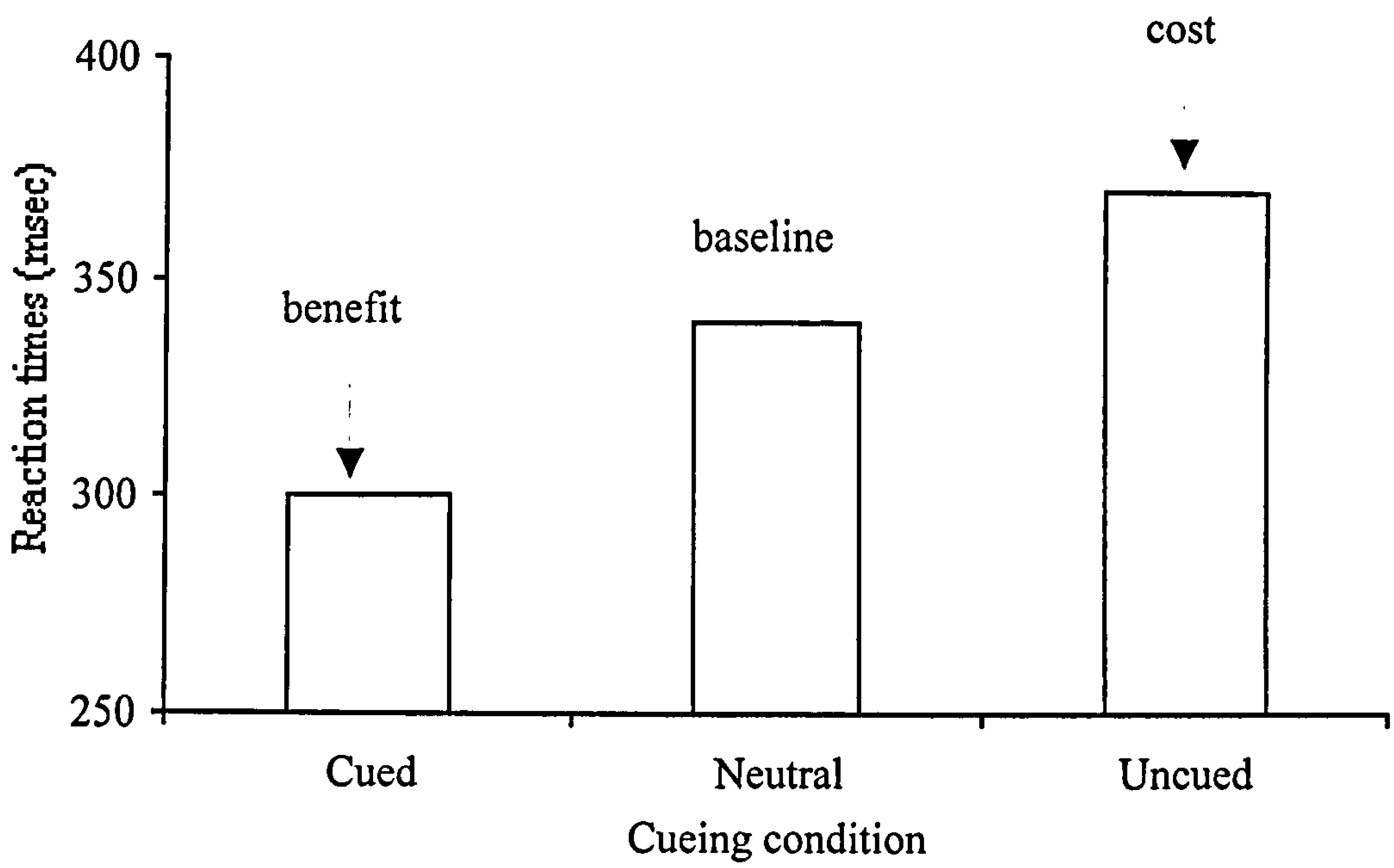


Figure 12: A graph illustrating the RT costs and benefits of attention in the spatial cueing task. In the 'neutral' cueing condition, no cue was presented prior to target presentation.

2.10.2 Object-based 'Inhibition of Return'

Recently studies in inhibition of return have shown that inhibitory components of attention can also operate on objects in the environment (Tipper, Driver & Weaver, 1991) as well as upon environmental locations, and objects under the same experimental conditions (Tipper, Weaver, Jerreat & Burak 1994; Abrams & Dobkin 1994; Gibson & Egeth, 1994; but see Muller & Muhlenen, 1996 and McAuliffe, Pratt & O' Donnell, 2001). Using moving (Tipper et al., 1991; Tipper et al., 1994) and static displays (Jordan & Tipper, 1999; Jordan & Tipper, 1998) Tipper and colleagues have consistently obtained an object based IOR effect a finding that supports the idea that, depending on the reference frame used, attention can be oriented to a region in space as well as to an object in space. Such a finding should not be surprising, if one considers the ecological validity of an inhibitory mechanism operating on objects as well as environmental locations. It is objects that in essence attract our attention (e.g. Yantis & Hillstrom, 1994) and whose successful recognition is important for elementary functions like survival and communication.

Object-based inhibition of return was first demonstrated by Tipper et al. (1991). They used displays containing three boxes, as most experiments that have elicited IOR, but instead of presenting them as stationary objects on the screen, they programmed them to appear to move in a clockwise fashion around the screen (their Experiment 2). In their procedure (Figure 13 below) the three boxes were initially diagonally aligned across the screen and then the two peripheral boxes started to 'move' clockwise, whilst the central box remained stationary serving as fixation. When the two peripheral boxes were horizontally aligned, one of them 'flickered' (as in the typical IOR experimental procedure) for a short time. Following the cueing event, the central box also 'flickered' to summon attention in the centre of the display, and the motion of the two peripheral boxes resumed. In one condition the boxes would stop moving at 90 degrees from the (previous) location of the cue and the target would appear in one of the peripheral boxes. Thus, the target would appear either within the box that was cued a few milliseconds earlier, or within the uncued box, both boxes being equidistant from the original cue location. In the other condition motion would cease at 180 degrees from the cue. In this condition the target appeared within the cued box that was now in the uncued location. Both conditions elicited an inhibition of return for the objects used. Interestingly, there was no interaction between degrees of box rotation and cueing, leading Tipper et al. (1991) to suggest that location-based inhibition (that would be expected in the 180 degrees rotation) did not confound the results. In later studies, however, Tipper and colleagues found evidence to suggest the co-existence of location and object-based reference frames within the same task.

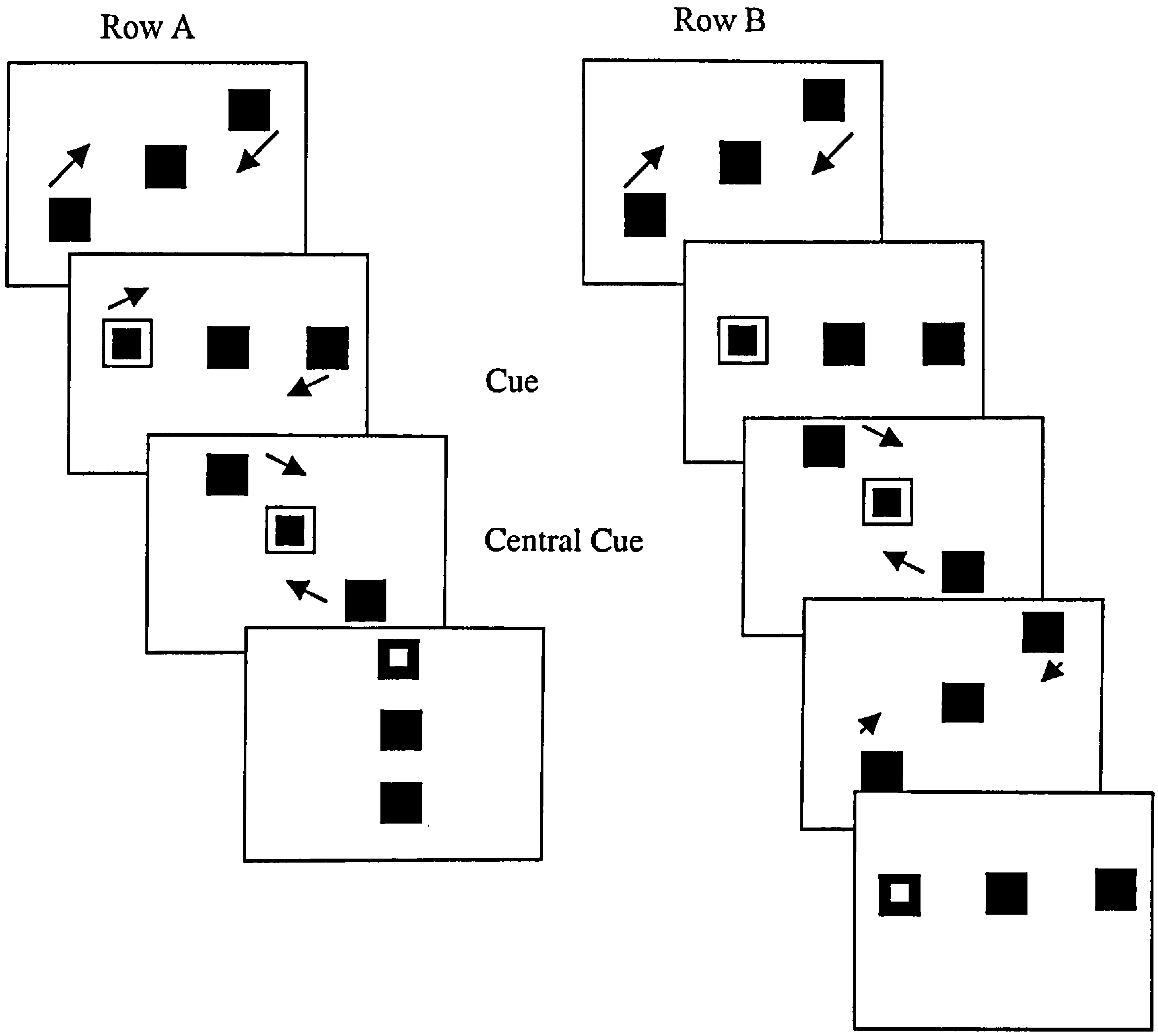


Figure 13. Illustrating the procedure used by Tipper, Driver and Weaver (1991). The two rows represent trials where the cue and the target were presented within the same box (cued object trials). Row A depicts a trial where the target appeared at 90 degrees from fixation. Row B depicts the target at 180 degrees from fixation. The direction of the arrows represent the direction of motion.

In particular, Tipper et al. (1994) replicated the Experiments 1 and 2 from the Tipper et al. (1991) study, manipulating, however, the display type (static and moving) as a within-subjects variable. They presented participants with displays consisting of four boxes, two of which were stationary on the screen and two appeared to be moving in a clockwise direction around the screen.

The target could have appeared in the cued box after 90 degrees rotation (Figure 13, Row A) or in the cued box after 180 degrees rotation (Figure 13, Row B, now in the same location where the cueing took place). They found that object-based IOR was larger in the 180-degree rotation condition, suggesting the interaction between object and location-based inhibitory mechanisms in the same experimental task. They also found that when an inhibited object moved into a new location inhibition stayed in the location where the cueing took place, independently of the existence of an object in that location. In contrast, object-based inhibition did require the existence of an object to draw attention. These findings led Tipper et al. (1994) to propose that location and object-based frames are utilised by observers within the same experiment.

Of more pertinence to the role of object-based representations in visual selective attention, Tipper, Jordan and Weaver (1999) investigated the issue of whether inhibition was associated with a certain part of the object or with the whole object. They used displays originally used by Weaver, Lupianez, and Watson (1998), consisting of three moving boxes to investigate *scene-based* (based on spatial relations between objects in a scene) and *object-centred* (based on spatial relations between parts within an object) frames of reference of IOR. In the object-centred condition the three boxes were connected with each other with straight lines, forming a triangle. In the scene-based condition the boxes were presented as separate objects (not connected with a line). They found object-based IOR in both conditions. Surprisingly, however, the location-based IOR (cued location minus uncued location) was replicated in the scene-based condition but *not* in the object-centred condition. Rather, in the latter condition there was a small *facilitation* for targets appearing in the same location as the cue. This result was consistent across three experiments and led Tipper et al. (1999) to propose that the location of a part within an object was not critical for search and action, and thus irrelevant to performance, therefore no inhibition was needed for that within-object location. This allowed facilitatory effects of the cue to be revealed that are normally hindered in scene-based co-ordinates, when the location of the cue is relevant and therefore likely to be inhibited.

2.11 Empirical Predictions

In the present investigation, the two hypotheses outlined above are examined. Using the object-based IOR paradigm, the experiments described in Chapter 3 examine whether the magnitude of object-based IOR effect is modulated by an internal structural discontinuity. If selection operates *solely* on global shape properties, then internal shape features would not be expected to modulate object-based IOR, supporting the *global structure* hypothesis. According to this hypothesis, local internal structure is not made explicit in the representation available to inhibitory object-based attention. Therefore, object-based IOR magnitude would not be significantly different between targets that appear on the same side of an internal structural discontinuity and targets that appear on different sides of the discontinuity. The pattern of results predicted on the basis of the global structure hypothesis are illustrated in the top graph of Figure 14.

Alternatively, if inhibitory object-based attentional selection operates over internally structured shape representations, as predicted by the *local structure* hypothesis, then it may be expected that the spread of inhibition will be modulated by internal structural discontinuities in the object, which are selected for attentional processing, even when the objects themselves are irrelevant to the task. This hypothesis does not allow any predictions about the specific pattern of object-based IOR modulation. In other words, if the internal boundary modulates the spread of inhibition across the object, it may do so by *reducing* inhibition for targets appearing on the opposite side of the discontinuity from the cue. In contrast, object-based IOR may *increase* for targets appearing on the opposite side of the discontinuity from the cue. These possible outcomes are depicted in the middle and bottom graphs in Figure 14.

Finally, it is important to consider the possibility that, if object-based IOR is sensitive to the internal structure of the objects, then this does not exclude the possibility that globally structured representations are also available for selection. Consider, for example, the diagram in Figure 8. The description of the human arm as the arrangement of the arm and the forearm includes information about the spatial relationship of these descriptions with the rest of the human body shape. In other words, the global description of the shape emerges from locally structured representations. Conversely, the global description of the human body shape need not include description of the shape's local structure, such as a detailed description of the arm (Marr, 1982).

Therefore, the local structure hypothesis allows that object-based inhibitory selection operates simultaneously over representations that specify object shape at both levels of local and global structure. Critically, the local structure hypothesis could be consistent with all three patterns shown in Figure 14, whereas the global structure hypothesis would only be consistent with the top predicted pattern.

Globally structured representations hypothesis

Locally structured representations hypothesis

Predicted Pattern 1

Predicted Pattern 2

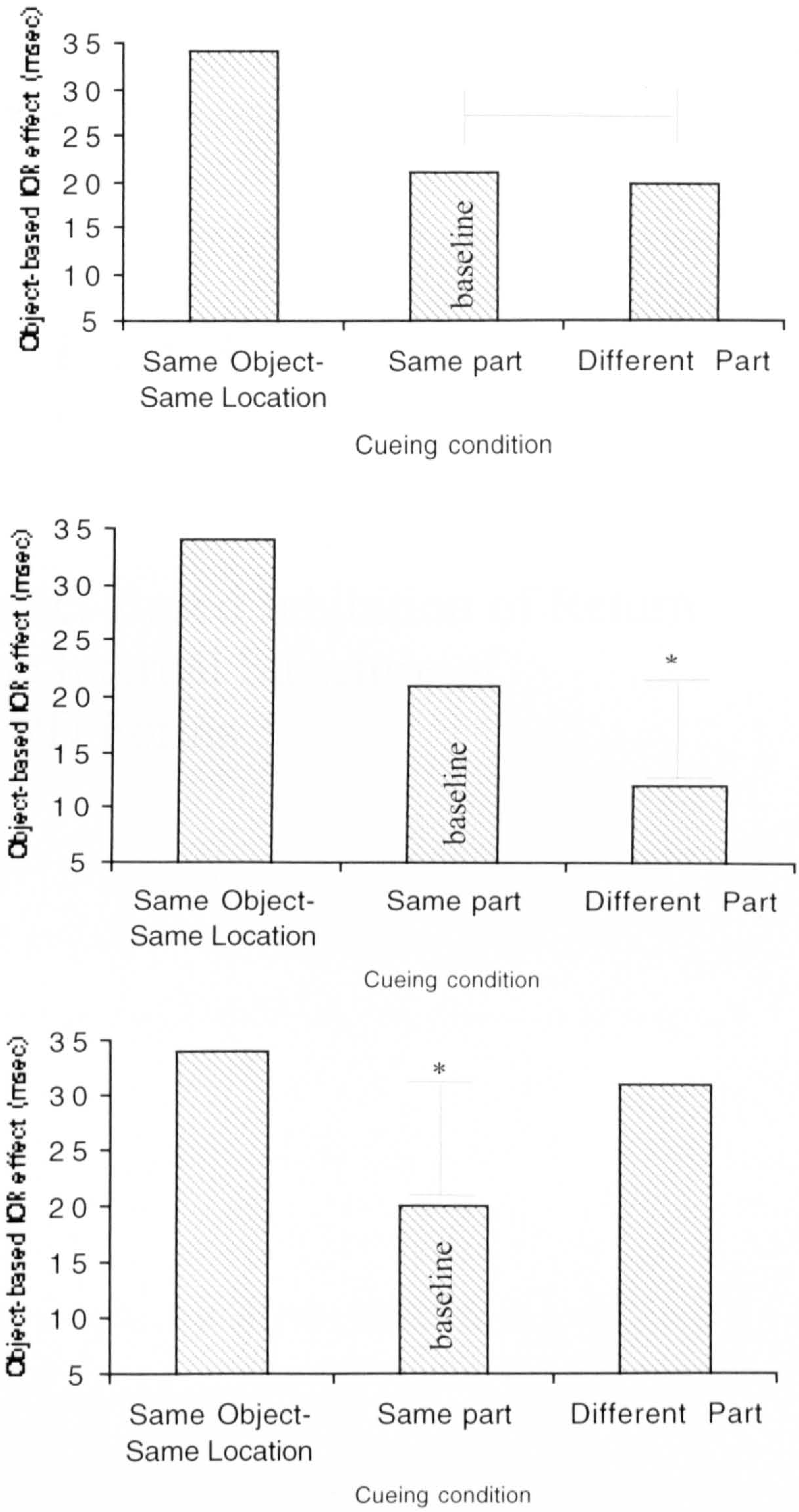


Figure 14: An illustration of the pattern of results predicted by the two hypotheses contrasted in this thesis. The predictions relate to the pattern of modulation of object-based IOR by an object s internal structure. The asterisk on the bars denotes that the difference between the two cueing conditions is expected to be significant.

Chapter 3

3. Modulation of Object-Based Inhibition of Return (IOR) by the Internal Structure of 2D Forms

Experiment 1

3.1 Introduction

The purpose of Experiment 1 was two-fold. First, to replicate the object-based IOR effect found by Jordan and Tipper (1999). Second, to establish a reliable set of experimental parameters for use in the experiments of the present thesis. The prediction was that if IOR is object-based then it should be observed at different locations along the surface of previously same objects.

3.1.1. Method

Participants

Twenty psychology (major) undergraduates, aged between 18 and 34, from the University of Wales, Bangor, participated in the experiment for one course credit. They all reported normal or corrected -to- normal vision and were naive to the purpose of the experiment.

Apparatus and Stimuli

Stimuli were presented on a 14-inch monitor connected to a Power Macintosh PC. Randomisation and presentation of the stimuli, as well as recording of the participants' reaction times, were controlled through PsyScope software (version 1.2.4; Cohen, MacWhinney, Flatt & Provost, 1993). Responses were made through a single letter key on a standard Apple keyboard connected to the computer.

The stimulus display consisted of two outline (black) rectangles, simultaneously presented on each side of a fixation cross (see Figure 15) against a light-grey background. The rectangles subtended 1.5° x 6.5° of visual angle when viewed from a 55 cm distance. The orientation of the rectangles varied randomly between trials, appearing either +45 or -45 degrees tilted from the vertical meridian. The fixation cross (+) subtended 0.8° x 0.8° of visual angle. The cue was an outline white square subtending 1.0° x 1.0°, whilst its white contours subtended .02°. The target was a filled white square measuring 0.8° x 0.8°. The central re-fixation cue was a white cross (+) sign subtending 0.5° x 0.5°, that replaced the black fixation cross. From end to end the display subtended 9.7° x 9.7° of visual angle.

Design

The experiment was based on a two factorial within-subject design with factors of Cueing, with four levels, and SOA with three levels. Levels of Cueing consisted of the following cue-target configurations. The target appeared (a) within the same object and at the same location as the cue (same object- same location), (b) within the same object but at a different location from the cue (same object-different location) (c) within a different object but at a location corresponding to the cue location in the same object (different object, baseline); and (d) within a different object but at a location diagonal to the cue (different object, diagonal). Figure 15 illustrates the four cueing conditions for the + 45 orientation in Experiment 1. As shown, the cue-target distance in the two critical, 'same object' and 'different object' cueing conditions was identical

(5.2° of visual angle), whilst the cue-target distance in the 'different object-diagonal' condition was greater than the distance in the 'different object' condition. Therefore, trials in the 'different object-diagonal' cueing condition were not used in any statistical analyses.

The second within-subjects factor was SOA, with three levels; 400, 820 and 1220 msec. Each cueing condition appeared 100 times over the three SOAs, randomly across all trials. Participants completed 10 practice trials followed by 600 experimental trials, of which 200 trials (33%) were 'no-target' trials.

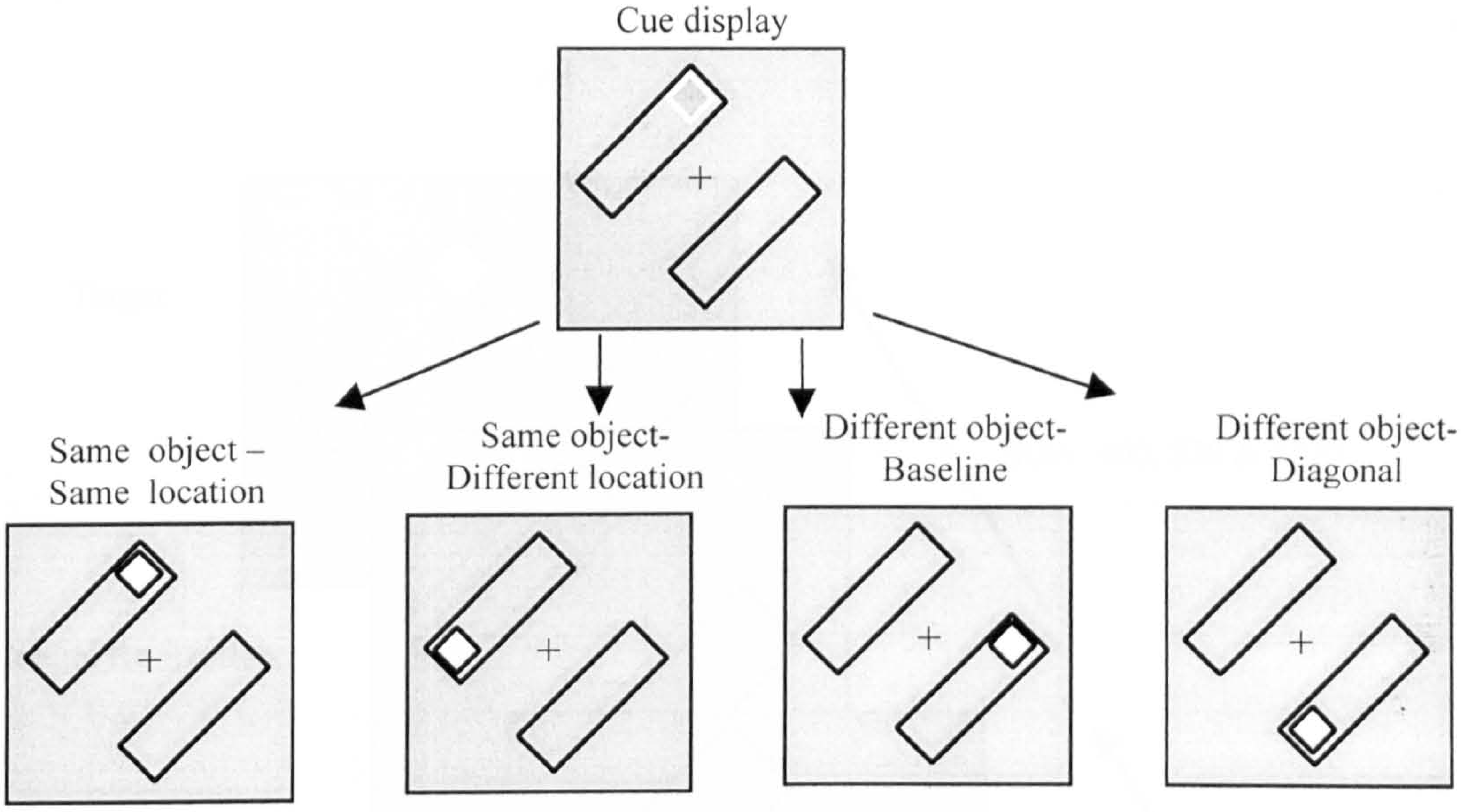


Figure 15: The diagram illustrates examples of the four possible cue-target configurations for each of the four cueing conditions in Experiment 1. Here the rectangles are presented in the +45 display orientation.

Procedure

At the beginning of each trial the fixation cross appeared in the centre of the monitor. After 1000 msec, two rectangles were simultaneously presented on each side of the fixation cross. Following a further 1000 msec delay the peripheral cue appeared within one of four random and equiprobable (above, below, on the left or on the right of fixation) locations within the inner ends of the rectangles (Figure 16). Cue duration was 90 msec. At intervals of either 90, 300 or 500 msec from cue offset, the central fixation cross changed from black to white for a period of 130 msec and then reverted to black until trial end (central re-fixation). After a further delay of 90, 300 or 500 msec the target was presented in one of same four random and equiprobable locations as the cue. The target remained visible for a 1000 msec, or until the 'b' key (response) was depressed.

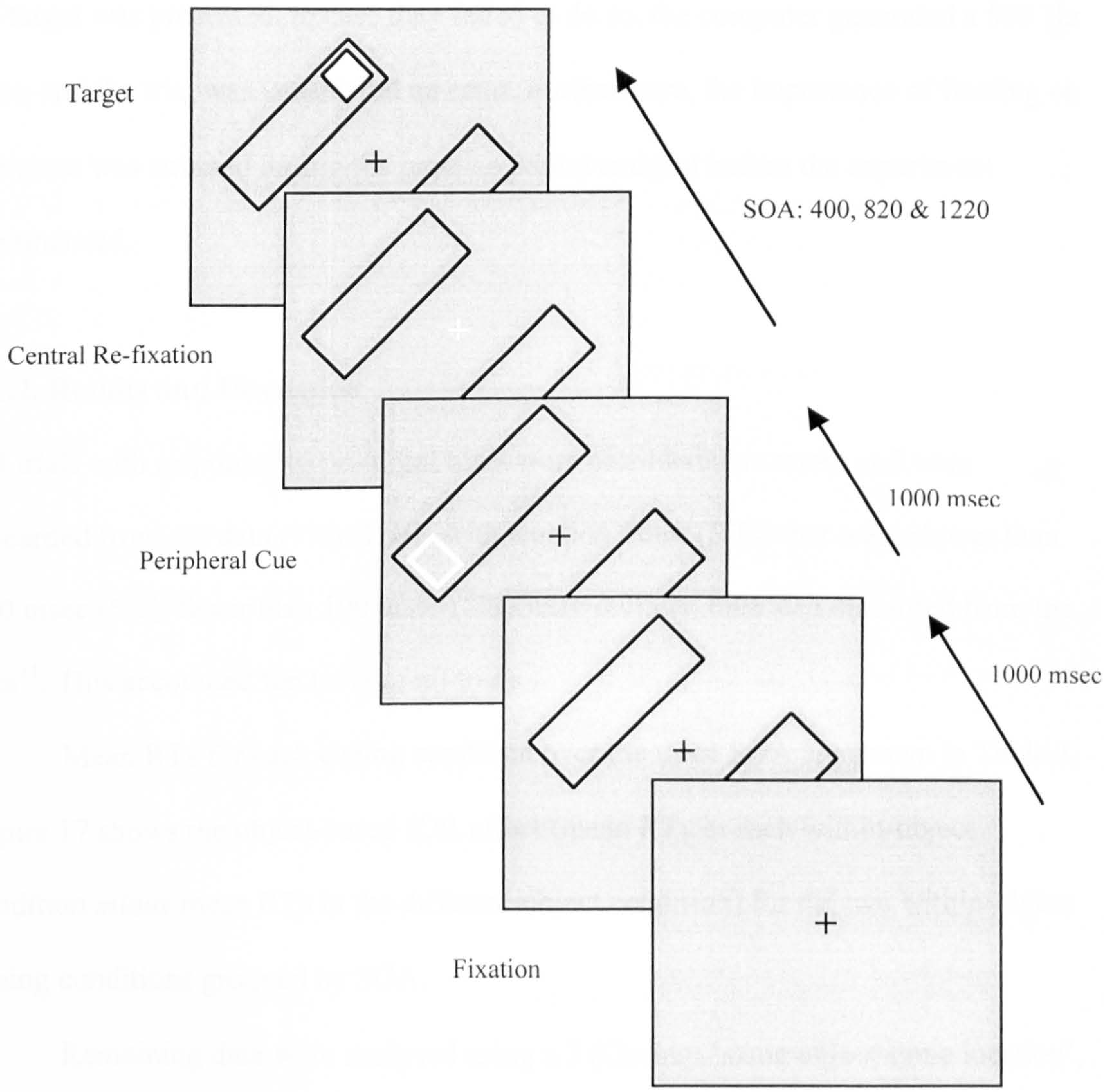


Figure 16. The diagram illustrates a 'same object-different location' trial, when the rectangles were presented at +45 degree orientation. The target was presented within the same object as the cue but in a different location.

Participants were informed that the white outline square (the cue) was not predictive of the location of the subsequent target. They were instructed to press the response key as soon as they detected the target and to withhold their response when no target was presented. In case they failed to do so, the computer generated a 500 Hz tone, and the trial was considered an error. Furthermore, the importance of fixating on the cross was stressed during the practice period and just before the experiment commenced.

3.1.2. Results and Discussion

All trials with responses to no-target trials were considered as errors and were discarded from the data (7%). Trials with reaction times (RTs) that were slower than 700 msec (5%), faster than 100 msec (3%) were outliers, thus also discarded from the data[11]. This accounted for 1.5% of all trials.

Mean RTs for each cueing condition over the three SOA are shown in Table 1. Figure 17 shows the object-based IOR effect (mean RTs in each within-object condition *minus* mean RTs in the different-object condition) for the two within-object cueing conditions grouped by SOA.

Remaining data were analysed using a 3 (Cueing: 'same object-same location', 'same object-different location', and 'different object-baseline') x 3 (SOA: 400, 820, and 1220 msec) repeated-measures ANOVA[12]. Results showed a significant main

[11] This cut-off procedure is followed in most investigations of object-based IOR. For reference also see Jordan & Tipper (1999); Weaver, Jordan & Tipper (1999).

[12] Assumptions of sphericity and homogeneity of variance were met for the three factors. Mauchley's test of sphericity produced a non-significant p value of .756 for the factor of Cueing, and a p value of .567 for the factor of SOA. Homogeneity of Variance between each condition was assumed as the largest variance(square route of SD) in each condition was not larger than three times the smallest variance (Dancey & Reidy, 1999).

effect of Cueing, F (2, 56) = 124, $p < 0.001$. There was also a significant main effect of SOA [F (2, 56) = 5.0, $p < 0.012$] with the longest RTs over all four cueing conditions at SOA 1220 (M = 388; SD = 35.7). The interaction between Cueing and SOA was not significant, F (4, 76) = 0.83, ns.

Table 1. Mean RTs for the four cueing conditions at each SOA. Calculation of IOR resulted by subtracting RTs in the 'different object' condition from each of the two within-object conditions ('same object-same location' and 'same object-different location'). Mean RTs from the 'different object-diagonal' cueing condition were not used in the analysis.

SOA (msec)	Same object – Same location		Same object - Different location		Different object- Baseline		Different object- Diagonal	
	M	SD	*M*	SD	*M*	SD	*M*	SD
400	404	32.5	375	31.6	363	36.7	359	33.6
820	396	33.4	373	35.3	364	31.7	355	30.6
1220	408	30.8	385	32.1	374	32.9	367	38.2

Planned comparisons were carried out for RTs in each cueing condition (collapsed across SOA). These showed a significant IOR effect for the 'same object - same location' condition collapsed over SOA of 36 msec [t (19) = -13.9, $p < 0.001$][13] and a significant IOR effect of 11 msec for the 'same object-different location' condition [t (19) = -4.0, $p < 0.01$][14].

In the remainder of the experiments, tests of sphericity and homogeneity of variance will be reported only if these assumptions are violated. However, it is noted that violation of the homogeneity of variance assumption is not catastrophic as long as there are equal numbers of participants in each experimental condition (Dancey & Reidy, 1999, p. 131).

[13] In this thesis, when reporting t-test results, the convention in most textbooks in statistics is followed, that is, only the degrees of freedom of the between-subject variable is reported.

[14] One of the examiners has pointed out that the IOR effects for the 'Same object/Different location'

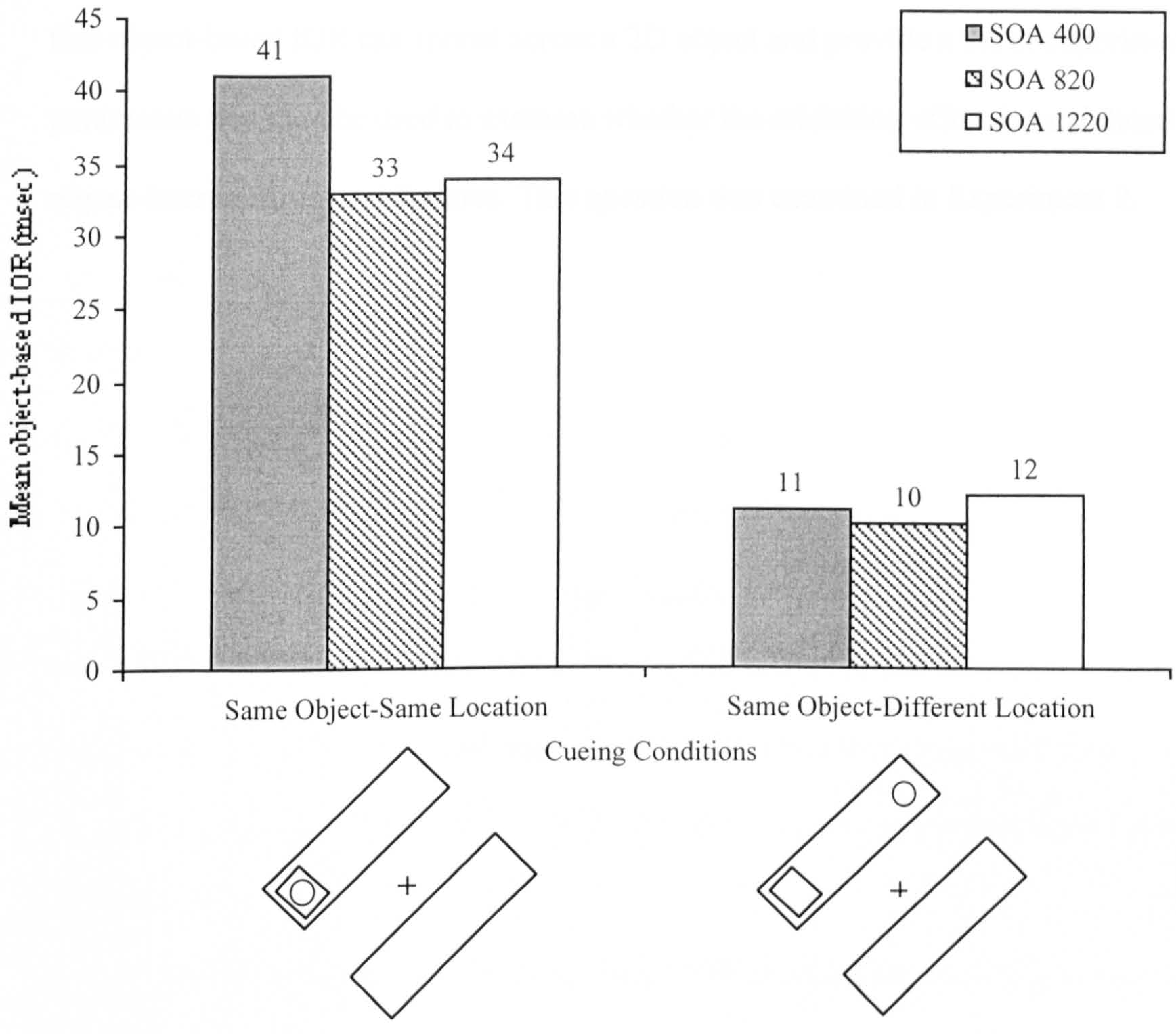


Figure 17. Object-based IOR effect for the two within-object cueing conditions grouped by the three SOA. The object-based IOR is calculated as RT different object minus RT same object (also see text). For illustration purposes cues are shown here as squares and targets as circles.

---

condition are rather small. However, planned comparisons showed that the RTs in that condition for all three SOA wre significantly slower from RTs in the 'different object' condition at each SOA [t (19) = -2.8, p = .009, t (19) = 2.73, p = .01, t (19) = -2.54, p = .02 , for SOA 420 msec, 820 msec and 1220 msec respectively.

The results from Experiment 1 confirm Jordan and Tipper's (1999) findings that object-based IOR can spread across a 2D object and provide a set of experimental parameters that may be used to examine whether the inhibition effect is modulated by object-internal structural features. This question was examined in Experiment 2.

Experiment 2

3.2 Introduction

The aim of Experiment 2 is first, to investigate whether object internal structure will modulate object-based IOR. More specifically, this experiment was designed to distinguish between the two hypotheses outlined in the Introduction, namely, the *global structure* and the *local structure* hypothesis, in their predictions about modulation of object-based attention by object internal structure. Second, this experiment aims to extend previous findings of the part-based facilitatory effects to inhibitory mechanisms of attention (e.g. Vecera et al., 2000).

In order to examine the effect of internal shape features on the magnitude of IOR, stimulus displays of 2D L-shaped forms that were either *segmented* by an internal structural discontinuity (a single edge contour), or *unsegmented* (see Figure 18) were used.

The predicted patterns of object-based IOR modulation according to each of the two hypotheses are as follows: If object-based selection operates over internally structured shape representations, as predicted by the *local structure* hypothesis, then the magnitude of object-based IOR will be modulated by the structural discontinuity in the segmented shapes, when the discontinuity separates the cue and the target. In contrast, if selection operates on global shape properties alone, i.e. the outline shape, as predicted by the *global structure* hypothesis, then the magnitude of the object-based IOR effect will not be modulated by the internal discontinuity. Task parameters were based on those used in Experiment 1, where statistically reliable object-based IOR effects were obtained.

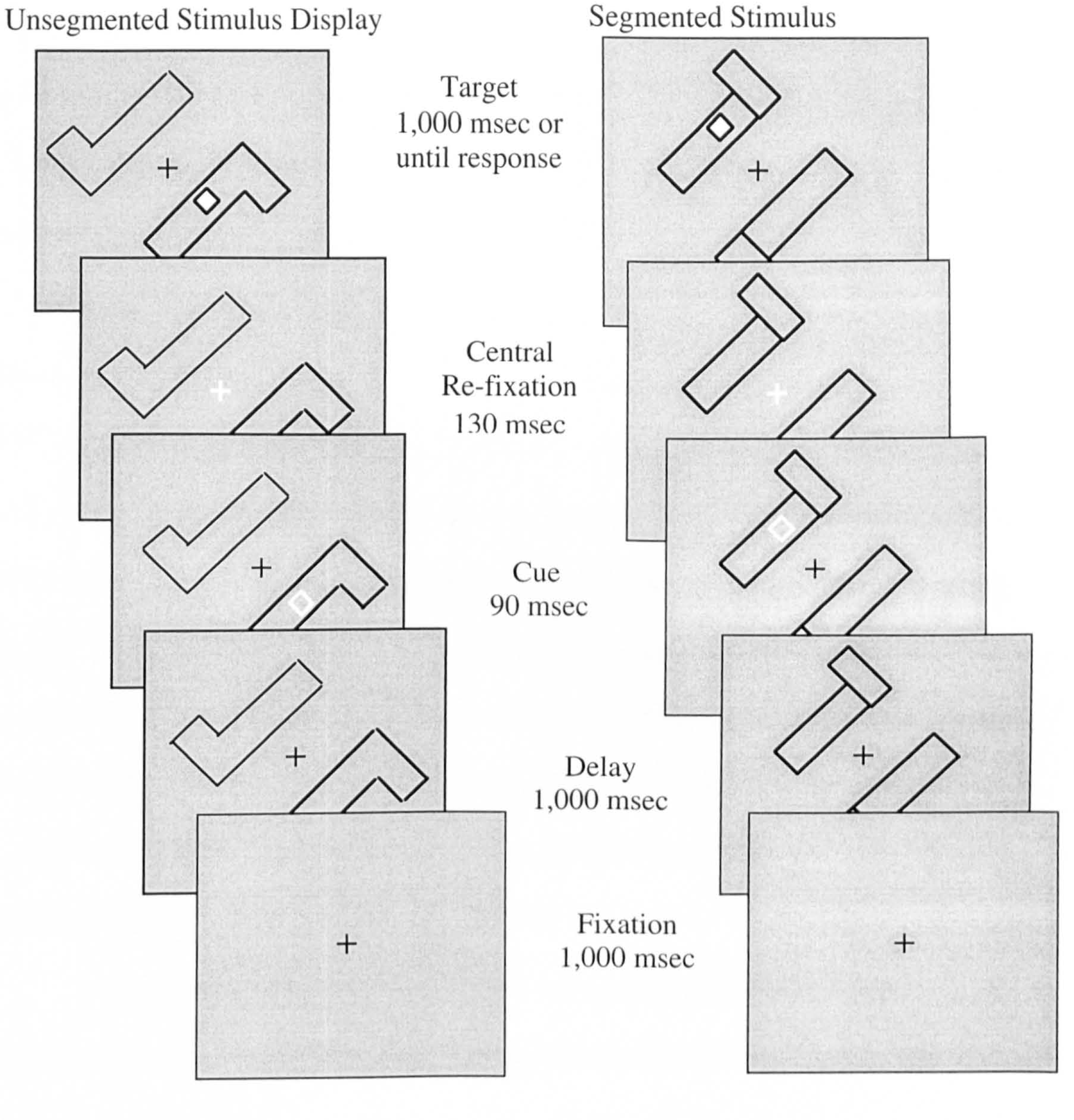


Figure 18. The diagram illustrates the main events in a 'target' trial in Experiment 2 (see text for further procedural details). The left column shows an example of a 'same object-same location' trial for the 'unsegmented' displays with the short rectangles of the two objects positioned on the left and right of fixation. The right column shows an example of the equivalent cueing condition in the 'segmented' displays, but with the short rectangles of the objects positioned above and below fixation.

3.2.1 Method

Participants

Ten participants, registered as major psychology students in the University of Wales, Bangor, took part in this 1-hour experiment, each receiving one course credit. All reported normal or corrected-to-normal vision.

Apparatus and Stimuli.

The same technical apparatus as in Experiment 1 was used. Displays consisted of two outline L-shapes appearing at both sides of a fixation cross at screen centre. Each L-shape was composed of one long and one short rectangle. Across all trials the two L-shapes were tilted to the right (+45 degrees) with the short rectangles positioned either on the left and on the right of fixation (e.g. Figure 18, leftmost column) or above and below fixation (e.g. Figure 18, rightmost column). The two rectangles comprising each L-shape were separated by an internal discontinuity in the 'segmented' condition but not in the 'unsegmented' condition (Figure 18).

At a viewing distance of 50 cm, the longer rectangle of each L-shape subtended 7.2° x 1.8° and the smaller rectangle 2.8° x 2.2° of visual angle. The fixation cross was black and measured 0.8° x 0.8° of visual angle. The cue was a white outline square subtending 0.6° x 0.6° (contours measuring 0.2° x 0.2°) and the target was a filled white square subtending 0.8° x 0.8° of visual angle. The whole display from end to end was 13.2° high and 10.8° wide.

Design

The experiment was based on a 6 (Cueing conditions) x 4 (Display type: segmented and unsegmented with the shorter rectangle positioned left and right or above and below fixation) x 2 (SOA: 820 and 1220 msec) within-subjects design. Figure 18 illustrates an example of an 'unsegmented' display with the short rectangles positioned to the left and right of fixation (Figure 18, left column), and an example of a 'segmented' stimulus display with the short rectangles positioned approximately above or below fixation (Figure 18, right column).

In both Display conditions ('segmented' and 'unsegmented') the cue was presented randomly and with equal probability at one of three possible positions within either L-shape, that is, at the end of the longer arm of either L-shape, in the middle of the longer arm of either L-shape or in the shorter arm of either L-shape. However, only trials where the cue appeared in the middle of the longer arm of either L-shape (see Figure 18) were used in subsequent analyses, in order to ensure equal cue-target distances in the 'same part', 'different part' and 'different object-baseline' cueing conditions (described below).

In the 'segmented' Display condition, targets appeared randomly, and with equal probability, in each of the following four cue-target configurations (also see Figure 19): The target appeared (a) on the same part and at the same location as the cue (same part – same location), (b) on the same part but at a different location from the cue (same part-different location), (c) on the same object but on a different part to the cue (different part), (d) on the corresponding part of a different object, but at the same corresponding location, as the cue (different object-baseline). The 'different object -baseline' condition was used as the baseline, between-objects condition as it was the only between-objects condition where the cue-target distance between the two L-shapes was identical to the cue-target distances in the 'same part' and 'different part' cueing conditions, when the cue appeared at the centre of either object.

There were also *filler trials*, that were not used in the analysis as the cue-target distance between different objects was not the same as the cue-target distances in the 'same part-different location' and 'different part' cueing conditions. These were the following cue-targets configurations: the target appeared (1) on the corresponding part of a different object, and at a different location (different object-same part), and (2) in a different object, and on a different part (different object-different part) relative to the cue. Trials in these filler trials were of no theoretical importance to this study and were excluded from any statistical analysis.

In the 'unsegmented' Display condition, cue-target configurations were exactly the same as in the 'segmented' display condition, the only difference being the absence of the internal discontinuity within the L-shapes. Figure 5 illustrates the four cueing conditions in the 'segmented' displays that were used in the analysis, when the cue was presented in the middle of either L-shape. The condition names were the same for both the 'segmented' and the 'unsegmented' displays.

Participants completed 10 practice trials followed by 680 experimental trials, of which 200 (30%) were 'no-target' trials. In the remaining 480 'target' trials, there were approximately 40 trials for each of the six cueing conditions, in each Display condition ('segmented' vs. 'unsegmented'), over the two SOAs[15].

[15] The break-down with respect to whether the cue appeared in the middle of the long rectangle, the end of the long rectangle or in the short rectangle is as follows: In 270 trials, where the cue appeared in the middle of the L-shape, there were 45 trials for the six each cueing conditions. In the remaining 210 trials where the cue appeared either at the end of the long arm or at the short arm of the objects, there

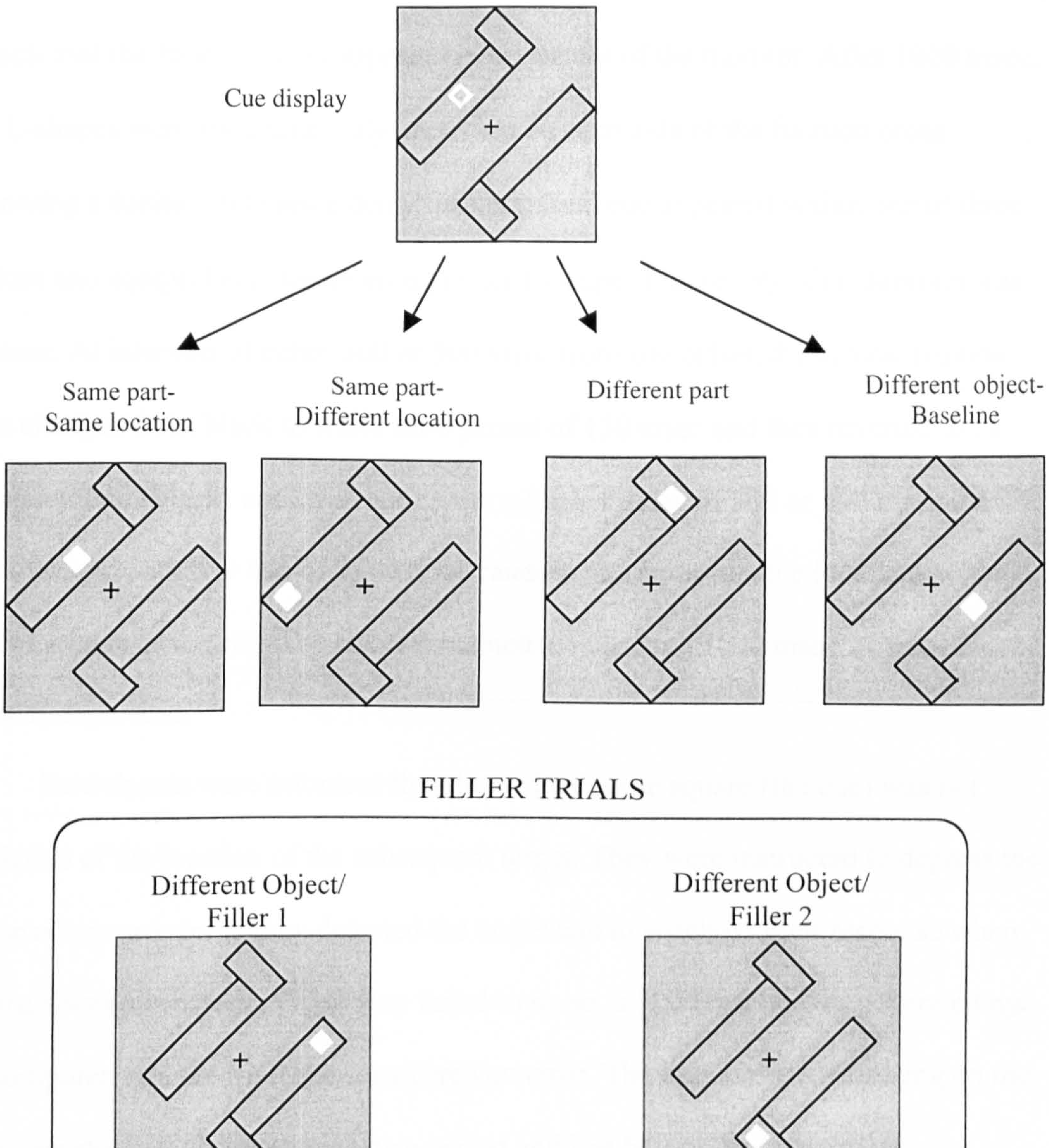


Figure 19: The figure illustrates examples of the four cueing conditions used in the analysis for the segmented displays, when the cue appeared in one of the two objects. In the unsegmented displays the location of the cues and targets were identical to those in the segmented displays. The bottom panel illustrates the two types of filler trials used (with respect to the cue at the top of the figure).

---

were 35 trials for each of the six cueing conditions. Those trials were not analysed.

Procedure

The procedure in Experiment 2 was almost identical to Experiment 1. At the beginning of each trial the fixation cross appeared in the centre of the monitor. After 1000 msec, two L-shapes were simultaneously presented on each side of the fixation cross. Following a further 1000 msec delay the peripheral cue appeared within one of three random and equiprobable locations on either L-shape (Figure 19). Cue duration was 90 msec. At intervals of either 300 or 500 msec from cue offset, the central fixation cross changed from black to white for a period of 130 msec and then reverted to its original colour (black) until trial end. After a further delay of 300 or 500 msec the target was presented in one of same three random and equiprobable locations within either L-shape as the cue. The target remained visible for a 1000 msec, or until the 'b' key was depressed.

Participants were informed that the white outline square (the cue) was not predictive of the location of the subsequent target. They were instructed to depress the response key as soon as they detected the target and to withhold their response when no target was presented. In case they failed to do so, a 500 Hz tone was generated by the computer, and the trial was considered an error. The importance of fixating on the cross was stressed during the practice period and just before the experiment commenced.

3.2.2 Results and Discussion

Only trials where the cue appeared at the centre of either L-shaped object (+45 and -45 degrees from fixation) were used in the analysis. As mentioned earlier, on these trials the cue-target distance was identical between the in the 'same object-same part', 'same object-different part' and 'different object-baseline' conditions. Furthermore filler trials ('different object-diagonal' and 'different object-different part') were excluded from any statistical analysis. This resulted in 41% of all trials (280 trials per participant) to be included in the analysis. Therefore, from the 18 possible cue-target configurations (three possible cue locations x six possible target locations – collapsed across object) only four (1 cue location x 4 target locations) configurations were used in the analysis. Table 2 shows the mean RTs for trials in each of the four cueing conditions over the two types of display (segmented and unsegmented) collapsed across the two SOAs.

Table 2. Mean RTs (msec) for each of the four experimental and two filler cue-target location conditions in Experiment 2. Mean RTs in the 'different object/filler 1' and 'different object/filler 2' conditions (grey cells) are only reported for the purpose of completeness but were not used in the statistical analysis.

	Display type	
	Segmented	Unsegmented
Cueing condition	Mean RT (msec)	
Same part /Same location	361	357
Same part /Different location	345	341
Same object / Different part	360	347
Different object / Baseline	318	326
Different object / Filler 1	321	333
Different object / Filler 2	324	332

Trials with RTs slower than 700 msec (.9%) or faster than 100 msec (.6%), and trials with responses to 'no-target' trials (2.0%) were excluded from the analyses (3.5% of all trials). Figure 20 illustrates the object-based IOR effect for the three within-object cueing conditions grouped by Display (segmented and unsegmented).

Reaction time data were analysed using a 4 (Cueing) x 2 (Display) x 2 (SOA) repeated measures ANOVA. Levels of Cueing were; a) same object-same location, b) same object-same part, c) same object-different part, and d) different object-baseline. Display was either; a) Segmented or b) Unsegmented; and SOA was either 820 or 1220 msec. Results showed a significant main effect of Cueing, $F(3,27) = 15.6, p < 0.001$. There were no significant main effects of Display and SOA, $F(1, 9) = 0.03$, ns and $F(3, 27) = 0.4$, ns respectively. The interaction between Cueing and Display was significant, $F(9, 81) = 2.0, p < 0.05$.

Separate planned comparisons were carried out between the 'same part–different location' and the 'different part' conditions in the two Display conditions to investigate the Cueing by Display interaction. It was predicted that if IOR is sensitive to the objects' internal structure, then the magnitude of object-based IOR would be modulated by the presence of the internal discontinuity, whilst no such modulation would be observed in the 'unsegmented' stimulus displays. In the 'segmented' display condition RTs were significantly larger for the 'different part' condition compared with the 'same part-different location' condition [$t(9) = 3.42, p < .001$]. In the 'unsegmented' condition, there was no significant difference in RTs [$t(9) = 0.62, p > 0.05$] between the two cueing conditions. Furthermore, a one-way ANOVA examining Cueing and SOA on the 'segmented' display condition revealed a significant main effect of Cueing, $F(3, 27) = 26.09, p < 0.001$ and no main effect of SOA, $F(1,9) = 2.1$, ns.

Object-based IOR for the 'segmented' displays in the 'same part-same location' condition was significantly different from the object-based IOR in the same condition for the 'unsegmented' displays (42 msec vs. 31 msec, $t(9) = 3.07, p = 0.01$).

Also significant was the difference in object-based IOR between the 'segmented' and the 'unsegmented' displays in the 'different part' cueing condition (41 msec vs. 23 msec, (t (9) = 4.01, p = 0.003)[16].

In order to ensure that the difference between the two critical cueing conditions ('same part-different location' and 'different part') are due to attentional modulation by the internal boundary and not to low level vision factors, post-hoc tests were carried out between the two filler cueing conditions (see Figure 19). In these trials the cue appeared in the centre of one of the two objects and the target subsequently appeared on either end of the other object. In the segmented displays, mean RTs for Filler 1 (different object-same part; $M = 325$, $SD = 34.2$) were not significantly different from mean RTs in Filler 2 (different object-different part; $M = 327$, $SD = 37.2$), $t(9) < 1$, ns. The difference between Filler 1 ($M = 327$, $SD = 33.6$) and Filler 2 ($M = 329$, $SD = 29.5$) were also non-significant for the unsegmented displays, $t(9) < 1$, ns.

The main finding of Experiment 2 was that the internal structural discontinuity significantly modulated the magnitude of object-based IOR. In addition, the results

---

[16] Whilst the chosen cueing condition is considered to be the best baseline condition for reasons outlined in the Design and the Results sections, further analyses were carried out on the object-based IOR for each within-object condition. This time, however, object-based IOR was calculated by subtracting from the group means of each of the three within-object conditions the equivalent between-object cueing condition. Thus, object-based IOR for the 'same part/same location condition was calculated as cueing condition ondition (a) minus condition (d); object-based IOR for the 'same part/different location' condition was calculated as Condition (b) minus Filler 1; and object-based IOR for the 'different part' condition was calculated as Condition (c) minus Filler 2.

Planned comparisons were carried out on the IOR effect (calculated using a different baseline for each within-object condition) in each within-object cueing condition. These showed that in the segmented displays, object-based IOR was significantly different between Conditions (b) and (c) (12 msec; t (9) = -2.3, p < 0.001. In the unsegmented displays, the difference in object-based IOR between the two aforementioned conditions was not significant (7msec; t (9) = 0.6, ns). These results show that using a different baseline to calculate the object-based IOR for each within-object cueing condition did not change the pattern of object-based IOR modulation.

show that the inhibition effect is larger when the target was presented on the *different* part of the same object; that is, when the cue and target are separated by an internal discontinuity. The implications of the findings for hypotheses about object-based selection, and the inhibitory mechanisms of attention, are discussed below.

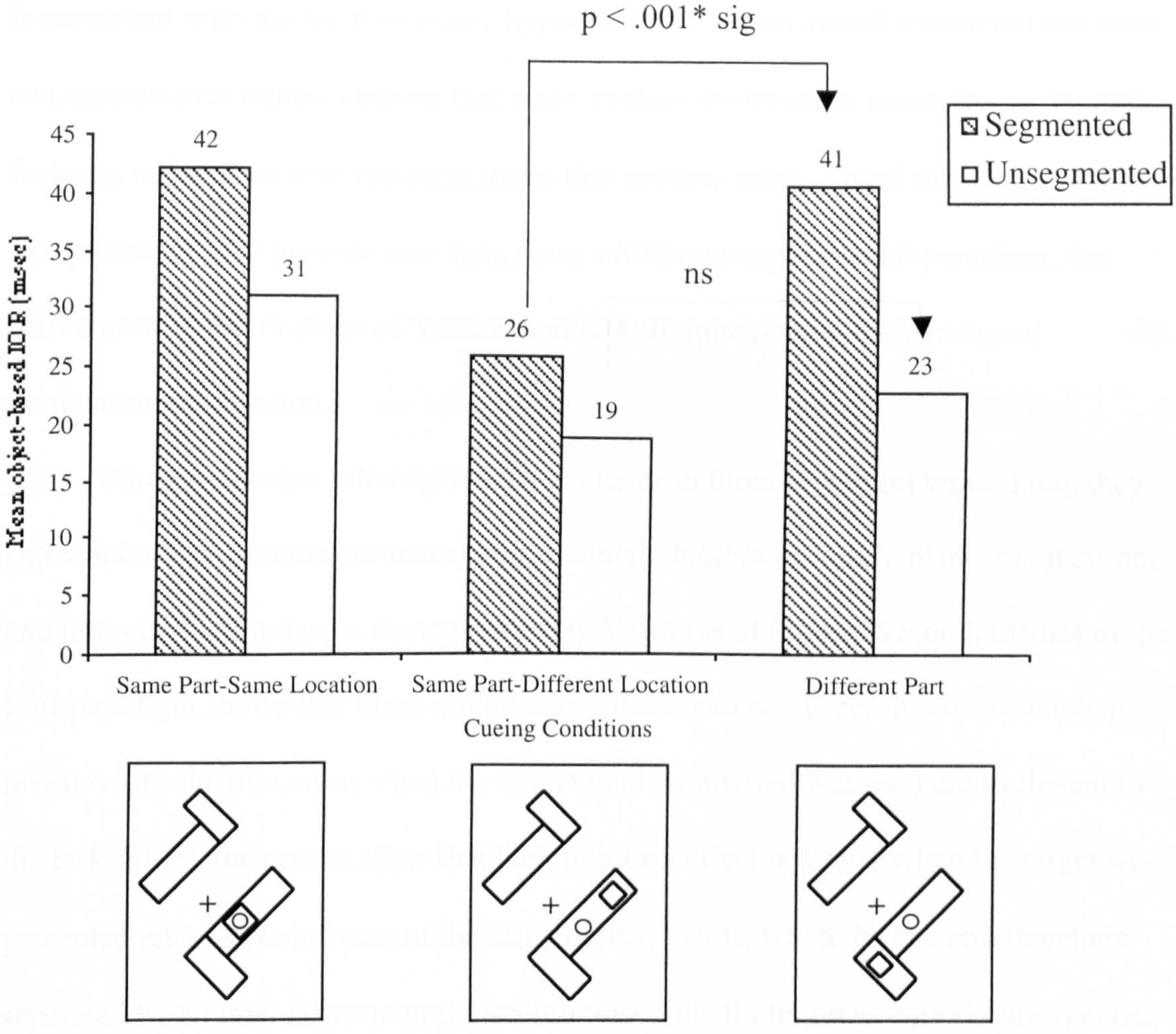


Figure 20. Object-based IOR effect for the three within-object cueing conditions for the 'segmented' and 'unsegmented' displays. The data are collapsed across the two SOAs. For illustration purposes here, cues are depicted as circles and targets as squares. In the 'same part-same location' condition targets appeared within the same location as the cue following a variable time interval.

3.2.3 Discussion

Experiment 1 showed that object-based IOR can generalise across a single 2-D object, replicating the object-based IOR effect reported by Jordan and Tipper (1999). Experiment 2 showed that the magnitude of this object-based inhibitory effect can be modulated by the presence of an object-internal structural discontinuity. This finding is consistent with the *local structure* hypothesis that object-based attentional selection can operate over representations that make explicit information about object-internal features, rather than over representations that contain *solely* global shape properties. The present results provide new data, from a different experimental paradigm, that converge with the findings of Vecera et al (2000), in revealing the part-based modulation of attention.

The results also extend previous evidence in three important ways. First, they show that object-internal structure may modulate *inhibitory* mechanisms of attention, and not only facilitation as demonstrated by Vecera et al (2000). Second, the use of the IOR paradigm shows that these modulatory effects can be observed with an *implicit* measure of selection, even when the object (and its internal features) are irrelevant to the task. Third, the results show that the inhibition effect is *larger* when the target was presented on an *different* part of the same object; that is, when the cue and target are separated by an internal structural discontinuity. This finding was entirely unexpected, and, is somewhat counterintuitive. One might assume that inhibition would be larger the closer a target appears to the cue, on the basis that, at these locations, there is a higher probability of summation of location-based and object-based inhibitory effects (e.g., Jordan and Tipper, 1999). On this account, the magnitude of object-based inhibition should decrease as an inverse function of cue-target proximity. Contrary to this prediction, it was found that object-based IOR increases for targets appearing on a different component of the same object, even when cue-target proximity is held constant. This finding will be discussed in detail in the General Discussion section (Chapter 5).

Another important aspect of the results is that object-based IOR is significantly larger for the segmented than for the unsegmented displays, even in the 'same object-same location' condition, where object internal structure should have played no or little role. Thus, it is possible that the visual complexity of the segmented displays compared to the unsegmented displays may have contributed to the observed cueing effects. One might argue that figural complexity (e.g. defined in terms of the number of edge segments), as opposed to attention to object parts, may have caused the observed differences in IOR modulation between the two types of display. However, despite the fact that visual complexity may have contributed to the overall difference in IOR magnitude between segmented and unsegmented displays, it should not have influenced the pattern of IOR modulation within each display. Furthermore, it was predicted that the two types of display would be equivalent if attention ignored internal object structure (global form selection hypothesis). The finding that they are not equivalent can be taken to mean that object-based IOR operates over representations other than solely global, supporting the *local structure* selection hypothesis. This point will be re-visited in the General Discussion. For now, it is contended that object-based IOR is modulated by representations that make explicit the internal structure of the object, which is inherently more 'complex' than a representation that does not.

Finally it is not clear from the results in Experiment 2, what is the relevant primitive mediating the observed effects. In other words, the internal discontinuity in the segmented object displays is 'interpreted' as a surface or volumetric component boundary. The experiments in the next Chapter approach this issue more directly.

To summarise, the results from Experiment 2 have shown that object-based IOR is modulated by internal features of objects, allowing to reject the global structure hypothesis. The results also extend previous findings of facilitatory part-based effects to inhibitory mechanisms of selection, by showing that inhibitory object-based selection can operate over representations that make explicit information about internal shape structure.

Chapter 4

4. Investigations into the Nature of Object Structure Modulating Object-Based IOR

4.1 Introduction

The four experiments reported in this Chapter are designed to further examine the role of object internal structure in the modulation of inhibition of return and to investigate the structural shape primitives responsible for such modulation. In Experiment 2 (Chapter 3), results suggested that for two-dimensional forms, object-based selection operates over internally structured shape representations. However, the results of Experiment 2 do not distinguish among the different hypotheses about the nature of shape representations over which these processes operate. In other words, the internal discontinuity in Experiment 2 could be interpreted as a surface or a volumetric component boundary. The aims of the experiments in this Chapter can be summarised as follows:

1. Establish the generality of the results from Experiments 1 and 2 to three-dimensional forms.
2. Investigate the nature of shape representations (shape primitives) modulating the IOR effect.

As mentioned earlier different theories of object recognition propose different intermediate stages of object-shape representation. Thus, whilst some theories propose that the representation of surfaces is a necessary intermediate stage for object recognition (e.g. Marr & Nishihara, 1978; Nakayama, He & Shimojo, 1995), others argue that the most efficient way of representing an object is by representing its component volumetric primitives (or *geons*) in certain spatial relationships (e.g. Biederman, 1987). Here the aim is to distinguish between the two types of shape primitives in their modulation of object-based IOR. Figure 21 illustrates the two alternative shape primitives explicit in the representation accessible to inhibitory component of object-based attention. In the following sub-sections, I present some evidence from studies of object recognition in support of surfaces or volumetric components as the *primal access*[17] primitive in object shape representation.

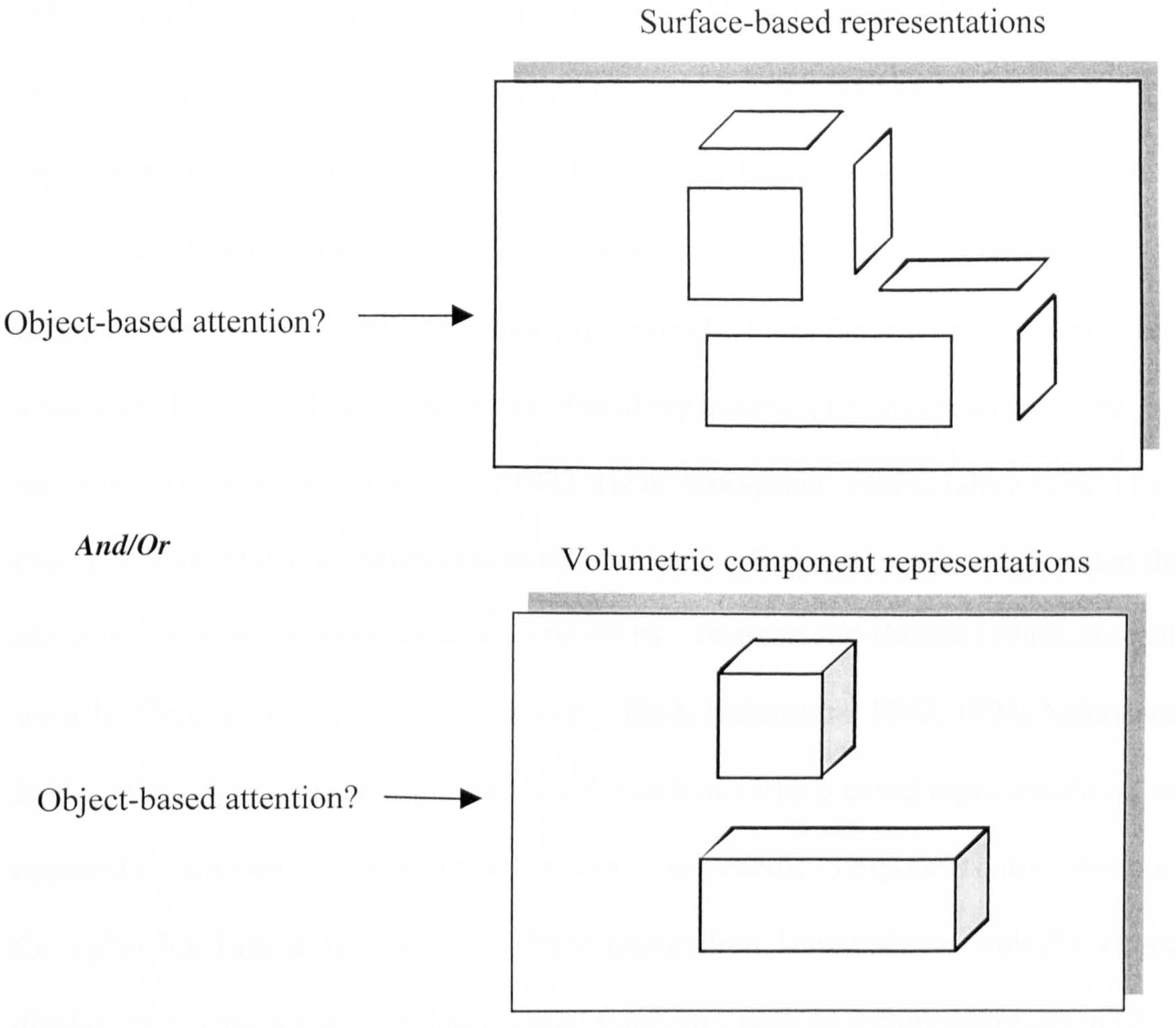


Figure 21: Schematic illustration of surface and volumetric component representations of object structure. Two alternative representations are considered to modulate the object-based attention (measured by object-based IOR). One making explicit surface primitives and the other making explicit volumetric primitives.

[17] "...the first contact between a single, isolated, undegraded, unanticipated object, and a representation in memory." (Biederman & Ju, 1988, p.40)

4.1.1 Surface-based representations in object recognition

Whereas in some object recognition theories (Biederman, 1987; Lowe, 1985; Ullman, 1989) surface information is of secondary importance, there exists a group of theorists that consider that surfaces are more immediately derived primitives from the image (Marr & Nishihara, 1978; Marr, 1980; Nakayama, He & Shimojo, 1995). These theories posit that surfaces are suitable features for the intermediate-level representations, formed *following* lower-level image-based representations and *prior to* the operation of higher-level, object-recognition processes, and as such are an indispensable part of visual processing. Accordingly, these theorists provide evidence suggesting that primal access to surface-based representations *alone* can facilitate recognition (Nakayama & Shimojo, 1992; He & Nakayama, 1992). Julesz (1981) for example claimed that attention is drawn to 'textons', which are primitives of vision that are more complex than the features proposed by Treisman and Gelade (1980). Recent work by He, Nakayama and colleagues (e.g. He & Nakayama, 1992, 1994; Nakayama & He, 1995; Nakayama, et al., 1995) has shown that surface-based representations, as opposed to representations of image features of volumetric components, are the input for higher-level visual functions, i.e. object recognition. Using stereoscopically viewed displays they have shown that basic visual functions, such as texture segregation or apparent motion perception are linked not so much to the retinal image, but to the formation of a surface representation of the input image.

In an elegant visual search experiment He and Nakayama (1992) showed that attention is directed to surface representations rather than to features of the image. They used a visual search task and manipulated the surface representation of features whilst leaving the featural representation (image features) intact. This was achieved by altering the depth relation of the stimuli on the display by means of shifts in disparity. That manipulation led to dramatic effects on visual search performance. They asked the participants to detect an inverted L-shape amongst upright L- shapes (Figure 22). In addition, they presented a black textured square that was adjacent to the upright and inverted Ls, so that when the squares were in front the Ls were perceived as amodaly[18] completed white squares, and when the squares were behind, the Ls were perceived as Ls. Reaction times to targets (e.g. the inverted Ls) were slower when the Ls were perceived as squares, than when they were perceived as Ls.

(A)

(B)

Figure 22. An example of the stimuli used by He and Nakayama (1992). Stimulus (A) was perceived as an L overlapped by the black square. Stimulus (B) was perceived as one white surface (created by amodal completion of the L behind the black square) and one black surface (Adopted from He & Nakayama, 1992).

This result indicated that leaving the features of a display intact but changing the surface representation (by means of binocular disparity) led to change in performance. He and Nakayama (1992) concluded that the input for visual search is a surface-shape representation rather than an edge-based representation.

He and Nakayama (1994) have also shown that the perceived direction of apparent motion can be manipulated by the layout of surfaces. Participants were more

[18] Here amodal completion refers to the formation of a larger surface from small regions that complete behind a visible surface (also see Nakayama, He & Shimojo, 1995, p.10)

likely to benefit from increased binocular disparity of two rows of items and perceived horizontal apparent motion, if the items were presented on two separate depth planes. However, no such benefit occurred when the two rows of items were perceived to rest on the same surface or comprise a single surface. They reasoned that the increasing bias towards perceiving horizontal apparent motion in their first experiment was due to the perception of two separate surfaces, as opposed to three-dimensional distance between the items (He & Nakayama, 1994).

Of more pertinence to the distinction between attention oriented to surfaces and attention oriented to volumetric components attempted in this thesis, is evidence - some already reviewed in Chapter One – showing that attention can generalise across perceived surfaces. He and Nakayama (1995) used a spatial cueing task, in which participants were asked to detect a target item in either of two rows of items, following an endogenous cue pointing with 80% probability to the row where the target would appear. The disparity between the two rows was manipulated and so was the relationship between the two rows of items in terms of surface representation. As shown in Figure 23, the rows appeared either in separate frontoparallel depth planes (A), in separate frontoparallel planes but 'resting' on the same depth plane (B), or in a single stereoscopic plane (C). He and Nakayama (1995) found that participants were faster at responding to the cued target when binocular disparity between the two rows (one and one uncued) was increased. However, that benefit of cueing by increasing disparity was only found in condition (A), where the two rows belonged to different surfaces. In contrast, when the items lied across a common surface (conditions B and C), then increasing disparity between the rows did not lead to a significant benefit for the cued row. This finding led He and Nakayama to conclude that "The visual system....can direct selective attention efficiently to any well-formed, perceptually distinguishable surface" (p. 1155).

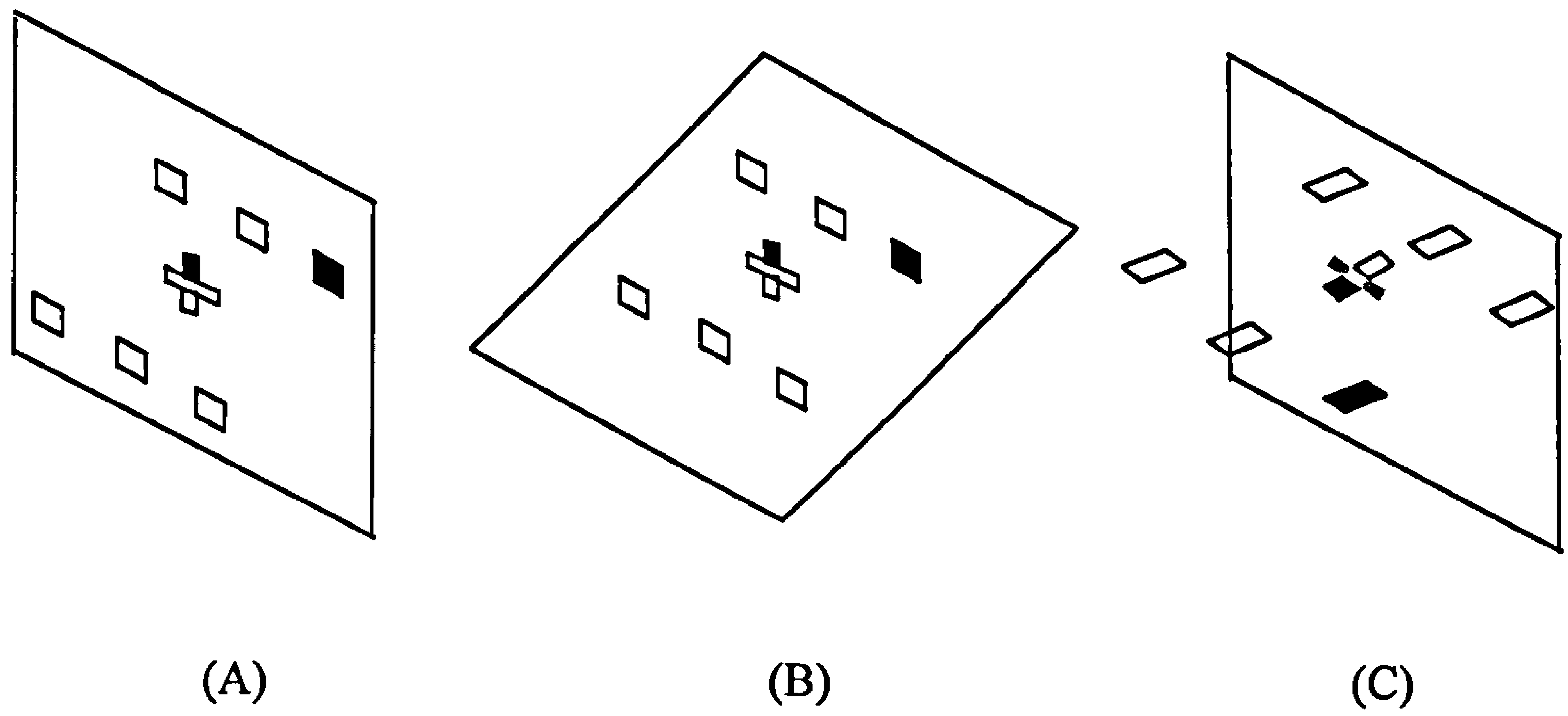


Figure 23: (Adopted from Nakayama et al., 1995). The two rows of items are presented with variable disparity (not depicted here) under three different stereoscopic depth conditions. In terms of cueing, all three panels depict a valid trial – the cue (black portion of a fixation cross) points to the row that contains the target.

Recently Xu and Nakayama (2001) reported a study in support of the notion that surfaces but not objects are important for selective attention processes. They presented participants with search arrays appearing on either of two visible faces of 16 cubes. The search items would be presented either on the same face of all cubes (consistent-face condition) or on either of the two faces of the cubes (mixed-face condition). They also used presented the search arrays on 16 2D stimuli with three surfaces to match the three visible surfaces of the (3D) cubes. Xu and Nakayama (2001) found that in the 3D stimuli, search for the target item was significantly slower in the mixed-face condition than in the consistent-face condition. This was not the case in the 2D search arrays. The reaction time advantage for targets in the consistent-face condition was interpreted as evidence for attention to surfaces as opposed to the objects these surfaces belong to.

Surface-based representations have also been shown to have a primary role in object representation in naming tasks. Humphrey, Goodale, Jakobson & Servos (1994) investigated the role of colour along with other surface properties in object recognition. They presented the visual agnosic patient D.F. and two normal controls with a set of real objects (some natural and some man-made) and a large range of coloured, black and white, and line-drawing pictures of natural and manufactured objects. Viewing conditions varied between full viewing, monocular viewing, monochromatic viewing (available luminance but not colour information) and monocular monochromatic viewing. In general, D.F.'s naming accuracy and naming latencies were better for real objects than for pictures or line-drawings of objects. Her naming latencies for naturally coloured objects (i.e. a green pepper) were faster than for achromatic (camera) or artificially coloured objects (a coloured wine bottle). Humphrey et al. (1994) suggested that D.F. was most reliant on surface information, i.e. texture, colour, and when this information was unavailable (under conditions of monocular or monochromatic viewing) her performance worsened. Thus, information about the object's contours or edges did not compensate for the loss of surface information. In the case of manufactured objects (i.e. cameras, telephones), however, naming latencies in the real object condition were the same as latencies in the coloured picture or line-drawing condition of the same object. Inappropriately coloured line drawings of objects, however, led to longer naming latencies than appropriately coloured ones, leading Humphrey at al. (1994) to suggest that surface information operates at a higher-level of visual perception where its representation facilitates or even mediates recognition. However, it is possible that the benefit for real objects is due to depth information available under normal viewing conditions, but otherwise lacking in photographs and line-drawings of objects. In this case, Humphrey et al. (1994) findings have little to do with surface-based representations but more to do with surface information through depth cues.

More recently Leek and Arguin (2000) used a part-whole matching task to investigate whether object shape representations are based on descriptions of edge contours, surfaces or volumetric components. Participants were briefly presented with an edge contour, a surface or a volumetric component prime and subsequently with a novel shape object, of which the prime was part of or not. Participants were faster and more accurate in matching volumetric component primes, than edge primes, to the novel shape. In their Experiment 2, however, when the amount of visible surface area in the volumetric component and the surface primes was matched, both types of prime led to faster and more accurate part-whole matching, than edge contour primes. Leek and Arguin (2000) argued that a surface-based representation must be computed when recognising objects.

4.1.2. Evidence for volumetric components in object recognition

The evidence reviewed above placed major emphasis on the representation of visible surfaces as an important intermediate stage of object shape representation. However, the theory put forward by Marr & Nishihara (1978) posit that, for recognition to occur, three-dimensional objects are represented by descriptions of axis-based *generalised cones* (also known as generalised cylinders), a proposal first put forward by Binford (1971). Binford proposed that all three-dimensional objects can be described on the basis of cylinders of variable size and shape according to the object's overall shape (as shown in Figure 8 in the Introduction). Generalised cylinders are constructed – online – by sweeping a two-dimensional shape along an axis (Figure 24).

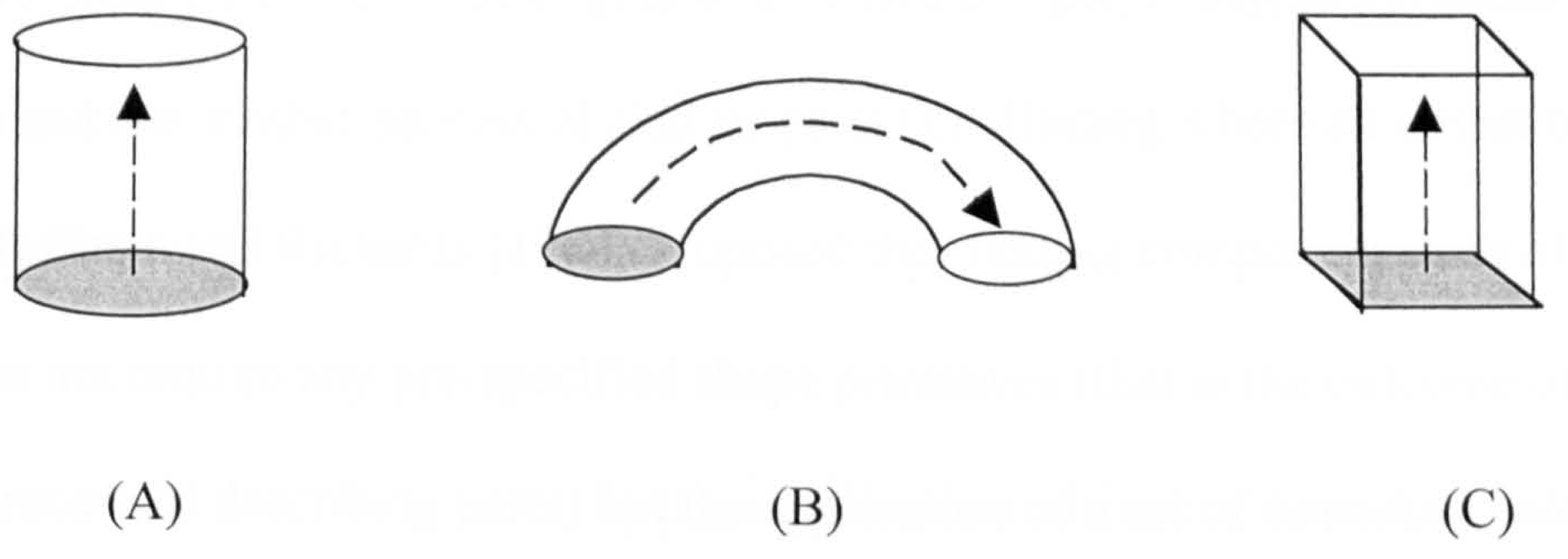


Figure 24. Examples of three different generalised cylinders constructed by sweeping a variable 2D shape (i.e. A and C) along a particular axis (i.e. A and B). (Adopted from Palmer, 1999).

These volumetric primitives were later conceptualised as simple, viewer-independent geometrical forms or *geons* (short for geometric ions) in Biederman's (1987) Recognition-By-Components (RBC) theory of object recognition. In Biederman's theory, an object is represented as a specific spatial arrangement of geons. Geons are recovered from low-level features of the image, such as edges and vertices as well as the image's non-accidental properties (properties of the image, such as symmetry or cotermination, that are unlikely to have emerged by accident; also see Lowe, 1985). According to RBC theory such low-level features and properties of the image alone cannot lead to the representation of an image (Biederman & Ju, 1988; Biederman & Cooper, 1991). Rather, they lead to the identification of the nearest geon that approximates these non-accidental properties on the image.

Biederman's (1987) proposal that a retinal image of an object is segmented into parts borrows an important principle for doing so, the *transversality* principle, from Hoffman and Richards' (1984) account of part representation. Hoffman and Richards (1984) proposed that part representation consists of two separate processes, one of *describing* parts (i.e. as belonging to a certain category, such as pyramids or cylinders) and the another process of *finding* parts (i.e. finding where an object can be parsed). Hoffman and Richards (1984) proposed that finding component parts of an object does not require any pre-specified shape primitives (that is the outcome of a separate process of describing parts) but the application of a set of *boundary rules* (or parsing rules). In this approach, sharp changes in surface orientation define the boundaries between component parts, i.e. when two parts intersect, they meet at points of concave discontinuities (see Palmer, 1999).

There is some empirical evidence to support the idea of the three-dimensional shape is represented in terms of volumetric primitives (e.g. Binford, 1971; Marr, 1982; Biederman, 1987; Biederman & Cooper, 1991). Biederman (1987) showed that naming times of line drawings of familiar objects were negatively affected by the number of deleted volumetric primitives from those images. This was taken as evidence in support of a volumetric component - based account for object shape representation. Subsequently other studies have supported this conclusion.

One study comparing the sufficiency of edges and volumetric components in representing object-shape was reported by Biederman and Cooper (1991), using the *priming paradigm* (Bartram, 1974; cited from Palmer, 1999). The basic priming procedure involves participants viewing two successive displays of objects. The idea behind this is that naming an object (i.e. "piano") is faster and more accurate when the same object (or object category) was previously viewed. In an image priming experiment, Biederman & Cooper (1991) investigated whether low-level features (i.e. edge boundaries, vertices) in their specific relations or component parts in their specific relations are responsible for object shape priming. They constructed two versions of line-drawings of objects, each depicting half of the contour or half of the component parts of an object, with the condition that once the two images were superimposed on each other they would reveal the picture of the whole object.

In two experiments they presented observers with images of objects with 50% of their contour removed either by deleting image features (Experiment 1) or image volumetric components (Experiment 2). In the first experiment, observers viewed the priming image, depicting objects with half their contour, or half of their component parts deleted. Subsequently they were presented with another set of pictures depicting either exactly the same image, an image with the complementary contours, or a same class exemplar. Biederman and Cooper (1991) found that priming effects were the same when they repeated the components of the prime (but not the same line segments as the prime; complementary contour condition) and identical images as the prime. This finding was taken to support that component parts of the object play an important role in the object's representation, as opposed to low-level features, such as the edges and vertices of the object. This conclusion was reinforced by the results of their second experiment, where observers were primed with images of objects with half their components deleted. When they repeated the same object but consisting of the other half of their parts, priming was much less, than in the condition where the image identical to the prime was repeated. Biederman & Cooper (1991) concluded that the intermediate representation of higher-order components and their spatial interrelations play an important role in object recognition and categorisation[19].

However, relatively few studies have attempted a direct comparison between surface-based and volumetric-based object shape representations (e.g. Biederman & Ju, 1988; Humphrey et al., 1994; Leek & Arguin, 2000). The issue in question when comparing surface-based and volumetric-based representations of objects is whether the presence of surface cues, such as changes in surface attributes like texture, luminance or orientation, facilitates the recognition of an object *over* information that is solely derived from depiction of the object's edges and vertices on the image (Biederman & Ju, 1988). Biederman and Ju (1988) compared response latencies in naming objects when these were depicted either as line drawings or as coloured photographs. They found no consistent difference in naming performance (their Experiments 1-3) between coloured photographs and line drawings. Biederman and Ju (1988) argued that this finding was in (indirect) support of the idea that edge-based descriptions (from which geons are derived) suffice for the representation of an object, whilst surface attributes, such as colour and texture, may play a secondary role in the activation of the object's shape representation in memory (i.e. sharp changes in surface attributes constitute the necessary edges and vertices that define points of convexity and concavity on the image, as Biederman & Ju have pointed out). Information about regions of concavity or curvature are readily available in an edge-based, volumetric description of the object.

---

[19] Biederman and Cooper (1991) note that in this specific investigation the term components corresponds to the geons as in Biederman's Recognition by Components (RBC) theory (p.413).

In summary, issue of the relevant primitives for object-shape representation for the purposes of recognition is a matter of an ongoing debate, with ample but inconclusive evidence on both candidate primitives. The question addressed in the following section is whether object-based inhibitory attention selects surface or volumetric component-based internally structured representations.

4.2. Experiment 3

4.2.1 Introduction

The aim of Experiment 3 was to investigate whether object-based IOR can be modulated by internal discontinuities in three-dimensional forms, thus allowing to examine the generality of Experiment 2 results to 3D forms. Furthermore, the results from Experiment 2 do not allow any inferences about the nature of internal structural primitive that modulates inhibitory effects. In this experiment, images of 3D forms are used to determine whether discontinuities that define surfaces as opposed to 2D regions of space (as in Experiment 2) can lead to modulation of IOR similar to that in Experiment 2.

The finding that object-based representations may influence IOR has been previously demonstrated (e.g. Gibson & Egeth, 1994; Houghton & Tipper, 1994; Tipper, Jordan & Weaver, 1999; Experiments 1 and 2 in this thesis). For example, Gibson and Egeth (1994) showed that location-based IOR can be obtained for locations that remain fixed with respect to an object. Gibson and Egeth (1994) used displays of a single brick rotating in depth to investigate whether locations defined with respect to the spatially invariant description of an object would be as inhibited as locations defined in environmental spatial co-ordinates (defined with respect to a landmark external to the viewer). They found IOR for locations defined both with respect to the (screen) environment *and* with respect to the object. Gibson and Egeth (1994) argued that the internal structure of the object can serve as a *spatial medium*, equivalent to the spatial medium of the environment (i.e. a location defined with respect to a landmark). Furthermore, they found that cueing an object's surface and subsequently attracting attention to a target on a different location of the same surface produced less IOR. Their results suggested that cueing a local region of an object – which in their case was the front surface of the brick - may not have entailed tagging the global description of that object[20] (also see Peterson & Gibson, 1991). This conclusion, however, is in contrast with the findings of Experiment 2 in this thesis, where it was found that not only is IOR observed across the object, but also that IOR was larger across the internal discontinuity, indicating that internal properties of the object affect the pattern of modulation of object-based IOR[21].

In this experiment displays of two images of 3D objects (bricks) are used. The general method was very similar to that of Experiment 2. Cues and targets appear with equal probability on either object. Following the offset of the cue, the target appears on the exact same location as the cue, on the same surface as the cue, on the same object but a different surface from the cue, or on a different object from the cue (baseline cueing condition).

It is predicted that IOR would (a) generalise across the cued surface of a 'brick' and (b) the object-based IOR effect for targets appearing on a *different surface* from the cue would be *larger* than the IOR effect for targets appearing on the *same surface* as the cue (as in Experiment 2). Such an outcome would indicate that

---

[20] Gibson & Egeth (1994) concede that their experiments do not permit any general conclusions about object-based IOR, in the sense that their results cannot differentiate between IOR that operates over spatially invariant description of an object and IOR that operates over spatially defined locations within an object (similar to location-based IOR).

[21] One crucial difference between Gibson & Egeth's set of experiments and Experiment 2 (as well as in the following experiments in this thesis) is that participants viewed displays of two separate objects. This allowed control over (a) which object is cued and (b) which location is cued within the object. Gibson & Egeth (1994) had suggested that such a design would allow one to make more conclusive suggestions as to whether IOR can 'accrue' to spatially invariant object-based representations (p.337).

inhibitory selection operates over representations of the object that differentiate between the object's surfaces, supporting the local structure selection hypothesis for 3D forms. In this experiment, a different surface is defined not only by a surface boundary (which was the case in Experiment 2), but also by the change in orientation in depth defining the three-dimensionality of the image.

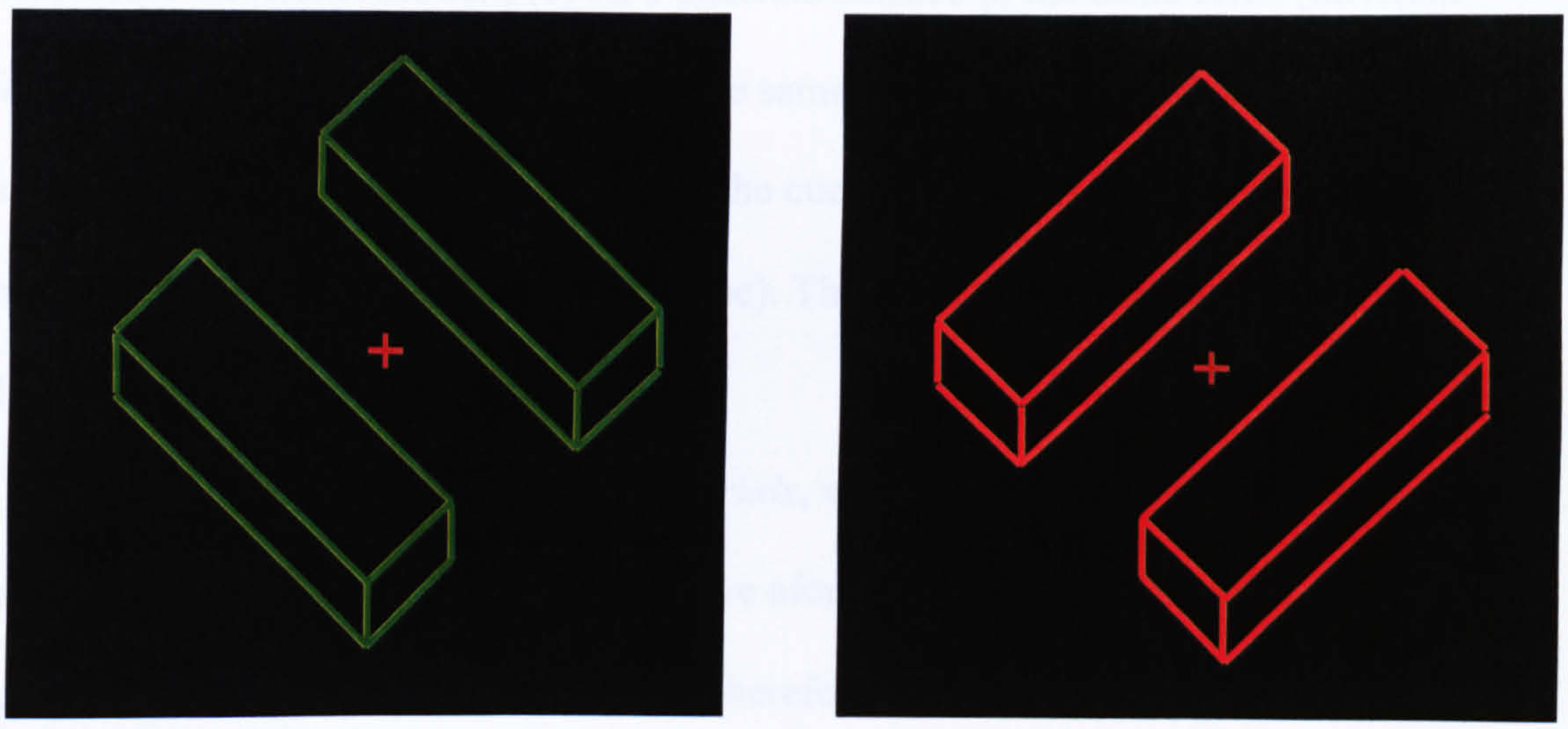
4.2.2 Method

Participants

Twenty psychology undergraduates aged between 19-32 years, recruited from the Student Participant Panel at the University of Wales, Bangor took part in this 1-hour experiment for two course credits. All participants reported normal or corrected-to-normal vision. None had participated in the previous experiments.

Apparatus/Stimuli

Stimuli were presented on a 14-inch monitor connected to a Power Macintosh PC. Randomisation and presentation of the stimuli, as well as recording of the participants' reaction times, were controlled through PsyScope software (version 1.2.4). Responses were made through a single letter key on a standard Apple keyboard connected to the computer. The stimulus display consisted of two outline (red or green) 'bricks', simultaneously presented on each side of a fixation cross (see Figure 25) against black background. The orientation of the 'bricks' varied randomly between trials, appearing either +45 or -45 degrees tilted from the vertical meridian. At each orientation, there were three visible surfaces, namely the *top*, the *side* and the *front* surface, each subtending 9.15° x 3°, 9.15° x 1° and 3° x 1° degrees of visual angle respectively when viewed from 50 cm distance. The fixation cross (+) subtended 0.8° x 0.8° of visual angle. The cue was an outline white parallelogram subtending 1.0° x 1.0° (its white contours subtending .02°). The target was a filled white parallelogram measuring 0.8° x 0.8°. The central re-fixation cue was a white cross (+) sign subtending 0.8° x 0.8°, that replaced the black fixation cross. From end to end, the display subtended 14° x 16° of visual angle.



(A) (B)

Figure 25. Stimuli used in Experiment 4.1. The stimuli are tilted -45 degrees from vertical in Panel A and +45 degrees from vertical in Panel B.

Design

The experiment was based on a three factorial within-subject design with factors of Cueing, with five levels, Display orientation with two levels (+45 and -45) and SOA, with two levels (820 and 1220 msec).

Design

The experiment was based on a three factorial within-subject design with factors of Cueing, with five levels, Display orientation with two levels (+45 and -45) and SOA, with two levels (820 and 1220 msec).

Levels of Cueing consisted of the following five cue-target configurations; the target appeared (a) on the same surface and at the same location as the cue (same surface-same location), (b) on the same surface but at a different location from the cue (same surface-different location) (c) on a different surface of the same brick (different surface/*large*) (d) on a different surface of the same brick (different surface/*small*) (e) on a different brick on a different surface to the cue but at a distance equal to the within-object targets (different object-baseline). The five cueing conditions used in the analysis (a-e) are shown in Figure 26.

There were also three types of *filler trials*, where the cue-target distance was greater than the cue-target distance for the five aforementioned within-object cueing conditions (4.2 degrees of visual angle) and therefore, due to these spatial distance constraints were excluded from the analysis. The cue-target positions in those trials were as follows: the target appeared (1) on a different brick, on the same corresponding surface as the cue (*different object 2*), (2) on a different brick, different surface and different corresponding location from the cue (*different object 3*) and (3) on the front surface of the different brick *(different object 4*).

Participants completed 10 practice trials followed by 480 experimental trials, of which 160 trials (33%) were 'no-target' trials. In the remaining 320 trials, there were 40 trials for each of the eight cueing conditions , 20 for each display orientation (+45 and –45).

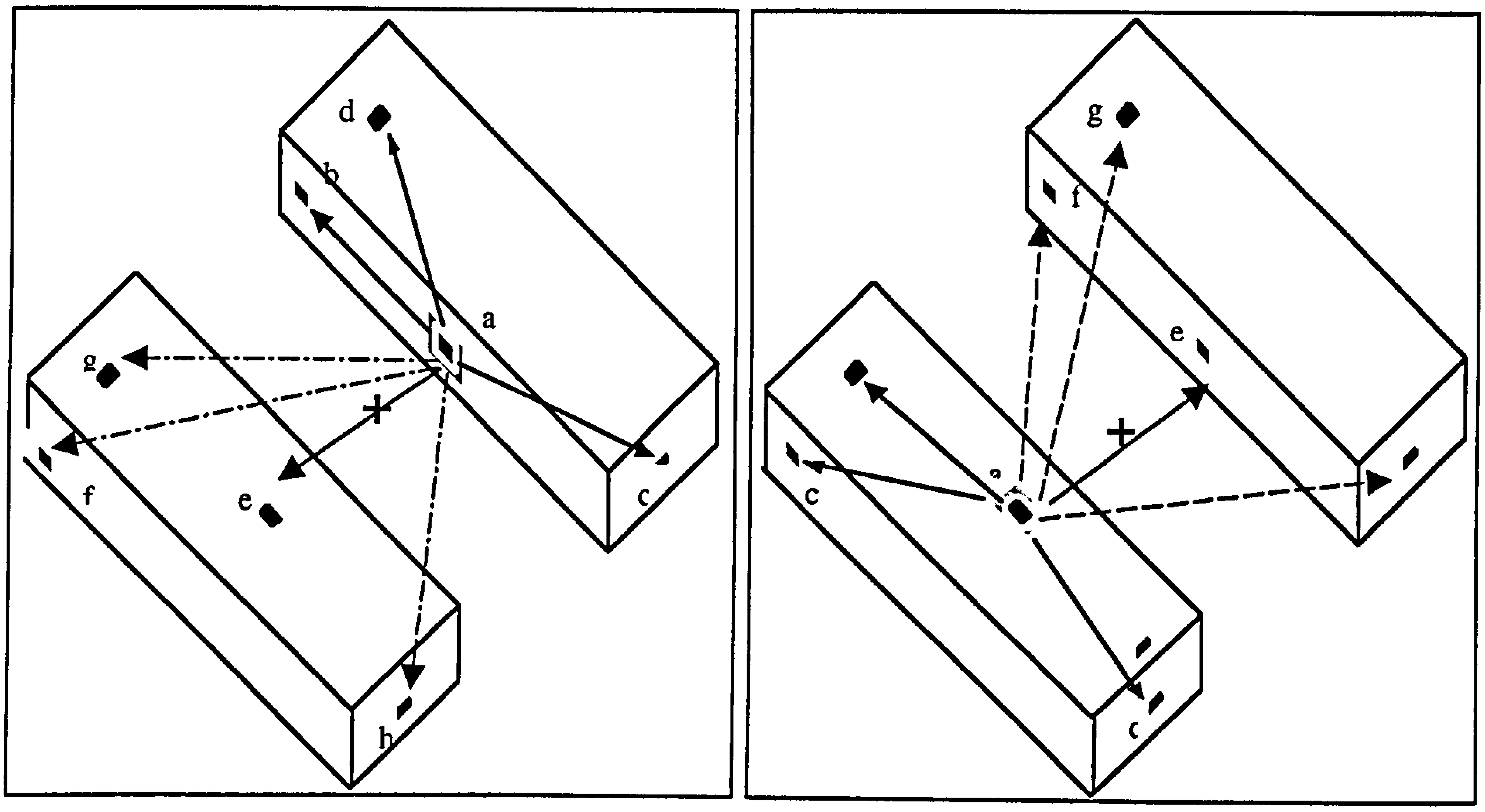


Figure 26 key:

(a) 'same surface-same location'
(b) 'same surface-different location'
(c) 'different surface/front'
(d) 'different surface/side'
(e) 'different object-baseline'

Filler trials (dashed lines)

(f) Filler 1: 'different object/same surface'
(g) Filler 2:'different object/different surface(large)'
(h) Filler 3: 'different object/different surface (small)'

Figure 26. The figure illustrates the position of cues and targets for each of the eight cueing conditions (see Design for details). Cues (that were white outline parallelograms in the experiment) are depicted in grey and targets (that were filled white parallelograms in the experiment) are shown in black. The cue-target distances in all cueing conditions (except in the 'same surface-same location') were 4.2° of visual angle.

Procedure

At the beginning of each trial, a fixation cross appeared at the centre of the monitor. After 1000 msec, the two 'bricks' were simultaneously presented on each side of the fixation cross. Following a further 1000 msec delay the peripheral cue appeared at one of two possible (on the left or on the right of fixation) locations within the 'bricks' (Figure 26). Cue duration was 90 msec, as was for Experiments 1 and 2. Following 300 msec from cue offset, the central fixation cross changed from red to white for a period of 130 msec and then reverted to red until trial end (central re-fixation). After a further delay of 300 or 600 msec the target was presented in one of eight random and equiprobable locations within either 'brick', yielding four cued-object and four uncued-object cueing conditions in total. The target remained visible for a 1000 msec, or until the 'b' key (response) was depressed.

Participants were informed that the white outline parallelogram (the cue) was not predictive of the location of the subsequent target. They were instructed to depress the response key as soon as they detected the target (filled white parallelogram) and to withhold their response when no target was presented. In case they failed to do so, the computer generated a 500 Hz tone, and the trial was considered an error. Furthermore, the importance of fixating at the cross was stressed during the practice period and just before the experiment commenced.

4.2.3 Results and Discussion

Trials with RTs that were greater than 700 msec or less than 100 msec, or incorrect ('no-target' trials) were discarded from the data. These exclusions accounted for 1.1% of all trials. The 'different surface/*large* and 'different surface/*small* cueing conditions were collapsed and formed the 'different surface' condition[22]. Only four cueing conditions were used ('same surface-same location', 'same surface-different location', 'different surface' and 'different object', also see Figure 26). Figure 27 shows the object-based IOR (mean RTs of cued targets minus mean RTs for 'uncued object' targets) for each of the three within-object cueing conditions in each of the two Display orientations collapsed across the two SOAs.

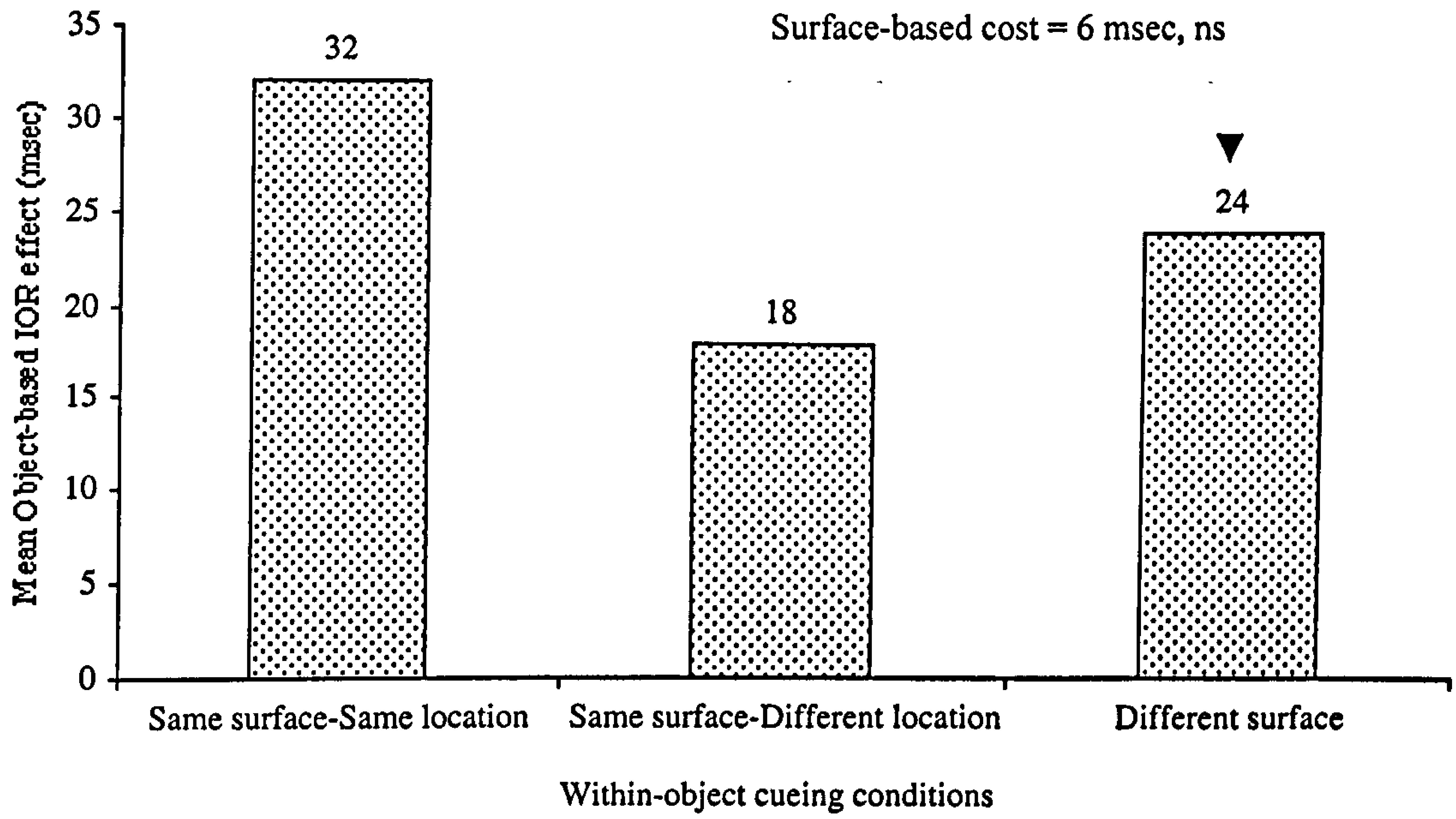
Reaction time data were entered into a 4 (Cueing) x 2 (Display orientation) x 2 (SOA) repeated-measures ANOVA. There was a significant main effect of Cueing, [F (3, 57) = 11.3, p<.0001] and a significant main effect of Display orientation, F (1,19) = 10.5, p< .004. There was no significant effect of SOA [F (1, 19) < 1, ns]. There were no significant interactions between any of the three factors, whilst the Cueing by Display interaction was only marginally significant, F (3, 57) = 2.1[23].

Planned comparisons were carried out in order to establish whether the internal structure of the object (defined by a surface boundary and change in surface slant) modulated IOR. RTs for the 'same surface-same location' (M=375, SD=55.0), 'same

[22] Post-hoc tests were carried in order to determine whether there were any differences between targets in the side and front 'uncued surface' condition. These showed that there was no significant difference between targets appearing in the 'different surface/large' and 'different surface/small' conditions, t (19) <1, ns.

[23] A 2 ('same surface-different location' and 'different surface') x 2 (+45 and -45) repeated-measures ANOVA showed a significant effect of Display orientation, F (1, 19) = 6.8, p < .01 and a significant Cueing by Display orientation interaction, F (1, 19) = 4. 1, p < .05. Planned comparisons showed that in the +45 orientation RTs for the 'same surface-same location' (M=372, SD=47.0), 'same surface-different location' (M=349, SD=57.3) and 'different surface' (M=363, SD=52.0) conditions were significantly different from the 'different object-baseline' condition (M=335, SD=49), t(19)=-4.0, p<.001, t(19)=-2.3, p<0.03, t (19)=-4.2, p<.001, respectively. Critically, the difference between RTs to 'same surface-different location' and 'different surface' conditions was also significant, t (19)=2.0, p<.05. In the -45 orientation, RTs for the 'same surface-same location' (M=379, SD=49.0), 'same surface-different location' (M=369, SD=56.2) and 'different surface' (M=367, SD=53.5) cueing conditions were significantly different from the 'different object' cueing condition (M=347, SD=48.7), t (19)=-3.8, p<0.001, t(19)=-3.86, p<0.001, t (19)=-3.40, p<0.003 respectively. However, there was no significant difference between RTs to 'same surface-different location' and 'different surface' conditions in the –45 display orientation, t (19) <-1.

surface-different location' (*M*=360, *SD*=62.4) and 'different surface' (*M*=366, *SD*=53.7) cueing conditions were significantly different from the 'different object' cueing condition (*M*=341, *SD*=53.8), $t(19)=-3.8, p<0.001$, $t(19)=-3.86, p<0.001$, $t(19)=-3.40, p<0.003$ respectively. However, the difference between RTs to 'same surface-different location' and 'different surface' conditions (6 msec) was not significant, $t(19) < -1$.



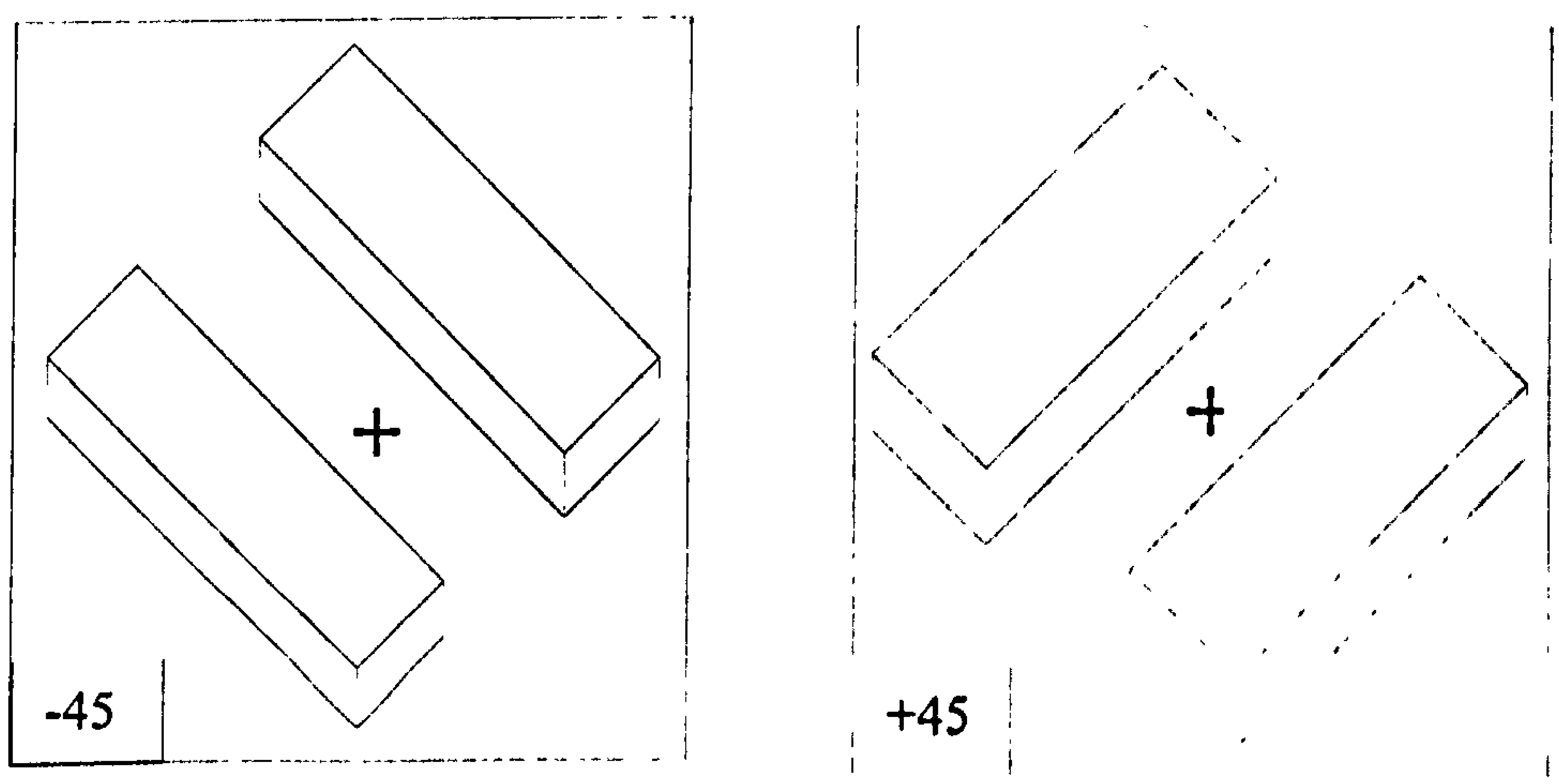


Figure 27. Illustration of the object based IOR effect observed for each within-object cueing condition (collapsed across the two orientations).

Post-hoc tests were also carried out between the Filler (different object) conditions, in order to establish (a) that is it not simply easier to attend to the centre of the object relative to the ends of the same object (hence faster RTs for the chosen baseline, which would increase the IOR effect for the within-object conditions), and (b) that the modulation of IOR by object-internal structure is not an artefact of low-level visual factors. These revealed no differences between the 'Different object' condition ($M = 341$, $SD = 53.8$) and Filler 1 ($M = 352$, $SD = 57.6$), Filler 2 ($M = 348$, $SD = 51.6$) and Filler 3 ($M = 352$, $SD = 52.6$) [$t\,(19) = 1.9$, ns; $t\,(19) < 1$, ns; $t\,(19) = 1.3$, ns, respectively]. The difference between the three Filler cueing conditions were also non-significant.

4.2.4 Discussion

The main findings from Experiment 3 can be summarised as follows. First, the finding that object-based IOR can generalise across the surface of an object was replicated using 3-D depictions of single component objects. The generalisation of IOR from one surface of the object to the other was similar for both small and large surfaces (see footnote 19). In this experiment surfaces are not only defined by a surface boundary but also by change of slant and tilt with respect to other surfaces of the object. This finding that IOR can accrue to a different surface of the cued object is consistent with the finding in Experiment 2. It is, however, inconsistent with previous findings from studies using visual search (e.g. Xu & Nakayama, 2001) and IOR methodologies (e.g.

Gibson & Egeth, 1994) which have shown that attention does not spread from one surface of a 3D depiction of an object to another. Further implications of this finding will be discussed in the General Discussion (Chapter 5).

Second, the modulation of IOR by the change of surfaces, whilst not significant, followed the same trend as in the previous experiments, where 2D forms were used. That is, object-based IOR was *larger* when cues and targets appeared on different surfaces of the same object. This pattern, however, only reached significance in the +45 display orientation. The issue of the pattern of IOR modulation, which will be explored in the General Discussion, was raised in Experiment 2, where targets appearing on a *different* surface from the cue were detected more slowly than targets appearing on the *same* surface as the cue.

Third, there was a significant effect of Display orientation, whilst the Cueing by Display interaction did not reach significance. One reason for this lack of interaction was that RTs in the 'same object/same location' condition were almost identical for the two orientations. When the two critical within-object conditions were examined, results confirmed a significant Cueing by Display interaction, revealing that object-based IOR was modulated by object-internal structure in the +45 but not in the -45 display orientation. How can this difference be accounted for? An interpretation of this finding may be in terms of the operation of local and global processing mechanisms. There are two possible accounts. One is that separate local and global structure representations are computed, and are both accessed by object-based inhibitory mechanisms. The other account, consistent with the local structure hypothesis, is that only one type of representation is computed, which makes explicit both the local and the global object structure. The present results are consistent with this second possibility allowing to exclude the global selection *only* account not only for 2-D forms (Experiment 2) but also for 3-D forms (Experiment 3).

However, the question still remains that, if object internal structure modulates object-based IOR in the +45 orientation, then how would this be manifested in the results? To illustrate, in the +45 orientation the two objects point rightwards; whilst in the-45 orientation the objects point leftwards (see Figure 12). This raises two possibilities that relate to global/local advantage for the left and right hemisphere respectively (e.g. Delis, Robertson & Efron, 1986). One possibility is that the *direction* of the objects may have biased processing of the display by the left hemisphere revealing an advantage for local object structure. The other possibility is that only the *right side object* was processed in terms of its local structure in both orientations, but perhaps most prominently in the +45 displays.

At present there is little direct evidence for the first possibility, that the leftwards or rightward direction of the object display biases processing by the right or left hemispheres. However, there is some evidence suggesting that the left-right hemisphere advantage for local and global information respectively (e.g. Sergent, 1982; Delis et al., 1986) may be associated with the internal representation of space that is not necessarily referenced relative to the viewer's midline (e.g. Robertson & Lamb, 1988). For example, Robertson and Lamb (1988) examined whether visual hemifield differences in global/local processing are affected by a dynamic perceptual property that would change the perceiver's perspective or frame of reference imposed on the display. Participants were required to indicate whether a group of letters were normal or mirror images of letters (see Figure 28). In half of the trials the letters were presented either on the left or on the right of fixation as either normally oriented or mirror images of letters. In the other half of the trials the whole displays (including the frame surrounding the letters and the fixation point) were rotated clockwise or counterclockwise. Clockwise rotation would result in stimuli originally presented on the right appearing in the lower visual field and perceived as being presented *right relative to the top part of the letters* (Figure 28, middle column). Counterclockwise rotation would result in stimuli originally presented on the right appearing in the upper visual field and, again, perceived as *right relative to the lower part of the letters* (Figure 28, last column).

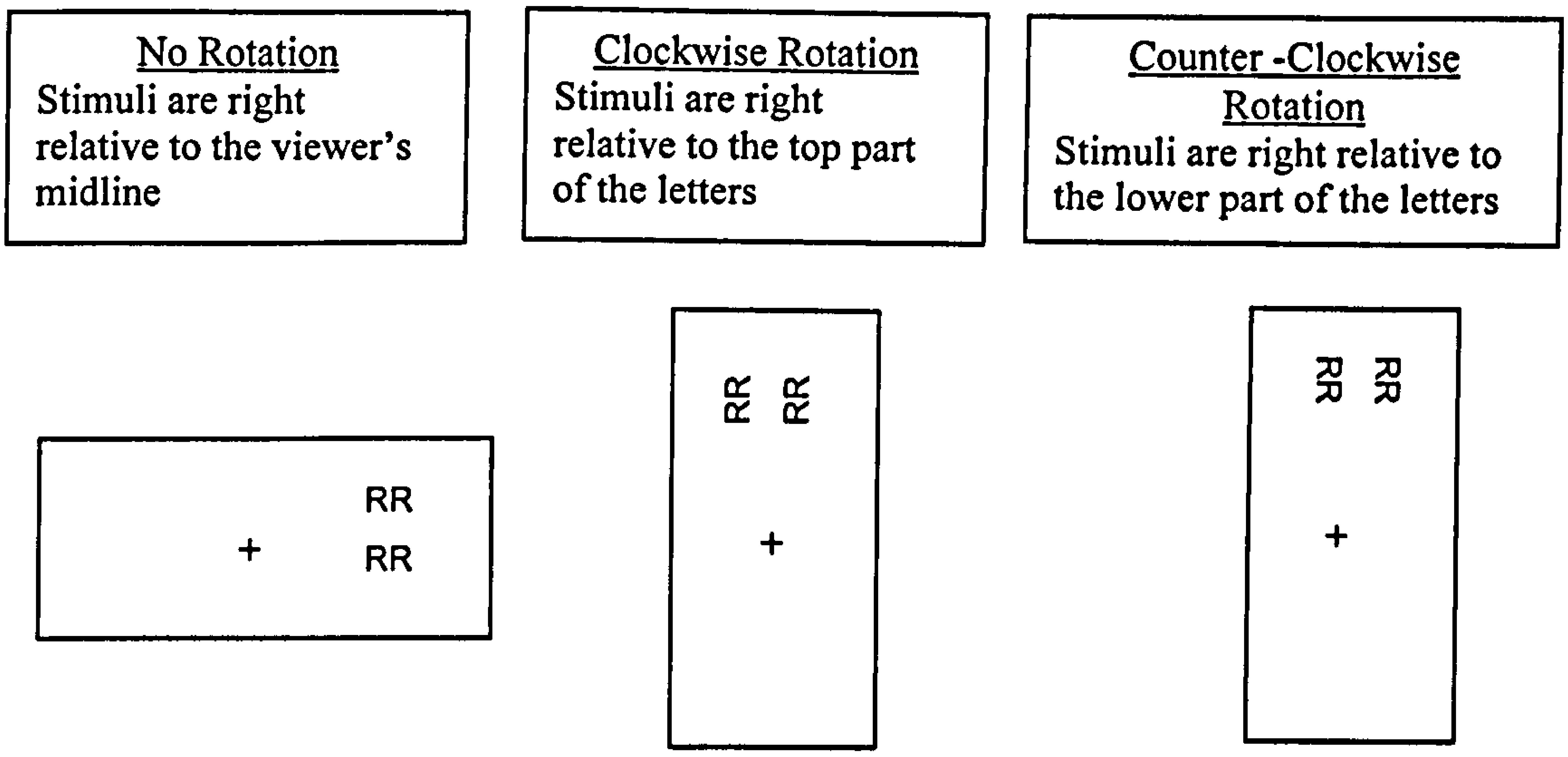


Figure 28: (Adopted from Robertson & Lamb, 1988). Examples of 'normal' letters presented either upright or rotated. Responses to whether the stimuli were normal or mirror images of letters were faster in the middle panel (clockwise rotation) than in the last panel (counterclockwise rotation).

Robertson and Lamb (1988) found that RTs "were faster to stimuli presented in the visual field that was 'right' relative to the reference frames defining the tops of the letters" (p. 149). They subsequently concluded that:

"It was the represented space defined by the rotation of the stimulus and not the absolute locations of stimulus presentation that determined the visual field advantage. Asymmetries can arise from several factors that are independent of the initially stimulated hemisphere. One of these factors is the reference frame the subject adopts" (p.150-151).

The relevant implication of Robertson and Lamb's (1988) study for the findings of Experiment 3, lies in their finding that visual field asymmetries can be modulated by the perceptual dynamics of a display, one of which is the orientation of the reference frame adopted during a task. Therefore, it is possible that object internal structure modulated the object-based IOR effect in the +45 display orientation by means of the rightward direction of the display, which facilitated local structure processing by the left hemisphere.

The second possibility is that modulation of IOR by the object's internal structure only occurred for the right side object in both orientations[24]. Some compelling evidence for the right visual field advantage (Left-hemisphere) for local structure processing comes from a pilot study reported by Delis, Robertson and Efron (1986). They presented left- and right-hemisphere damaged patients images of hierarchical stimuli (i.e. a large M made of small z) and asked them to draw the stimulus. Left-hemisphere damaged patients drew the larger form (i.e. M) but not the small forms that compose it (i.e. z). The opposite pattern was observed in right-

[24] Here, 'left' and 'right' objects are defined in terms of the objects' relative position in relation to fixation, as opposed to the display's vertical meridian.

hemisphere damaged patients. Delis et al. (1986) suggested that the left hemisphere "is superior in processing the smaller forms in hierarchical stimuli while the RH [right hemisphere] is superior in processing the larger forms..." (p.208).

On this account, there should be a difference in IOR modulation between the right and the left object in both orientations, with the right-side object processed locally and the left-side object processed globally. However, inspection of Figure 29 reveals that this is not the case. The figure shows that in the +45 orientation, the difference between targets appearing on the *same surface* as the cue and targets appearing on a *different surface* from the cue was significant at approximately 13 msec for the left object and 17 msec for the right object [25]. On the other hand, in the –45 display orientation the difference between the two within-object cueing conditions is approximately 3 msec for the left object and 9 msec for the right object[26].

[25] For the left object (in the +45 orientation) the difference between same surface (M=342) and different surface/large (M=368) and different surface/small (M=366) was significant, t (19) = 2.3, $p<0.05$ and t (19) = 2.1, $p<0.05$. For the right object the difference between the same surface (M=349) and different surface/large (M=371) and difference surface/small (M=359) was also significant, t (19) = 2.1, $p<0.05$ and t (19) = 2.0, $p<0.05$, respectively.

[26] For the left object (in the -45 orientation) the difference between same surface (M=377) and different surface/large (M=372) and different surface/small (M=370) was *not* significant, t (19) = .48 and t (19) = .34 respectively. For the right object the difference between the same surface (M=368) and different surface/large (M=361) and difference surface/small (M=356) was also *not* significant, t (19) = -1.2, and t (19) = -.7, respectively.

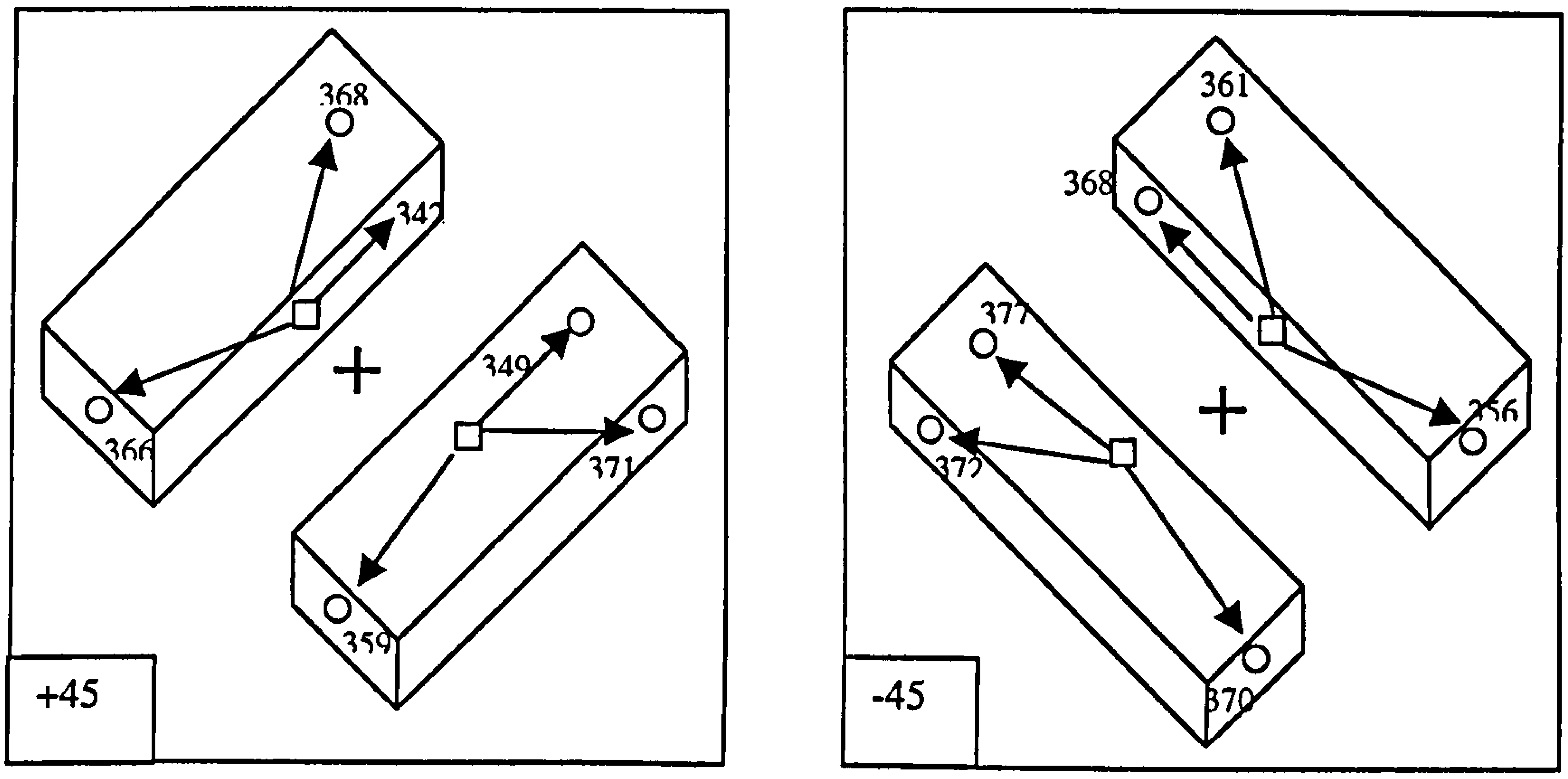


Figure 29: Showing mean RTs for the two cueing conditions (1. 'same surface-different location' and 2. 'different surface') in each object separately for the two display orientations.

Taking these differences into consideration the possibility that there was a bias for local structure processing of the object on the right of fixation can be excluded, as the magnitude of IOR modulation by within-object structure is equivalent for the both right and left (relative to fixation) objects in the +45 and the –45 orientation.

4.2.5. Summary

Results from Experiment 3 show that (a) object-based IOR can be modulated by surface-based representations of 3D images of objects, extending the pattern of results for 2D forms in Experiment 2, (b) object-based IOR was *larger* for targets appearing on a different surface (of the same object) from the cue, and (c) object-based IOR was modulated by object internal structure in the +45 but not in the -45 display orientation.

At present it is unclear what may account for the effect of display orientation. One possibility that was considered is that the difference in IOR modulation for the two orientations was instigated by hemifield differences in object processing and may reflect the formation of a single representation making explicit both local and global structure information. However, this issue requires further research and is beyond the scope of the present investigation.

Finally, the effect of display orientation, or else, the presence versus absence of local/global structure effects, is consistent with the local structure hypothesis, as outlined in Chapter Two. On the other hand, *both* (observed) effects cannot be accounted for by the global structure hypothesis. Finally, it is important to note that although the results are consistent with the local structure hypothesis, this hypothesis cannot be used to account for the display orientation effect.

3.4. Experiment 4

4.3.1 Introduction

Results from Experiments 2 and 3 suggest that object-based attention has access to representations that make explicit internal object properties, and that these properties can modulate the IOR. Furthermore, it was found that this is the case for both two- and three-dimensional forms. The results from Experiment 3 also showed that this modulation of IOR by internal object structure can interact with the stimulus display orientation.

Thus, one issue arising from the results of Experiment 3 is that significant within-object IOR modulation was only observed in the +45 orientation, whilst no such modulation was observed when the objects were presented in the –45 display orientation. The effect of object orientation on object-based IOR modulation raises the question of whether the within-object IOR differences observed are *solely* due to the object or whether some other factors relating to stimulus display properties, such as the cue-target configuration, may contribute to these differences.

Whilst the notion of a two-component IOR is well researched and established by a number of different laboratories (e.g. Abrams & Dobkin, 1994; Gibson & Egeth, 1994; Jordan & Tipper, 1998; 1999; Tipper et al, 1991; 1994; 1997) the generality (e.g. Muller & Muhlenen, 1996) or the independence of the object-based IOR from the location-based IOR (e.g. McAuliffe, Pratt & O' Donnell, 2001) has been questioned. In other words, it has been argued that object-based IOR effect in fact mediated by location-based frames of reference. For example, in a recent study McAuliffe et al.

(2001) examined the location- and object-based components of IOR[27]. They manipulated the displays so that the cue and the subsequent target would appear either on objects (outline squares) or on empty locations. They predicted that if the object-based component of IOR is indeed separate from the location-based component then there should be a difference between trials where there was an object on the screen and trials where there was no object. If, however, the two components were demonstrations of a single inhibitory mechanism, then they predicted no difference between in the amount of inhibition for objects and empty locations.

McAuliffe et al. (2001) found that when the two types of trials were randomly presented, then IOR for targets appearing on objects was larger than for targets appearing on empty locations. However, when the two types of trials were blocked (i.e. placeholder-present and placeholder-absent trials were different displays) the magnitude of IOR for cued objects was *not* significantly larger from the IOR effect for cued locations. They proposed that these results question the idea of separate additive components of IOR, one for locations and one for objects. Instead, they proposed that much of the observed object-based IOR effects can be explained by a single inhibitory mechanism that inhibits objects or locations, depending on the context.

The results from Experiment 2 in this thesis showed that not only attention selects from higher-order representations of objects but also from representations of different component parts of the object. Therefore, it is important to establish that the mechanism by which this is accomplished is dedicated to objects, and is not simply an artefact of the spread of attention across the (empty) visual field.

---

[27] I thank Steve Tipper for bringing this study to my attention.

The general issue addressed in Experiment 4 is whether the IOR effects, reported in Experiments 2 and 3, are due to the object *per se*. It is important to determine:

(2) Whether IOR is generally greater for targets appearing on objects than for targets appearing on empty locations.

(3) Whether a similar modulation of IOR, as that observed in Experiment 2 (for segmented objects) would *also* be observed in displays with identical cue-target configurations but no object stimulus.

In order to address these questions, a within-subjects design was employed, where participants performed a simple, target detection task under two conditions. In one condition the cue and the target were presented at locations of an L-shape object; this was termed the *Object Present* condition. In the other condition the cue and the target were presented on an otherwise empty screen around the fixation cross; this was termed the *Object Absent* condition. In both Object conditions the location of the cues and targets were identical. Furthermore, the two types of trials (Object present and Object absent) were blocked. In such a blocked design, McAuliffe et al (2001) would predict that the magnitude of IOR in the two conditions (Object Present and Object Absent) would be equivalent.

4.3.1.1 Predictions

1. The predictions with respect to the first issue outlined above are as follows:

If object-based and location-based IOR reflect a *single* mechanism that inhibits objects and locations in the same way, then the magnitude of IOR should be equivalent for both Object Present and Object Absent trials. In contrast, if IOR operates over *two* separate frames of reference, one associated with locations and another associated with objects, then the upper limit of IOR in the Object Present displays will be significantly higher that the upper limit of IOR in the Object Absent displays.

2. The predictions with regard to the second issue addressed in this experiment are as follows:

   If the difference in IOR modulation between the two critical within-object cueing conditions is due to the object, then IOR modulation should be observed only in the Object Present, but not in the Object Absent condition. If, however, this modulation is due to factors *independent* of the presence of objects, then there should be a significant difference between the two within-object cueing conditions both in the Object Present and Object Absent conditions.

4.3.2 Method

Participants

Thirty-four psychology undergraduates, aged between 18 and 33 years from the University of Wales, Bangor participated in this experiment for two course credits. They had not participated in any of the previous experiments. They all reported normal or corrected-to-normal vision.

Apparatus and Stimuli

The same technical apparatus as in Experiment 2 was used. Stimuli for the *Object Present* condition were the L-shaped objects used in Experiment 2. These were presented in two possible orientations, tilted either +45 or -45 from the vertical meridian. When the L-shapes were tilted +45 degrees, the short rectangles were positioned above and below fixation, whilst when they were tilted –45 degrees, the short rectangles were positioned to the left and right of fixation (Figure 30). The two rectangles comprising the L-shapes were separated by an internal discontinuity in the 'segmented' (Figure 30, Panels A and C) but not in the 'unsegmented' (Figure 30, Panels B and D) display condition. At a viewing distance of 50 cm, the longer rectangle of each L-shape subtended 7.2° x 1.8° and the smaller rectangle 2.8° x 2.2° of visual angle. The fixation cross was black and measured 0.8° x 0.8°. The cue was a white outline square subtending 0.6° x 0.6° (contours measuring 0.2° x 0.2°) and the target was a filled white square subtending 0.8° x 0.8° of visual angle. The whole display from end to end was 13.2° high and 10.8° wide. The cue-target distance in each cueing condition was 4.5°, irrespective of whether the targets appeared within the same or within a different object from the cue.

In the *Object Absent* condition participants viewed sequences of the fixation cross, the cue and the target appearing on an otherwise empty screen. The position of the cues and targets were presented in exactly the same positions as in the *Object Present.*

Design

The experiment was based on a within-subject design with factors the Object, Cueing, Display orientation, Segmentation and SOA. Object had two levels: Object Present and Object Absent. Display orientation had two levels: +45 and -45; Segmentation had two levels: segmented and unsegmented; SOA had had two levels: 820 and 1220 msec; and Cueing had four levels.

In the *Object Present* condition targets appeared randomly, and with equal probability, in each of the following four cue-target configurations: The target appeared (a) within the same part and at the same location as the cue (cue-target (CT) location 1), (b) within the same part but at a different location from the cue (CT location 2), (c) within the same object but on a different part to the cue (CT location 3), (d) on the corresponding part of a different object, and at the same corresponding location, as the cue (CT location 4). Figure 31 illustrates the four cueing conditions used in the analysis.

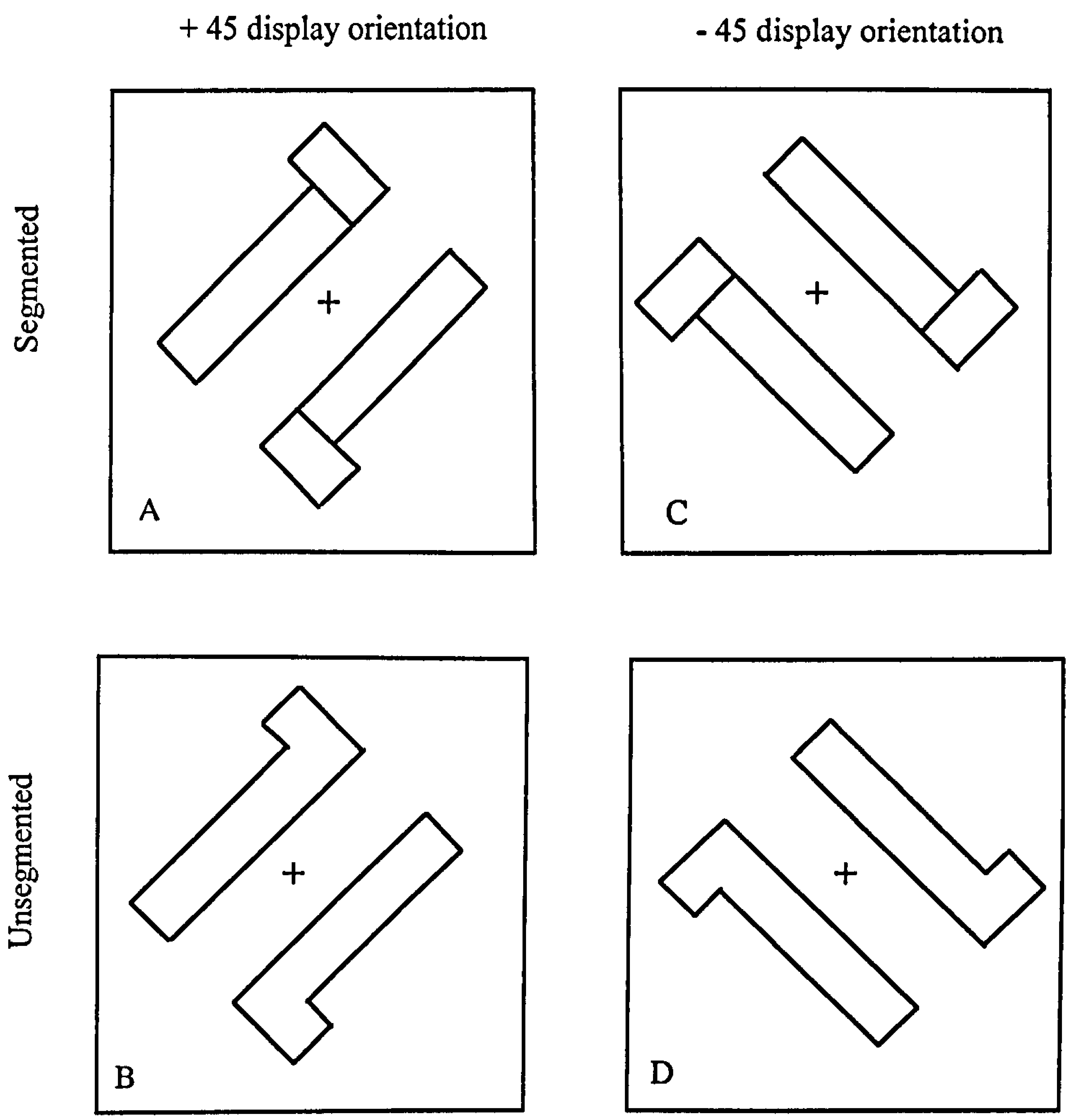


Figure 30. Examples of the L-shaped stimuli used in the Object Present condition in Experiment 4.

There were also *filler* trials. In those the target appeared (1) on the corresponding part of a different object, and at a different location (Filler 1), and (2) in a different object, and on a different part (Filler 2) relative to the cue. These trials were not used in the analysis as the cue-target distance was larger than the distance in the two within-object conditions (c) and (d).

All the above cue-target (CT) configurations were identical in both types of Display (segmented and unsegmented), except that in the unsegmented displays there was no internal discontinuity in the L-shapes.

In the *Object Absent* condition cues and targets appeared exactly in the same physical locations as in the Object Present condition with the only difference that there was no object stimulus on the screen. Thus, the cueing conditions were identical for the Object Present (Figure 31, left column) and Object Absent conditions (Figure 31, right column). The task, identical for both Object conditions, was to respond by a simple key-press to the onset of a pre-specified target on the screen.

Participants completed 10 practice trials followed by 340 experimental trials in the Object Present condition and 160 trials in the Object Absent condition. Thirty percent of all trials (100 trials in the Object Present and 40 trials in the Object Absent condition) were 'no-target' trials. In the remaining, ‘target’ trials, there were 10 trials for each cueing condition, for each type of Display (+45 segmented, +45 unsegmented, -45 segmented and -45 unsegmented) and over (or collapsed across) the two SOAs for the Object Present condition. In the Object Absent condition participants viewed each cueing condition 10 times for each Orientation (+45 and –45) over (or collapsed across) the two SOAs.

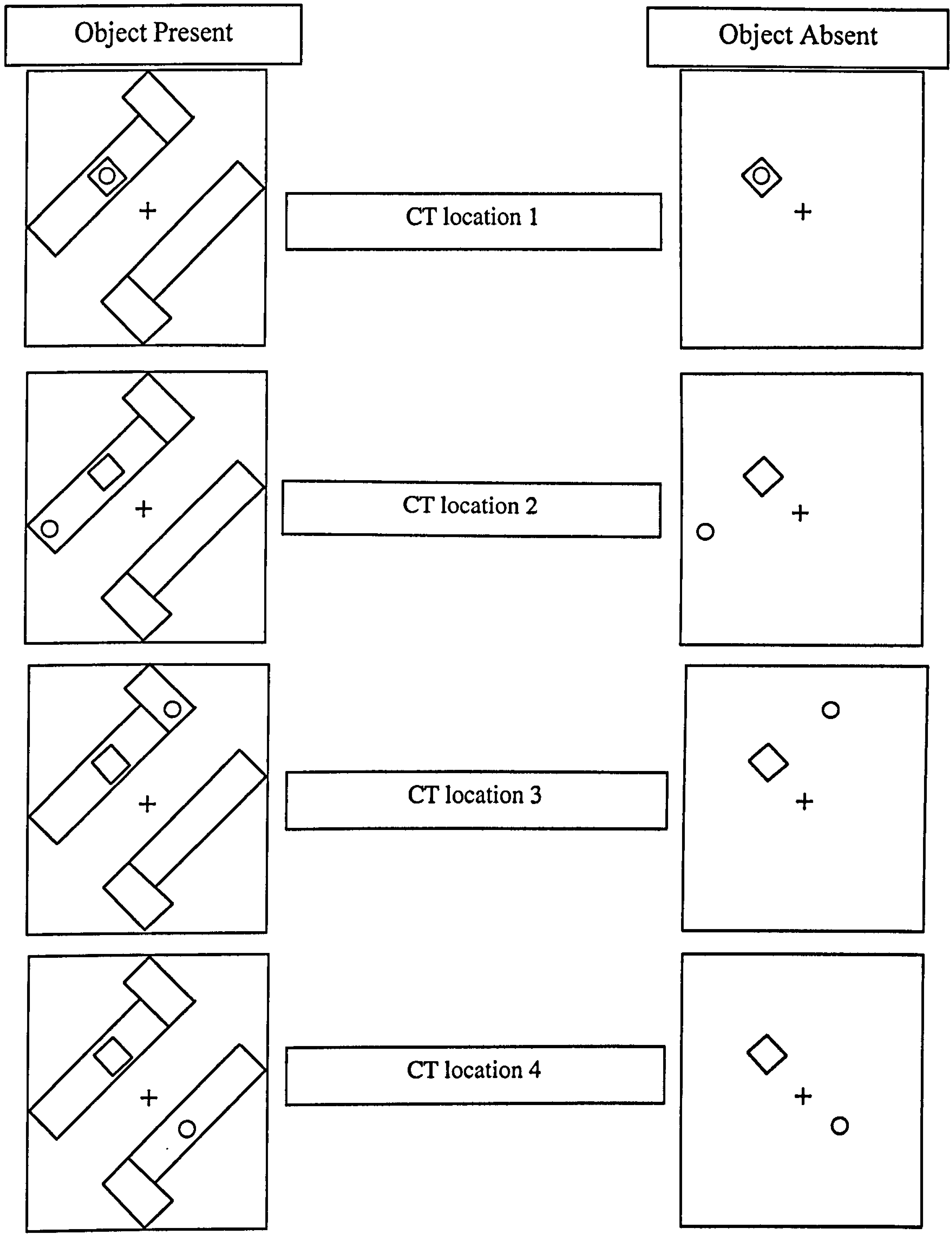


Figure 31. An illustration of the four cueing conditions (three within-object and one between-object) used in the analysis in Experiment 4. The display orientation depicted here is +45 degrees from vertical. Cues are depicted as squares and targets as circles. The cueing conditions were the same for the Object Present and Object Absent conditions.

Procedure

The procedure for the Object Present displays was identical to Experiment 2, except for the following modifications: The cue was presented at the centre of either object (on either side of the fixation cross). The target subsequently appeared at one of six possible locations on either L-shaped object, above, below, left, right, plus 45 degrees or minus 45 degrees from fixation. Three of these positions were at locations on the same object as the cue, and three were at locations on the other object, yielding a total of three same–object and three different-object conditions (as per Design).

For the Object Absent displays the sequence of events was identical except that there were no objects on the screen and the cues and targets were presented at the same locations as in the Object Present condition. The order of the Object Present and Object Absent conditions was counterbalanced between participants, so that half the participants viewed the Object Present condition first and the other half viewed the Object Absent condition first.

4.3.3 Results and Discussion

4.3.3.1. Object Present

Trials with RTs greater than 700 msec (slow) or less than 200 msec (anticipatory), as well as trials with responses to 'no-target' trials were discarded from the data. These exclusions made up only 1.4% of the data.

A repeated-measures ANOVA was carried out with Task Order as the between-subject factor and Cueing, Display orientation and Segmentation as the within-subject factors. Task Order had with two levels (1. Object Present was first and 2. Object Present was second). Cueing had four levels ('CT Location 1', 'CT Location

2', 'CT Location 3', and 'CT Location 4') and Display had two levels (+45 and –45); Segmentation had two levels (segmented and unsegmented). The effect of Task Order was not significant, $F(1, 33) <1$, ns. There was a significant main effect of Cueing [$F(3, 99) = 84.8, p < .001$]. The main effect of Display orientation was not significant, $F(1, 33) < 1$, ns, and neither was the main effect of Segmentation, F (1, 33) = .9, ns and main effect of SOA, $F(1, 33) < 1$, ns. There were no significant interactions between any of the factors. Due to the lack of significant main effect of Order of Object the two sets of data (1. Object Present was first and 2. Object Present was second) were collapsed and analysed as homogeneous.

Mean RTs for the four cueing conditions (three within-object and one between-objects) grouped by Object Present type (segmented vs. unsegmented) are shown in Table 4. The object-based IOR effect for each cueing condition over the two display types (segmented and unsegmented) is shown in Figure 32.

A 3 (CT Location 1-3) x 2 (Segmentation: segmented vs. unsegmented) repeated measures ANOVA was carried out on the mean IOR effects for the three same-object conditions (CT location 1-3). This showed a significant main effect of CT location [$F(2, 66) = 27.7, p < .0001$], and a significant main of Segmentation, $F(1, 33) = 4.97, p < .03$. The CT Location by Segmentation interaction was not significant, $F(1, 33) = 1.2$, ns.

Table 4: Mean RTs for the four cueing conditions for Object Present and Object Absent trials. RTs in the 'CT location 2' and 'CT location 3' cueing conditions are shaded to emphasise the difference between the two conditions.

Object		CT Location 1	CT Location 2	CT Location 3	CT Location 4
Present	Segmented	384	358	368	326
	Unsegmented	385	366	367	331
Absent		365	348	353	337

On the basis of Experiment 2 findings, showing that the magnitude of IOR can be modulated by the presence of internal structural features, planned comparisons were carried out between mean IOR effects at CT locations 2 and 3 for segmented versus unsegmented displays. In the *segmented* displays the IOR effect in the 'CT location 3' condition (42 msec) were significantly slower than in the 'CT location 2' condition [32 msec; t (33) = 2.1, p <0.05]. There was no significant difference between the two CT location conditions in the *unsegmented* displays, t (33) = 1.2, ns.This replicates previous findings of Experiment 2[28].

[28] An important issue raised by the results of Experiment 3 was the difference in modulation of object-based IOR as a function of the display orientation (+45 or -45). In the +45 segmented displays RTs in the 'CT location 3' condition were significantly slower than RTs in the 'CT location 2' condition [t (33) = 2.3, p <0.02]. There was no significant difference between the two cueing conditions in the +45 *unsegmented* displays, t (33) = .9, ns. However, the difference between the 'CT location 2' and 'CT location 3' cueing conditions was not significant in the –45 segmented displays [t (33) = .7, ns] or in the –45 unsegmented displays [t (33) < .5, ns]. This finding is consistent with previous findings in the same orientation in Experiment 3.

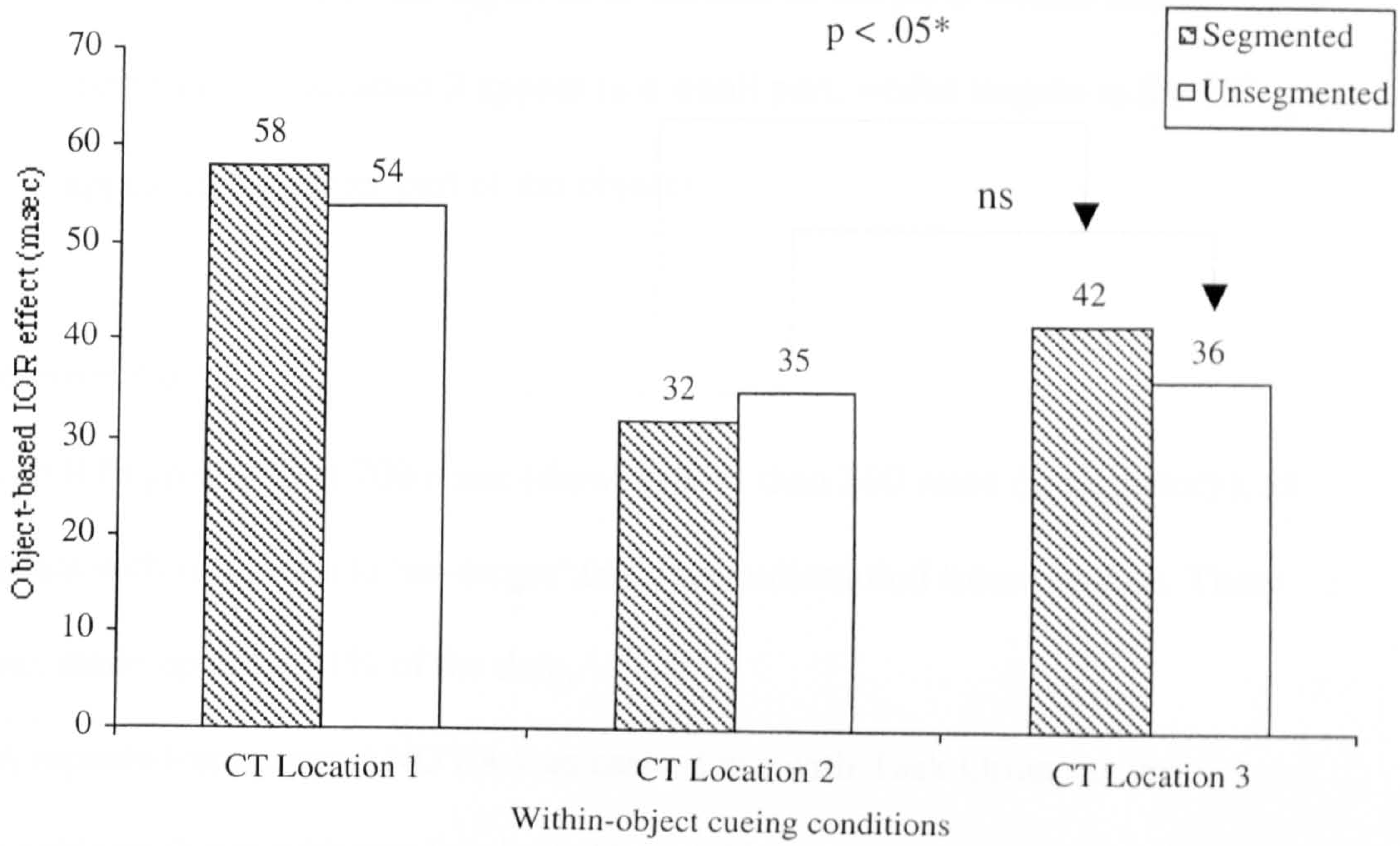


Figure 32: *Object Present*: Object-based IOR in the three within-object cueing conditions in the two display orientations. Results in the +45 display orientation replicate the pattern of results of Experiment 2.

Post-hoc comparisons between the filler trials in the segmented displays were also carried out (for the same reasons as those outlined in Experiments 2 and 3). These revealed no significant difference between the between-object baseline condition (CT location 4; $M = 326$, $SD = 51.2$) and the Filler 1 ($M = 321$, $SD = 50.6$) and Filler 2 ($M = 331$, $SD = 48.7$) conditions, $t\ (33) < 1$, ns and $t\ (33) = 1.09$, ns respectively. The difference between the two Filler conditions was also not significant, $t\ (33) < 1$, ns.

The results from the Object Present condition replicate the pattern of results from Experiments 2 and 3, in that the object-based IOR was significantly larger when cues and target appeared on *different surfaces* defined by an internal edge boundary, compared to when they appeared on the *same surface*. Analysis of the filler trials confirm that the results are not simply due to differences in attending to the ends of the object relative to the centre of the object or to the size of the parts within which targets appear (i.e. targets in CT location 3 appear in a small part, whilst targets in the CT location 2 appear in the larger part of the object).

4.3.3.2 Object Absent

Trials with RTs greater than 700 msec (slow) or less than 200 msec (anticipatory), as well as trials with responses to 'no-target' trials were discarded from the data. These exclusions made up only 1.1% of the data.

A repeated-measures ANOVA was carried out with Task Order as the between-subjects factor with two levels (1. Object Absent was second; 2. Object Absent was first), and Cueing (CT location 1-4), Display orientation (+45 and –45) and SOA (820 and 1220) as the within-subjects factors. There was a significant main effect of Cueing [F (3, 99) = 14.0, $p < 0.001$]. The effect of Task Order, Display orientation, and SOA were not significant [F (1, 33) = 3.2, $p = 0.08$; F (1, 33) = 1.2, $p = 0.2$; F (1, 33) < 1, ns respectively]. None of the interactions between the three factors were significant. Therefore the two sets of data (1. Object Absent was second; 2. Object Absent was first) were collapsed and were analysed as homogeneous.

Mean RTs for the four cueing conditions in each display orientation are shown in Table 4. The IOR effects (CT location1-3 minus CT location 4) for each CT location are shown in Figure 33.

A 3 (CT location 1-3) x 2 (+45 and –45) repeated-measures ANOVA on mean IOR effects revealed no significant main effect of CT location, F (2, 66) < 1, ns and no significant main effect of Display Orientation, $F(1, 33) < 1$, ns. Planned comparisons were also carried out to investigate IOR between the CT locations that were significantly different in the Object Segmented displays. There was *no* significant difference in the amount of IOR between CT locations 2 and 3 [$t(33) < -1.0$, ns]. Therefore the difference in IOR between the same locations in the segmented object displays was not related to the spatial locations of the cue-target pairs independently of the objects in which those locations are probed. On the basis of this finding, one can rule out the possibility that the modulation of IOR by the internal boundary in the segmented object displays was simply due to the differential distribiution of IOR across the visual field.

Planned comparisons were carried out to investigate the main effect of Cueing. There was a significant difference (28 msec) between the 'CT location 4' condition and the 'CT location 1' condition [$t(33) = -5.1$, $p < 0.000$], replicating the location-based IOR effect. There was also a significant difference between the 'CT location 4' and the 'CT location 2' conditions [$t(33) = -2.6$, $p = .01$] and the 'CT location 4' and 'CT location 3' conditions, $t(33) = -2.1$, $p = .04$. Finally, there was *no* significant difference between the 'CT location 2' and 'CT location 3' cueing conditions, $t(33) < -1.0$, ns[29].

---

[29] In the +45 display orientation there was significant difference between the 'CT location 4' condition and the 'CT location 1' condition [$t(33) = -4.9$, $p<0.000$], replicating the location-based IOR effect. The difference between the 'CT location 4' and the 'CT location 2' 'CT location 3' conditions, was also significant, $t(33) = -2.1$, p = .04 and $t(33) = 2.2$, $p = .05$. There was no significant difference between the 'CT location 2' and 'CT location 3' cueing conditions, $t(33) = 1.64$, ns.

In the –45 display orientation there was a significant difference between the 'CT location 4' condition and the 'CT location 1' condition [$t(3) = -2.7$, $p < 0.01$], again, replicating the location-based IOR effect. There was no significant difference between the 'CT location 4' and the 'CT location 2' conditions [$t(33) = -1.5$, ns] and the 'CT location 4' and 'CT location 3' conditions, $t(33) = -1.4$, ns. Finally, there was no significant difference between the 'CT location 2' and 'CT location 3' cueing conditions, $t(33) < 1$, ns.

The main findings from the Object Absent condition can be summarised as follows. First, the original location-based IOR effect was replicated, with RTs to targets appearing at the same location as the cue (CT location 1) slower than RTs to targets appearing at a different location from the cue and at the opposite side of fixation (CT location 4).

Second, there was no significant RT difference between the 'CT location 2' and 'CT location 3' cueing conditions, indicating that the significant difference between these conditions in the Object Present condition, *was* due to the objects' internal structural discontinuity, and not to the cue-target spatial relationship.

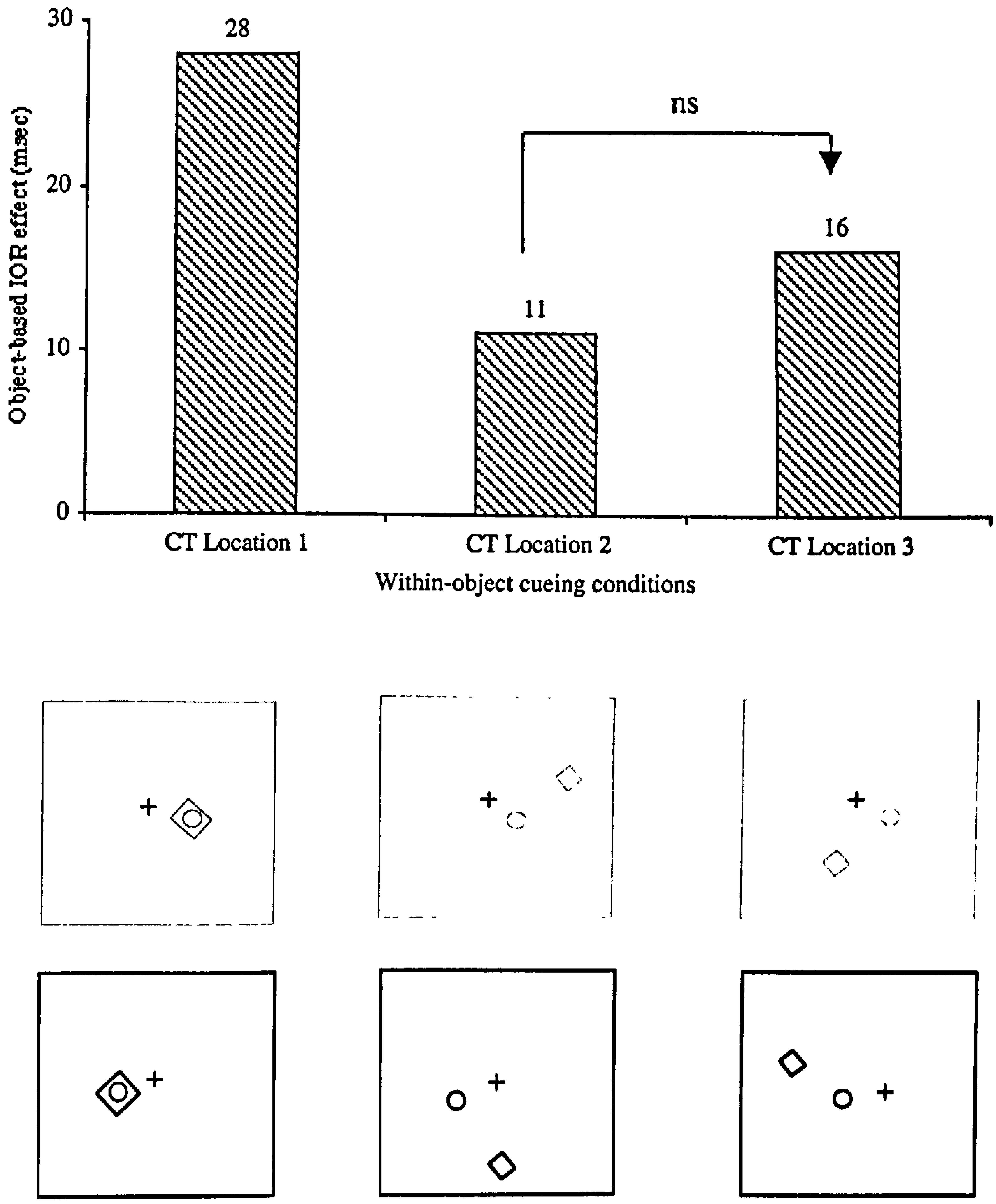


Figure 33: *Object Absent*: The IOR effect in the three within-object cueing conditions in the two display orientations. Here *circles* depict the cues and *squares* the targets.

4.3.3.3. Comparison between Object Present and Object Absent displays

One of the predictions in this experiment was that if IOR is indeed modulated by the objects' internal structural properties rather than by purely spatial factors, then there should be a significant difference between the Object Present and Object Absent conditions in IOR modulation. Thus, the IOR effect for each CT location condition was compared for the Object Present and Object Absent conditions.

For the comparison between the Object Present and Object Absent conditions data from Object segmented and unsegmented displays were used separately. This was necessary as it was predicted that the two types of segmentation procude different patterns of IOR modulation. Figure 34 shows the mean IOR effect in each Cueing condition for the Object Present and Object Absent conditions.

A 2 (Object Present Segmented and Object Absent) x 4 (CT Location 1-4) x 2 (Display orientation +45 and –45) x 2 (SOA of 820 and 1220 msec) repeated measures ANOVA revealed a significant main effect of Object, $F\,(1, 33) = 4.8, p = .03$ with larger IOR for Object Present (42 msec) than Object Absent (18 msec) displays. The main effect of CT Location was also significant, $F\,(3, 99) = 50.0$, $\underline{p} = .0001$. The main effects of SOA and Display Orientation were not significant [$F\,(1, 33) < 1$, ns and $F\,(1, 33) = 1, 33) < 1$, ns, respectively]. The only significant interaction was between CT Location by Object, $F\,(3, 99) = 10.9, p = .0001$.

A 2 (Object Present Unsegmented and Object Absent) x 4 (CT Location 1-4) x 2 (Display orientation +45 and –45) x 2 (SOA of 820 and 1220 msec) repeated measures ANOVA. There was a significant main effect of Object, $F\,(1, 33) = 7.1, p =$ .008 with larger IOR for Object Present unsegmented displays (38 msec) than Object Absent dipslays (18 meac). Also significant was the main effect of CT Location, *F* (3,

99) = 45.0, p = .0001. The main effects of SOA and Display Orientation were not significant [F (1, 33) < 1, ns and F (1, 33) = 1, 33) < 1, ns, respectively]. The only significant interaction was between CT Location by Object, F (3, 99) = 7.0, p = .0001.

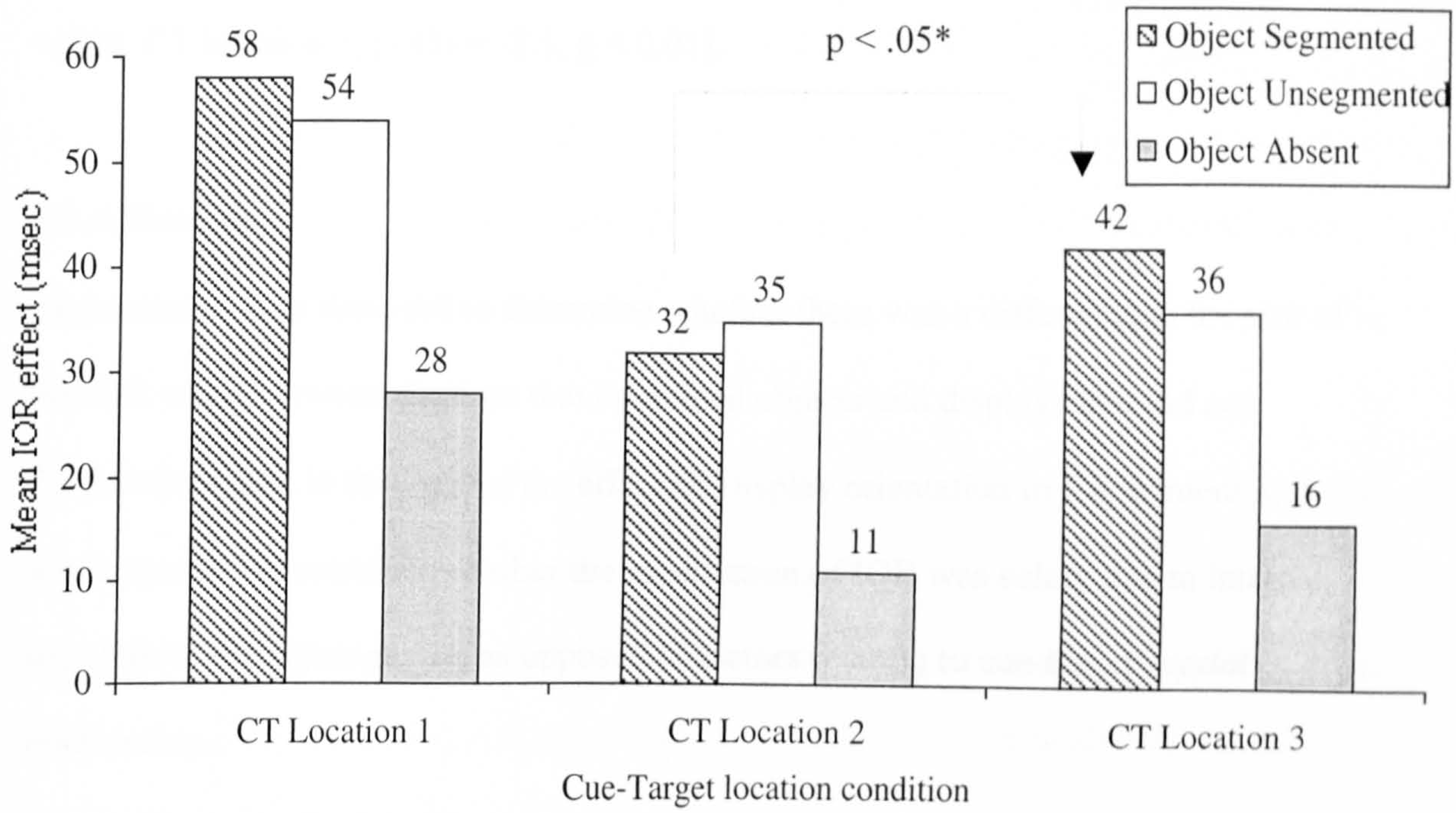


Figure 34: Comparing the mean IOR effect for the Object Present and Object Absent conditions, in each cueing (cue-target location) condition.

Planned comparisons[30] were carried out to investigate the Object by Cueing interaction. These showed a significant difference in the IOR effect (RT cued minus RT uncued) between the Object Present Segmented and Object Absent conditions in

17. The comparison between the two Object conditions (Present and Absent) was between the *IOR magnitude* in each within-object cueing condition (RT cued minus RT uncued). The reason for not comparing RT in each within-object cueing condition was that the two parts of this Experiment had a different baseline. Typically, for conditions that share the same baseline (uncued object) planned *RT* comparisons between cueing conditions are carried out.

the 'CT location 1' [t (33) = 2.5, $p < 0.01$], in the 'CT location 2' condition [t (33) = 3.4, $p < 0.002$] and in the 'CT location 3' condition, t (33) = -3.2, $p < 0.02$. The same pattern of results was found contrasting Object Present Unsegmented and Object Absent displays [CT location 1, t (33) = -3.0, $p < 0.004$; CT location 2, t (33) = -2.8, p =.008; CT location 3, t (33) = -2.5, $p < 0.01$].

4.3.4 Discussion

Experiment 4 was designed to determine whether there was a difference in the size of the IOR effect between displays that contained objects and displays that did not. Furthermore, and in the light of the effect of Display orientation in Experiment 3, it was important to establish whether the modulation of IOR was solely due to internal structure of the *object* per se, as opposed to factors relating to cue-target *spatial* relationship.

The main findings from Experiment 4 can be outlined as follows. First, IOR was significantly larger in the Object Present than in the Object Absent displays. This finding indicates that object-based inhibition is separable and independent from location-based IOR, contrary to the proposal put forward by McAuliffe et al (2001). Furthermore, the finding of an IOR component that is *dedicated* to objects, indicates that the previous findings of IOR modulation (Experiments 2 and 3) were not simply artefacts of the spatial distribution of IOR.

Second, in the Object Present condition object-based IOR was modulated by the objects' internal structure in the segmented displays, replicating the findings from Experiment 2. In contrast, in the Object Absent condition there was no significant difference between targets appearing on the same side of fixation (in locations that corresponded to the Object Present CT Location conditions).

Third, there was a significant interaction between the Object Present and the Object Absent conditions and the Cue-target location. This interaction is a clear indication that modulation of IOR was a function of the object and its internal structural properties. Absence of the objects resulted in significantly less IOR not only for targets appearing within the same side of fixation as the cue ('CT location 2' and 'CT location 3') but also for targets that appeared at the same location as the cue ('CT location 1').

The present findings are important for at least two reasons. First, they replicate the pattern of results in previous studies that used static displays to examine location- and object-based IOR effect (e.g. Jordan & Tipper, 1998). More specifically, they further endorse the notion that object-based and location-based IOR components are observed independently and can combine to produce an overall larger IOR effect (e.g. Tipper et al., 1991; Jordan & Tipper, 1998). Second, they indicate that the obtained modulation of IOR was a genuine object-based effect rather than an artefact of the cue-target spatial relationship.

In conclusion, the findings from Experiment 4 demonstrate that object-based IOR is indeed a distinct attentional component, separable from location-based IOR, and allow us to rule out the possibility that factors other than the objects themselves and their internal structural properties (i.e. discontinuities) modulate the IOR effect.

4.4. Experiment 5

Modulation of object-based IOR by volumetric component representations

4.4.1 Introduction

In Experiment 4 it was established that object-based IOR modulation is related to object internal structure and stimulus orientation. Experiment 5 further examines the *nature* of shape representations that modulate object-based IOR in 3D forms.

The primary aim of this thesis is to investigate whether internal object properties modulate object-based IOR, and to determine the relevant shape primitives that mediate this attentional modulation. As noted earlier, theories of shape representation posit different types of primitives used in computing a 3D representation of object shape representation. Some theorists (e.g. Biederman, 1987; Marr, 1982) posit a role for higher-order groupings or *volumetric components* in the representation of 3D shape for recognition and action purposes; whilst others (e.g. Pentland, 1989; He & Nakayama, 1992; Nakayama, He & Shimojo, 1995) posit that a surface-based representation of an image is a necessary initial stage for recognition of 3D shape.

In Experiment 3, surface properties of objects, such as internal surface boundaries, modulated object-based IOR across an image of a 3D object. The present experiment is the first attempt to investigate whether volumetric components can modulate object-based IOR *in addition* to surfaces. (Hereafter the abbreviation VC will be used in the place of the term 'volumetric component').

Findings from the previous experiments have shown that object-based selection may operate over representations that make explicit surface discontinuities.

The question behind Experiment 5 is whether IOR can be modulated by VCs *in addition* to any modulation of object-based IOR by surfaces; and whether these two types of representations are significantly different in the way they modulate IOR. It is predicted that if 3D objects are represented by VCs, then object-based IOR would be modulated by the boundary separating the two VCs. Therefore, if VCs are used to represent object shape, it is assumed that a representation that makes such boundaries explicit is available for inhibitory attentional selection.

4.4.2. Method

Participants

Sixteen psychology undergraduates aged between 19-36, from the School of Psychology at University of Wales, Bangor took part in this experiment for two course credits. None had participated in the previous experiments. They all reported normal or corrected-to-normal vision.

Apparatus and Stimuli

The experiment was presented on a 17" SONY monitor connected to a Power Macintosh G3 computer. Stimulus randomisation and presentation as well as recording of responses and reaction times were controlled through PsyScope (version 1.2.4.). Responses were made through the letter key 'b' on a standard Apple keyboard connected to the computer.

The displays were images of two objects, each composed of two VCs, presented in dark grey colour on either side of a fixation cross (Figure 35) against light grey background. At a viewing distance of 55 centimetres the larger VC was a brick shape measuring 6.2° long, 2.6° tall and 3.1° wide. The smaller VC was a forward leaning cube that was 1.5° long, 3.1° tall and 1.5° wide. The two objects were presented in two possible orientations (Figure 35). At each orientation there were three visible surfaces of each VC. The fixation cross was a red 'plus' (+) sign and subtended 0.7° x 0.7°. The cue was a small white-filled parallelogram with black outline, measuring 0.4° x 0.3° (contour was 0.2° x 0.2°). The target was a dark grey parallelogram subtending 0.4° x 0.3° of visual angle. The distance between the cue and the targets was 4.6° of visual angle. The whole display from end to end subtended 18° x 10.3° of visual angle.

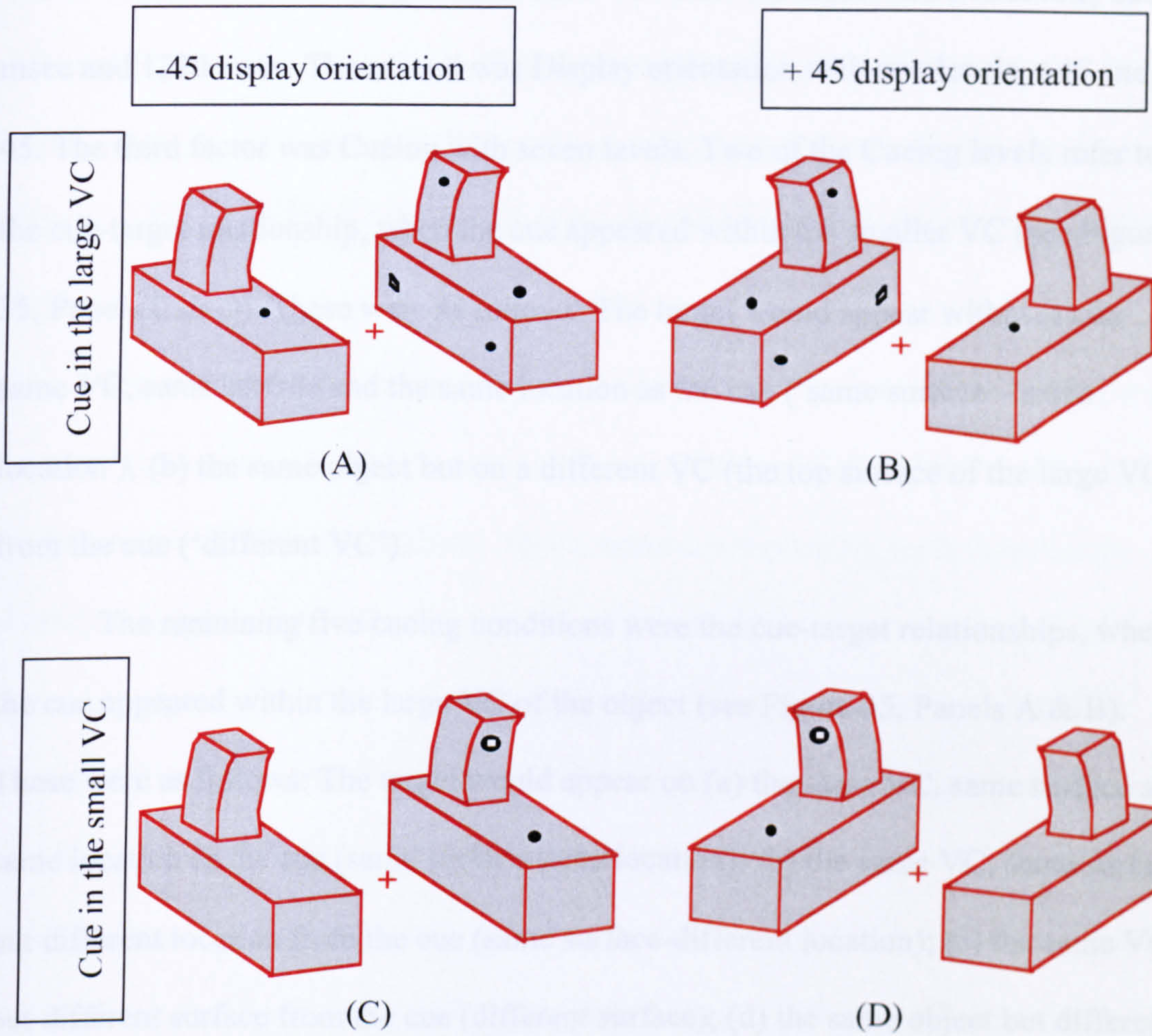


Figure 35. Displays of the objects used in Experiment 5. Cues are depicted as white-filled boxes and targets as black circles. Panels A and B depict the four within-object and one between-objects cueing conditions used in the statistical analysis. In the same surface-same location condition cue and target overlap. Panels C and D depict two of the cueing conditions (same surface-same location and different VC when the cue was in the smaller VC. In the different VC condition the cue-target distance was the same as that in the same condition when the cue was in the large VC. Only the means of these conditions are reported (see text for details).

Design

There were three factors in this experiment. The first was SOA with two levels, 820 msec and 1220 msec. The second was Display orientation with two levels, +45 and -45. The third factor was Cueing with seven levels. Two of the Cueing levels refer to the cue-target relationship, when the cue appeared within the smaller VC (see Figure 35, Panels C & D). These were as follows: The target would appear within (a) the same VC, same surface and the same location as the cue ('same surface – same location'*)*, (b) the same object but on a different VC (the top surface of the large VC) from the cue ('different VC').

The remaining five cueing conditions were the cue-target relationships, when the cue appeared within the large VC of the object (see Figure 35, Panels A & B). These were as follows: The target would appear on (a) the same VC, same surface and same location as the cue (same surface-same location); (b) the same VC, same surface but different location from the cue (same surface-different location); (c) the same VC but different surface from the cue (different surface); (d) the same object but different VC from the cue (different VC); (e) different object from the cue but same VC and same location as the cue (different object-baseline).

There were also filler trials, where the cue-target distance was larger than the cue-target distance for the within-object cueing conditions (apart from the 'same surface-same location' condition) outlined above (greater or shorter than 4.5 degrees of visual angle). These trials were not used in any subsequent analyses.

The *filler* trials when the cue was in the small VC were as follows: the target would appear (1) the same VC but different surface from the cue, (2) on the same object but on a different VC (the side surface (near) of the large VC) from the cue, (3) on the same object and different VC (on the side surface (far) of the large VC) from the cue, (4) on a different object, same surface and same location as the cue, (5) on a different object same VC different surface and different location from the cue (6) on a different object and different VC from the cue (near side), (7) on a different object and different VC from the cue (far side), and finally (8) on a different object and different VC from the cue.

The *filler* trials when the cue appeared in the large VC were as follows: the targets would appear (1) on the same object but different VC from the cue (front side of the small VC); (2) on a different object, same corresponding surface and same corresponding location as the cue (3); on a different object, same corresponding VC, same corresponding surface and different location from the cue; (4) on a different object, different corresponding VC and different corresponding surface (side surface of the small VC) from the cue; (5) on a different object and different VC from the cue (front surface of the small VC).

Each participant completed a block of 550 trials, 150 of which were 'no-target' trials (approx. 30%). In the remaining 400 trials each cueing condition appeared 10 times in each of the two orientations (-45 and +45) over both SOAs. A computer generated message appeared every 100 trials advising participants to take a short break.

Procedure

Each trial started with the presentation of the fixation cross at the centre of the screen. After 1000 msec from fixation onset the two objects were presented on either side of the cross. Following a 1000 msec delay, the cue would appear on either the large or the small VC of either object. It would appear either on the side surface of the large VC (the far end of it) or on the front surface of the small VC. The cue remained visible for 90 msec and was then extinguished. Following a delay of 100 or 300 msec the red fixation cross turned to white for 130 msec and then reverted to red again. Finally, after a further delay of 100 or 300 msec, the target appeared in ten possible locations on either object in either of the two orientations. The target remained on the object until the participant had pressed the key 'b' or until 1000 msec had elapsed. Participants were reminded just before the experiment commenced to respond as quickly as possible and to withhold their response in trials where no target was presented.

4.4.2.1. Data analysis

First, only trials where the cue appeared in the large VC were analysed. These were trials in the following cue-target configurations: (1) 'same surface – same location', (2) 'same surface-different location', (3) 'different surface', (4) 'different VC', (5) 'different object'. Trials where the cue appeared within the small VC were not analysed as the distance between cues and targets in the within-object conditions was not equal. However, the means in the 'same surface-same location' and 'different VC' conditions are reported in Table 6. Filler trials were not analysed.

Second, data from only one of the objects in each orientation were used. This meant that in the -45 display orientation only RTs to targets appearing on the right object were used, because it was only that object for which the 'different object' cueing condition (baseline) was equidistant to the within-object conditions. The same rule applied in the +45 display orientation for the left object. This ensured that the 'different object' condition was equidistant with all the within-object object conditions.

4.4.3. Results and Discussion

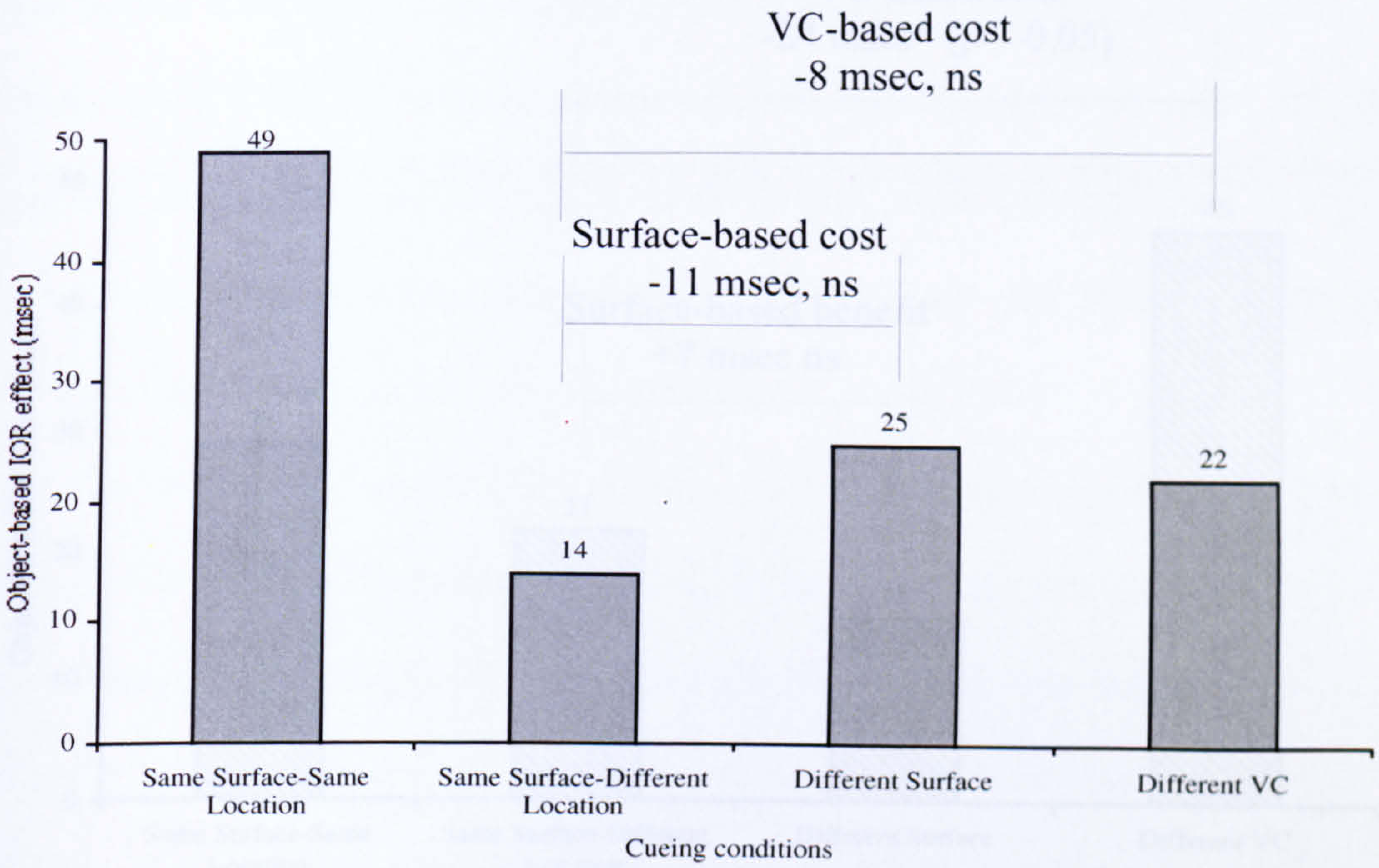
All trials with RTs that were less than 100 msec (anticipatory), greater than 700 msec (slow), or incorrect ('no-target' trials), were discarded from the data (as in the previous experiments). This accounted for 2.9% of all trials. Table 5 shows the mean RTs for each of the seven (five for large VC and two for the small VC) cueing conditions used in the analysis in each display orientation.

Figures 36 and 37 show the object-based IOR effect (RTs in each within-object condition *minus* RTs in the 'different object' condition) for five of the object-cued conditions (which represent trials where the cue appeared in the large VC), separately for each display orientation and collapsed across SOA.

The data were analysed using a 2 (Display orientation) x 2 (SOA) x 5 (Cueing conditions) repeated-measures ANOVA. Results showed a significant main effect of Cueing, $F(4, 60) = 5.5, p < .001$. There was no effect of Display orientation [$F(1, 15) < 1$] and no significant effect of SOA, $F(1, 15) < 1$. The interaction between Display orientation and Cueing was significant [$F(4, 60) = 2.3, p < .05$].

Table 5: Mean RTs for each cueing condition used in the analysis in each of the two display orientations. Mean RTs in the three within-object cueing conditions are in grey cells to emphasise their theoretical importance.

Cue location	Same Surface Same location	Same Surface Different location	Different Surface	Different VC	Different Object
		+ 45 display orientation			
Large VC	375	340	351	348	326
Small VC	371	-	-	348	326
		- 45 display orientation			
Large VC	353	346	339	370	324
Small VC	358	-	-	351	324



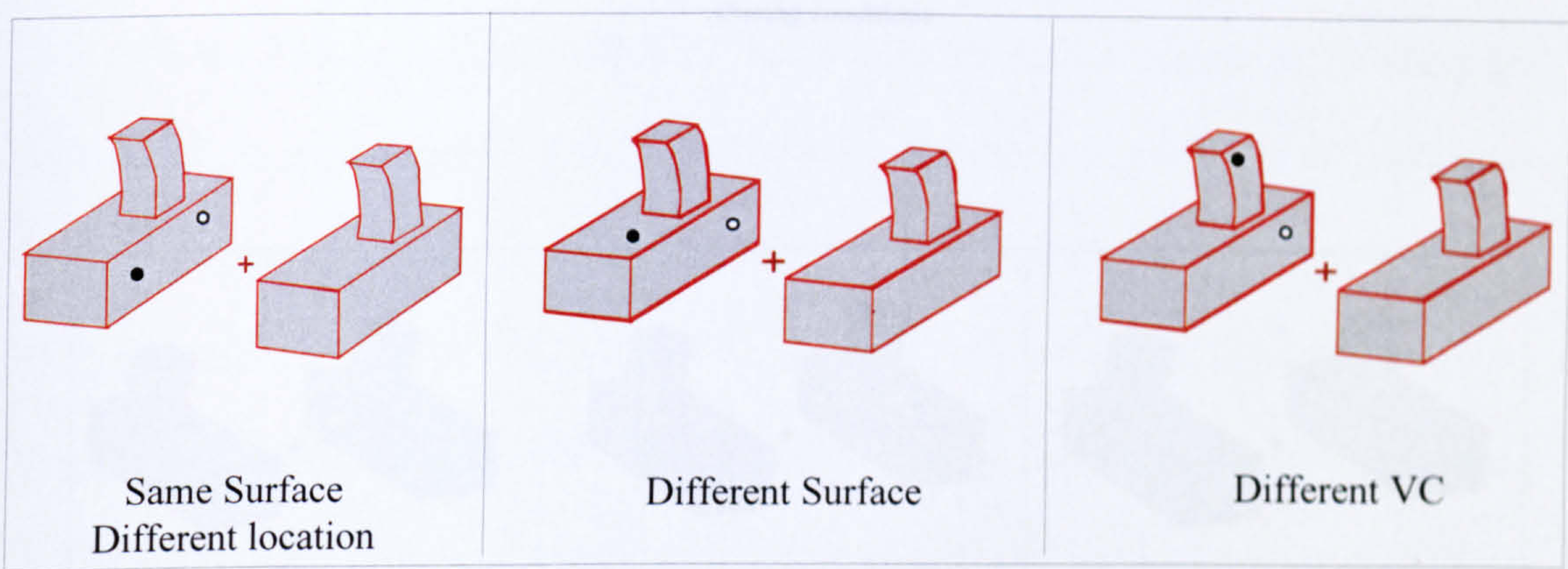
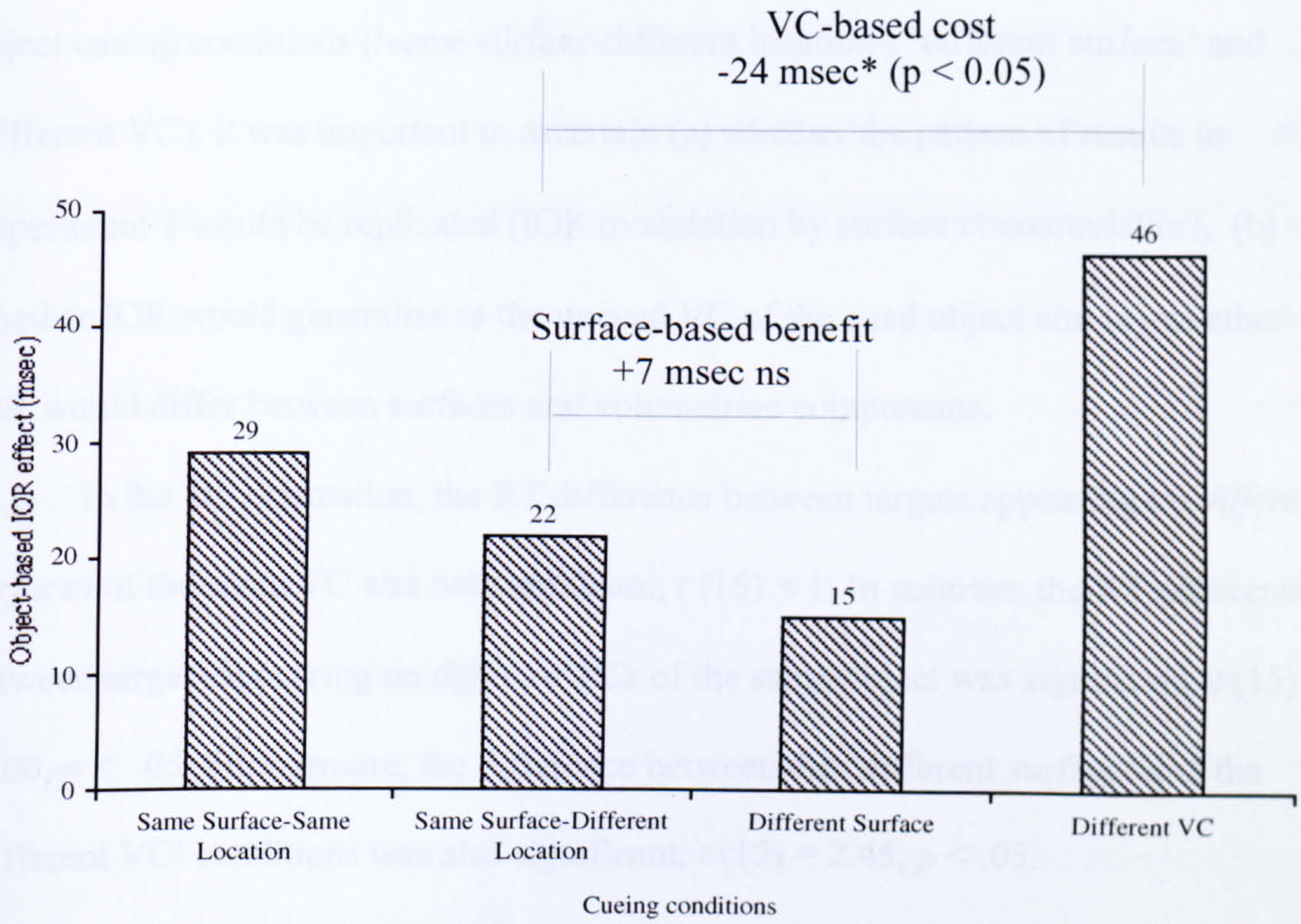


Figure 36. Mean object-based IOR effect for each of the five cueing conditions (where the cue appeared in the large component) the dual-component objects in the *+45* display orientation.



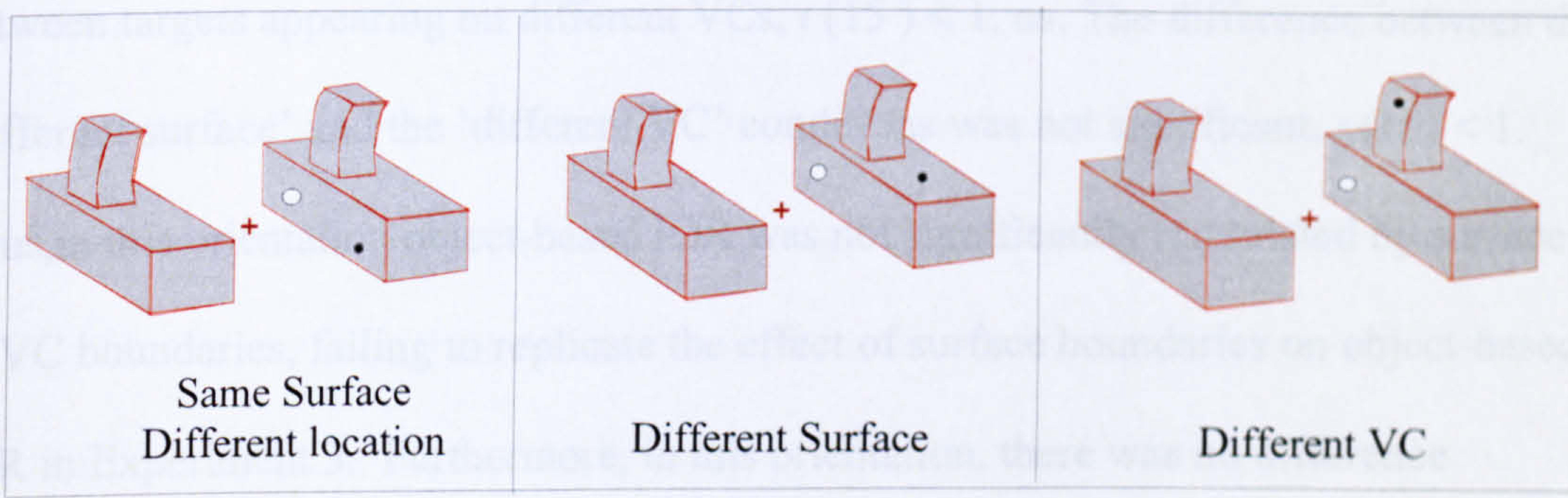


Figure 37. Mean object-based IOR effect for each of the five cueing conditions (where the cue appeared in the larger component) the dual-component objects in the *-45* display orientation.

Planned comparisons were carried out between each within-object cueing condition and the 'different object' condition, as well as between the critical within-object cueing conditions ('same surface-different location', 'different surface' and 'different VC). It was important to ascertain (a) whether the pattern of results in Experiment 3 would be replicated (IOR modulation by surface characteristics), (b) whether IOR would generalise to the uncued VC of the cued object and (c) whether IOR would differ between surfaces and volumetriac components.

In the -45 orientation, the RT difference between targets appearing on *different surfaces* of the same VC was not significant, $t\,(15) < 1$. In contrast, the RT difference between targets appearing on *different VCs* of the same object was significant, $t\,(15) = -2.00$, $p < .05$. Furthermore, the difference between the 'different surface' and the 'different VC' conditions was also significant, $t\,(15) = 2.45$, $p < .05$.

In the +45 display orientation the RT difference between targets appearing on different surfaces was not significant [$t\,(15) < 1$, ns] and neither was the RT difference between targets appearing on different VCs, $t\,(15) < 1$, ns. The difference between the 'different surface' and the 'different VC' conditions was not significant, $t\,(15) < 1$. Thus,in this orientation object-based IOR was not significantly modulated by surface or VC boundaries, failing to replicate the effect of surface boundaries on object-based IOR in Experiment 3. Furthermore, in this orientation, there was no difference between VCs and surfaces in the way they modulated object-based IOR[31].

[31] Planned comparisons between the four between-object filler conditions were carried out for each orientation. In the +45 displays there were no significant differences between the 'different object-baseline' (M = 326) and the 'different object-same surface-different location' (M = 318) condition, t (15) <1; the 'different object-different surface'(M = 313) condition, t (9) = 1.7, and the 'different object-different VC' (M = 322) condition, t (9) <1. The differences between the three filler conditions were not significant.

In the -45 displays, there were no significant differences between the 'different object-

4.4.4. Discussion

The aim of this experiment was to ascertain whether (a) the pattern of results in Experiment 3 (IOR modulation by surface characteristics) would generalise to two-component objects, (b) IOR would generalise to the uncued VC of the cued object and (c) the two types of shape primitives – surfaces and VCs - would differ in the way they modulate object-based IOR. It was predicted in the Introduction to this chapter that if object-based IOR operates over representations that make explicit information about surface characteristics, then IOR for targets that appear on a different surface from the cue would be different (larger) from IOR for targets that appear on the same surface as the cue. It was also hypothesised that if, in addition to the effect of surfaces, VC boundaries are *also* made explicit in the representations for attentional selection, then IOR would also be modulated by boundaries between VCs.

The main findings of Experiment 5 can be summarised as follows. First, any differences in IOR between different *surfaces* of the same VC were attenuated (but not absent). Instead, IOR was significantly modulated by the *VC boundaries* in the objects. However, this result was observed only in the -45 display orientation.

Second, the observed significant differences in IOR between the VCs of the object are always marked with an increase of object-based IOR (larger RTs) when attention is orientated across an internal structural boundary. This pattern of results

---

baseline' (M = 324) and the 'different object-same surface-different location' (M = 315) condition, t (9) =1.6, ns; the 'different object-different surface'(M = 329) condition, t (9) < 1, and the 'different object-different VC' (M = 331) condition, t (9) =1.4, ns. None of differences between the three filler conditions were significant.

replicates the pattern in the previous experiments for surface boundaries in simple 2-D forms and will be further discussed in the General Discussion section.

The third point for discussion concerns the significant interaction between Display orientation and Cueing. Comparing the difference between the 'same surface-different location' and 'different surface' conditions in Experiments 3 and 5 it is clear that the pattern of IOR modulation by surfaces, despite non-significant in Experiment 5, is similar.

What would cause the modulatory effect of surfaces to attenuate? The answer may lie in the presence of a second VC in the object. It is possible that in this experiment, the local descriptions accessible to IOR are those of the object's VCs, whose surfaces are now parts of these (perhaps more relevant or prominent) local descriptions. If the objects are perceived as consisting of two distinct VCs, then attention may be modulated by the boundaries separating these components. In this case, a representation of the object was computed based on an algorithm for parsing the object shape at points of sharp concave discontinuities (e.g. Hoffman & Richards, 1984). Surfaces within these components are not accessed individually by inhibitory attentional mechanisms. In other words, the relevant primitives (surfaces of VCs) may change as a function of object complexity (i.e. surfaces in single component objects and VCs in dual component objects).

The second explanation for the attenuation of the surface-based modulation of IOR is more methodological. The two VCs comprising the objects in this experiment are separated by *two* internal structural boundaries, whilst surfaces of each VC are separated by a *single* structural boundary. In this case, it is premature to draw any conclusions with regard to whether IOR is modulated by the presence – and the number - of internal structural boundaries per se (in which case the number of boundaries would produce even greater difference between targets appearing on either side of them) or alternatively, whether IOR is modulated by a more stable representation of volumetric shape primitives. This issue is addressed in Experiment 6.

Finally, the significant interaction between Cueing and Display orientation may hold an important clue for the observed results. As pointed out in the Design and Results sections, in the +45 orientation only data from the left object were analysed, whilst in the –45 orientation only data from the right side object were analysed. Therefore, it is possible that the difference in the two orientations reflects a difference in the way that the left and the right side object are represented, with the Left-side object being represented on the basis of its surfaces (despite not reaching significance IOR was modulated by surfaces in the +45 displays), whilst the Right -side object being represented by its components. This is only a viable possibility if one assumes (a) that the objects are perceived as, in a sense, being graspable by the small VC, and (b) that an action-based representation can be automatically evoked for an object (irrelevant to the response). According to such a hypothesis, attention would be directed to the 'handle' (small VC) of the right object resulting in larger IOR for targets that subsequently appear within that component (different VC condition). In contrast, the object on the left of fixation would not evoke the same representation for action as the right side object, thus being encoded in a different context and for a different purpose (perhaps for recognition). This possibility will be explored in Experiment 6.

4.5. Experiment 6

4.5.1 Introduction

In Experiment 5 results showed that object-based IOR was modulated by internal object properties that separated VCs in multi-component objects. In contrast, no significant modulation of object-based IOR was observed by surface boundaries. One factor that was considered to account for the elimination of any surface-based effect in Experiment 5 was that, whilst surfaces were separated by a single edge boundary, VCs were separated by two edge boundaries. This difference may have resulted in larger IOR for VCs than for surfaces. It was, therefore, premature to conclude that multi-component 3D objects are represented by VCs but not by surfaces.

In this experiment participants viewed displays of objects consisting of two VCs, that were separated by a single boundary (see Figure 38). It was hypothesised that if modulation of IOR was due to the number of edge boundaries separating the two VCs, then removing the second edge boundary would reduce the VC-based effect on IOR modulation.

Furthermore, instead of presenting the objects in two different display orientations relative to the vertical meridian (+ 45 and –45 from vertical) participants viewed displays of two objects that were either upright or inverted. There are two methodological benefits of this manipulation. The first is that RTs from both objects can be used for the analysis, as opposed to RTs from only one of the objects, as in Experiment 5. The second benefit of using data from both objects, is the opportunity to explore the issue raised in Experiment 5, with regard to the difference between left and right objects in the nature of the primitives (surfaces or VCs) by which they are represented.

4.5.2 Method

Participants

Ten volunteers aged between 22 and 34 took part in this experiment for a fee of £5.00. They were recruited through the UWB Intranet site and fitted the criteria of normal or corrected-to-normal vision and naivety to the purposes of the experiment.

Apparatus and Stimuli

This experiment utilised the same apparatus as that used in Experiment 5. The stimulus displays were images of two objects presented in dark grey colour on either side of a fixation cross (Figure 38) against light grey background. Each object consisted of two VCs. At a viewing distance of 55 centimetres the larger VC was a brick shape measuring 5.2° long, 2.6° tall and 3.1° wide. The smaller VC was a forward leaning cube that was 1.6° deep, 3.6° tall and 3.1° wide. The two objects were presented facing outwards from fixation. They were either upright or inverted (Figure 38). The fixation cross was a red 'plus' (+) sign and subtended 0.7° x 0.7°. The cue was a small white-filled parallelogram with black outline, measuring 0.4° x 0.3° (contour was 0.2° x 0.2°). The target was a light grey parallelogram subtending 0.4° x 0.3° of visual angle. The whole display from end to end subtended 18° horizontally and 8.6° vertically.

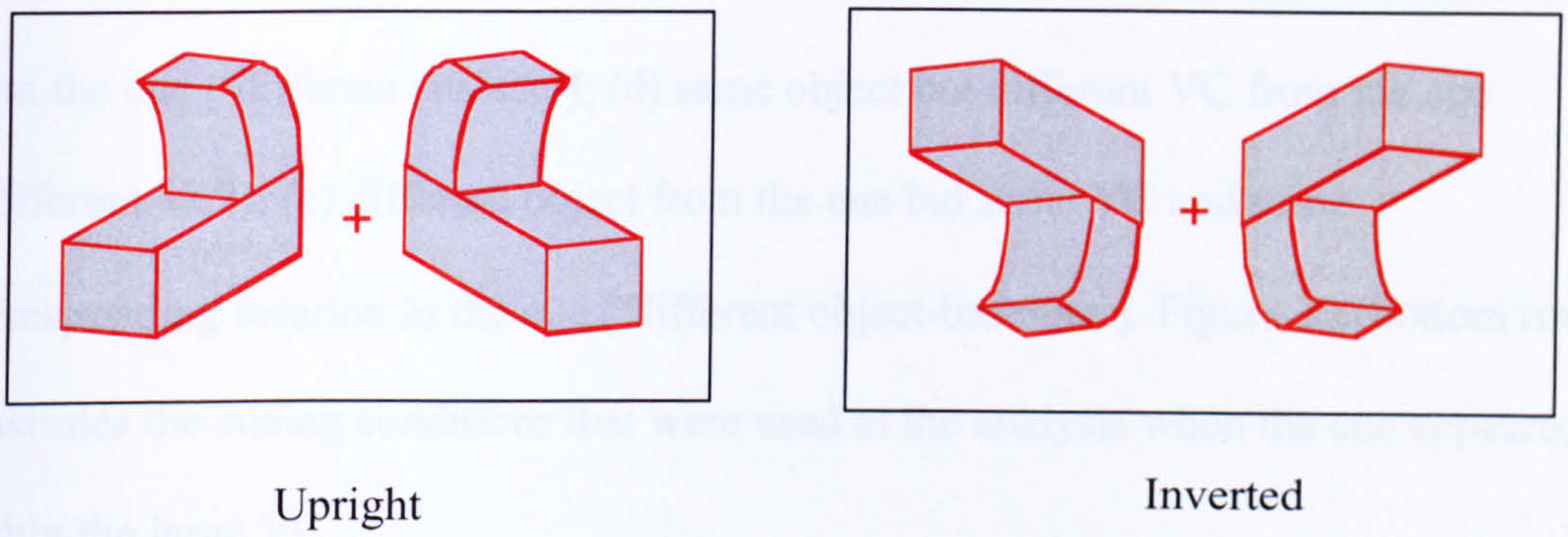


Figure 38. Images of the objects used in Experiment 6. The background was light grey during the actual experiment.

Design and Procedure

There were three factors in this experiment. The first was SOA with two levels, 820 msec and 1220 msec. The second was Display with two levels, Upright and Inverted. The third factor was Cueing with seven levels. The cues appeared in the following four positions on the objects (a) the side of the right large VC, (b) the side of the left large VC, (c) the side of the right small VC, and (d) the side of the left small VC.

Two of the Cueing levels refer to the cue-target relationship, when the cue appeared within the smaller VC. They were as follows: The target would appear within (a) the same VC, same surface and the same location as the cue (same surface-same location), (b) the same object but on a different VC (the side surface of the large VC) from the cue (different VC). Figure 39 (top row) illustrates the cueing conditions that were used in the analysis when the cue appeared within the small VC.

The remaining five cueing conditions were the cue-target relationships, when the cue appeared within the large VC of each object. These were as follows: The target would appear within (a) the same VC, same surface and same location as the cue ('same surface-same location'); (b) the same VC, same surface but different location from the cue ('same surface-different location'); (c) the same VC but different surface from the cue ('different surface'); (d) same object but different VC from the cue ('different VC'); (e) different object from the cue but same VC and same corresponding location as the cue ('different object-baseline'). Figure 39 (bottom row) illustrates the cueing conditions that were used in the analysis when the cue appeared within the large VC.

There were also *filler* trials, where the cue-target distance was larger than the cue-target distance for the within-object cueing conditions (larger than 4.0°). These trials were not used in any subsequent analyses. The filler trials when the cue was in the small VC were as follows: the target would appear (1) on the same VC but different surface from the cue; (2) on the same object but on a different VC (the top surface of the large VC) from the cue; (3) on the same object and different VC (on the side surface (near) of the large VC) from the cue; (4) on a different object, same surface and same location as the cue, (5) on a different object same VC different surface and different location from the cue (6) different object and different VC from the cue (top surface); (7) on a different object and different VC from the cue (side far); and (8) on a different object and different VC from the cue (side near).

The *filler* trials when the cue was in the large VC were as follows: the targets would appear (1) on the same object but different VC from the cue (front surface of the small VC); (2) on a different object, same corresponding surface but different corresponding location from the cue; (3) on a different object, same corresponding VC but different corresponding surface (top) from the cue; (4) on a different object and different corresponding VC from the cue (target appeared on the side surface of the small VC); and (5) on a different object and different corresponding VC from the cue (target appeared on the front surface of the small VC).

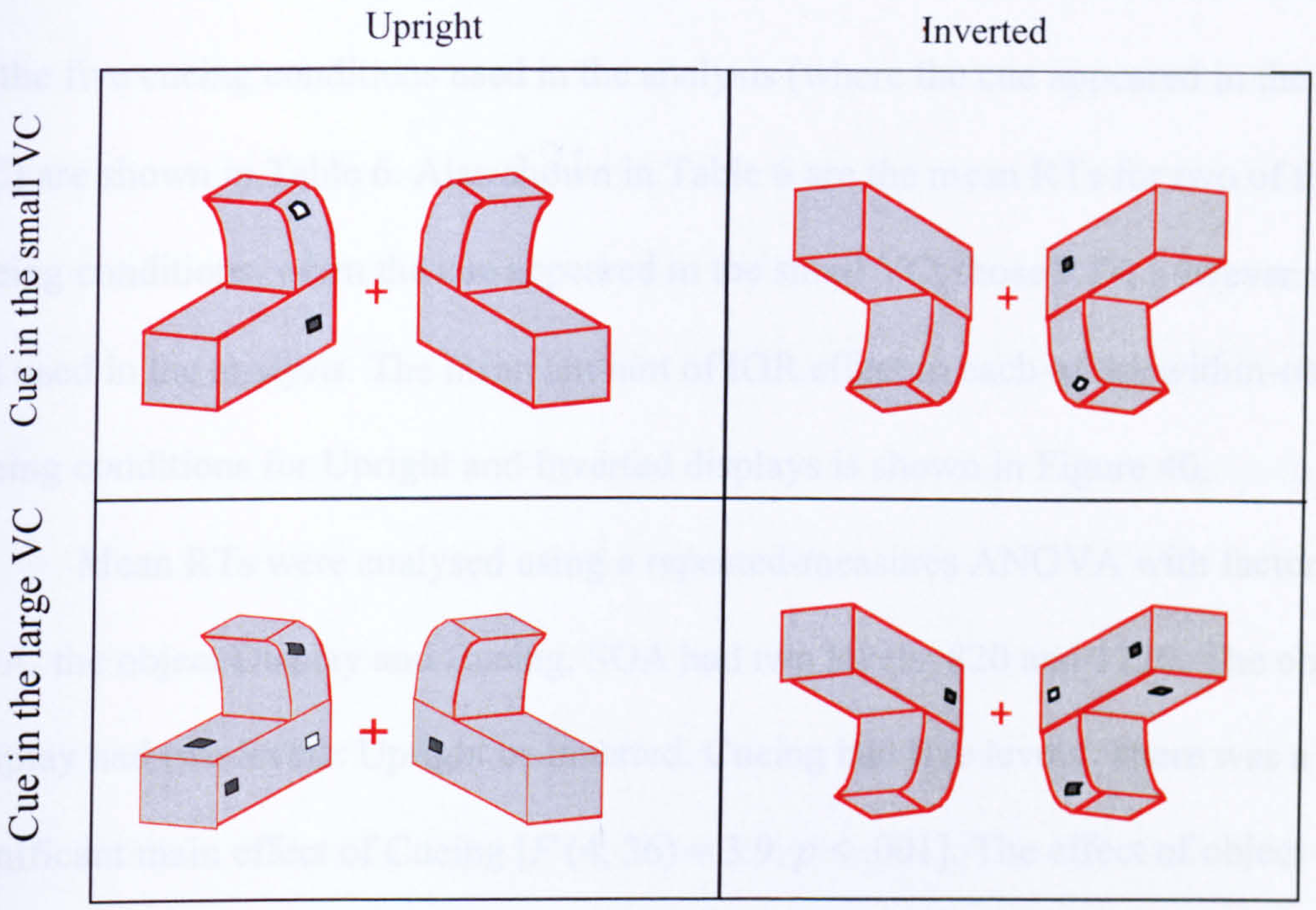


Figure 39. Examples of the positions of cues and targets in trials that were *included in the analysis*. Cues are depicted as white parallelograms with black contours and targets as grey parallelograms. Here only cue-target combination from one of the objects are depicted. The same cue-target combinations occurred for the other object as well.

Each participant completed a block of 480 trials, 160 of which were no-target trials (33%). In the remaining 320 trials each cueing condition (including the filler trials: 20 in total) appeared 10 times in each of the two displays (upright and inverted). There were short breaks every 100 trials, which participants were encouraged to take by a computer generated message. The procedure was identical to that in Experiments 4 and 5.

4.5.3 Results and Discussion

As in the previous analyses, trials with very fast (less that 100 msec) or very slow (longer than 700 msec) responses or with responses to no-target trials (incorrect) were excluded from the analysis. Filler trials were also excluded. Mean RTs for each of the five cueing conditions used in the analysis (where the cue appeared in the large VC) are shown in Table 6. Also shown in Table 6 are the mean RTs for two of the cueing conditions, when the cue appeared in the small VC; those RTs, however were not used in the analysis. The mean amount of IOR effect in each of the within-object cueing conditions for Upright and Inverted displays is shown in Figure 40.

Mean RTs were analysed using a repeated-measures ANOVA with factors the SOA, the object Display and Cueing. SOA had two levels: 820 and 1220. The object Display had two levels: Upright or Inverted. Cueing had five levels. There was a significant main effect of Cueing [F (4, 36) = 3.9, $p < .001$]. The effect of object Display was not significant [F (1, 9) <1, ns, and neither was the effect of SOA, F (1, 9) < 1, ns. The Display x Cueing interaction, however, was significant [F (4, 36) = 2.3, $p < .05$].

Table 6. Mean RTs (msec) for each of the cueing conditions for the Upright and Inverted displays.

Cue location	Same Surface Same location	Same Surface	Different Surface	Different Volume	Different Object
		Upright			
Large VC	324	307	324	316	288
Small VC	326	-	-	348	319
		Inverted			
Large VC	312	300	308	307	293
Small VC	321	-	-	325	293

Planned comparisons were carried out on RTs in trials where the cue appeared in the large VC. In those trials the cue-target distance in each within-object cueing condition (except the 'same surface-same location' condition) and in the 'different object-baseline' condition was identical. In the Upright displays RTs in the 'same surface-same location', 'same surface-different location', ' different surface', and 'different VC' cueing conditions were significantly slower than RTs in the 'different object' condition [t (9) = -2.4, p = 0.03, t (9) = 2.0, p < .05, t (9) = 2.0, p < .04, t (9) = 2.1, p < .05 respectively]. Critically, the difference between the 'same surface-different location' and 'different surface' cueing conditions was significant (17 msec), t (9) = -2.0, p <.05, replicating previous results. However, there was no significant difference between the 'same surface-different location' and 'different VC' conditions, (1 msec), t (9) < 1, ns. Finally, there was no significant RT difference between the 'different surface' and the 'different VC conditions, t (9) < 1, ns[32].

In the Inverted displays RTs in the 'same surface-same location' condition were significantly larger than RTs in the 'different object' condition, t (9) = 2.1, p < .05, replicating the original object-*plus* location-based IOR. RTs in the 'different surface' and 'different VC' cueing conditions were significantly different from the 'different object' condition [t (9) = 2.0, p < .04 and t (9) = 2.2, p < .05 respectively]. The difference between the 'same surface-different location' and 'different surface'

---

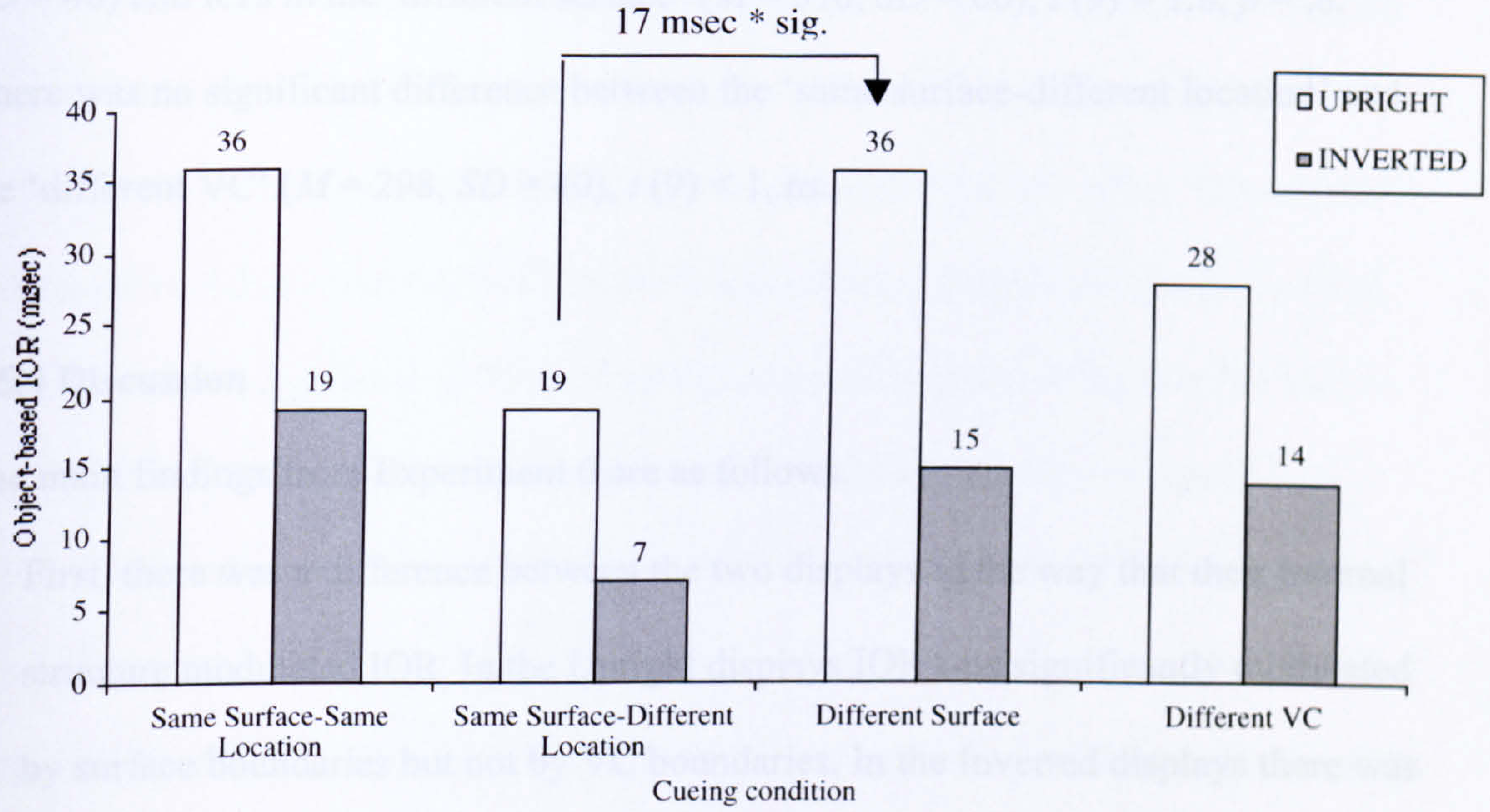
[32] Planned comparisons between the four between-object filler conditions for the Upright displays revealed no significant differences between the 'different object-baseline' (M = 288) and the 'different object-same surface-different location' (M = 303) condition, t (9) <1; the 'different object-different surface'(M = 294) condition, t (9) = 1.1, and the 'different object-different VC' (M = 277) condition, t (9) <1. None of the differences between the three filler conditions were significant.
In the Inverted displays, there were no significant differences between the 'different object-baseline' (M = 293) and the 'different object-same surface-different location' (M = 280) condition, t (9) =1.6, ns; the 'different object-different surface'(M = 285) condition, t (9) < 1, and the 'different object-different VC' (M = 298) condition, t (9) <1. The differences between the three filler conditions were not significant.

conditions (8 msec) was not significant, $t(9) < 1$; and neither was the difference between the 'same surface-different location' and 'different VC cueing conditions (7 msec), $t(9) < 1$.

Furthermore, post-hoc comparisons between the IOR (RT cued minus RT uncued) effects for each cueing condition in the two Display orientations were carried out. These did not reveal any significant differences between the two orientations. The difference in IOR between the two orientations in the 'same surface-same location' and the 'different surface' conditions were only marginally significant, $t(9) = 2.0, p = .07$ and $t(9) = 1.7, p = .09$.

As previously described, in Experiment 5 data from only one object were used to investigate the effects of object-internal structure on Cueing. In short, in the *+45 displays* only data from the *left object* were used, and in the *–45 displays* only data from the *right object* were used. The results showed that in the +45 displays there was no significant modulation of IOR by surfaces or by VCs (whilst there was a trend for a surface-based cost of 11 msec). In contrast, in the –45 displays there was a significant modulation of IOR by VCs. It is, therefore, possible that the difference in display orientation reflected a difference between left and right objects.

In order to investigate the patterns of IOR modulation for the left and right objects, further analyses were carried out for each object in each display orientation. In the Upright –Left object, RTs for the 'same surface-different location' trials (M = 305, SD = 66) were significantly different from RTs in the 'different surface' trials (M = 323, SD = 71), $t(9) = 2.3, p = .03$; but not significantly different from the 'different VC' trials (M = 283, SD = 56), $t(9) < 1$, ns. This pattern of results replicates the trend for a surface-based cost in the +45 displays of Experiment 5, where only the *left* object was used. In the Upright-Right object, only RTs between the same surface-different location (M = 311, SD = 50) and the different VC (M = 327, SD = 72) conditions were marginally significant, t (9) = 1.99, p = .08. This trend of a VC-based cost replicates the pattern of results in the —45 displays of Experiment 5, where only the *right* object was used.



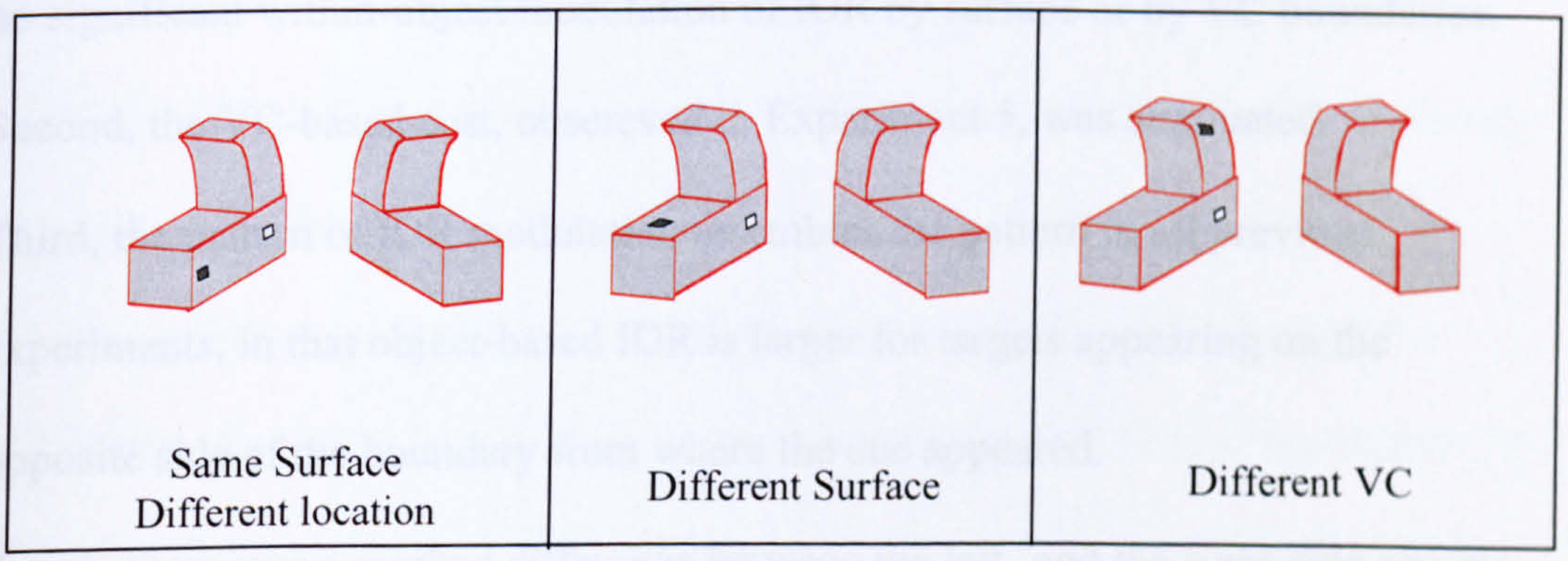


Figure 40. Object-based IOR for each within-object cueing condition in each display grouped by Orientation. Only trials where the cue appeared in the large VC are shown here, as these are the only trials where cue-target distance in each of the within-object cueing conditions (except the same surface-same location) was identical. The only significant difference was between the Upright same surface-different location and Upright different surface conditions.

In the Inverted - Left object, RTs for the 'same surface-different location' trials (M = 305, SD = 60) were not significantly different from RTs in the 'different surface' trials (M = 305, SD = 58), t (9) < 1, ns, or RTs in the 'different VC' trials (M = 312, SD = 55), t (9) < 1, ns. In the Inverted - Right object there was a marginally significant difference between RTs in the 'same surface-different location' (M = 290, SD = 48) and RTs in the 'different surface' (M = 310, SD = 60), t (9) = 1.8, p = .6. There was no significant difference between the 'same surface-different location' and the 'different VC' (M = 298, SD = 40), t (9) < 1, ns.

4.5.4 Discussion

The main findings from Experiment 6 are as follows.

- First, there was a difference between the two displays in the way that their internal structure modulated IOR. In the Upright displays IOR was significantly modulated by surface boundaries but not by VC boundaries. In the Inverted displays there was no significant within-object modulation of IOR by surface or by VC boundaries.
- Second, the VC-based cost, obsereved in Experiment 5, was attenuated.
- Third, the pattern of IOR modulation resembles the pattern in all previous experiments, in that object-based IOR is larger for targets appearing on the opposite side of the boundary from where the cue appeared.
- Fourth, there was a marked difference between the left- and the right-side objects in the primitives that significantly modulated the IOR effect.

Two important issues are raised from the results of the Experiment 6. The first relates to the elimination of the VC-based effect previously observed in Experiment 5. The other relates to the Cueing by Display interaction.

With regard to the first issue, it was unclear from the results of Experiment 5, whether the VC-based cost was due to the number of boundaries separating the cue-target locations or to the representation of the VCs per se. It was hypothesised that if modulation of object-based IOR operates over VC-based as opposed to surface-based representations, then equating the number of VC boundaries with the number of surface boundaries, would not reduce the VC effect found in Experiment 5. If, on the other hand, the VC effect was due to the number of boundaries separating the two volumes, then removing the extra boundary would reduce the VC effect. Indeed, the results showed that when the two VCs were separated by a single boundary, the IOR effect for targets appearing in the uncued VC was reduced, indicating that the VC-based cost in Experiment 5 was most likely due to the number of boundaries separating the two components. In contrast, the object-based IOR effect for targets appearing on the *other side* of a surface boundary (36 msec; on the same VC) was significantly different from the IOR effect for targets appearing on the *same side* of the surface boundary as the cue (19 msec). This surface-based cost was the *only* statistically reliable effect and replicates the surface-based cost observed in previous experiments in this thesis.

The different pattern of modulation for the left and the right object of the upright displays may also help explain the attenuation of the VC-based modulation of IOR. The results showed that surface-based modulation of IOR was observed predoninantly for the left object, whilst there was a (marginally significant) trend for VC-based modulation for the right object. This pattern is consistent with the pattern of IOR modulation for each object in Experiment 5. One possible explanation for this finding is that the small volumetric component of the right object may be the most prominent (and perhaps graspable) feature of the object, especially for right-handed participants (eight out of the ten participants were right-handed). Thus, it is possible that a representation is created for the right side object that makes its volumetric parts, one of which is a graspable part, explicit. Whilst this issue which lies beyond the specific scope of this investigation, it bears important implications for the interepretation of the present results with regard to the nature of shape primitives used to represent the perceived objects.

It is therefore proposed that surface-based representations are customarily, and across many different experiments, computed and modulate object-based IOR. When, however, the perceived object posseses an action-associated component, then this component determines the representation computed for attention. In this case it is a representation of the object's volumetric components, one of which is the component affording action.

One possible confound in Experiment 6 is that cues and targets in the 'different' surface condition appeared on different planes, whilst cues and targets in the 'different VC' condition appeared on the same plane. Therefore, this difference in planes in the 'different surface' condition may have accentuated the parsing of the object into different surfaces and caused larger object-based IOR for this condition. In contrast, the two surfaces in the 'different VC' condition appeared along the same plane, only separated by an edge boundary. It is not clear at present, why the edge boundary separating the two VCs did not significantly modulate object-based IOR (as was the case in Experiments 2 and 4). Further work is needed to clarify the way that 3D forms modulate object-based attention compared with modulation of object-based attention by 2D forms.

The second issue raised by the present results relates to the interaction between Display and Cueing. Once again, two different displays of the same stimulus set resulted in contrasting patterns of object-based IOR. What accounts for such difference in the modulation of object-based IOR? Take, for example, the predictions of the *local structure* hypothesis. According to this hypothesis inhibitory attentional selection can operate over representations of an object's local structure, *without* excluding the representation of the object's global shape. How would these predictions materialise in the data? One way perhaps would be in the formation of a single representation which makes explicit both the object's global and local structure. The present results indicate that one type of representation for the objects is computed that makes explicit both the local and the global structural infomration. This is consistent with the local structure hypothesis, as it is the only hypothesis that can account for *both* observed effects.

This explanation is also plausible on the grounds of previous proposals, that have posited that an object and its parts can be represented simultaneously. For instance, in a connectionist model for 'decomposition' of visual scenes, Mozer (1999) found that an object can be parsed into different parts and each part can be assigned a different *tag*. Despite this the structure of the object is still represented in the network, which would allow the object to be represented as a whole, whilst each of its parts is also represented separately.

Despite the fact that the local structure hypothesis can account for the present pattern of results, it cannot *explain* the effect of display orientation. One reason why the two orientations would show different patterns of IOR modualtion may lie in the physical plausibility of the objects. Upright objects may be perceived as probable or 'comfortable' objects, unlike the inverted objects which may appear to be standing, rather uncomfortably on a hypothetical horizontal plane. If object-based IOR selects from higher-order representations of objects, i.e. after categorasation, then it is possible that the inverted objects may not attract attention the same way that the upright objects do, since they are encoded as improbable or 'uncomfortable' objects. This is an interesting possibility and its verification in future studies will have serious implications for theories of object-based attention.

The present results replicate previous work in the following ways. First, they indicate that IOR operates over shape representations that make explicit boundaries between surfaces, replicating the surface-based effects for 3D forms in Experiment 3. Second, they support previous work showing that 3D shape is represented by surface primitives, whose representations are subsequently matched in long-term memory (e.g. Leek & Arguin, 2000; He & Nakayama, 1992, 1995). Third, the present results raise the possibility that volumetric components can also computed depending on the context in which the object was encoded. This is largely consistent with evidence that even objects that are irrelevant to the task in hand (i.e. press a button at the onset of an imperative target), can evoke a mental representation for action (e.g. Tucker & Ellis, 1998; Phillips & Ward, 2002). Future work is needed to systematically explore this possibility.

Finally, it is clear that the present results do not support the *global structure* hypothesis, according to which selection operates *solely* over representations of an object's global form. Instead a variant of the *local structure* hypothesis, where the object's global form is simultaneously selected with that object's local structure, is more likely to account for the present results.

Chapter 5

5. General Discussion and Conclusions

5.1 Introduction

In two sets of experiments this thesis examined whether internally structured shape representations can modulate the object-based IOR effect, and the nature of the shape primitives that are responsible for such modulation. In the first set of experiments, two hypotheses with contrasting predictions about the kinds of object shape representations were examined. The *global structure* hypothesis predicted that inhibitory attentional selection, (expressed as the magnitude of object-based inhibition of return, IOR) operates over representations of the object's global shape, without making explicit the object's internal structural information. The *local structure* hypothesis predicted that inhibitory selection mechanisms operate over representations that make explicit the object's internal structure.

In the second set of experiments, this thesis examined the units or primitives of object-shape representation – surfaces, or VCs - that modulate object-based IOR. Previous empirical evidence has shown that attention can be oriented not only towards whole objects in the environment but also towards component parts of objects. More specifically, a number of spatial cueing studies (e.g. Egly, Driver & Rafal, 1994; Nakayama, He & Shimojo, 1995; He & Nakayama, 1995; Jordan & Tipper, 1999) have shown that attention can spread across the surface of a two-dimensional form (object); whilst studies in divided attention (e.g. Vecera, Behrmann & McGoldrick, 2000; Vecera, Behrmann & Filapek, 2001) have shown that there is a benefit in accuracy and response latency, when two to-be-reported attributes belong to the same component part of the object.

Although theories and models of object-based attention and object-shape representation have co-existed for some time, few attempts have been made to integrate conclusions from the two areas. This is unfortunate if any predictions about the units of object-based attention are to be made. With only a few exceptions (e.g. Baylis and Driver, 1995), studies on object-based attention have intuitively pre-defined objects on the basis of Gestalt principles of grouping, similarity, proximity, common fate, and so forth.

In this final chapter I present the main findings from the experiments in this thesis and discuss their theoretical importance for (a) accounts of object-based attention, (b) accounts of object-shape representation, and (c) the constraints they posit on the distinction between facilitatory and inhibitory components of attention.

5.2 Object-internal structure modulates object-based IOR

As reviewed in Chapter 2 recent work in object-based visual attention has identified two ways in which attentional selection of objects may operate. Thus, evidence exists that attention may be directed to the whole object (e.g. Fuentes et al., 1998; O' Craven et al., 1999), or to individual parts of an object (e.g. Vecera et al., 2000; Vecera et al., 2001). This evidence give rise to at least two hypotheses that make contrasting predictions about the level of representation available for inhibitory attentional selection. According to the *global structure* hypothesis, object-based attention only selects objects as global forms, whilst the *local structure* hypothesis predicts that object-based attention selects from locally structured representations.

The results of Experiment 2 clearly showed that object-internal structure modulates object-based IOR in objects containing an internal structural discontinuity. This finding was in direct support of the local structure hypothesis and was the first demonstration of 'part-based' attentional modulation of *object-based IOR*. It was also the first demonstration of 'part-based' attentional effects using an *implicit* measure of attention. These points will be addressed in more detail later.

Experiment 4 was designed to (a) replicate the findings of Experiment 2 and (b) to address the issue of the effect of display orientation in Experiment 3. As mentioned before, the effect of stimulus display orientation on IOR modulation was entirely unexpected and of little theoretical importance in this series of experiments. However, display orientation did sometimes interact (Experiment 5) or was a main effect (Experiment 3) at the same time as the, theoretically relevant, main effect of cueing. Therefore, it was difficult to ignore the fact that the way the objects appeared on the screen had an impact on the part-based modulation of IOR. Possible explanations for the display orientation effect will be discussed later in section 5.6. For now, and with respect to the two contrasting hypotheses, the data suggest that *both* local and global shape information being computed. This is inconsistent with the global structure hypothesis and can only be accounted for by the predictions of the local structure hypothesis.

The implications of the results from Experiment 4 are very important. First, they replicate the finding that object-based IOR is modulated by internal object structure (Object Present). Second, the effect of IOR in the Object present trials was significantly larger than the IOR effect for the Object absent trials. Third, when there was no object present (Object Absent), there was no within-hemifield modulation, indicating that any IOR modulation in the Object Present conditions was not just due to the spread of IOR across the cued hemifield (e.g. Pratt, Spalek & Bradshaw, 1999; Collie, Maruff, Yucel, Danckert & Currie, 2000). Furthermore, the absence of any significant effects of orientation in the Object Absent condition points to the possibility that any effect of, or interaction involving orientation, is related to properties of the *object itself in that orientation.*

These results are in agreement with previous studies on IOR showing that object-based IOR can operate over representations that code within-object locations. Recent studies on representations mediating object-based IOR have shown that, depending on the task, IOR can be associated with the whole object or with its parts (e.g. Tipper et al., 1999), or with a particular location within an object (Gibson & Egeth, 1994). The present findings are also in agreement with other recent studies showing part-based IOR effects using divided attention tasks (Vecera et al., 2000; Vecera et al., 2001). Finally, the results form an important body of evidence towards a current debate on the location- and object-based inhibition-of-return. Unlike recent evidence suggesting that IOR is a single-component orienting mechanism (e.g. McAuliffe, et al., 2001) that operates on objects and location in a similar manner, results from Experiment 4 are largely consistent with the notion of seperate location-based and object-based components of IOR (e.g. Jordan & Tipper, 1998). This is extremely important, as it indicates that the modulation of IOR by internal-object structure is not simply an artefact of the spatial distribution of IOR across the visual field.

5.3 Nature of object-shape representations in object-based attention.

Having obtained evidence in favour of the local structure hypothesis, the second issue addressed in the thesis concerned the *nature* of internal representations that modulate IOR. As briefly reviewed in Chapter 2, there are several competing proposals about the nature of shape representation in the object recognition literature. Some hypotheses assume that object shapes are represented in terms of low-level features such edges and shape contour (e.g. Lowe, 1985; Ullman, 1989), in terms of surface-based shape descriptions (e.g., Leek & Arguin, 2000; He & Nakayama, 1992; Nakayama & He, 1995; Marr & Nishihara, 1987; Pentland, 1989) or in terms of structural descriptions of volumetric shape components (e.g. Biederman, 1987; Marr & Nishihara, 1987). Here, two possible candidate shape primitives, that is, surfaces and VCs, were investigated as primary candidates for the representation of 3D forms over which object-based IOR operates.

In short, Experiments 3, 5 and 6 showed that object-based IOR is also modulated by internal structural properties, such as surface discontinuities, of 3D forms. The pattern of this modulation followed the same pattern as in experiments using 2D stimuli, with object-based IOR being larger for targets appearing on a different surface of the same object from the cue.

Results from Experiment 3 suggested that IOR was modulated by an internal representation that made explicit information about the surfaces in three-dimensional forms. This finding is consistent with a number of accounts that posit the importance of surfaces in selective attention (e.g. He & Nakayama, 1995) and divided attention tasks (e.g. He & Nakayama, 1992) as well as in object recognition tasks, such as 'whole-part matching' (e.g. Leek & Arguin, 2000). However, this finding contradicts the idea that attention is bound to the surface of the object it was initially deployed to, without generalising to other surfaces of the same object (e.g. Xu & Nakayama, 2000).

Experiment 5 examined the issue of the nature of object-shape representations available for selection more directly, using images of objects consisting of two VCs. The design allowed the direct comparison between surfaces and VCs in the way they modulate object-based IOR. In accordance with a volumetric-component account of object-shape representation, the IOR effect for targets that appeared on the same VC were significantly larger from those targets that appeared on a different VC. However, there was no significant difference between targets that appeared on different surfaces of the same VC. Importantly, the VC-based effect was only observed in the +45 orientation, where data from the right object were used (for reasons explained in the Design section of Experiment 5). This finding was in stark contrast with results in all previous experiments that underlined the importance of surfaces in the representations available to object-based IOR.

One possible explanation for the attenuation of the surface-based effect is that the VCs in Experiment 5 were separated by two edge boundaries, whilst surfaces within each VC were separated by a single edge boundary. Therefore, cues and targets were separated by two edge boundaries in the condition where the cue and target appeared on different VCs ('different VC' condition), but only by a single edge boundary in the 'different surface' condition. This may have biased observers to 'parse' the objects into two separate perceptual groups, the two VCs, a representation of which did not make surfaces explict. Furthermore, the VC-based effect was observed in the +45 displays, where data from only the right object were used. This contrsasted the (statistically unreliable) surface-based effect for the –45 displays, where data from the left side object were used. Therefore, there was a possibility that there was visual field component underlying the VC-based effect.

Experiment 6 was a step towards the resolution of this contrasting pattern of results across experiments. Two important findings emerged from this experiment, relating to the nature of primitives that can modulate object-based IOR. The first relates to the attenuation of the previously observed VC-based effect. The second, relates to the difference between the left- and right-side objects in the primitives that can modulate IOR.

The first most important finding in Experiment 6 was the VC-based effect gave way to a reliable surface-based effect on the modulation of object-based IOR. Thus, it is possible that in Experiment 5 the extra boundary between VCs encouraged the 'parsing' of the object into two separate perceptual groups belonging to the same object. Thus, on the one hand, IOR for targets that appeared on the same surface as the cue was attenuated because cue and target belonged to the same perceptual group (same VC)– resulting in facilitation (Fox, 1995; Fuentes et al., 1998). On the other hand, targets that appeared on a different VC from the cue were 'free' from the facilitatory effect of perceptual grouping and were therefore subject to object-based IOR. Therefore, it is possible that objects in Experiment 6, where surfaces and VCs were separated by a single boundary, were 'parsed' into surfaces, encouraging the perception of each surface as a separate perceptual group.

The second important finding of Experiment 6 was the difference in the pattern of IOR modulation and the relevant primitives for the left and the right object. Results from separate analyses of the two objects in the Upright displays, showed that modulation of IOR by surface boundaries is more predominant and statistically reliable for the left object, whilst modulation of IOR by VC is observed for the right object (even though it does not always reach significance).

On the basis of the present findings, it is premature to exclude the role of VCs in representations available for attentional selection. The present results only allow to conclude that VCs are *not* encoded in the shape representations that modulate object-based selection, whilst surfaces are encoded in these representations. However, one cannot exclude the possibility that VCs *are* made explicit in other forms of shape representations, for instance those representations computed for perceptuo-motor tasks.

Therefore, whilst the present results consistently suggest that surfaces are more readily available for attentional selection, it appears that visual object properties, other than shape internal structure, such as the object's possibilities for action, may be encoded in the representation available to object-based attentional mechanisms. One way to take these results further, would be to systematically manipulate the salience of an object's VC observing its effect on the pattern of IOR modulation. Finally, consider differences in object-based attentional modulation between familiar and unfamiliar objects. At the moment the existing literature offers little ground for making sound predictions on any of these issues. An illustration of the present findings with regard to the nature of object-shape representations in perception/recognition is outlined in Figure 41.

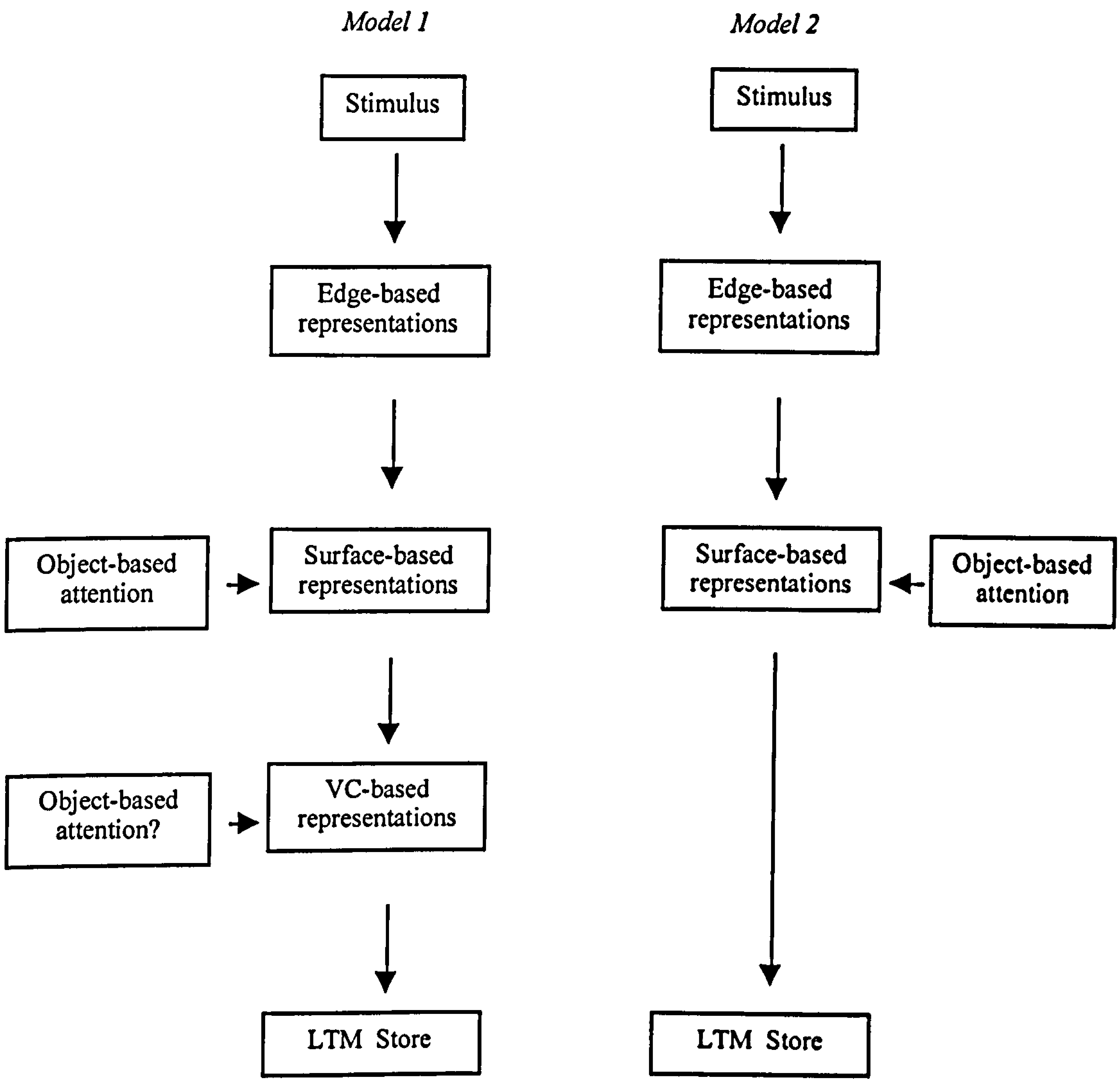


Figure 41: Schematic illustration of the two models for the derivation of object shape representation in object-based attention. The data in the present thesis support the idea that object-based inhibitory mechanisms operate over representations of internal structure. Furthermore, the present data suggest that surfaces are the relevant shape primitive for these representations (Model 1). The data from experiments 5 and 6 suggest that a representation of VCs is not computed for object-based attention (Model 2). However, the role of VC-based representations is not excluded for the purposes of recognition or perceptual-motor control.

The present set of experiments is only a stepping stone to an exciting area of research that will bridge the gap between the areas of object-based attention and object shape representation. Future work can take these findings further in a number of ways. One way is by systematically controlling surface and VC characteristics. Does a change of texture, colour or orientation modulate IOR, without or in addition to internal surface boundaries? For example, what would be the pattern of results for targets that appeared on different regions of a shaded cylinder? Furthermore, perceptuo-motor tasks, such as reaching and grasping, can reveal a lot about the shape representations computed for action-oriented attention (e.g. Milner & Goodale, 1995). Representations computed for the purpose of *perceiving* the world, i.e. recognising one's own coffee mug in a cluttered staff kitchen, are very different from those computed for the purpose of *acting* in the world, i.e. guiding our hand to reach and grasp the object in question, whilst avoiding surrounding obstacles.

The following section reviews two of the issues raised in these experiments relating to object-based attention, that is (a) the generalisation of IOR across the whole object, and (b) the effect of stimulus display orientation on modulatory effects on object-based IOR.

5.4 Generalisation of attention across the whole object

The results from Experiments 3 and 6 suggest that (a) the object-based IOR effect generalises from one surface to another surface of the same object and that (b) a representation that makes explicit surface properties modulates this effect.

These findings are consistent with the body of literature, which posits that attention can be deployed to surface-based representations of a scene as opposed to low-level image features, such as edges and vertices (e.g. He & Nakayama, 1992; Nakayama & Shimojo, 1992; He & Nakayama, 1995). However, these results are inconsistent with previous work in selective and divided attention tasks using images of 3D objects, showing that attention does not generalise to the whole object (e.g. Gibson & Egeth, 1994; Xu & Nakayama, 2001). For example, in a recent study Xu and Nakayama (2001) reported that once attention is deployed on a surface of an object it does not generalise to the other surfaces of that same object.

In contrast to the above, the present findings support the idea that object-based attention can generalise across the whole object, whilst having access to representations of that object that make surfaces explicit. The theoretical implications of this finding are important, since it indicates that a part-based representation of the object need not exclude processing of global shape properties, as predicted by the local structure hypothesis. This point is addressed in the following section.

5.5. Comparison between the present findings and previous evidence on part-based modulation of attention

The present data form a body of evidence in support of the idea that attention may select from internal representations that make explicit local structural information of an object. Thus, together with other paradigms, where 'part-based' effects were observed in observers' judgements about object properties, this set of findings endorse the notion of part-based modulation of attention.

The present findings differ from previous work in the area in three important ways. First, a measure of *inhibition* was used instead of facilitation. Second, previous work on part-based attention has utilised divided attention tasks, in which participants are required to explicitly report attributes of object parts. In the present experiments the object were irrelevant to response, allowing object-based IOR to act as an *implicit measure* of attention to objects and object parts. Third, evidence that part-based effects may co-exist with object-based effects, in the form of larger object-based IOR for uncued surfaces of the cued object. These differences are discussed in turn.

5.5.1. IOR versus facilitation

Vecera and his colleagues (Vecera, Behrmann & McGoldrick, 2000; Vecera, Behrmann & Filapek, 2001) have shown that individual parts of a single object can be selectively attended to and lead to facilitation in performance, when two attributes of the same part are to be reported. The present data show that object-internal structure may also modulate object-based IOR (see Houghton & Tipper, 1994).

5.5.2. Implicit versus explicit measure of selective attention

Previous work on part-based attention has utilised divided attention tasks, in which participants are required to explicitly report attributes of object parts (e.g. Vecera et al., 2000; Vecera et al., 2001). A potential limitation of this paradigm is that it requires participants to attend to separate parts, thus encouraging the coding of the stimulus not as an object but as a aggregation of separate parts[33]. In the experiments in this thesis

[33] This suggestion was also raised by Vecera et al., 2000. In their second experiment they used two object displays in order to address this potential confound. However, the task still required participants

the use of the IOR paradigm shows that these modulatory effects can also be observed with an implicit measure of attentional selection, when the object and its component parts are irrelevant to the task.

5.5.3. Pattern of modulation of object-based IOR

An important difference between the present findings and previous reports on part-based modulation of attention lies in the pattern of such modulation. The results showed that object-based IOR was larger when the target was presented within the *different* part of the cued object; that is, when the cue and target were separated by an internal structural discontinuity. The reduction in the magnitude of IOR observed in the present experiments for cue-target pairs appearing on the same object feature – surface or VC - may represent the summation of a facilitatory and an inhibitory effect. As such, this finding adds to other evidence showing that object-based attention may operate on grouped locations within objects (Egly et al., 1994; Gibson & Egeth, 1994; Lavie & Driver, 1996; Avrahami, 1999). Furthermore, the finding of increased inhibition across a surface discontinuity is inconsistent with a view of inhibition operating on a within-object space-based way, where inhibition should *decrease* as a converse function of cue-target proximity. On these grounds, this finding may provide an important constraint on hypotheses about the distribution of facilitation and inhibition across shape representations during object-based selection.

Two possible explanations for this finding are considered here. One relates to effects of the possibility of simultaneous operation of inhibitory and facilitatory

---

to attend to and report properties of the objects' parts.

components of attention on the different representations of the objects. The other relates to the visual complexity of the stimuli.

It is possible that this effect reflects the differential spread of facilitation and inhibition across representations of the object shape. This is consistent with previous reports have repeatedly shown examples of the simultaneous operation of distinct facilitatory and inhibitory components of attention (e.g. Tassinari, Aglioti, Chelazzi, Peru & Berlucchi, 1994; Tipper, Rafal, Reuter-Lorenz, Starrveldt, Ro, Egly, Danzinger & Weaver, 1997; Mari-Beffa, Houghton, Estevez & Fuentes, 2000). For example, Mari Beffa et al (2000), using a letter search task, found that letter-word grouping reduced negative priming effects, when the task was to report a letter (target) within a word (distracter). Negative priming was found when the letter to be reported appeared outside the distracter word. Thus, targets appearing <u>within</u> the cued part of a task-relevant object may be subject to both facilitation through grouping, and inhibition, depending on the cue-target interval (SOA). In contrast, targets appearing <u>outside</u> the cued part of an object appear to be subject to only inhibitory effects. Thus, in the present case, it is likely that facilitation and inhibition operate in parallel across object structure, but spread differentially across cued and uncued object parts. While inhibition may spread equally across visible components of object shapes, facilitation may accrue only on cued object components. The summation of these differential effects would produce a modulation in the pattern of IOR consistent with the present results; that is, greater IOR for cues and targets appearing on different object components.

As such, this observation adds to other evidence showing that object-based attention may operate on grouped locations within objects (Egly et al., 1994; Gibson and Egeth, 1994; Lavie and Driver, 1996), and may provide an important constraint on hypotheses about the distribution of facilitation and inhibition across shape representations during object-based selection.

Another factor that may have contributed to this pattern of results is the *visual complexity* of the segmented displays compared to the unsegmented displays, suggesting that visual complexity as opposed to attention to object parts may have caused the observed differences between the two types of display. However, as already argued in the discussion of Experiment 2, if visual complexity had contributed to the overall difference in IOR magnitude between segmented and unsegmented displays, it would not have influenced the pattern of IOR modulation within each display. Irrespective of whether the difference between the segmented and unsegmented displays is characterised in terms of part structure or in terms of visual complexity, the fact remains that object internal structure modulates object-based inhibitory attention.

Is it possible that the difference between segmented and unsegmented displays is not really due to the internal discontinuity but to the same factors that differentiate one orientation (i.e. +45 or Upright) from the other (i.e. –45 or Inverted) in the other experiments of this thesis? Most probably not, since the *only* difference between the two types of display (segmented and unsegmented) in Experiment 2 was the internal edge boundary separating the two parts. In contrast, in the other experiments internal structure remained the same but the stimulus display orientation changed, introducing a factor that is external to (or other than) the objects' internal structure.

This pattern of modulation of object-based IOR has important theoretical implications relating to issues of object-based and part-based attention. One consideration of previous effects of part-based modulation of attention has been the reconciliation between these part-based effects and object-based effects. In other words, if the object's individual parts are represented and subsequently modulate attention, then is the object's identity also represented and subsequently able to modulate attention? The present findings go some way towards answering this question. The finding that object-based IOR is always larger for uncued surfaces of the cued object indicate that attention generalises to the whole object *on the basis* of surface-based representations. Therefore, modulation of attention by object parts does not exclude modulation of attention by the whole object. This is consistent with Mozer's (1999) proposal that an object and its parts can be represented simultaneously. In a connectionist model of processing of visual scenes, Mozer found that an object can be parsed into different parts and its part can be assigned a different *tag*. Despite this the structure of the object is still represented in the network, which would allow the object to be represented as a whole, whilst each of its parts is also represented separately.

This finding is also consistent with proposals that attention to an object entails attention to all of its features (Kahneman & Henik, 1991; Duncan, 1993; O' Craven et al., 1999). It is possible that the object's features (i.e. surfaces or VCs) were selected as part of the whole object, leading to object-based attention (IOR for different parts of the same object) but their individual representations were also able to modulate attention, leading to findings of part-based attention (IOR was *larger* for different parts of the object).

5.6 The effect of display orientation on object-based IOR modulation

One issue raised in the 3D form experiments relates to the finding that stimulus display orientation played a role – either as a main effect or by interacting with the effect of cueing – in the modulation of object-based IOR. The results from Experiment 4 confirmed that the orientation effect on within-object modulation was closely linked with the presence of the object, rather than with simply directing attention towards the right or the left hemifield. In this section I explain how the effects of orientation may reflect (a) left and right visual field differences for Experiment 5 and (b) the effect of higher-order object-shape representations on the the modulation of IOR in the case of Experiment 6.

One account for the unexpected difference in within-object modulation of IOR between the two orientations, is the difference between the two orientations in terms of left and right visual field (LVF and RVF) asymmetries. The objects in the +45 orientation (in which the part-based effect were observed) were pointing towards the RVF, stimuli in which are processed in terms of their local structure, whilst the objects in the –45 orientation were pointing towards the LVF, stimuli in which are processed as whole forms (e.g. Sergent, 1982; Delis et al., 1986; Robetrson & Lamb, 1988). For instance, Robertson and Lamb (1988) showed that a dynamic property of the display, such as the reference frame adopted during a task, can modulate visual field asymmetries (as discussed in Experiment 3). Taking their interpretation a step further, it is possible that, in Experiments 3 and 5, object internal structure modulated the object-based IOR effect in the +45 display orientation by means of the rightward direction of the display, which facilitated local structure processing by the left hemisphere.

A different account, along the same line of argument, lies in the difference in modulation of IOR between the Left and the Right object in the Upright dispays. In Experiments 5 and 6 the left- and right-side objects modulated IOR differently. That is, a (non-significant) surface-based effect was observed for the left side object, whilst a (significant) VC-based effect for the right side object. In the case of Experiment 5, as previously dicussed, it is possible that the differences in display orientation reflected differences in the way the left and right objects were represented. However, whilst the difference between right and left object in Experiment 5 may explain the different pattern of results betwen the two orientations (i.e. VC-based effect for the right object and surface-based effect for the left obejct) , it does not explain the lack of significant modulation of IOR in one of these orientations (-45). I will return to this point later.

Finally, in Experiment 6, the difference in modulation of IOR between the two orientations may lie in an altogether different source. In Experiment 6 two types of displays were used, upright and inverted, both of which were identical in terms of their left-right orientation. Despite this modification, once again internal object properties significantly modulated object-based IOR only in the Upright orientation. This finding suggested that the difference in within-object modulation of IOR between two different orientations in Experiment 5 was not necessarily due to the leftward or rightward orientation of the objects. Instead, in the case of Experiment 6 the difference in the two orientations may reflect the level of object shape processing prior to selection. Thus, if object-based selection occurs at the higher level of object processing, i.e. after categorisation, then the two displays, upright and inverted, may be attended differently, on the grounds that upright objects are more probable and real, whilst the inverted objects are physically improbable or unbelievable. If inverted objects are not perceived as real or probable objects then their representations may not modulate object-based attentional mechanisms in the same way as representations of real or physically possible objects.

As mentioned above, the difference between the right- and left-side objects in the primitives that modulate object-based IOR does not in iteself explain (especially in the case of Experiment 5) why internal-object structure *significantly* modulates object-based IOR in only one of the two possible orientations. In other words, if the left vs right object hypothesis is correct then one would expect significant IOR modulation by surfaces in the left-side object *and* by VCs in the right-side object. However, this was not the case. It is, therefore, proposed that object-based inhibition has access to a representation of objects at different hierarchical levels of object structure, that is, both in terms of their *local structure description* and in terms of their *global form description*, consistent with the local structure hypothesis. On speculative grounds, presentation of the same objects in different orientations allows these two forms of selection to operate simultaneously on different object orientations. This is particularly true on the basis of Experiment 5 results and is consistent with previous evidence suggesting that the two levels, global and local, can affect performance, i.e. speeding of response times, at approximately the same time (i.e. Hoffman, 1980; Miller, 1981).

One other possibility is that the two types of information, global and local, emerge at different time courses. Previous evidence (Navon, 1977) suggests that attention is biased towards global object properties and has documented the advantage of global over local information, known as *global precedence*. Indeed, conflicting information between global and local properties (i.e. a large letter H made of many small 's' shapes) interferes with the identification of the local property (letter 's') but not with the identification of the global one (letter H). Navon's results suggested that perceptual processes proceed from global to local hierarchical level. Furthermore, other evidence suggests (e.g. Lamb & Yund, 1996; Sergent, 1982; Sergent & Hellige, 1986) that the processing of global stimuli is facilitated by the activation of low spatial frequency visual channels, whilst the processing of local stimuli is facilitated by the activation of high spatial frequency channels. Low spatial frequency channels (global information) have a temporal advantage over high spatial frequency channels (local information; e.g. Breitmeyer, 1975, c.f. Lamb & Yund, 1996), which may explain Navon's (1977) global precedence hypothesis.

Thus, one line of future work may be to pursue the possibility that at different cue-target onset asynchronies emerges a different hierarchical information, global or local. In the global precedence hypothesis is correct, then modulation of IOR by internal object properties only emerges at longer SOA. Future work, perhaps systematically controlling display orientation and cue-target onset asynchronies, is needed to identify the factor(s) that determine whether an object is processed in terms of its global form or in terms of its local structure.

5.7. Epilogue

In conclusion, the present data have shown that internal features of objects modulate object-based inhibition of return. This result provides a constraint on hypotheses about the nature of attentional selection, and suggests that object-based IOR can operate over representations that make explicit information about internal shape structure. The pattern of such modulation, that is larger object-based IOR for uncued components (surfaces or VCs) compared with the cued components, may be the result of coexisting effect of facilitatory and inhibitory components of attention. This observation may provide an important theoretical constraint on hypotheses about the distribution of facilitation and inhibition across shape representations in object-based selection. Another important theoretical implication of the pattern of object-based IOR modulation is that object-based attention and part-based attention may not be mutually exclusive, but rather co-exist and influence performance simultaneously. This point is reinforced by the finding that for the same objects part-based effects are observed in some display orientations but not in others.

Furthermore, and in relation to the nature of these object-shape representations, this thesis provided evidence for a representation that consistently makes explicit surface characteristics. On the other hand, it was shown that factors other than the object's internal geometry, such as the number of VCs that comprise the object, may result in the creation of a representation that makes explicit information about VCs. At the moment it seems that either of these representations can modulate object-based IOR depending on the type of object attention is oriented towards. It is proposed that object-based IOR operates over internally structured representations of objects that make explicit information about the object's surfaces.

References

Abrams, R.A. & Dobkin, R.S. (1994). Inhibition of return: Effects of attentional cueing on eyemovement latencies. *Journal of Experimental Psychology: Human Perception & Performance, 20*, 467-477.

Allport, A. (1993). Attention and control: Have we been asking the wrong questions? A critical review of twenty-five years. In D.E. Meyer &S. Kornblum (Eds.). (1993): *Attention and performance XIV: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience* (pp. 183-218). Cambridge, MA, US: The MIT Press.

Avrahami, J. (1999). Objects of attention, objects of perception. *Perception & Psychophysics, 61*, 1604-1612.

Backer - Cave, C. & Kosslyn, S.M. (1993). The role of parts and their spatial relations in object identification. *Perception, 22*, 229-248.

Baylis, G.C., Driver, J., Baylis, L. & Rafal, R.D. (1994) Reading of letters and words in a patient with Balint's syndrome. *Neuropsychologia, 32*, 1273-1286.

Baylis, G.C. & Driver, J. (1992). Visual parsing and response competition: The effect of grouping factors. *Perception & Psychophysics, 51*, 145-162.

Baylis, G.C. & Driver, J. (1993).Visual attention and objects: evidence for hierarchical coding of location. *Journal of Experimental Psychology: Human Perception & Performance, 19*, 451-470.

Behrmann, M. & Tipper, S.P. (1994). Object-based attentional mechanisms: Evidence from patients with unilateral neglect. In C. Umilta & M. Moscovitch (Eds) *Attention & Performance XV: Conscious and nonconscious information processing* (pp. 351-375). Cambridge MA: MIT Press.

Biederman, I. (1987). Recognition by components: A theory of human image understanding. *Psychological Review, 94*, 65-96.

Biederman, I. & Cooper, E.E. (1991). Priming contour-deleted images: Evidence for intermediate representations in visual object recognition. *Cognitive Psychology, 23*, 393-419.

Biederman, I. & Ju, G. (1988). Surface versus edge-based determinants of visual recognition. *Cognitive Psychology, 20*, 38-64

Binford, T.O. (1971). Visual perception by computer. *IEEE Conference on Systems and Control*, Miami.

Bisiach, E. (1993). Mental representation in unilateral neglect and related disorders: the twentieth Bartlett Memorial lecture. *Quarterly Journal of Experimental Psychology, 46A*, 435-462.

Bisiach, E. & Vallar, C. (1988). Hemineglect in humans. In F. Boller & J. Grafman (Eds.) *Handbook of Neuropsychology: Volume 1.* Amsterdam: Elsevier.

Bower, G.H. & Glass, A.L. (1976). Structural units and the reintegrative power of picture fragments. *Journal of Experimental Psychology: Human Learning & Memory, 2*, 456- 466.

Cave, K.R. & Bichot, N.P. (1999). Visuospatial attention: beyond a spotlight model. *Psychonomic Bulletin & Review, 6*, 204-223.

Cave, K.R. & Pashler, H. (1995). Visual selection mediated by location: selecting successive visual objects. *Perception & Psychophysics, 57*, 421-432.

Cherry, E.C. (1953). Some experiments on the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America, 25*, 975-979.

Cohen J.D., MacWhinney B., Flatt M., & Provost J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, & Computers, 25(2)*, 257-271.

Collie, A., Maruff, P., Yucel, M., Danckert, J & Currie, J. (2000). Spatio-temporal distribution of facilitation and inhibition of return arising from the reflexive orienting of covertattention. *Journal of Experimental Psychology, Human Perception & Performance, 26*, 1733-1745.

Daney, C.P. & Reidy, J. (1999). *Statistics without Maths for Psychology.* Pearson Education Ltd.

Delis, D.C., Robertson, L.C. & Efron, R. (1986). Hemispheric spacialization of memory for visual hierarchical stimuli. *Neuropsychologia, 24*, 205-214.

Driver, J. & Baylis, G.C. (1989). Movement and visual attention: The spotlight metaphor breaks down. *Journal of Experimental Psychology: Human Perception & Performance, 15*, 448-456.

Driver, J. & Baylis, G.C. (1995). One-sided edge assignment in vision. II. Part decomposition, shape description, and attention to objects. *Current Directions in Psychological Science, 4*, 201-206.

Driver, J. & Baylis, G.C. (1998). Attention and visual object segmentation. In R. Parasuraman (Ed.) *The attentive brain.* Cambridge MA: MIT Press.

Driver, J., Baylis, G.C., Goodrich, S.J. & Rafal, R. D. (1994). Axis-based neglect of visual shapes. *Neuropsychologia, 32*, 1353-1365.

Driver, J. & Halligan, P.W. (1991) Can visual neglect operate in object-centred co-ordinates: An affirmative study. *Cognitive Neuropsychology, 8*, 475-496.

Driver, J., Mattingley, J. B., Rorden, C. & Davis, C. (1997). Extinction as a paradigm measure of attentional bias and restricted capacity following brain injury. In P. Thier & H – O. Karnath (Eds.) *Parietal Lobe Contributions to Orientation in 3D Space*, 401-429.

Duncan, J. (1984) Selective attention and the organisation of visual information. *Journal of Experimental Psychology: General, 113*, 272-300.

Duncan, J. (1993). Similarity between concurrent visual discriminations: Dimensions and objects. *Perception & Psychophysics, 54*, 425-430.

Duncan, J., Humphreys, G. & Ward, R. (1997) Competitive brain activity in visual attention. *Current Opinion in Neurobiology, 7*, 255-261

Edelman, S. & Bulthoff, H.H. (1992). Orientation dependence in the recognition of familiar and novel views of three-dimensional objects. *Vision Research, 32*, 2385-2400.

Egly, R., Driver, J. & Rafal, R.D. (1994). Shifting visual attention between objects and locations: Evidence from normal and parietal lesion subjects. *Journal of Experimental Psychology: General, 123 (2)*, 161-177.

Eriksen, C.W. & Eriksen, B.A. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics, 16*, 143-149.

Eriksen, C.W. & Hoffman, J.E. (1972). Temporal and spatial characteristics of selective encoding of visual displays. *Perception & Psychophysics, 12*, 201-204.

Eriksen, C.W. & St James, J.D. (1986). Visual attention within and around the field of focal attention: a zoom lens model. *Perception & Psychophysics, 40*, 225-240.

Farah, M.J. (1990) *Visual agnosia: Disorders of object recognition and what they tell us about normal vision*. Cambridge, MA: MIT Press.

Farah, M.J., Wallace, M.A., Brunn, J.L. & Madigan, N. (1989) Structure of objects in central vision affects the distribution of visual attention in neglect. *Society for neuroscience Abstracts, 15*, 481.

Friedman-Hill, S.R., Robertson, L.C. & Treisman, A. (1995) Parietal contributions to visual feature binding: Evidence from a patient with bilateral lesions. *Science, 269*, 853-855

Fuentes, L.J., Humphreys, G.W., Agis, I.F., Carmona, E. & Catena, A. (1998). Object-based perceptual grouping affects negative priming. *Journal of Experimental Psychology: Human Perception & Performance, 24*, 664-672.

Fox, E. (1995). Negative priming from ignored distractors in visual selection: A review. *Psychonomic Bulletin & Review, 2*, 145-173.

Gibson, J.J. (1979). *The ecological approach to visual perception.* Boston: Houghton, Mifflin.

Gibson, B.S & Egeth, H. (1994). Inhibition of return to object-based and environment-based locations. *Perception & Psychophysics, 55*, 323-339.

Goldsmith, M. (1998). What's in a location? Comparing object-based and space-based models of feature integration in visual search. *Journal of Experimental Psychology: General, 127*, 189-219

Goodale, M.A. (1995). The cortical organization of visual perception and visuomotor control. In M. Kosslyn & D.N. Osherson (Eds), *An Invitation to Cognitive Science: Visual Cognition*, MIT Press, (p. 167-213).

Halligan, P.W., Marshall, J.C. & Wade, D.T. (1990). Do visual field deficits exacerbate visuo-spatial neglect? *Journal of Neurology, Neurosurgery & Psychiatry, 53*, 487-491.

He, Z.J. & Nakayama, K. (1992). Surfaces vs features in visual search. *Nature, 359,* 231-233.

He, Z.J. & Nakayama, K. (1994). Perceived surface shape not features determines correspondence strenth in apparent motion. *Vision Research, 34,* 2125-2135.

He, Z.J. & Nakayama, K. (1995). Visual attention to surfaces in 3-D space. *Proceedings of the National Academy of Sciences USA, 92,* 11155-11159.

Hoffman, J. (1980). Interaction between global and local levels of a form. *Journal of Experimental Psychology: Human Perception & Performance, 6,* 222-234.

Hoffman, J.E. & Nelson, B. (1981). Spatial selectivity in visual serach. *Perception & Psychophysics, 30,* 283-290.

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

Houghton, G. & Tipper, S.P. (1994). A model of inhibitory mechanisms in selective attention. In D. Dagenbach & T. Carr (Eds.), Inhibitory Processes in attention, memory and language (p. 53-112). San Diego, CA: Academic Press.

Humphrey, G.K, Goodale, M.A., Jakobson, L.S. & Servos, P. (1994). The role of surface information in object recognition: studies of a visual form agnosic and normal subjects. *Perception, 23,* 1457-1481.

Humphreys, G.W. & Riddoch, M.J. (1993). Interactions between object and space systems revealed through neuropsychology. In D.E. Meyer & S. Kornblum (Eds.) *Attention & Performance XIV* (p. 143-162). Hillsdale, NJ: Erlbaum.

Humphreys, G.W., & Riddoch, J.M. (1994). Attention to Within-object and Between-object Spatial Representations: Multiple Sites for Visual Selection. *Cognitive Neuropsychology, 11(2),* 207-241.

Jordan, H & Tipper, S.P. (1998). Object-based inhibition of return in static displays.

Psychonomic Bulletin & Review, 5, 504-509.

Jordan, H. & Tipper, S.P. (1999). Spread of inhibition across an object's surface. *British Journal of Psychology, 90*, 495-507.

Julesz, B. (1981). Textons, the elements of texture perception, and their interaction. *Nature, 290*, 91-97.

Kahneman, D. & Henik, A. (1981). Perceptual organisation and attention. In M. Kubovy & J. R. Pomerantz (Eds.). *Perceptual organisation.* Hillsdale, NJ: Erlbaum.

Kahneman, D., Treisman, A. & Gibbs, B.J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology, 24*, 175-219.

Klein, R. (1988). Inhibitory tagging system facilitates visual search. *Nature, 334*, 430-431.

Kosslyn, S.M., Flynn, R.A., Amstredam, J.B. & Wang, G. (1990). Components of high-level vision: A cognitive neuroscience analysis and accounts of neurological symptoms. *Cognition, 34*, 203-277.

Kramer, A.F. & Jacobson, A. (1991). Perceptual organisation and focused attention: The role of objects and proximity in visual processing. *Perception & Psychophysics, 50*, 267-284.

Lamb, M.R. & Yund, E.W. (1996). Spatial frequency and interference between global and local levels of structure. *Visual Cognition, 3*, 193-219.

Lamy, D. & Tsal, Y. (2000). Object features, object locations, and object files: which does selective attention activate and when? *Journal of Experimental Psychology: Human Perception & Performance, 26*, 1387-1400.

Lavie, N. & Driver, J. (1996). On the spatial extent of attention in object-based visual selection. *Perception & Psychophysics, 58(8)*, 1238-1251.

Leek, E.C. & Arguin, M. (2000). Surface-based shape descriptions in object recognition: Preliminary evidence from part-whole matching and priming. *41st Annual Meeting of the Psychonomic Society,* New Orleans.

Logan, G. (1997). The CODE theory of visual attention: An integration of space-based and object-based attention. *Psychological Review, 103*, 603-649.

Lowe, D.G. (1985). *Perceptual organisation and visual recognition.* Boston, MA:Kluwer-Nijhoff.

Luria, A.R. (1959). Disorders of 'simultaneous perception' in a case of bilateral occipito-parietal brain injury. *Brain, 83*, 437-449.

Mari-Beffa, P., Houghton, G., Estevez, A.F. & Fuentes, L.J. (2000). Word-based grouping affects the prime-task effect on semantic priming. *Journal of Experimental Psychology: Human Perception & Performance, 26(2)*, 469-479.

Marr, D. (1982). *Vision.* New York: Freeman.

Marr, D. & Nishihara, H.K. (1978). Representation and recognition of the spatial organisation of three dimensional shapes. *Proceedings of the Royal Society of London, 207*, 187-217.

Maylor, E.A. & Hockey, R. (1985). Inhibitory component of externally controlled covertorienting in visual space. *Journal of Experimental Psychology: Human Perception & Performance, 11*, 177-187.

McAuliffe, J., Pratt, J. & O' Donnell, C. (2001). Examining location-based and object-based components of inhibition of return in statis displays. *Perception & Psychophysics, 63*, 1072-1082

McCloskey, M. (2001). Spatial representation in mind and brain. In B. Rapp (Ed.) *The*

Handbook of Cognitive Neuropsychology: What Deficits Reveal About the Human Mind, (p. 101-132). Psychology Press.

Miller, J. (1981). Global precedence in attention and decision. *Journal of Experimental Psychology: Human Perception & Performance, 7*, 1161-1174

Milner, A.D. & Goodale, M.A. (1995). *The visual brain in action.* Oxford University Press, Oxford.

Moore, C., Yantis, S. & Vaughan, B. (1998). Object-based visual selection: Evidence from perceptual completion. *Psychological Science, 9,* 104-110.

Moran, J. & Desimone, R. (1985). Selective attention gates visual processing in the extrastiate cortex. *Science, 229*, 782-784.

Mozer, M.C. (1999). A principle for unsupervised hierarchical decomposition of visual scenes. In M.S. Kearns, S.A. Solla, & D. Cohn (Eds.), *Advances in neural information processing systems II.* Cambridge, MA: MIT Press.

Muller, H.J. & Muhlenen, A. (1996). Attentional tracking and inhibition of return in dynamic displays. *Perception & Psychophysics, 58*, 224-249.

Nakayama, K. & He, Z.J. (1995). Attention to surfaces: Beyond a Cartesian Understanding of Focal Attention. In T.V. Papathomas, C. Chubb, A. Gorea & E. Kowler (Eds), *Early Vision & Beyond.* MIT Press.

Nakayama, K., He, Z.J. & Shimojo, S. (1995). Visual surface representation: a critical link between lower-level and higher-level vision. In S.M. Kosslyn & D. Osherson (Eds.), Visual Cognition: Volume 2. *An invitation to cognitive science.* Cambridge, MA: MIT Press

Nakayama, K & Shimojo, S. (1992). Experiencing and perceiving visual surfaces. *Science, 257*, 1357-1363.

Navon, D. (1977). Forest before the trees: The precedence of global features in visual perception. *Cognitive Psychology, 9,* 353-383.

Nissen, M.J., (1985). Accessing features and objects: Is location special? In M.I. Posner and O.S.M. Marin (Eds), *Attention & Performance XI* (pp. 205-219). Hillsdale, NJ: Erlbaum.

Neisser, U. (1979). The control of information pickup in selective looking. In A. Pick (Ed.), *Perception and its development.* Hillsdale, NJ: Erlbaum.

Neisser, U. & Becklen, R. (1975). Selective looking: attending to visually specified events. *Cognitive Psychology, 7,* 480-494.

O' Craven, K., Downing, P. & Kanwisher, N. (1999). fMRI evidence for objects as the units of attentional selection. *Nature, 401,* 584-587.

Palmer, S.E. (1975). Structural aspects of perceptual organisation. *Dissertation Abstracts International, Volume 35 (11-B),* p.5672.

Palmer, S.E. (1977). Hierarchical structure in perceptual representation. *Cognitive Psychology, 9,* 441-474.

Palmer, S.E. (1999). *Vision Science: Photons to phenomenology.* MIT Press

Pashler, H. (1998). *Attention.* Psychology Press Ltd.

Pentland, A. (1989). Shape information from shading: A theory about human perception. *Spatial Vision, 4,* 165-182.

Peterson, M.A. & Gibson, B.S. (1991). Directing spatial attention within an object: Alerting the functional equivalence of shape descriptions. *Journal of Experimental Psychology: Human Perception & Performance, 17,* 170-182.

Phillips, J.C. & Ward, R. (in press). S-R correspondence effects of irrelevant visual affordance: Time course and specificity of response activation. *Visual Cognition, 9.*

Posner, M. & Cohen, (1984). Components of visual orienting. In H. Bouma & D. Bouwhuis (Eds) *Attention & Performance X* (pp. 531-556). Hillsdale, NJ: Erlbaum.

Posner, M..I. & Peterson, S.E. (1990). The attention system of the human brain. *Annual Review of Neuroscience, 13*, 25-42.

Posner, M.I, Walker, J.A., Friedrich, F.J. & Rafal, R.D. (1984). Effects of parietal injury on covert orienting of attention. *Journal of Neuroscience, 4*, 1863-1874.

Pratt, J., Spalek, T.M. & Bradshaw, F. (1999). The time to detect targets at inhibited and non-inhibited locations: Preliminary evidence for attentional momentum. *Journal of Experimental Psychology: Human Perception & Performance, 25*, 730-746.

Rafal, R. (1997). Balint syndrome. In T. Feinberg & M. Farah (Eds.), *Behavioural neurology & Neuropsychology*. New York: McGraw-Hill.

Rafal, R. (1998). Visual attention: Converging Operations from neurology and psychology. In A.F. Kramer, M.G.H. Coles and G.D. Logan (Eds) *Converging Operations in Visual Selective Attention*. American Psychological Association.

Reed, S.K. & Johnsen, J.A. (1975). Detection of parts in patterns and images. *Memory & Cognition, 3*, 569-575.

Robertson, L.C., Treisman, A., Friedman-Hill, S. & Grabowecky, M. (1997). The interaction of spatial and object pathways: evidence from Balint syndrome. *Journal of Cognitive Neuroscience, 9*, 295-317.

Robertson, L.C. & Lamb, M.R. (1988). The role of perceptual reference frames in visual field asymmerties. *Neuropsychologia, 26*, 145-152.

Scholl, B.J. (2001). Objects and attention: The state of the art. *Cognition, 80*, 1-46.

Scholl, B.J., Pylyshyn, Z.W. & Feldman, J. (2001) What is a visual object? Evidence from target merging in multiple object tracking. *Cognition, 80*, 159-177.

Sergent, J. (1982). The cerebral balance of power: Confrotation or cooperation? *Journal ofExperimental Psychology: Human Perception & Performance, 8*, 253-272.

Sergent, J. & Hellige, J.B. (1986). Role of input factors in visual-field asymmetries. *Brain & Cognition, 5*, 174-199.

Stroop, J.R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643-662

Tarr, M.J. & Bulthoff, H.H. (1995). Is human object recognition better described by geon-structural-descriptions or by multiple-views? Comment on Biederman & Gerhardstein (1993). *Journal of Experimental Psychology: Human Perception & Performance, 21*, 1494-1505.

Tarr, M.J. & Pinker, S. (1989). Mental rotation and orientation-dependence in shape recognition. *Cognitive Psychology, 21*, 233-282.

Tassinari, G, Aglioti, S., Chelazzi, L., Peru, A., & Berlucchi, G. (1994). Do peripheral non-informative cues induce early facilitation of target detection? Vision Research, 34, 55-72.

Tipper, S.P. (1985). The negative priming effect: Inhibitory effects of ignored primes. *Quarterly Journal of Experimental Psychology, 37A*, 571-590.

Tipper, S.P. & Behrmann, M. (1996). Object-centred not scene-based visual neglect.

Journal of Experimental Psychology: Human Perception & Performance, 22, 1261-1278.

Tipper, S.P., Driver, J. & Weaver, B. (1991). Short report: Object-centered inhibition of return of visual attention. *Quarterly Journal of Experimental Psychology, 43A*, 289-298.

Tipper, S.P., Jordan, H. & Weaver, B. (1999). Scene-based and object-centered inhibition of return: Evidence for dual orienting mechanisms. *Perception & Psychophysics, 61:1*, 50-60

Tipper, S.P., Weaver, B., Jerreat, L.M. & Burak, A.L. (1994). Object-based and environment-based inhibition of return of visual attention. *Journal of Experimental Psychology: Human Perception & Performance, 20 (3)*, 478-499.

Tipper, S.P., Rafal, R.D, Reuter-Lorenz, P., Starrveldt, Y., Ro, T., Egly, R., Danzinger, S & Weaver, B. (1997). Object facilitation and inhibition from orienting in the human split-brain. Journal of Experimental Psychology: Human Perception and Performance, 23, 1522-1532.

Treisman, A. (1988). Features and objects: the Fourteenth memorial Bartlett lectures. *Quarterly Journal of Experimental Psychology, 40*, 201-237.

Treisman, A. (1999). Feature binding, attention and object perception. In G.W. Humphreys, J. Duncan & A. Treisman (Eds.) *Attention, Space and Action: Studies in Cognitive Neuroscience*. Oxford University Press.

Treisman, A & Gelade, G (1980). A feature integration theory of attention. *Cognitive Psychology, 12*, 97-136.

Tsal, Y. & Lavie, N. (1988). Attending to colour and shape: The special role of location in selective visual processing. *Perception & Psychophysics, 44*, 15-21.

Tucker, M. & Ellis, R. (1998). On the relations between seen objects and components of potential actions. *Journal of Experimental Psychology: Human Perception & Performance, 24*, 830-846

Ullman, S. (1989). Aligning pictorial descriptions: An approach to object recognition. *Cognition, 32*, 193-254.

Umilta, C. (2001). Mechanisms of attention. In B. Rapp (Ed.) *The Handbook of Cognitive Neuropsychology: What Deficits Reveal About the Human Mind*, (p. 135-158). Psychology Press.

Van der Heijden, A.H.C., Kurvink, A.G., de Lange, L., de Leeuw, F & van der Geest, J.N. (1996). Attending to color with proper fixation. *Perception & Psychophysics, 58*, 1224-1237.

Vecera, S.P. (1994). Grouped locations and object-based attention: comment on Egly, Driver & Rafal (1994). *Journal of Experimental Psychology: General, 123*, 316-320.

Vecera, S.P., Behrmann, M. & McGoldrick, J. (2000). Selective attention to the parts of an object. *Psychonomic Bulletin & Review, 7*, 301-308.

Vecera, S.P., Behrmann, M. & Filapek, J. (2001). Selective attention to the parts of an object. *Perception & Psychophysics, 63*, 308-321.

Vecera, S.P. & Farah, M.J. (1994) Does visual attention select objects or locations? *Journal of Experimental Psychology: General, 123*, 146-160.

Ward, R. & Goodrich, S.J. (1994). Differences between objects and non-objects in visual extinction: a competition for attention. *Psychological Science, 3*, 177-180.

Ward, R., Goodrich, S.J. & Driver, J. (1994). Grouping reducses visual extinction: neuropsychological evidence for weight-linkage in visual selection. *Visual Cognition, 1*, 101-129.

Watson, S.E. & Kramer, A.F. (1999). Object-based visual selective attention and perceptual organisation. *Perception & Psychophysics, 61*, 31-49.

Weaver, B., Lupianez, J. & Watson, F.L. (1998). The effects of practice on object-based, location-based and static-display inhibition of return. *Perception & Psychophysics, 60 (6)*, 993-1003.

Xu, Y. & Nakayama, K. (2001) Surface-based attention is separable from object-based attention. *Abstracts from Psychonomic Society, 42nd Annual meeting.*

Yantis, S. & Hillstrorm, A.P. (1994). Stimulus-driven attentional capture: Evidence from equiluminant visual objects. *Journal of Experimental Psychology: Human Perception & Performance, 20*, 95-107.