Perceptual organization and its influence upon attention

DE-WIT, LEE,HOWARD

DE-WIT, LEE,HOWARD (2010) *Perceptual organization and its influence upon attention*, Durham theses, Durham University. Available at Durham E-Theses Online: http://etheses.dur.ac.uk/103/

Perceptual organization and its influence upon attention

Lee de-Wit

Thesis submitted for the degree of Doctor of Philosophy

Durham University, Department of Psychology

2009

Declaration

I confirm that no part of the material offered has previously been submitted by me for a degree in this or in any other University. If material has been generated through joint work, my independent contribution has been clearly indicated. In all other cases material from the work of others has been acknowledged and quotations and paraphrases suitably indicated.

Copyright

Acknowledgement

I would like to thank Bob Kentridge and David Milner for their supervision over the last three years, in particular I am grateful for Bob's technical knowhow and ingenious ideas (and dislike of sentences starting with however) and David's professionalism and his ability to 'highlight' my sometimes somewhat imprecise writing style.

I would like to thank Geraint Rees and Catherine Tallon-Baudry for hosting me in their labs at various points during my PhD. I learnt an enormous amount at the ICN (including an unexpected appreciation for you-tube in Geraint's lab meetings), and started a fruitful collaboration with Ayse Sygin, which (also with some assistance from Carmen Lefèvre) led to the results reported in Chapter 7. My visit to Catherine Tallon-Baudry's lab in Paris led to the data reported in Chapter 4, for which I am also grateful to Claire Sergent and Antoine Decorps. Within Durham I would also like to thank Geoff Cole, who collaborated in the work reported in Chapter 3. I would also like to thank Thomas Schenk, both for many stimulating conversations, and for giving me the opportunity to help organise David Milner's Festschrift during my PhD.

I'm also particularly grateful to Bob Metcalfe, without whose programming tutorials and (Direct X) class, I would not have been able to complete a lot of the work reported in this thesis.

I also could not have finished this thesis without the many participants who have sat in the dark pressing buttons for me over the last 3 years, I am very grateful to all of them, but in particular I would like to thank Debbie whose time and effort provided the basis for the last 3 empirical chapters of this PhD.

Although we have not formally worked together I would also like to thank Dr CCP, who has gently (for an Italian) highlighted some of my weaknesses as a scientist during my PhD, I hope I have learnt from her advice.

I would also like to thank many friends and colleagues for stimulating conversations over the past few years, in particular Dr Marco Hirnstein and Dr Callum Durnford with whom I've shared offices, houses, labs and the process of doing a PhD. Finally I would like to thank my undergraduate tutor Tom Troscianko for encouraging me to pursue a career in science.

Abstract

Humans are able to control so much of their environment not through brute strength or enhanced sensory receptors, but through our ability to understand the world around us. In order to make sense of the world around us we need to organize the information that our sensory systems receive. One of the most fundamental steps in this organizational process lies in the construction of objects. By breaking down our sensory input into objects the mind provides a basis upon which it can begin to scaffold our understanding of the world.

This thesis therefore explores the basic stages at which the visual system organizes our sensory input into distinct objects. It explores these stages by exploiting the fact that the brain's limited processing resources can be selectively allocated on the basis of 'object-hood'. This allocation of processing resources, or attention, on the basis of these early stages of segmentation is commonly referred to as 'object based attention'. 'Object based attention' and perceptual organisation are explored in three sections in this thesis:

Understanding the Phenomenon of Object Based Attention.

The first three chapters of this thesis seeks to further our understanding of the phenomenon of 'object based attention', for example, chapter 3 explores whether the visual system can simultaneously parse several objects as potential units of attention, or whether it can only segment one or two objects at a time.

Object Based Attention, a Tool to Explore the Nature of Perceptual Organisation

The second section of this thesis attempts to use the phenomenon of 'object based attention' as a tool to explore the nature of perceptual representations, for example chapter 5 tests whether different modalities (in particular vision and touch) are able to directly share information about objects in order to build up an integrated model of the external world.

Object Based Attention, Perceptual Organisation and Shape Processing Area LO.

In the final section of this thesis the nature of perceptual organization is explored in a patient with a very specific form of brain damage that enables us to ask what areas of the brain are critically required for different aspects of perceptual organization.

Table of Contents

Chapter One - General Introduction:

Object Based Attention – An empirical handle on the philosopher's axe

This thesis seeks to understand how the early stages of visual information processing are organised, and more specifically how that organisation influences the later allocation of visual processing resources. Emmanuel Kant (1781) argued that in order for humans to understand their sensory experience they have to bring to that experience certain 'a-priori' assumptions regarding how the world is organised. Kant's philosophical argument regarding the logical necessity of 'a-priori' assumptions of 'objecthood' and causality bears an interesting parallel to the principles outlined by early European psychologists. For instance, whilst Kant argued that some notion of 'causality' had to be brought to our experience in order for it to be understood, Michotte (1963) highlighted that even highly impoverished displays of colliding shapes were sufficient for humans to automatically construct causal interpretations. Furthermore Kant's argument that some form of 'objecthood' needs to be brought to bear on sensory experience before it can be understood is consistent with the catalogue of principles outlined by Gestalt psychologists by which visual information is organised such that the individual parts of a scene are automatically interpreted to generate wholes or groups that are more than the sum of their parts (Koffka 1935). It is not clear exactly how these principles manifest, or whether they truly accord with the notions suggested by Kant, nevertheless their automatic influence upon perception adds weight to the basic notion that sensory experience has to be organised before it can be understood.

Kant's position regarding the necessity of certain a-priori assumptions was motivated by the argument that properties such as objecthood or causality did not exist as inherent features of the physical world. The challenge faced in trying to provide an adequate physical definition for an object is often brought to light with the thought experiment of the 'philosopher's axe'. In this thought experiment we are asked to imagine a lumber jack who, if asked, says he has used the same axe for the past five years. On further inspection however we learn that over the years he has had to replace different parts of the axe as they have worn down. In fact when forced the lumber jack admits that he has replaced so many parts of the axe that he is not sure whether it still contains any physical remnants of the axe he first started using. From a physical perspective there is no clear definition by which one can define the object as the same, yet psychologically we can entertain this possibility. In this sense the mind creates an entity that is more than the sum of its parts.

This notion that we can maintain an enduring, albeit, abstract representation of something as the same, even as its parts change, has a very direct parallel in modern cognitive science. Blaser, Holcombe and Pylyshyn (2000) demonstrated that on top of our ability to track an object as it moves through space, we can also maintain an independent representation of two overlapping objects at the same spatial location even as the individual features of those objects (their colour, orientation and spatial frequency) continually change. Thus just as the lumber jack can maintain that his axe is the same after having changed all of its parts, so too the early process of perceptual organisation can establish an enduring object tag that can be maintained despite changes to the individual features that make up that object.

These parallels between the nature of early perceptual organisation and higher level cognition, are just that, parallels. Indeed the organisation of early perception and the organisation of higher cognition undoubtedly involve different modes of information processing (Driver, Davis, Russell, Turatto & Freeman, 2001). It is possible however that the early stages of perceptual organisation provide a framework that enables the development of these higher forms of cognition. In fact Pylyshyn (1999) has argued exactly that, that the organisation, and in particular the ability to selectively attend to the units or proto-objects established by this early process of perceptual organisation, provide a scaffold upon which conceptual knowledge can develop. Indeed Carey and Xu (2001) argue that the 'mid-level' perceptual organisation studied in adulthood may in fact reflect exactly the same underlying resources responsible for the primitive understanding of 'objecthood' evident in infants from as young as 2 and a half months; from this early age therefore the process of perceptual organisation could play a role in shaping our higher level notions of object-hood.

Whilst undoubtedly attractive, the hypothetical role of early perceptual organisation in shaping higher level cognition is a somewhat intractable empirical problem. If we are to one day come to a position whereby we can test this role we first need to understand the nature of perceptual organisation per se. This thesis will therefore focus on one of the most basic levels at which perceptual organisation influences further cognition. That is in terms of its influence upon attention, a process commonly referred to as 'object based attention'. Before exploring the phenomenon of object based attention this thesis will first seek to clarify what is meant by attention.

Attention – One word, many processes

If we want to inspect some aspect of our visual environment, we look at it. That is to say we point our eyes, and focus our lens to ensure that light received from that aspect of our environment falls upon the densely pact receptors of the fovea. This is, however, not the only means by which we can enhance the processing of different aspects of our environment. Even in the absence of overt eye movements the visual system is able to bias the activations associated with different stimuli. This ability to alter the activations associated with different stimuli is one of the most studied aspects of human cognition.

This focus on the visual system's ability to selectively alter the activations associated with different processes has not gone without criticism. Findlay and Gilchrist (2001) have argued that in everyday tasks the subtle internal facilitation of information processing plays little (if any) role. Findlay and Gilchrist maintain that the primary means by which we select aspects of our visual environment for more detailed processing is via overt eye movements.

This debate, regarding the relative importance of overt eye-movements vs internal shifts in processing resources (or covert attention), is an important one. Focusing on the relative importance of overt vs covert changes in information processing can however lead to a neglect of a very important reason for studying purely internal shifts of processing resources. In order to appreciate this factor we need to step back and remember that progress in science depends upon experiments: experiments that offer insights into causal relationships by keeping certain things constant, whilst changing others. Often the causal relationships one wants to investigate are hard to observe directly. This is most certainly the case with the conglomerate of mechanisms that cause our behaviour. In this context, purely internal shifts in processing resources offer an excellent scientific tool, because one can maintain the input received by a participant and study changes in out-put that must have been caused by, and thus must be informative regarding, internal processing mechanisms.

This thesis will therefore focus on this internal process of covert selection, not out of any belief that it is more (or less) important than overt eye movements, but because of the insights it might offer into the nature of the causal mechanisms that sit between the stimuli we receive and the behaviour that results.

The common reference to the phenomena of 'covert attention' is perhaps misleading[1]. It is misleading because the succinct nature of the term implies that there is a simple unitary mechanism responsible for alterations in the resources allocated to different information processing mechanisms. Attention, however, should not be taken as synonymous with a unitary cognitive or neural structure, indeed rather than identifying 'attention' with a unitary cognitive or neural structure this thesis will employ a behavioural definition of attention, in terms of an improvement (or decrement) in the accuracy or reaction time with which participants process a selected (or inhibited) stimulus. This is definition should not be taken to imply that we have gained no insight into the underlying causes of this behavioural phenomena. Desimone and Duncan (1995) for instance have offered a frame-work in which attention can be understood in terms of a competition for representation. This competition for representation is influenced both by the nature of the stimulus, our experience with that stimulus in the past, and the participant's current task demands. The selective enhancement of visual information is, in this framework, not something that sits out-side the process of perceptual representation, but something that is integral to, and operates at many stages of this process.

The interrelated nature of perceptual representations and the allocation of processing resources is illustrated par excellence by the phenomenon of object based attention (Duncan 1984).

Object Based Attention – Theory or Phenomenon?

Early research emphasized the ability of the visual system to use the spatial location of a stimulus as the means by which to select that stimulus for more detailed information processing (Treisman and Gelade 1980, Posner 1980). However although spatial location is probably the most rapidly extracted feature by which the visual system can select stimuli for more detailed information processing (Martinez et al. 2006) it is clear that the visual system can also bias the allocation of processing resources on the basis of a number of other stimulus dimensions. Indeed it has become uncontroversial to say that the allocation of processing resources is not simply based on spatial location but that it can also be object based. In fact this form of selection is commonly referred to as 'object based attention' (Scholl 2001).

---

[1] Indeed, this could well be regarded in terms of a bias towards essentialism, a notion that is explained later in this introduction.

The term ‘object based attention’ is however a troublesome one, because it refers both to a large cluster of empirical phenomena and also implies a certain theoretical position. What underlies the commonality between the array of phenomena that come under the umbrella term ‘object based attention’ is the fact that the allocation or competition for more detailed information processing is determined not simply by the spatial arrangement of targets in the display, but the manner in which that display is organised or grouped into distinct ‘objects’. However the extent to which this cluster of empirical phenomena reflect the same underlying mechanisms has not been systematically explored. Moreover the extent to which some of these phenomena live up to the theoretical status implied by the term ‘object based attention’ has also been questioned (Driver, et al. 2001).

This thesis will largely focus on one example of ‘object based attention’ namely the Egly, Driver and Rafal (1994) object based cuing paradigm. As will become clear whilst this paradigm is one of the most cited and explored examples of ‘object based attention’ the extent to which this phenomenon involves a distinct form of non-spatial attention (Martinez et al. 2006) or indeed the extent to which it can meaningfully be described as ‘object’ based has come under question (Driver, et al. 2001). Before turning to investigate this phenomenon in more detail however we shall briefly review the range of paradigms that could be included under the umbrella of ‘object based attention’, first because some of them will become relevant at different times in this thesis, but secondly because they are very rarely (if ever) all outlined in one place.

The first example of ‘object based attention’ was provided by John Duncan in 1984. Duncan set out to understand how the limited information processing resources of the human visual system could select important aspects of the sensory input for more detailed processing. As noted above previous work had suggested that this process of selection occurred on a fundamentally spatial basis, whereby the visual system would enhance the activation for stimuli appearing at a given location in space (Posner 1980, Treisman & Gelade, 1980). Duncan tested the explanatory power of this spatial model of selection by presenting two overlapping objects at the same location. On each trial the shape of these objects would vary in different ways and participants were asked to report either one shape feature from each object, or two shape features of the same object. Even though these two objects were spatially overlapping participants could more efficiently report two features of the same object. Thus judging one part of an object attracted processing resources in a manner that rendered other aspects of that object more readily available to the visual system. The fact that these objects were overlapping meant that this allocation of processing resources could not have been spatially based, because such a spatial allocation would have equally enhanced the features of both objects.

Duncan's work launched one of those classic dualistic debates in cognitive science (Newell 1973), regarding whether the allocation of processing resources (or attention) was based on a spatial or object based selection mechanism. Egly, Driver and Rafal (1994) sought to draw some resolution to this debate by demonstrating both spatial and object based selection within the same paradigm. In this paradigm, which forms the basis for most of the experiments in this thesis, Egly et al. revealed that when one end of one object is cued, not only do targets appearing at that location receive an attentional enhancement, targets appearing elsewhere on the same object also receive an advantage, relative to equidistant targets appearing on another object. This finding demonstrated that processing resources were preferentially allocated both to the location of the cue, and the cued object. This thesis will focus upon this phenomenon because it is one of the most commonly cited and explored examples of object based attention. In addition to the 352 citations it has received over the last 15 years it has also been used as an experimental paradigm in 46 publications, a list of these replications/extensions of this paradigm is provided in Appendix A.

Whilst the Duncan and Egly et al. paradigms are two of the most well-known pieces of evidence for the influence of objects in attention, there are a wide variety of other phenomena that could also come under the heading of object based attention. The Egly et al. paradigm depends upon the attraction of processing resources following a cue; it has become clear however that the extent to which a 'cue' is effective in attracting attention can vary depending upon the extent to which that cue is interpreted as a novel object, or a change to an existing object (Cole, Kentridge & Heywood, 2004). Furthermore the presentation of a cue not only leads to a facilitation, or attraction of processing resources, but also leads to some form of inhibition over longer time scales (Klein 1988). This phenomenon, known as Inhibition of Return, selects (or more accurately de-selects) not only the spatial location at which the cue was presented but will also manifest on a cued object, even if that object moves to a new location (Tipper, Driver, Weaver 1999). Therefore the effectiveness of visual cues, and their role in the facilitation and later inhibition of attention can be said to be influenced by objects.

The role of objects in attention is however by no means limited to ‘cuing’ paradigms. The classic cognitive phenomena of ‘flanker interference’ (Erikson and Erikson 1974), where participants take longer to recognise a target surrounded by incongruent distracters, is also sensitive to object structure, such that the extent of interference is enhanced or reduced depending on whether or not the target and distracters form a perceptual group (Driver & Baylis 1999). This suggests that the extent of interference caused by distracters is influenced by the amount of attention they receive and that the amount of attention they receive is contingent upon the manner in which they group with the task relevant target. Furthermore directly comparing two visual items, also reveals an object based effect, such that objects on the same object are more easily categorised as the same or different, in contrast to two items appearing on different objects (Watson & Kramer, 1999).

In the domain of binocular rivalry, in which incongruous images are presented to each eye such that only one of them can become conscious, the competition for conscious representation is also influenced by objecthood (Mitchell, Holcome, Reynolds, Maruya & Blake 2004). In patient neuropsychology there are also two syndromes that could be included under the umbrella of object based attention. The neuropsychological syndrome of Neglect has been argued to reflect a disruption to the normal allocation of attention, in which, despite intact low level visual information processing, the representations developed on one side of space do not reach a sufficient representational status such that they can influence explicit behaviour. Driver, Baylis and Rafal (1992) have shown that in some instances the presence of objects in the visual scene defines what side of visual space is ignored. The related neuropsychological disorder known as Extinction has also been shown to be influenced by objects, such that although these patients can only become aware of one ‘thing’ at a time, previously unrecognised aspects of the visual input can be recognised if they are linked to form a single object or group with an attended item (Humphreys, Romani, Olson, Riddoch & Duncan, 1994).

The tracking of moving targets is also clearly influenced by the manner in which those targets group to form different objects. Indeed Scholl, Pylyshyn and Feldman (2001) have argued that the role of object-hood in visual tracking is so clear that this phenomenon can be used as a tool to answer the question ‘what is an object?’ The Multiple Object Tracking (MOT) paradigm was developed by Pylyshyn and colleagues (Pylyshyn & Storm, 1988) to illustrate the visual system’s capacity to simultaneously index multiple visual objects (or ‘proto-objects'). Pylyshyn (1999) hypothesised that a phenomenon like MOT must be possible in order to provide a scaffold upon which a higher level representation and conceptualisation of those indexed proto-objects could develop. Pylyshyn's notion of an indexed proto-object is closely related to the notion of 'object-files', which are slots or tags, in which information regarding a given aspect of the environment can be built up. Kahneman, Treisman and Gibbs (1992) argue that the existence of object files is supported by the 'Object Specific Preview Benefit', in which briefly presented 'pre-view' information has a larger facilitatory effect on later target recognition when that preview information is presented on the same object as the target, even as those objects moved to new locations in space.

These slightly more nuanced notions of 'object files' or 'proto objects' introduced in the last two examples of 'object based attention' opens the somewhat thorny question regarding the exact nature of the underlying units of selection in the various 'object based' paradigms outline above. As noted in the context of previous philosophical work, objects do not form an inherent part of our environment, objects emerge as a result of an interaction between the sensory data we receive and the processes that our central nervous system applies to that data (for a more recent articulation of this position see Feldman, 2003). The hierarchical nature of visual information processing forces us to question at exactly what level of visual representation the units of attentional selection are extracted. One means of framing this hierarchy is in terms of an initially distinct or modular processing of different stimulus features, such as colour, shape, texture and motion, which becomes integrated into a united percept including all of these attributes. Framed in this way the phenomena above focus on selection via one particular attribute of an object, that is its shape.

Indeed one could argue that rather than selecting unified objects, attention itself actually leads to the integration of the various features making up an object. Treisman and Gelade proposed just such a role for spatial attention. More recent functional imaging has however revealed that when two objects are presented at an overlapping spatial location, the selection of one feature of an object automatically leads to the activation of the other features of that object even if they are not task relevant (O'Craven, Downing & Kanwisher 1999). An exploration of the time course of this effect suggests that attention initially selects one feature of an object, and then rapidly spreads to all other features of that object (Schoenfeld et al. 2003). One could therefore argue that the cluster of 'object based attention' phenomena listed above do not quite live up to the theoretical status implied by the term ‘object’ based attention because they are really based on the selection of one feature of an object.

However whilst the cluster of phenomena outlined above seem to revolve around the selection of the shape feature of an object, any attempt to pin down the exact nature of the shape features that influence attention in general is hampered by the lack of direct comparison across these different phenomena. Indeed to our knowledge there are no attempts in the literature to directly address whether the representations that influence the allocation of processing resources in one paradigm are the same as those that influence attention in another. Not only is there an absence of positive evidence regarding the extent to which these different phenomena reflect the same underlying process, there are some striking differences between the limitations seen in different paradigms. For example in the multiple object tracking paradigm participants are able to track somewhere in the range of four items, whilst in the Egly et al. paradigm participants seem to be able to selectively enhance representations on only one object.

Chapter 3 will attempt to provide some reconciliation of this discrepancy by exploring the number of objects that can simultaneously be parsed as units of selection in the Egly et al. paradigm. In general however this thesis will not attempt to resolve the relationships within the range of phenomena coming under the umbrella term ‘object based attention’. Rather this thesis will focus (although not exclusively) on one particular phenomenon, the Egly et al. within-vs-between-object cuing advantage. The thesis will for example attempt to use this phenomenon as a tool to ask questions about the representations required to guide attention in patient neuropsychology (Chap 7), the nature of cross modal interactions between the shape representations in vision and touch (Chap 5) and the representational status of shadows in the human visual system (Chap 6). This phenomenon was selected because it has been cited, replicated, extended and explored extensively in the 15 years since its publication. Indeed the extensive literature on this phenomenon brings into focus some important questions regarding the nature of ‘object based attention’ in general. As noted above, principal among these issues is the nature or level at which ‘object’ representations can influence attention. Before addressing this issue however this introduction will describe the paradigm in more detail and then focus on the question of whether within-vs-between-object advantage in this paradigm truly reflects a distinct form of object based attentional selection, or whether it can be understood in terms of a fundamentally spatial mechanism that is heavily influenced by object structure.

Object Based Attention and the Egly et al. Within-vs-Between-Object Advantage

- *The Basic Effect*

The original paper by Egly, Driver and Rafal (1994) has been cited over 350 times, and within those citations 46 actively use this paradigm either to further advance our understanding of this phenomenon, or to use it to ask some broader questions about the nature of attention and object representations. Cataloguing the 46 papers that replicate the paradigm forces one to consider exactly what features are critical to this paradigm. In the original version Egly et al. presented two rectangles as the objects, and cued one of the 4 potential corner locations on those rectangles, and then asked the participant to perform a target detection task. Some of the properties of this original task are however more important than others, for example the exact shapes or objects used (rectangles or wrenches) the type of the cue (exogenous or endogenous) and the type of task (detection or recognition) are not essential. What is critical however is that the participant has a set of equally spaced target locations that either lie within the same 'object' or on different 'objects' and that after some form of cue to one of these locations the participant should be presented with some form of detection or discrimination task, either in the same location as the cue, the same object as the cue, or in another object. What is critical to the results of this paradigm is that one finds that targets paired with cues within the same 'object' are responded to more rapidly or more accurately when compared with equally spaced cue target pairings that cross between different objects. Unless otherwise stated, all the object based attention effects described below employ a paradigm that conforms to this definition of the Egly et al. within-vs-between-object task.

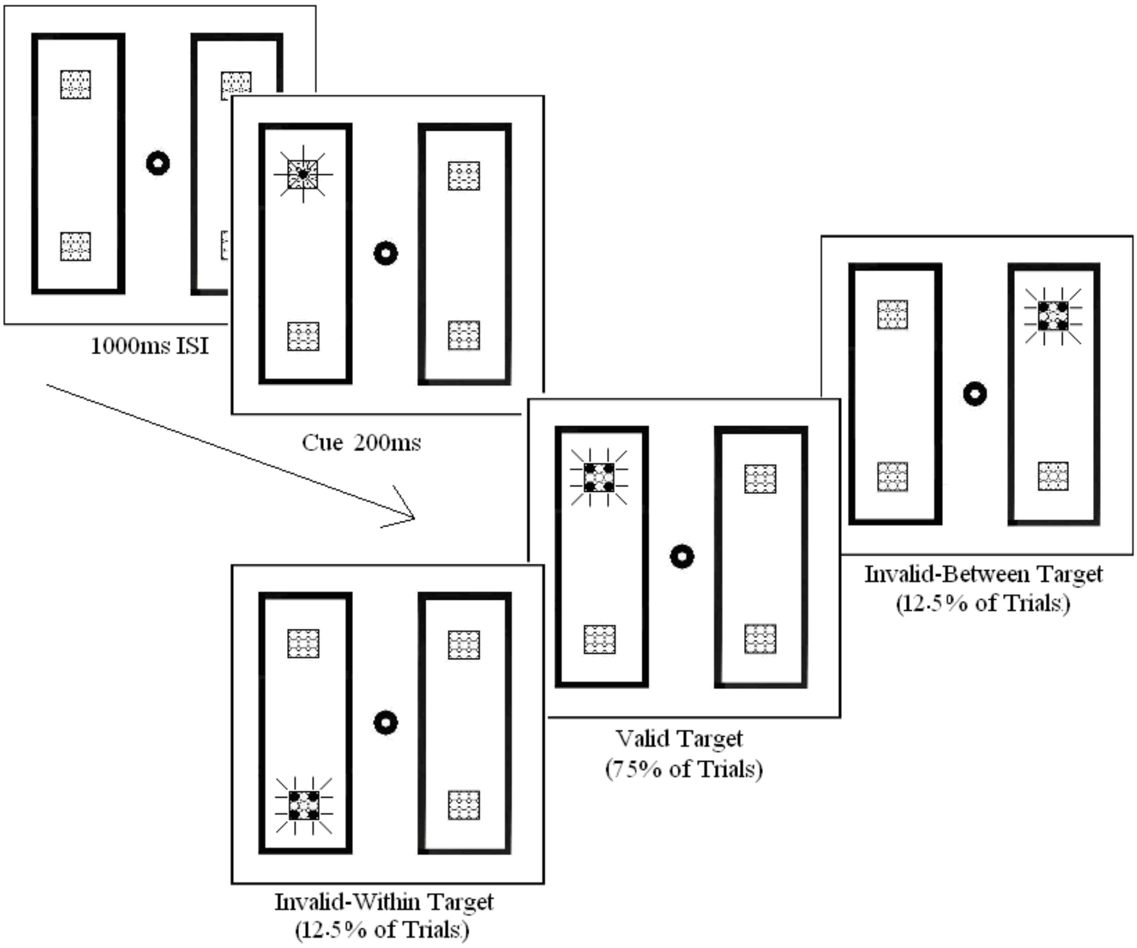


Figure 1, an illustration of the basic Egly et al. paradigm adapted from Chapter 5. Participants are first presented with a pair of rectangles, with four cue/ target locations and a central fixation cross. After viewing the rectangles for approximately 1000ms the participant is presented with some form of cue (here a brightening) at one of the four locations. After the cue the participant is presented with some form of target, either to detect (present, yes/no) or to discriminate (colour, red/green). The target can appear in three locations relative to the cue, targets at the valid location are responded to most quickly, followed by invalid targets on the same object, followed by invalidly cued targets on another object.

Object Based Attention and the Egly et al. Cuing Paradigm

– Object 'based' selection or objects 'biasing' spatial selection

Although commonly cited as a key piece of evidence for object based attention, the Egly et al. within-vs-between-object advantage is in fact often explained in terms of a fundamentally spatial process of selection. In this spatial explanation, the cue in one corner of one shape does not lead to the selection of that shape per se, but rather leads to an application of spatial attention that then preferentially spreads along that object. In this sense the selection in this paradigm is not object 'based', but rather the shape of the object 'biases' what is fundamentally a spatial process of selection.

This 'spatial' approach to the Egly et al. cuing paradigm is concisely and explicitly summarised in the title of a paper by Martinez et al. (2006): "Objects are Highlighted by Spatial Attention". In this paper Martinez et al. argue that this model of "object guided spatial selection" (p. 298) is supported by electrophysiological recordings showing that although they have a different time course, purely spatial attention and 'object based' attention effects have a similar modulatory effect on the N1 component of visual processing. Indeed Martinez et al. note that this fundamentally spatial process of selective enhancement is further consistent with fMRI data from early retinotopic areas of the primary visual cortex, highlighting that spatial locations within a cued bounded shape become more active (Muller & Kleinschmidt 2003). Martinez et al.'s bold statement of this 'spatial spreading' account of object based attention is, if anything the norm, rather than the exception. Davis, Driver, Pavani and Shepard (2000) for example argue that the within-vs-between-object advantage is explained by the fact that "attention tends to spread from the task-relevant parts of objects to their task-irrelevant parts" (p. 1324). Indeed Davis et al. (2000) go on to apply this fundamentally spatial approach to other 'object based attention' paradigms. Another example of this 'spatial spreading' account in the Egly et al. paradigm is endorsed by Richard, Lee and Vecera (2008) who explain that "An attentional spreading view of object-based attention proposes that the rate and efficiency of perceptual processes are improved by spreading attention through an attended object" (p. 843). In one final example Goldsmith and Yeari (2003, p. 916) argue that although it may seem paradoxical, 'object based attention' may in fact reflect the spread of spatial attention 'within object boundaries and groupings'. In chapter 2 this spatial spreading account of the Egly et al. paradigm is explicitly tested by exploring the time-course over which the ‘within-vs-between object’ advantage manifests for different locations on the cued shape.

The possibility that spatial attention could underlie the ‘within-vs-between object’ advantage was originally questioned by the fact that the spatial vs ‘object based’ components of this paradigm were differentially sensitive to different types of cues: in general when manipulating the allocation of processing resources one can use two distinct types of cues: exogenous and endogenous. The cue type already described in the Egly et al. paradigm is a prime example of an ‘exogenous cue’, a brief flash that in and of itself attracts attention to the location in which it appears. One can also employ endogenous cues that are more symbolic in nature, for example one could use an arrow, which does not draw attention to itself, but to the direction in which it points. Macquistan (1997) argued that although a basic spatial attention effect could be found with both exogenous and endogenous cues, the within-vs-between-object advantage only manifested when an exogenous cue was used. This dissociation could be taken as evidence that the spatial and object based components of the Egly et al. paradigm have very different causes. Work by Goldsmith and Yeari (2003), however, demonstrated that the apparent insensitivity to endogenous cues did not reflect a fundamentally different form of attention, but reflected a confound in terms of the differential initial spread of processing resources when using endogenous vs exogenous cues. Goldsmith and Yeari note that when waiting for an endogenous cue, one’s initial focus of attention is typically centred on the location where the cue will appear, whereas exogenous cues, which might appear anywhere on the display, engender a more dispersed spread of attention. Furthermore they demonstrated that endogenous cues are, in fact, effective when the participant is forced to maintain a more distributed initial allocation of attention. Indeed independently of cue type they found that the ‘within-vs-between object’ advantage was attenuated with a focused (rather than dispersed) spread of attention. Goldsmith and Yeari’s result therefore not only eliminates a potential distinction between object and spatial based attention, but also highlights the influence of the initial spatial spread of attention upon object based effects.

This does not in itself prove that spatial attention is responsible for the ‘within-vs-between object’ advantage, but it does suggest that some degree of processing resources have to be allocated to an ‘object’ before that object can either act as a unit of selection, or bias the spatial allocation of attention. This observation throws into question an assumption that 'object based attention' phenomena reflect a pre-attentive level of perceptual representation (Egly et al.). Indeed one of the arguments for the functional importance of 'object based attention' effects in the human visual system revolves around the argument that the visual environment is too complicated to be understood all at once, and that not only must one select critical pieces of information, but also this initial selection needs to be guided in some manner so as to pick out elements of a scene that are meaningfully related to each other (Driver et al. 2001). The fact however that some degree of spatial attention is required before 'objects' are parsed to an extent such that they influence attention (Goldsmith and Yeari, 2003), could seem to undermine the potential utility of these phenomena. Furthermore the segmentation demands of the standard Egly et al. paradigm, in which the participant is presented with only two outline objects, are trivial in comparison to the rich complicated pattern of information present in a real world scene. Chapter 3 therefore seeks to test the extent of 'pre-attentive' processing in the Egly et al. paradigm, and the extent to which it can scale up to and still operate with more complicated scene dynamics.

A further means of thinking about the relationship between spatial and 'object based' attention that will be considered in this thesis involves their relationship to conscious awareness. Recent neuropsychology (Kentridge, Heywood, Weiskrantz 1999), cognitive neuroscience (Wyart & Tallon-Baudry 2008), and cognitive science (Kentridge, Nijboer & Heywood 2008) has demonstrated that one can bias the processing of information via purely spatial means without becoming conscious of either the cues that elicit attentional effects or the stimuli whose representational status is enhanced. It remains, however, something of an open question whether one can demonstrate 'object based attention' effects when one is not conscious of those objects. Indeed Mole (2008) has argued that there are good reasons to believe that access to conscious representation might be related to specifically object based forms of attention. This argument is of course predicated on the assumption that spatial and object based attention phenomena reflect truly distinct underlying processes. An investigation of the potential for 'object based attention' to operate in the absence of awareness would tell us not only about the relationship between this phenomenon and awareness but could also be informative in terms of the more general relationship between spatial and object based attention. The potential for objects to influence the competition for processing resources when those objects are otherwise unconscious is explored in Chapter 4.

Object Based Attention – Is the Egly et al. cuing effect driven by Objects, Proto-objects, Shapes, or Edges?

There are not only multiple stages at which the human visual system organises information there are parallel 'dorsal' and 'ventral streams' in which that information is organised. Whilst 'objecthood' was once regarded as a feature solely extracted by the ventral stream (Ungerlieder & Mishkin, 1982) it has become clear that shape information is extracted at different levels of information processing in both the dorsal and ventral visual pathways (see Konen & Kastner 2008 for evidence from human fMRI and Lehky and Sereno 2007 for evidence from primate electrophysiology). Any attempt to pin down the nature of the 'objects' that influence attention has to be addressed not only at the level at which these representations are extracted (Driver et al. 2001) but also with respect to the stream within which these representations develop.

Goodale and Milner (1992, Milner and Goodale 1995) argued that representations developed in the dorsal stream, are extracted for the purpose of guiding manual action, and that representations in the ventral stream are responsible for forming our explicit perception of objects and space. 'Object based attention' effects are to some extent defined in terms of their influence on the efficiency of perceptual responses. In the context of Milner and Goodale's framework therefore one would predict that it is the representations extracted in the ventral stream that are responsible for influencing the allocation of object-based attention.

One indication (although not from the Egly et al. cuing paradigm) regarding which stream is critical comes from the work of He and Nakayama (1995). He and Nakayama explored the factors that influenced the grouping of targets and distracters at different depths. In a parallel of the Driver & Baylis (1989) object based flanker distracter task, He and Nakayama found that flanking distracters were more likely to be grouped with (and therefore cause a greater interference effect upon) targets when they appeared on a surface at a common depth. Interestingly however they found that when targets and distracters were presented at a common depth, defined purely on the basis of absolute disparity (without a common surface), then distracters were not grouped with the target. This finding is interesting to frame within Neri's (2005) conclusion that although the dorsal stream can compute depth on the basis of absolute disparity, only areas in the ventral stream are sensitive to the relative disparity between different objects. Taken together Neri's and He & Nakayama's results suggest that relative depth relationships computed within the ventral stream are required in order to influence the allocation of attention. It is worth highlighting however that this conclusion is limited by the fact that Neri's review of depth sensitivities within the dorsal stream does not focus on those areas recently recognised to show significant visual shape sensitivities (Konen & Kastner 2008).

More direct evidence for the role of the ventral stream comes from a combination of fMRI and EEG neuroimaging to pin down the first area to reveal a significant 'within-vs-between object' target facilitation (Martinez et al. 2006). Previous work (Muller & Kleinschmidt 2003) had highlighted that the 'within-vs-between object' advantage manifested as a pattern of increased spatial activation within the primary visual cortex, however Muller and Kleinschmidt argue that the time course of this activation meant that it was not possible to test whether this activation in fact reflected feedback from the dorsal or ventral stream. Martinez et al. attempted to circumvent this problem using a combination of EEG to detect the first activation sensitive to the within vs between object distinction, and fMRI to localise this EEG signal. Using this combined method Martinez et al. argue that the first site at which the 'object' based effects manifest in the Egly et al. paradigm can be localised to the Lateral Occipital area (LO), within the ventral stream. This work is however itself problematic because Martinez et al. required participants to attend to the display in order to perform a shape based discrimination task. The increased activation in LO might therefore not have reflected a fundamental role for this area in acting as the level at which objects are selectively enhanced, but rather may have reflected the known role of LO in discriminating shapes in general (James et al. 2003). In other words the enhanced LO activation may have reflected the fact that the representations in this area were critical to this particular task, not that they are critical to object based attention in general. The Martinez et al. result would single out LO much more clearly if it could be replicated with a discrimination task that did not in itself require some analysis of shape (for example the participant could be asked to discriminate between two colours). Thus although there is some evidence consistent with the idea that the representations that are required to drive attention are located within the ventral stream (and within LO in particular) the evidence does not conclusively single out this explanation.

In parallel to the neuro-imaging data explored above, the nature of the objects that influence attention can also be explored in purely behavioural terms. The behavioural evidence highlights an interesting juxtaposition whereby ‘object based attention’ is both contingent upon ‘high-level’ experience-dependent interpretations of the nature of perceptual groupings (Moore, Yantis & Vaughan, 1998, Zemel, Behrmann, Mozer & Bavelier 2002) and yet can be influenced by the most basic of segmentation cues that do not clearly relate to what one would think of as objects (Avrhami 1999). The contribution of internal principles of perceptual organisation in the construction of the units that influence attention was most clearly established by Moore et al. (1998), who showed that a within-vs-between-object advantage could be found for partially occluded objects that had to be constructed on the basis of amodal completion. Indeed Zemel et al. (2002)[2] demonstrated that whether or not an object was constructed such that it could influence attention was dependent upon previous experience of that object; when presented with the shapes in figure 2 A, for example, a within object advantage was only found for the amodal completion of these shapes after the participant had previous visual experience of how they could be grouped into a single shape.

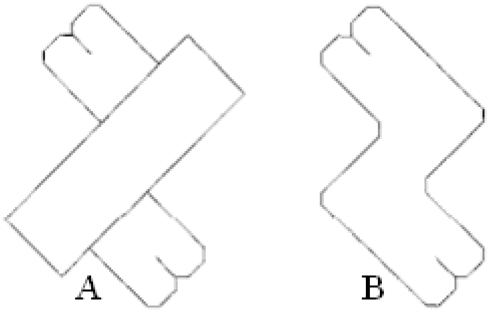


Figure 2, *stimuli from Zemel et al., the two fragments shown in figure A only generated a within-object advantage for pairs of targets on those shapes after the participant had seen how those two fragments could group into one shape (B).*

---

[2] Zemel et al. in fact use a two item comparison paradigm, rather than the Egly cuing paradigm, but this effect makes an interesting logical extension of those found with amodal completion in the Egly et al paradigm by Moore et al..

Despite the involvement of previous organisational principles (Moore et al.) and recent visual experience (Zemel et al.) object based attention can also be influenced by visual structure that does not clearly relate to a clear notion of a coherent object. The clearest demonstration of this is provided by Avrahami (1999) who found that the Egly et al. cuing paradigm could be replicated when there were no 'objects' present at all, but simply a series of edges (see figure 3). Stimuli cued within a single pair of lines were more rapidly responded to than ones cued by an equi-distant cue located between a different pair of lines. In the light of this demonstration it could be argued that attention is not being influenced by objects, but merely by the edges that necessarily make up the shape of those objects.

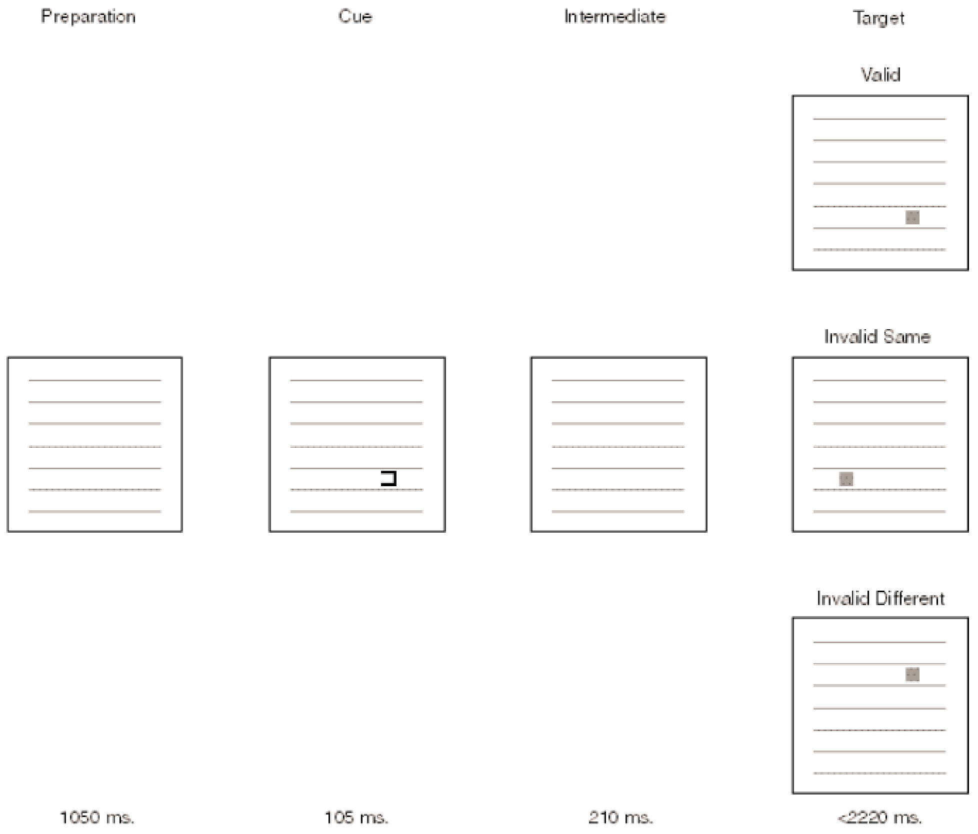


Figure 3, Stimuli and cueing procedure from Avrahami (1999). See text for details.

Ben-Shahar, Scholl and Zucker (2007) recently demonstrated however that the presentation of oriented edges, was not in itself sufficient to influence the allocation of attention[3]. Ben-Shahar et al. set out to use within-vs-between-object attention effects as a tool to explore how different textures are segmented into distinct parts of a visual scene. In doing so they found that when two textures had differently oriented elements, these different textures acted as units of attentional selection, such that there was a ‘within object’ advantage for two items appearing within the same texture. Interestingly however they found this ‘within object’ texture advantage manifested whether or not the edges of that texture where oriented ‘along’ or ‘against’ the grain in which attention had to move (see figure 4). Indeed they further extended this effect to show that if the display merely contained a series of oriented edge elements the facilitating effect was insensitive to the fact that those edges were oriented in line with, or at right angles to, the necessary movement of attention.

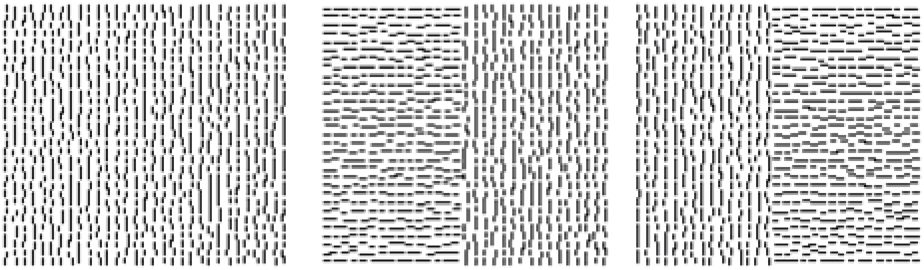


Figure 4, adapted from Ben-Shahar, Scholl and Zucker (2007). Edges at right angles to the direction in which attention moves did not influence attention, whilst the texture based segmentation boundaries in the second two stimuli did.

Thus whilst simple edges can influence the allocation of attention (Avrahami 1999), this effect is critically dependent upon the manner in which those edges are interpreted. If those edges are interpreted as belonging to a coherent texture then they do not, in and of themselves, influence the allocation of attention. Ben-Shahar et al. themselves note however that the textures in their experiment are intuitively speaking somewhat short of a genuine real world ‘object’. Stepping back however it is clear that even in the original Egly et al. paradigm

---

[3] The paper actually focuses on a two item comparison object based attention task, however they indicate in a footnote (which was confirmed in personal communication) that the effects also replicated with the Egly et al. within-vs-between-object effect.

the simple outlined objects bear very little relationship to the objects we experience in our daily lives. As the brain attempts to organise the wash of sensory information received upon the retina however it could be argued that the factors that influence the allocation of attention relate to the early stages of perceptual organisation in which potential or candidate distinctions between different objects are segmented (Driver et al. 2001). Indeed Driver et al. have argued that object based attention does not reflect either the initial extraction of edges, or the meaningful recognition of objects, but an intermediate, or 'mid-level' stage of visual perception in which the visual system tries to organise the information it receives so as to facilitate subsequent perception and recognition of that information.

Further evidence that object-based attention effects are dependent upon more than the basic registration of edges has recently been provided by Albrecht, List and Robertson (2007) who investigated the ability of 'holes' to act as the units of selection in a replication of the Egly et al. paradigm. Albrecht et al. found that when stereoscopic depth cues where employed to ensure that two rectangular gaps were perceived as holes, the surfaces behind these 'rectangular holes' did not generate a within-vs-between-object advantage. Given that the edges making up the rectangular figure vs the rectangular hole were identical this effect further emphasizes that it is not edges, but the interpretation applied to those edges in terms of their figure ground significance that influences the allocation of processing resources. Integrating this behavioural finding with the potential role for the Lateral Occipital area in generating the representations required to influence attention reveals another suggestive point of overlap, given that representations in LO are known to be differentially sensitive to patterns of edges that form a hole vs a figure and are less active when presented with the former (Vinberg & Grill-Spector 2008). This suggestive correspondence again adds weight to the potential role of the Lateral Occipital area in generating the representations that influence attention. This correspondence is further explored in Chapter 5, where the known role of LO in developing cross-modal shape representations are exploited to test whether or not an object that is only perceived via non-visual or tactile information can influence the allocation of visual attention. Furthermore the role of LO in developing the representations required to influence attention is further explored in Chapter 8, where the existence of spatial and object based attention effects are examined in a patient with visual form agnosia who has a clear bilateral lesion to LO (James et al. 2003).

Object Based Attention – what can it reveal about perceptual organisation and attention?

- Setting the direction of this thesis

This thesis started with the possibility that early forms of perceptual organisation may play an important role in structuring later cognition (Pylyshyn 1999, Xu & Carey 2001). It moved on to catalogue the many paradigms by which visual perceptual organisation can influence the allocation of attention. Unfortunately the extent to which these different forms of 'Object Based Attention' reflect the same processing or level of representation has not been systematically examined. This thesis therefore sought to focus on one particular object based attention phenomenon, namely the Egly et al. cuing paradigm. The thesis explores the nature of perceptual organisation and attention in relation to this phenomenon in three sections.

The first section of the thesis seeks to explore this phenomenon in more detail. As reviewed above the exact contributions of 'object' vs 'spatial' attention to this phenomenon have come under question, Chapter 2 will therefore seek to shed some light on this debate, by exploring the time course over which the within-vs-between-object advantage manifests at different locations. Indeed given the known role of a distributed allocation of spatial attention in developing the representations required to influence attention (Goldsmith & Yeari 2003), Chapter 3 will explore the extent to which the 'proto-objects' that drive the within-vs-between-object advantage reflect a parallel or pre-attentive stage of visual scene segmentation. Finally in this section Chapter 4 will explore the potential association between object based attention and awareness, in the context of the known dissociation between spatial attention and awareness (Kentridge et al. 1999).

The second section of the thesis turns away from an exploration of object based attention as a phenomenon for its own sake, and seeks to use object based phenomena to explore the nature of perceptual organisation. More specifically Chapter 5 explores the potential cross modal nature of shape representations and Chapters 6 and 7 explore the potential differential representation of stimuli such as shadows and biological motion, which, given their differential ecological importance, might have very different roles in the eyes of attention. In the final section of this thesis the nature of object based attention (Chapter 8), proto-object formation (Chapter 9) and perceptual organisation on the basis of features other than shape

(Chapter 10), are explored in the context of a visual form agnosia patient with a severe lesion to the Lateral Occipital area.

"If the object based theory is correct, then the study of visual attention and of perceptual organisation must proceed together". Duncan (1984).

In the first paper on object based attention John Duncan argued that our understanding of perceptual organisation and attention must proceed hand in hand. Twenty five years later this thesis sits exactly on this interface between our understanding of attention and the nature of perceptual organisation.

Section One: Understanding the Phenomenon of Object Based Attention.

This section will detail three sets of experiments that seek to advance our understanding of the phenomenon of object based attention. In chapter 2 this thesis explores more closely the role ‘objects’ play in influencing attention in the Egly et al. paradigm by examining the time course over which the within-vs-between-object advantage manifests at different locations. In the context of the debate regarding the relationship between spatial and object based attention, the extent to which this phenomenon is truly thought to be ‘object based’ (independent of the question of the level at which those ‘objects’ are extracted by the visual system) has been questioned. A number of researchers assume that the phenomenon does not reflect an object based form of selection, but a spatial selection that is merely guided by object structure (Davis et al. 2000, Martinez et al. 2006, Goldsmith & Yeari 2003, Richard et al. 2008). Indeed this assumption is so widespread that chapter 3 was in fact motivated not by any attempt to test this hypothesis, but just to explore how this ‘spatial spreading account’ would manifest itself. As will become clear however the results in fact proved quite counter to expectation.

As outlined in the introduction a key motivation for studying object based attention follows from the possibility that this early stage at which perceptual organisation influences attention may have implications for the manner in which higher cognition is organised. If this possibility is to hold any validity it must be the case that object based phenomena manifest in our daily life interaction with the world. In the Egly et al. paradigm (upon which this thesis will focus) however, this assumption needs to be tested, especially if one considers the highly over-simplified shapes typically employed as ‘objects’ in this paradigm. There are many directions in which one could take this paradigm in order to test the possibility that the within-vs-between-object advantage still manifests with more complicated scene dynamics. Chapter 3 tests the extent to which the results found by Egly et al. are still manifest when the visual system is presented with multiple pairs of objects presented in constantly changing and unpredictable locations. As will become clear this analysis not only provides a first step in testing whether the phenomenon can scale up to more complicated scene dynamics, but also addresses some more specific questions about the potential pre-attentive nature of proto-object extraction, the segmentation limits driving the Egly et al. within object advantage and other object based paradigms, and the potential lateralisation of the object based attention effects manifest in the Egly et al. paradigm.

In chapter 4 the thesis explores the role of awareness in object based attention. This investigation was motivated not only in terms of our understanding of the role of awareness in attention, but also by means of contrasting the known ability of purely spatial attention to operate in the absence of awareness (Kentridge et al. 1999). This chapter exploits a phenomenon called Continuous Flash Suppression (Tsuchiya & Koch 2005) to render the ‘objects’ unconscious. This technique exploits the more general phenomenon of binocular rivalry, in which incongruous images are presented to each eye using a system of mirrors, resulting in a competition such that only the information reaching one eye is represented to the level at which it becomes conscious at any one time. Continuous Flash Suppression in particular seeks to bias this competition between the two eyes, such that stimuli presented to one eye can remain unconscious for long periods of time[4]. This chapter initially set out to explore whether the Egly et al. paradigm would manifest under conditions of Continuous Flash Suppression: after some pilot work however this aim was modified such that a much simplified paradigm could be employed. This novel paradigm is detailed and explained in the chapter. As will become clear this simplified paradigm in fact proved more informative than the Egly et al. paradigm would have done, and in fact questions the ability of the visual system to extract proto-objects per se during Continuous Flash Suppression, let alone the ability of those objects to influence attention.

---

[4] Although not detailed in the chapter itself it is worth noting here that a significant amount of pilot work was required in order to make the traditional continuous flash suppression mask effective for the objects employed; although not explored systematically it seemed to be the case that in order to develop an effective suppression-mask the mask-stimuli had to be composed of a lower spatial frequency relative to the to-be-masked stimuli.

Chapter 2 - Object ‘based’ selection or objects ‘biasing’ spatial selection?

Abstract

One of the most commonly cited object based paradigms (Egly, Driver and Rafal, 1994) has been explained in terms of a fundamentally spatial mode of selection in which ‘objects guide spatial selection’ (Martinez et al. 2006). Motivated by this possibility the current work explores the manner in which the ‘within-vs-between object’ advantage in the Egly et al. paradigm emerges over space and time. The results however do not provide any evidence consistent with an ‘object guided’ form of spatial attentional spreading. In fact the results highlight an object based switch cost that is sensitive to the distinction between ‘figures’ and the ‘ground’ from which they are segmented.

Introduction

The allocation of visual processing resources was initially conceived as a spatial mechanism (Posner 1980, Treisman & Gelade 1980); this spatial mode of attention was epitomized in the notion of a visual spotlight. The influence of object structure, upon both the facilitation (Egly, Driver & Rafal 1994), and later inhibition (Tipper, Jordan & Weaver 1999), of visual information processing however questioned this ‘spotlight’ metaphor (Driver & Baylis, 1989). Indeed the capacity for a fundamentally spatial system of attention to select for different objects is contingent upon those objects existing at different locations; when objects are presented at an overlapping location, however, it is clear that some form of ‘object based’ selection takes place (O’Craven, Downing & Kanwisher 1999, Duncan 1984). Dominant models of attention in fact argue that information processing can be biased towards different aspects of the environment on the basis of a number of stimulus dimensions, be they space, color, motion or shape (Desimone & Duncan 1995).

The role of objects in attention is now uncontroversial[5] (Scholl 2001). It is therefore perhaps surprising that one of the most cited examples of ‘object based’ attention (Egly, Driver &

---

[5] Although what exactly constitutes an ‘object’ or the unit of selection, in this context, remains a matter of much debate (see Driver et al. 2001).

Rafal, 1994[6]) has been argued to reflect a fundamentally spatial selection mechanism, which is merely influenced by object structure. Egly et al. revealed that when one end of one object is cued, not only do targets appearing at that location receive an attentional enhancement, but targets appearing on the same object also receive an advantage relative to equidistant targets appearing on another object. This within-vs-between-object advantage cannot be explained by a purely spatial allocation of attention. Unlike a number of other object based paradigms, however, there is a spatial separation between the two objects, which means that the effect in itself cannot be taken as exclusive evidence for a fundamentally object based form of selection. Indeed one could propose a number of hybrid models in which objects are not the units of selection per se, but merely influence what is a fundamentally spatial allocation of attention. One could argue for example that the cue, presented to the end of one of the shapes does not lead to the selection of that shape per se, but rather leads to an application of spatial attention that then preferentially spreads along the cued shape. In this sense the selection in this paradigm is not object 'based', but rather the shape of the object 'biases' what is fundamentally a spatial process of selection. Indeed such a hypothesis is concisely and explicitly summarized in the title of a paper by Martinez et al. (2006): "Objects are Highlighted by Spatial Attention". Martinez and colleagues argued that this model of "object guided spatial selection" (p. 298) is supported by electrophysiological recordings showing that although purely spatial attention effects manifest more rapidly they have a similar neural profile, particularly in terms of their modulation upon the N1 component, to neural correlates of the 'within-vs-between object' advantage. Martinez et al. note that this fundamentally spatial process of selective enhancement is further consistent with fMRI data from early retinotopic areas of the primary visual cortex, which show that the spatial locations within a cued bounded shape become more active (Muller & Kleinschmidt 2003). Martinez et al.'s bold statement of this 'spatial spreading' account of object based attention is, if anything, the norm rather than the exception. Davis, Driver, Pavani and Shepard (2000) for example argue that the cuing effects seen in the Egly et al. paradigm are explained by the fact that "attention tends to spread from the task-relevant parts of objects to their task-irrelevant parts" (p. 1324). Richard, Lee and Vecera (2008) also endorse an 'attentional spreading' account which they summarize thus: "An attentional spreading view of object-based attention proposes that the

---

[6] On the 20th August 2009, Duncan (1984) had been cited 781 times, Driver and Baylis (1989) 200 times, Tipper, Jordan and Weaver (1999) 44 times, and Egly, Driver and Rafal (1994) 251.

rate and efficiency of perceptual processes are improved by spreading attention through an attended object" (p. 843).

This project sets out to explore whether the time course of the within-vs-between-object advantage in the Egly et al. cuing paradigm is consistent with a fundamentally spatial mode of attentional selection. If the within-vs-between-object advantage reflects a process by which spatial attention is attracted to, and gradually spreads along, the cued object, one should find a stronger within object advantage for targets closer to the cue, although this might only be evident at certain time points. In the standard Egly et al. cuing paradigm within object (invalid) targets are presented at the extreme end of the cued objects, and compared to targets on another object that are equally spaced from the cue. The current study therefore introduced a new set of target locations midway between the traditional standard distances to test whether a within-vs-between-object advantage would manifest earlier for this 'middle' target pairing (see figure 1). This potential differential emergence of a within-vs-between-object advantage for different distances from the cue was explored at a range of times after the presentation of the cue.

Although principally motivated by the fundamentally spatial explanations for the within-vs-between-object advantage in the Egly et al. paradigm, this project also enables one to test for a potential differential advantage for targets located at middle locations on the Egly et al. rectangles that could be predicted on the basis of recent results from Kravitz and Behrmann (2008). Kravitz and Behrmann explored the pattern of enhanced activations in the spatial locations surrounding a cued object, and found a gradient of enhanced activation that in fact centered on the middle of the object. The authors maintain an 'object based' explanation for the fundamental process of selection, but argue that once the object is selected all other features of the object become active, including its spatial location. Kravitz and Behrmann therefore argue that the 'object based' selection leads to a form of feedback that creates a spatial gradient of enhanced activation, emanating from the center of the object. Kravitz and Behrmann data only concerned the space surrounding the selected object, but in principle, it could also explain the enhanced processing of within-object targets. If the proposed object centered feedback to locations outside the object also extends to locations within it, one would predict an advantage for targets presented at the mid-object locations employed in the present experiment. Note however that although this model leads to a similar prediction, in terms of a larger object based effect for mid-location targets, it manifests for different reasons.

In the spatial spreading account the guidance of spatial attention by the object leads to attentional enhancements first registering (or first passing through) the centre of the object. In this second account however the potential enhanced differential advantage for mid location targets results from a form of selection that is fundamentally object based but that leads to a pattern of spatial integration (as could be considered consistent with Treisman & Gelade 1980) that emanates primarily from the centre of the selected object.

To summarize, although commonly cited as an example of ‘object based attention’, the within-vs-between-object advantage in the Egly et al. cuing paradigm is open to explanations in terms of fundamentally spatial mechanisms both at the level of selection itself (Martinez et al., 2006) and at the level of integration following object based selection (Kravitz and Behrmann, 2008). This study set out to test these potential spatial explanations by exploring the time course over which cues presented on outline shapes influence the allocation of processing resources for targets located at different positions within and between those shapes. More specifically it hypothesized that if the within-vs-between-object advantage seen in this paradigm reflects an underlying spatial mechanism, either at selection or integration, there should be an advantage for targets appearing at the center of those outline shapes that would emerge before an advantage for within-vs-between target pairs presented at the standard full length target locations traditionally employed in this paradigm.

Methods

Participants

Fifty four participants completed this experiment, either in exchange for a small financial reimbursement or course credits. All had corrected or corrected to normal vision. There were 8 males and the mean age was 19.5 (SD=1.73).

Design

The experiment employed a within participant design. There are three experimental factors, Cue Type, Target Distance and Cue-Target-Onset-Asynchrony. There were three Cue Types, ‘valid’ (where the target appears at the same location as the cue, which was only included to replicate the standard paradigm and was not included in the final analysis), ‘invalid within’ (where the target does not appear at the cued location, but is still on the same rectangular shape) and 'invalid between' (where the target appears either on another rectangle or in the space between those rectangles). The two invalid conditions were subdivided into two further Target Distances (see figure 1). The 'standard' target locations used by Egly et al. (1994), are located either at the other corner of the cued shape ('invalid within') or the equidistant corner of the other shape ('invalid between'). These were supplemented by two new 'middle' locations, either halfway along the cued shape ('invalid within') or at the equidistant location between the two objects ('invalid between'). Each of these conditions was further subdivided into eight different Cue-Target-Onset-Asynchronies (or CTOAs), defined as the length of time following the cue before the target appeared.

Procedure

Throughout the experiment participants were instructed to fixate a central cross. On each trial two rectangles were presented either horizontally, above and below the fixation cross, or vertically, to either side of the cross. In a counterbalanced order participants were presented either with horizontal or vertical rectangles for the first half of the trials, and then after a short break the other orientation for the remaining trials. The short break was required to ensure that any after-images built up from the shapes in the previous orientation had faded before starting with the new orientation type. The data from each orientation type will not be analysed separately but collapsed into the appropriate cue conditions. On each trial one of four cue locations at the ends of each rectangle would brighten for 53ms. After the offset of this cue a red or green target would appear, and the participant would have to report its color, using the Z (red) or N (green) key. The DirectX library DirectInput was used to improve the temporal precision with which key presses were recorded, resulting in a maximum variability of ±5ms on each trial .The length of time between the offset of the cue and the appearance of the target (CTOA) for each trial had 8 different values: 0, 27, 53, 80, 107, 133, 160 and 187 ms (the values here are rounded on the basis of a monitor re-fresh rate of 75Hz). The target could appear in three locations relative to the cue, either at the same location as the cue ('valid' 1/3 of all trials), on the same object ('invalid within' 1/3 of trials) or not on the same object ('invalid between', 1/3 of trials). On the invalid trials half the targets were presented at the 'Standard' locations and half of them at the new 'Middle' locations. All participants completed 768 trials; because of this large number of trials they had a break every 200 trials. At the start of the experiment, and after each break, four non-recorded trials were added to reduce the introduction of 'noisy' responses as the participant got back into the rhythm of responding. Cue types and distances in these ‘practice’ trials were randomly sampled from the other trial types. The participants were unaware that these trials were not recorded, and the experiment ran straight on from the practice trials to the actual recorded trials. There was a 1000ms interval between trials, during which the fixation cross and outline rectangles remained in view. The reaction times reported in the results are based only on accurate trials, on which the participant responded between 100-850ms after the onset of the target and within 2 standard deviations of the mean for that participant for that condition.

Stimulus

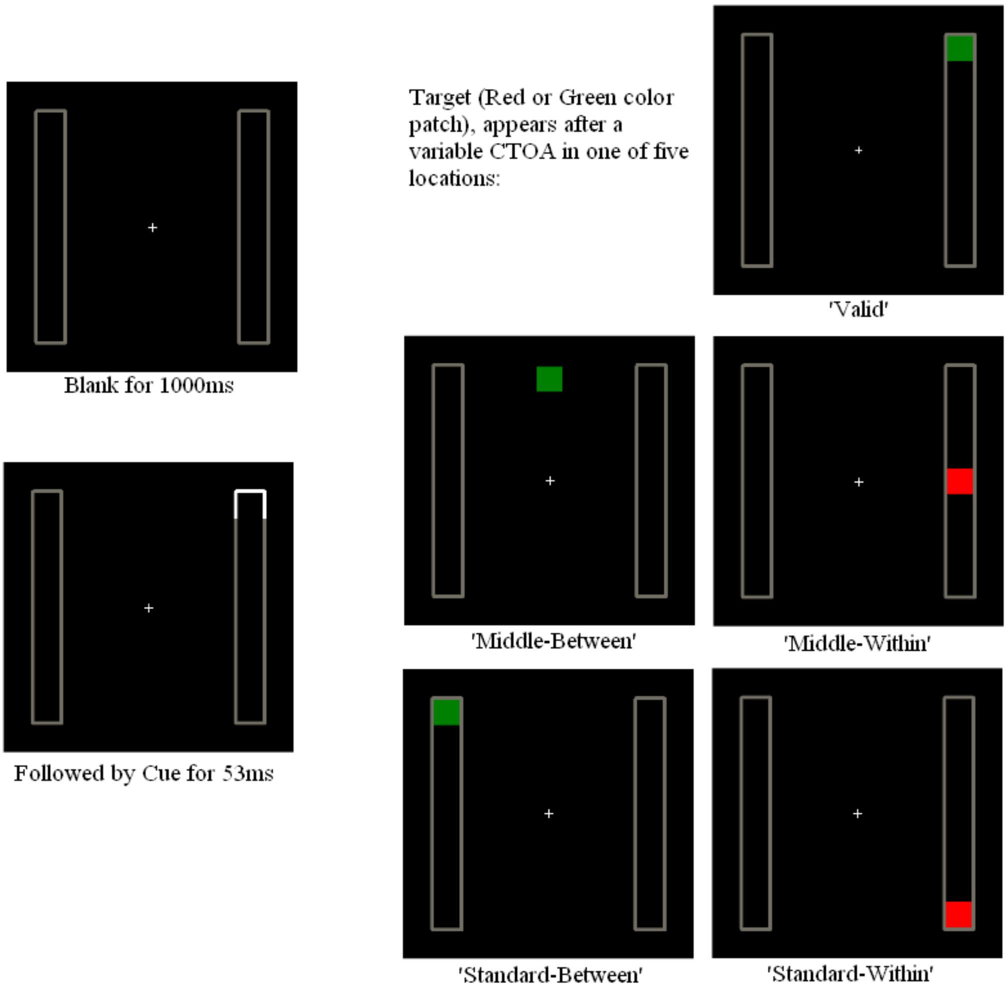


Figure 1: The above figure is based on ‘screen shots’ from the different conditions in the actual experiment, see text for details.

The stimuli were presented on a 17inch monitor running at 75Hz, at a viewing distance of 57cm. The central fixation dot was 0.4 by 0.4 degrees of visual angle, one pixel thick. The lines making up the two rectangles were 0.2 degrees thick. The target color patch was 1.5 by 1.5 degrees square. The width of the rectangles was 1.7 degrees and the distance between either the thin top edge within each rectangle or between the two outer edges between each of the rectangles was 13.7 degrees. This large rectangle length/separation was employed to help bring to light any potential differential effects of attentional spreading to different locations.

Results

The analysis was simplified in two ways. First, although illustrated in the graphs below, the data from the valid trials were not included in the ANOVAs[7]. The analysis of the eight Cue-Target-Onset-Asynchronies was also simplified, by taking the average across the first 4 and last 4 values. The reaction time data for the results simplified in this manner are presented in figure 2.

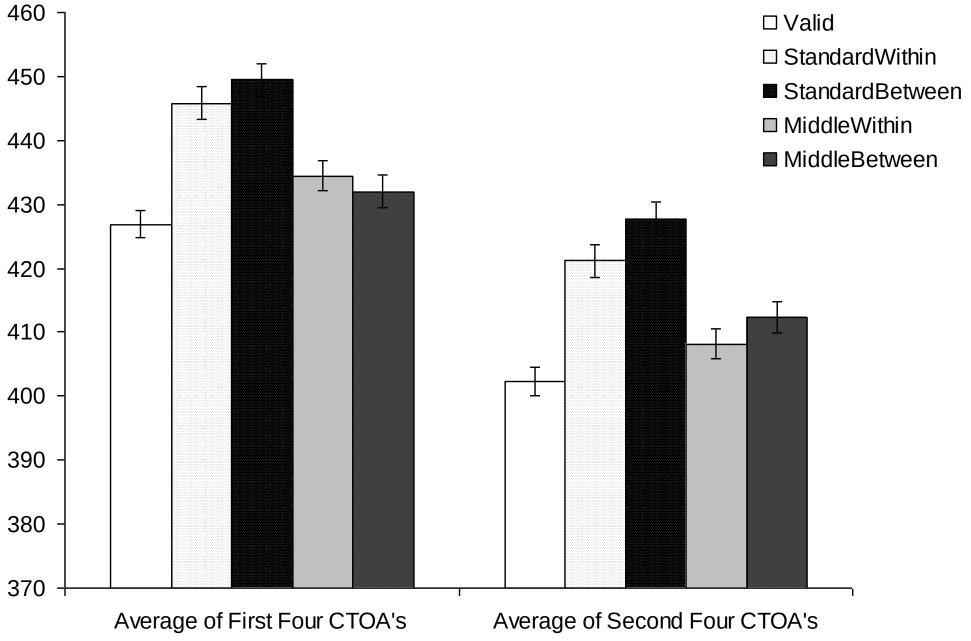


Figure 2, Reaction time (ms) across the five different cuing locations for the average of the first and last four time points (CTOAs).

The reaction time data reveal a clear influence of the distance of the targets from the cue, with targets in the 'middle' locations being detected faster than the spatially more distant

---

[7] Valid trials are important in the Egly et al paradigm, to ensure that the participant comes to rely upon the cue, they are however not critical to understanding the time course over which the within-vs-between-object advantage develops.

‘standard’ locations (‘middle’, 421ms, ‘standard’, 436ms, F(1,53)=164.02, p<0.001). This influence of the spatial separation between the cue and the targets shows no signs of interacting with CTOA over the time course explored (F(1,53)=0.006, p=0.936). There is, however, a main effect of CTOA (F(1,53)=423, p<0.001), which a brief inspection of figure 2 reveals reflects lower reaction times in the last four time points.

The critical within-vs-between-object effect is significant overall (F(1,53)=6.83, p=0.012), although when collapsed across all time points and the two pairs of distances this effect reflects only a 3ms advantage for within object trials (within, 427ms, between 430ms)[8]. The strength of the within-vs-between-object effect interacts with the location of the targets (F(1,53)=4.44, p=0.04). The strength of the within-vs-between-object effect also varies over time (interaction with CTOA, F(1,53)=3.971, p=0.051).

In order to pin down how the within-vs-between-object effect varies over time and distance, these factors were explored separately for the ‘middle’ and ‘standard’ target locations. The reaction time data shows no overall within-vs-between-object main effect at the middle target locations (F(1,53)=0.36, p=0.549), but this factor does interact with CTOA (F(1,53)=5.31, p=0.025). This interaction reflects the fact that the within-vs-between-object difference is not significant at the first four time points (within 434ms, between 432ms, t(53)=0.99, p=0.325) but is significant for the average of the last four time points (within 408ms, between 412ms, t(53)=2.60, p=0.011). In the standard condition there is again a clear overall effect of within-vs-between-object cuing (F(1,53)=10.0, p=0.003) but still a trend similar to that seen in the ‘middle’ conditions, whereby the within-vs-between-object difference was less clear over the first four time points (Within 446ms, Between 449ms, t(53)=1.85, p=0.07) than over the final four time points (Within 421ms, Between 428ms, t(53)=2.78, p=0.007). This similar trend in the ‘standard’ distance data did not however reach significance as an interaction between cue type and time (F(1,53)=1.03, p=0.316).

---

[8] With a CTOA of around 200ms one would typically expect a 10-15ms within object advantage.

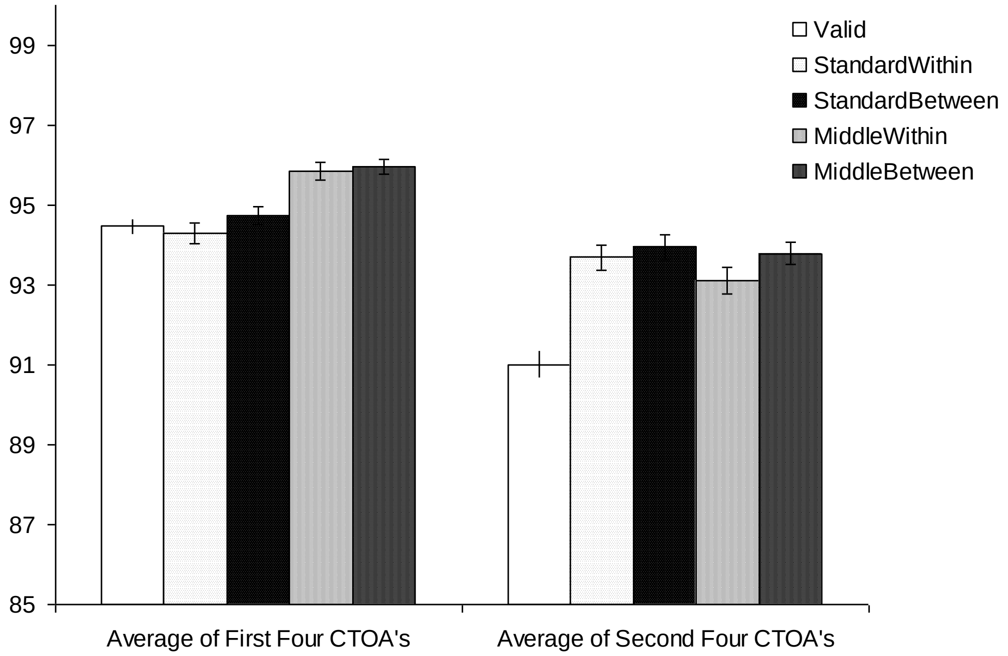


Figure 3, Accuracy (as percentage correct) across the five different cuing locations for the average of the first and last four time points.

The accuracy data show a trend towards higher scores at the middle than the standard locations (F(1,53)=3.50, p=0.067), consistent with the spatial advantage seen in the reaction time data. Responses were significantly less accurate at longer CTOAs (95.2% vs 93.6%, F(1,53)=21.6, p<0.001), indicating a speed-accuracy trade-off against the decreased reaction times seen during those last four time points. There was no within-vs-between-object (Cue Type) effect (F(1,53)=1.60, p=0.211) no interaction between Cue Type and Distance (F(1,53)=0.01, p=0.934) and no interaction between Cue Type and CTOA (F(1,53)=0.171, p=0.681). There was however an interaction between Distance and CTOA (F(1,53)=11.78, p=0.001). This interaction reflects an advantage for ‘middle’ over ‘standard’ targets only in the first four time points (Middle 95.3%, Standard 94.5%, t(53)=2.93, p=0.005) but not over the last four (Middle 93.5%, Standard 93.8%, t(53)=0.945, p=0.349). This demonstrates that the marginal spatial attention advantage for middle-located targets in the accuracy data in fact reflects an advantage apparent only in the average of the first four time points.

In Summary, there was a main effect of within-vs-between-object cues in the reaction time data. This effect was clearly stronger at the ‘standard’ distances: indeed a reliable within-vs-between-object effect only emerged for the ‘middle’ locations over the last four time points.

There was a general spatial advantage for the middle locations in both the reaction times and the accuracy scores; in the reaction time data this did not change across CTOAs, but the effect was only evident in the accuracy data for the first four time points. The participant's reaction time decreased with increasing CTOAs, but this seemed to reflect a speed-accuracy trade-off with longer CTOAs being associated with lower accuracies.

Discussion

This experiment was motivated by the possibility that fundamentally spatial mechanisms might explain the now-classic within-vs-between-object cuing advantage reported first by Egly et al. (1994). If this within-vs-between-object advantage reflected a spreading of spatial attention, preferentially guided within an object, one would predict that the within-vs-between-object advantage would manifest itself first at target locations midway between the standard locations employed in the standard Egly et al. paradigm. This did not occur; in fact not only did the advantage first manifest itself at the *standard* locations, but the within-vs-between-object advantage was stronger at the 'standard' locations throughout the timecourse of the experiment. The stronger effect at the standard locations probably reflects the fact that the 'middle-between' target lay in the space between the two objects, rather than on another object; on this interpretation, switching attention *away* from an object may be achieved less efficiently when that move involves the selection of another object or figure.

The Egly et al. cuing paradigm is one of the most commonly cited and replicated examples of 'object based attention'. It is therefore perhaps somewhat surprising that many authors assume that 'objects' do not play a fundamental role in allocating attention in this paradigm. Rather 'objects' have been thought to guide or influence the allocation of a fundamentally spatial attentional process, a notion embodied most boldly in the title of Martinez and colleagues' recent neuroimaging paper "Objects Are Highlighted by Spatial Attention" (2006; also see Davis et al., 2000, and Richard et al., 2008). Indeed even if one does not assume that the initial mode of selection is spatial, there is evidence consistent with the possibility that once a shape is selected, other features appearing on that shape (such as the colored surfaces used in this experiment) could become more active via a fundamentally spatial form of integration, consistent with the hypothesized role of spatial information in feature binding (Treisman & Gelade, 1980). This evidence is provided by Kravitz and Behrmann (2008) who explored the spread of processing resources around a cued object, and found that this gradient of processing resources centered on the middle of the object, not the location of the cue. The spatial integration phenomenon identified by Kravitz and Behrmann could in principle therefore also explain the advantage for targets located within an object.

In this context it is somewhat counter-intuitive that the results in fact reveal a stronger object-based advantage at the standard full-distance locations. Although not predicted, this result does have quite a straightforward explanation. In order to create a target location that was spatially matched to the mid-within shape location it was necessary to place the target in the space between the two objects. When switching attention between the standard full-distance locations, one not only needs to move attention off of one shape, but also to switch to a target location on another shape. When switching to the new mid-target location, one still has to move attention off the cued shape, but attention does not have to move over to a new shape; rather one switches to the background on which both of the shapes are located. In other words, 'figures' have a different status for attention in comparison to the 'ground' from which such figures are segregated. Recent neuropsychology has suggested that the visual area LO may be responsible for generating the figural representations that influence the allocation of attention (de-Wit, Kentridge, & Milner, 2009a) in the Egly et al. paradigm. The differential sensitivity to figure vs ground structure in terms of the movement of attention in this experiment is accordingly consistent with the fact that representations in LO are themselves sensitive to figure-ground relationships (Zourtzi & Kanwisher 2001, Vinberg & Grill-Spector, 2008). This interpretation in terms of the figure ground status of the target is however not the only one available, for example when shifting attention between objects at the middle location the re-allocation of attention might be easier, because the target is not 'crowded' by surrounding lines. Further experiments in which this confound is controlled will be required before one can conclude for certain that it is the figure ground status of the new item to be selected that modulates the between object switch cost.

In summary, the present results do not provide any evidence consistent with a spatial mode of selection or integration in the Egly et al. within-vs-between-object advantage. Rather they suggest that a critical part of the standard effect does not reflect an advantage for targets located within an object, but rather reflects a shift cost that may be differentially sensitive to the figure ground status of the next target of selection. The results therefore contradict a common assumption in terms of the underlying interpretation of the within vs between object advantage in a prominent object based attention paradigm.

Chapter 3 - Is the extraction of visual proto-objects achieved pre-attentively?

Abstract

It is commonly assumed that the influence of visual proto-objects upon attention reflects a pre-attentive stage of visual scene segmentation. In the context of the trivial segmentation demands required in many ‘object based attention’ phenomena and recent evidence regarding the role of attention in early visual interactions, this work set out to explore the extent to which the representations that influence attention could truly be achieved in parallel across the entire visual scene. To achieve this aim the Egly, Driver and Rafal (1994) object based attention paradigm was adapted to test whether a ‘within object advantage’ would still manifest when the visual system was presented with multiple objects in unpredictable locations. The results reveal that the within object advantage emerged in a manner that was independent of both the number of objects presented and the hemifeild to which those objects were presented. The results suggest that the representations that drive the Egly et al. within object cuing advantage are indeed extracted in parallel across the visual scene, furthermore this result provides a first step in establishing that the Egly et al. within object advantage can still operate with more complex scene dynamics more akin to those faced in our interaction with the real world.

Introduction

Following Neisser (1967), the notion of a pre-attentive stage to visual scene perception was brought to prominence by Treisman and Gelade (1980), who employed evidence from visual search to argue that scene perception could be decomposed into two distinct stages. In this model visual search commences with a parallel stage of processing in which the entire scene is segmented into basic ‘features’. If target detection is not achieved on the basis of these features then second serial stage in which features are bound into conjoined objects. Wolfe (2003) has subsequently argued that a number of alterations to this model are required in order for it to explain the full range of visual search data. Wolfe has argued for example that the second stage does not simply select items extracted at the first stage arbitrarily, rather this selection is guided by the results of the first stage. Furthermore, Wolfe has also argued that the second stage may not be a truly serial process, but rather allow for the parallel processing of items that are selected sequentially from the initial stage of segmentation.

Whilst questioning the nature of the second stage, and its connection to the first, Wolfe and colleagues (Wolfe, Oliva, Horowitz, Butcher & Bompas 2002) have found evidence consistent with the notion of an initial parallel pre-attentive stage of segmentation. Wolfe et al. (2002) demonstrated that if the initial stage is made more difficult (by adding noise to the background) there is an increase in the latency of target detection. This increasing difficulty did not, however, increase with the number of potential targets, demonstrating that potential targets were not segmented one by one, rather the segmentation of features from their background appeared to occur in one parallel sweep for all items simultaneously. This therefore provides evidence for an initial all or none processing stage of segmentation in which any number of candidate (or proto) objects across the entire visual scene are parsed and ready to feed into the second stage of recognition.

This initial ‘preattentive’ stage opens up the possibility that later attentive processing will be influenced by the manner in which the visual scene is initially organised. Exactly such a phenonmenon is argued to be evident in examples of ‘object based attention’ (see Scholl 2001 for a review). In a number of different paradigms the allocation of processing resources has been shown to be influenced by the initial organisation of visual stimulation. These object based attention effects are thought to be preattentive, because they themselves influence the allocation of attention, and must therefore have been extracted ‘pre’attentively.

A classic demonstration of object-based attention was reported by Egly, Driver and Rafal (1994). These authors presented observers with two large rectangles, one either side of fixation. One end of one of the rectangles was illuminated (i.e., cued) immediately before the appearance of a target. Targets could occur either at the location of the cue, within the same object as the cue but at the opposite end, or in the other rectangle. Replicating the effect reported by Posner (1980), reaction time (RT) was shorter when the target appeared at the cued location. However, the crucial finding was that RTs were reduced when targets were presented within the same object as the cue relative to when the target was presented in the non-cued object. Egly et al. argued that the ability of these basic shapes to influence the allocation of attention demonstrated their pre-attentive extraction as objects (or proto objects, Driver et al. 2001).

There are, however, reasons to question the extent to which the visual grouping that drives Egly et al's cuing paradigm can truly be thought of as 'pre-attentive'. Goldsmith and Yeari (2003), for example, found that the initial spread of visual attention was critical to obtaining a within- versus between-object cuing advantage. More specifically they found that the within object cuing advantage only manifested when the participants' initial spread of attention was distributed across the scene, and was abolished when the participant had to focus on the centre of the display. This suggests that some degree of diffuse attention across the entire scene was required in order for objects to be constructed to a level that would allow them to influence the later allocation of attention. Goldsmith and Yeari posited that this evidence explains why object based attention effects had not been obtained for endogenous cues that draw one's focus to the centre of a display (Macquistan 1997). In the present context however their result indicates that the visual system might have to allocate some degree of attention to an object before that object can in turn influence the spread of attention.

This interactive relationship between attention and grouping accords with Driver, Davis, Russell, Turatto and Freeman's (2001, p. 90) contention that "most (perhaps all) of the literature on 'object-based' attention is in fact concerned with how segmentation processes constrain attentional processes, and vice-versa". Driver et al further note that although it represents a useful heuristic the distinction between a pre-attentive and attentive stage of visual processing represents a "gross oversimplification of biological reality". Elegant proof of this over-simplification is highlighted in another publication by this group in which they demonstrate that interactions in the early visual system are influenced by attention (Freeman, Sagi, Driver, 2001).

If the segmentation required to influence attention in the widely employed Egly et al cuing paradigm is not seen as an automatic bottom-up process but is viewed in this more interactive framework it raises the question as to whether this supposed object-based attention effect can play a genuine role within more complex dynamic scenes. This question is all the more pertinent if one considers the relatively trivial segmentation demands of the typical Egly et al paradigm, in which participants are presented with two outline shapes in predictable locations. Furthermore, the within-object advantage is also assumed to occur as a result of preattentive processes because the rectangles do not require attention and are task irrelevant. However, the fact that stimuli are task-irrelevant does not mean they are unattended. Indeed, as with many experiments assumed to index preattentive processes, the typical Egly et al. displays are so impoverished that it would be almost impossible not to attend to the rectangles. Thus, the predictable segmentation demands of the display make it unclear whether the within-object advantage would manifest with less predictable, more complex displays. A further aspect of the original experiment that might have encouraged the rectangles to be attended is the relatively long presentation time of the objects prior to the cue (i.e.1000 ms),

In the present study we sought to investigate whether the within vs between cuing advantage would manifest when the visual system was presented with multiple pairs of rectangles in unpredictable locations. This addition of multiple object pairs in varying locations not only allows us to question whether the visual system is able to perform more demanding levels of segmentation but also affords a comparison between the potential limits of this and other ‘object based’ phenomena. A number of other object based phenomenon are thought to be limited to something in the order of four objects (see Cowan 2001 and commentaries). In the Multiple Object Tracking paradigm (Scholl & Pylyshyn, 1999) for example participants track not just the location of four targets, but four grouped objects (Scholl, Feldman & Pylyshyn 2001). If the visual system is able to parse objects with unpredictable scene locations it may be the case that there is a fundamental constraint on the number of objects that can be represented that is commensurate with the limitations seen in other paradigms. Indeed if at least four objects can be constructed and tracked in the context of MOT, one could well predict that at least four objects should be parsed as potential units of selection in the context of the Egly et al. cuing paradigm. This comparison of the representational limitations in different object based paradigms should prove informative regarding whether or not these different paradigms tap fundamentally similar levels of representation.

In summary the current work seeks to address three questions: can objects be parsed automatically from a visual scene such that they can influence the allocation of attention when the number of objects and their location is unpredictable? If this is possible, is there a limit in terms of the number of objects that can be pre-attentively parsed? Finally, if there is such a limit will it be in line with the ‘magic number four’ that is apparent in other object based paradigms?

2. Experiment One

2.1 Method

Thirty-one students of the University of Durham completed Experiment 1 in exchange for course credits. Each participant completed 256 trials. On each trial the participant was first presented with 2, 4 or 6 rectangles (1, 2 or 3 pairs: see figure 1). The locations of the rectangles were randomly distributed over a 18 by 18 deg area of the screen. The location of each rectangle was determined separately, so it was possible for the rectangles to overlap. The rectangles were present for 500ms, after which a cue was presented at the end of one rectangle of a particular pair. This cue was present for 175ms, followed by a 50ms gap, followed by a red or green target either at the same location as the target (valid condition, 75% of trials), at the other end of the same rectangle (invalid within, 12.5% of trials) or at the same end of the other rectangle in that pair (invalid between, 12.5 % of trials). After the participant's response there was a 500ms pause before the next trial. Throughout the experiment participants were instructed to maintain fixation on the cross in the centre of the display.

Trials with inaccurate responses, or with reaction times two standard deviations above or below the mean for each subject in each condition were removed from further analysis. Stimulus presentation was controlled from C++ using the software package *3D State*. *DirectX* was used to record keyboard responses. Stimuli were presented on a Windows PC, using 17 inch monitors.

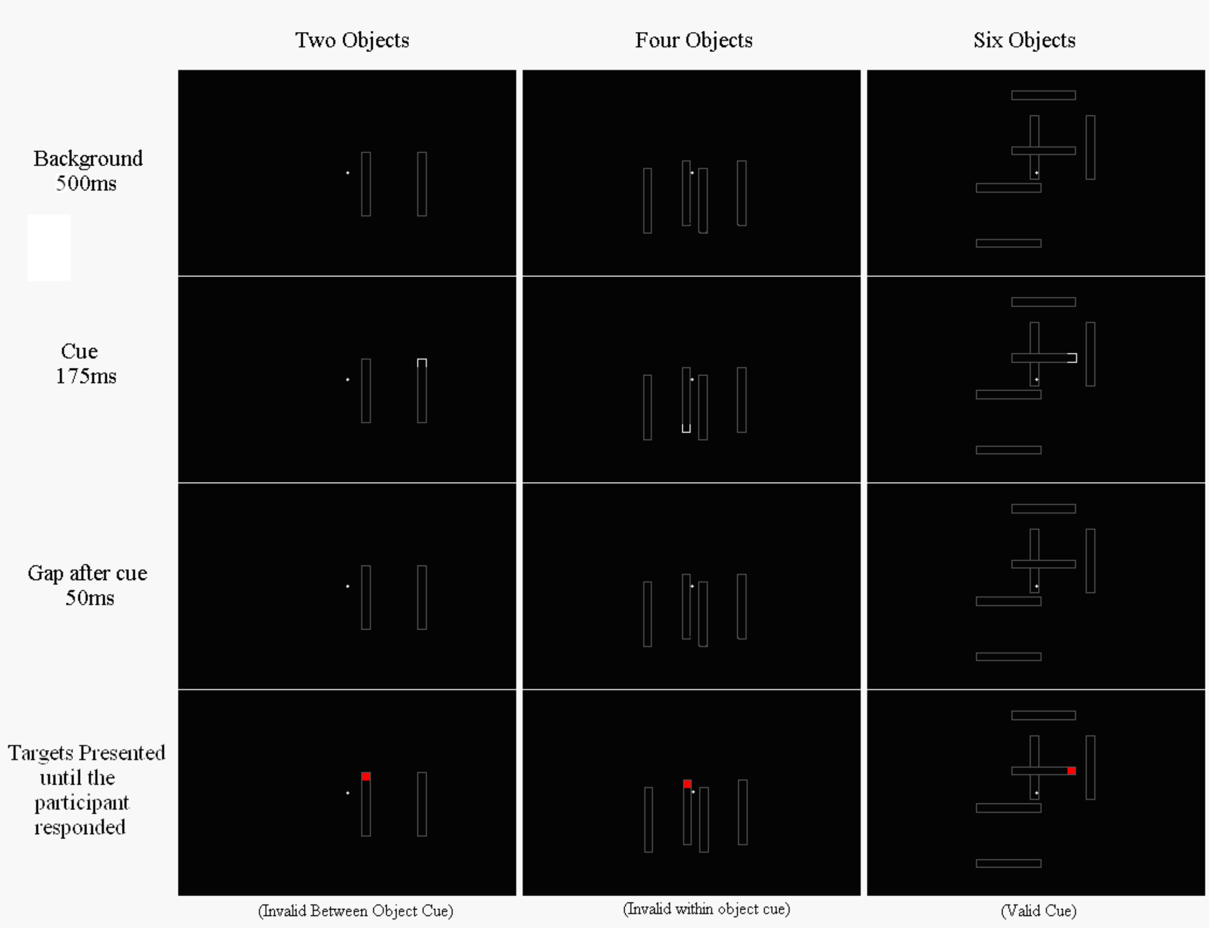


Figure One:

The rectangles (0.9 by 6.4 deg, and 0.1 deg thick) were presented in a 18 by 18 deg area in the centre of the screen. Cues appeared as a lightening of the end of one of the objects, followed by targets (0.7 by 0.7 deg), the angular distance between the centre of the cue and target on invalid trials was 5.5 degrees. Throughout the experiment a 0.3 by 0.3 deg fixation cross was presented, one pixel thick, in the centre of the display.

2.3 Results and discussion

Analysis was performed on 28 of the 31 subjects tested because three of the participants performed at chance (50%) in one of the conditions. The RTs for the valid, invalid-within and invalid-between object trials with two, four or six objects present are plotted in Figure 1. The critical question for the purpose of this experiment is whether the within vs between object difference interacts with the number of objects presented. As expected, the RT results reveal a straightforward replication of the classic Egly et al finding, with targets cued within the same object being recognised faster than targets following cues on another object (F(1,27)=8.2,p=0.008). There is however no interaction between this effect and the number of objects presented (F<1). Analysis of accuracy revealed no within between object sensitivity (F<1) and no interaction between cue type and number of objects (F<1).

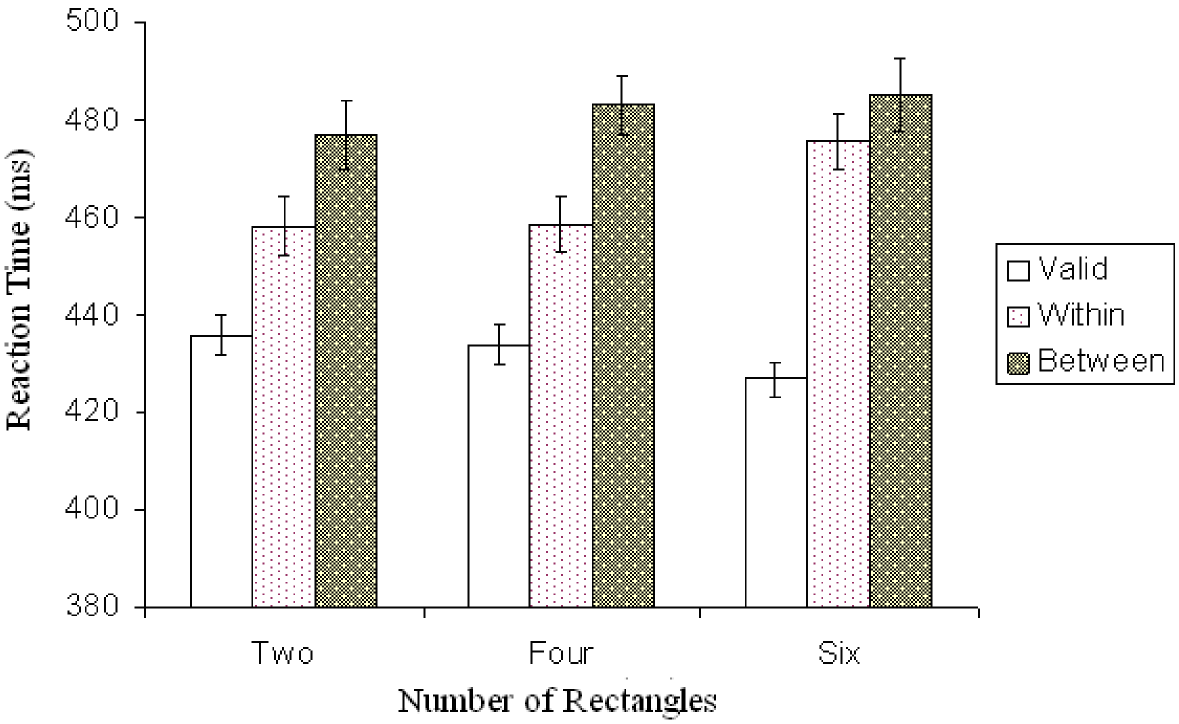


Figure 2, Reaction times (ms) and stardard error's for the Valid, Invalid Within and Between conditions with two four or six objects present.

The absence of any interaction between the number of objects presented and the strength of the object based attention effect suggests that the visual system can not only parse pairs of objects in unpredictable locations, but it can achieve this for up to six objects. The RT means however suggest that the strength of the object based effect begins to reduce with the presentation of more than 4 objects. Although not supported statistically from the results of Experiment 1, there could be a number of reasons why an interaction did not occur. First the experiment only employed a modest increase in the number of objects. If the visual system is able to automatically parse up to 4 objects, then it would still be able to parse two thirds of the objects in the six object condition, and this could be sufficient to drive an object based effect in this condition. Second, the locations of all rectangles were randomised independently such that they could overlap. It follows therefore that in the trials with three pairs of rectangles there is a greater probability that additional segmentation boundaries will be present between the rectangles on which the targets are presented. Thus even if the 'target' rectangles have not themselves been parsed the boundaries in between them may well act as a sufficient segmentation cue to reduce the efficiency with which attention can move (Avhrami 1999). This additional effect of the distracting rectangles could therefore have masked any potential decrement in the strength of the object based attention effect in the six-object condition.

Experiment 2 sought to address both of these issues, first by increasing the range of objects from 2, 4 and 6, to 4, 6 and 12, and second by stipulating that the rectangles could never overlap. In addition to these changes we also reduced the time for which the rectangles were presented from 500 to 300ms. As we suggested in the general introduction the 1000 ms employed by Egly et al., together with the presentation of only two objects, may have encouraged the objects to be attended. A large number of objects presented for only 300 ms before the cue would have reduced the possibility of the objects being attended.

In Experiment 1, the locations of the 'target' and 'distracting' rectangle pairs were not recorded. However if the location of the target pair was available this would afford an additional analysis of any potential lateralisation of the Egly et al cuing effect. The original papers outlining object based cuing effects emphasise that the effect is lateralised. Indeed, the basic experiment has been undertaken by patients with left or right hemisphere lesions (Egly, et al. 1994) and split brain patients (Egly, Rafal, Driver & Starrveveld 1994). Results showed that the object based cuing effect is dependent chiefly upon neural substrates in the left hemisphere. This focus on the potential lateralisation of object based cuing effects has received relatively little attention in the work that has followed. As highlighted by Egly, et al, the potential differential lateralisation of object based attention could have important implications for the clinical understanding of why certain patients might develop deficits in specifically object based forms of attentional orienting. In Experiment 2 therefore the location of the target pair of rectangles was recorded so that the strength of the within vs between object advantage could be calculated independently for stimuli in each visual hemifield.

3. Experiment Two

3.1 Methods

Except for those aspects detailed below the methods were the same as those for Experiment 1. Each participant completed 4 practice and 288 trials, the four practice trials were randomly selected from the different trial conditions. On each trial the participant was first presented with 4, 8 or 12 rectangles (2, 4 or 6 pairs: see figure 3). The locations of the rectangles were distributed over a 16 by 16 degree area of the screen. To avoid any overlap between each pair of rectangles each pair was assigned (with a small random offset) to one of 16 locations in a virtual 4 by 4 grid across the screen. The rectangles were present for 300ms, after which a cue was presented at the end of one rectangle of a given pair of rectangles. This cue was present for 75ms, followed by a 100ms gap, followed by a red or green target at the same location as the target (valid condition, 33.3% of trials) at the other end of the same rectangle (invalid within, 33.3% of trials) or at the same end of the other rectangle in that pair (invalid between, 33.3 % of trials). The higher ratio of invalid trials in Experiment 2 was employed to increase the number of trials contributing to the mean in each of these conditions without increasing the overall length of the experiment. After the participant's response there was a 500ms pause before the next trial. Throughout the experiment participants were instructed to maintain fixation on the cross in the centre of the display.

Trials with inaccurate responses, or with reaction times two standard deviations above or below the mean for each subject in each condition were removed from further analysis. Stimulus presentation was controlled from C++, and *DirectX* was used to control stimulus presentation and response collection. Stimuli were presented on windows PCs, using 17inch monitors.

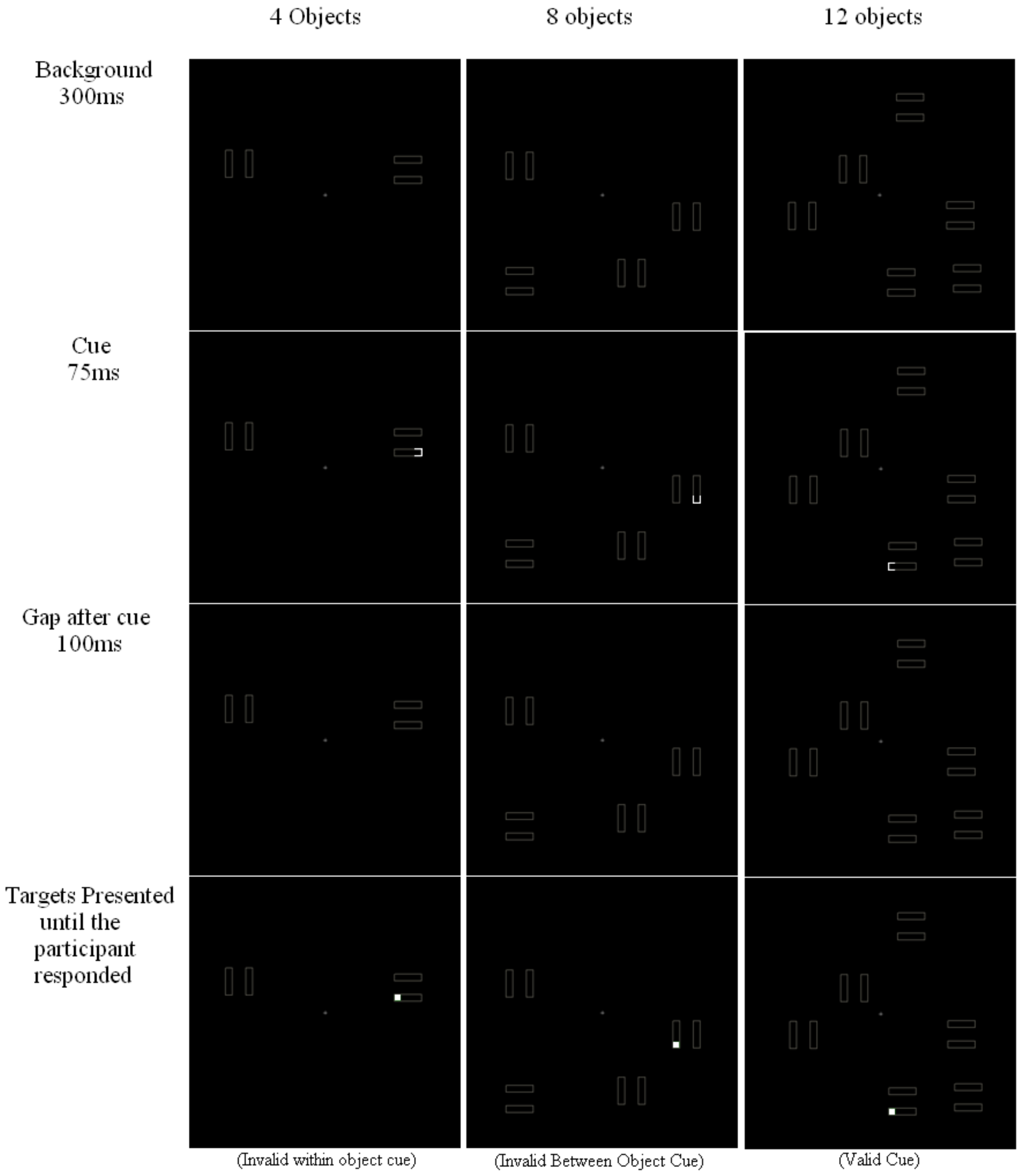


Figure 3. Presentation, cuing and targets on 4 8 and 12 rectangles. Each rectangle measured 0.7 by 2.9 degrees of visual angle and was 1 pixel thick. The target square was 0.6 by 0.6 degrees and the central fixation cross was made up of two lines 0.4 degrees long and 1 pixel thick. The target is illustrated with a white square, but was red or green in the experiment.

3.2 Results and Discussion

The reaction times for the different cuing conditions with differing numbers of objects are presented in Figure 3. As with Experiment 1, a repeated measures ANOVA revealed a robust difference between within vs between object trials (F(1,39) = 8.93, p=0.005). Again as with Experiment 1 this effect did not interact with the number of objects present (F<1). Also, the accuracy of participants' responses was insensitive to cue type (F(1,39)=1.28,p=0.264) and again showed no interaction between cue type and the number of objects (F<1).

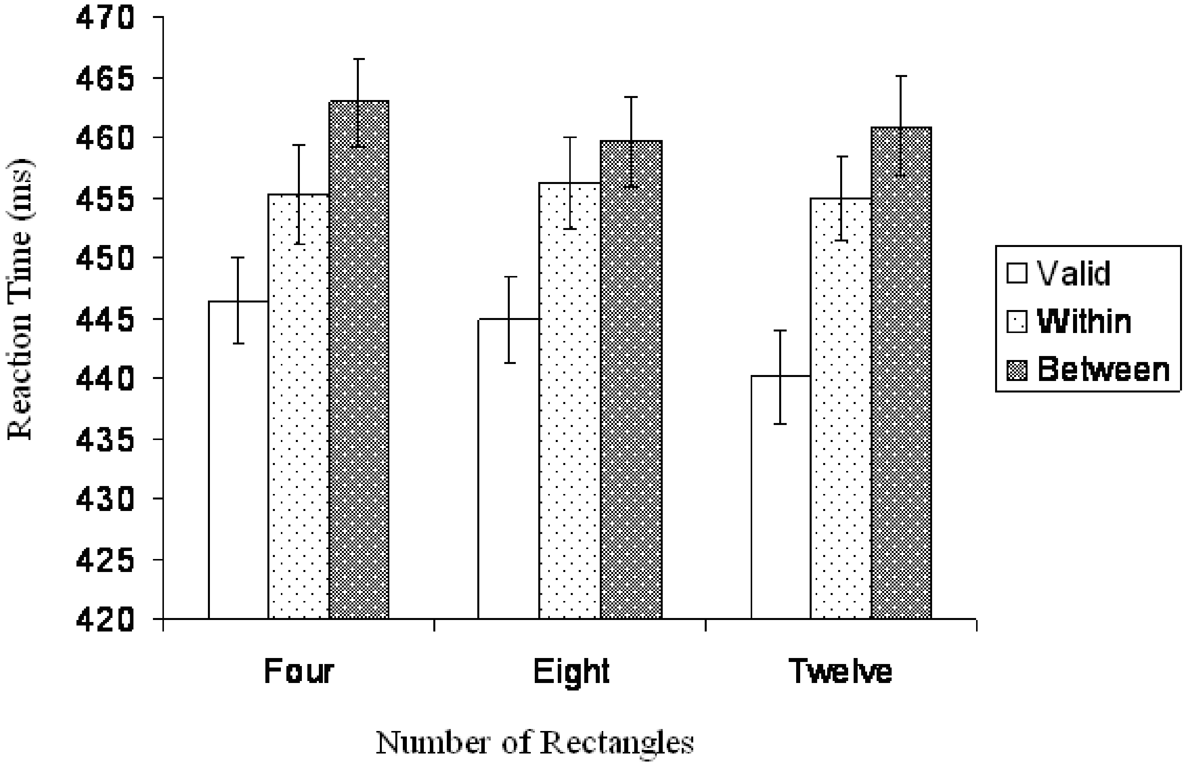
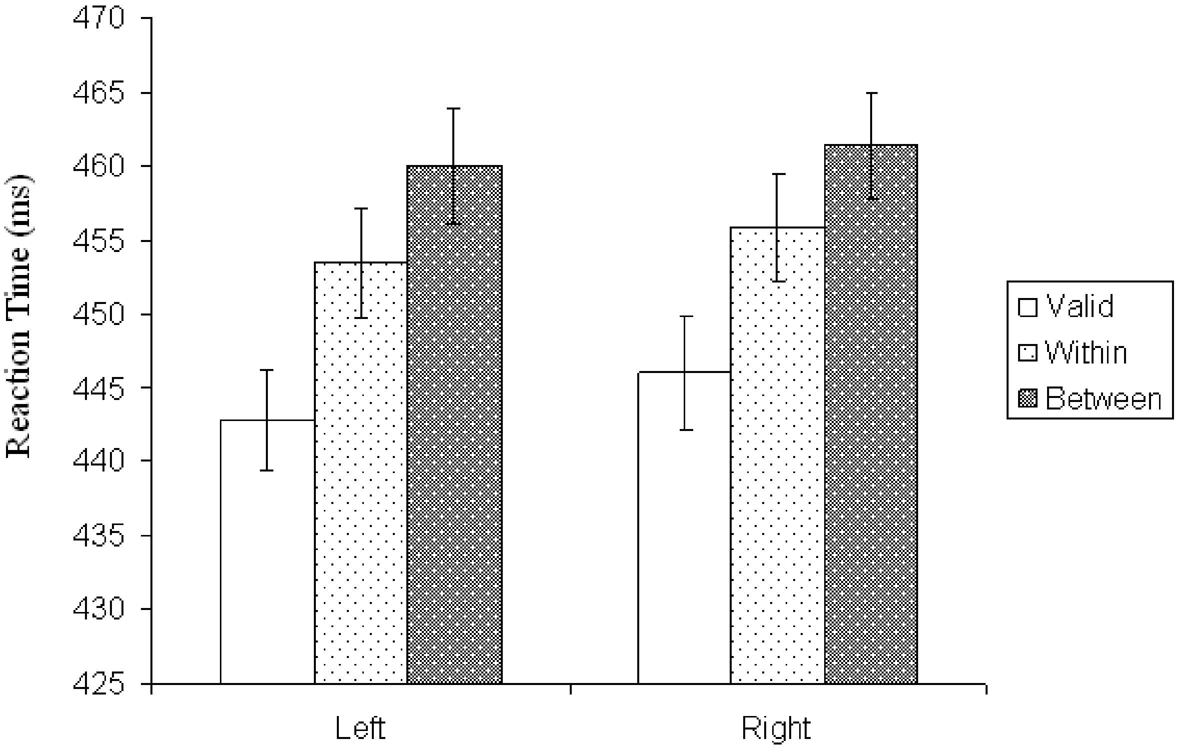


Figure 4, reaction times for the valid, invalid within and between cuing conditions with four eight and twelve rectangles.

When the RT data are analysed according to the hemispheric location of the rectangles upon which the targets are presented, the sensitivity to the within vs between object cuing conditions (F(1,39)=11.67, p=0.002) does not interact with the left vs right hemifield location of the target pair (F<1). The accuracy of participants responses is insensitive to the distinction between within and between object cues (F<1) and there is no interaction between within and between object cues and the hemifield of presentation (F<1).



In summary the results from Experiment 2 provide evidence that the distinction between within vs between object cuing can manifest for up to twelve objects (six pairs) and this effect manifests equally for object pairs presented to the left or right of fixation.

4. General Discussion

The object based advantage in the Egly et al. cuing paradigm is traditionally interpreted as reflecting a pre-attentive segmentation of the visual scene. The extent to which the representations that drive object based attention (Goldsmith and Yeari 1999) and early visual interactions (Freeman et al. 2001) are independent of attention has however come under question. Given both the potential attentional demands of low level segmentation and the trivial segmentation requirements of the typical object based attention experiment we sought to investigate whether the within-object cuing advantage could manifest with more complicated scene dynamics. Furthermore we sought to compare the potential limitations in the number of objects that could be segmented in the Egly cuing paradigm with the known limits of approximately four objects in other paradigms (Cowan 2001).

In two experiments we found that the within vs between object cuing advantage still manifested when the location of the relevant pair of rectangles was varied randomly from trial to trial. Moreover, even when the participant was presented with multiple objects (up to 12), any of which could serve as the site at which the participant was presented with cues and targets, the advantage for targets that followed cues within the same object was still apparent. This result suggests that the visual system is able to rapidly parse multiple shapes, in a manner that is sufficient for these to influence the later spread of attention. This result sits in clear contrast with other object based phenomenon in which the ability use objects to guide the allocation of attention seems to be limited to approximately four objects (Cowan 2001). This result could be used to suggest that the with-object advantage in the Egly et al. paradigm in fact reflects an earlier level of segmentation, which provides the proto-objects used in other, perhaps more cognitively complex, paradigms. Before exploring this interpretation however it is necessary to head off a number of possible counter explanations.

Perhaps the most critical of these is the possibility that the number of objects in the scene does not influence the strength of the effect because the representation that drives the effect is actually achieved after the presentation of the cue. Given that only one cue is ever presented on each trial this would mean that in effect only the two ‘objects’ proximal to the cue would need to be extracted on each trial. Thus one could argue that the objects are not simultaneously processed before the cue, rather it could be that the presentation of the cue prompts both the extraction of, and allocation of attention within, the objects proximal to it. This interpretation however does not accord with previous findings that demonstrate that objects have to be present for a certain amount of time if they are to influence the later allocation of attention. Chen and Cave (2008) found that when followed by a 120ms cue, object based attention effects were evident with objects presented for 1005ms, but not for presentation times of 120ms. Chen and Cave’s result highlights that even with a 120ms duration cue, more than 120 ms of prior processing is needed before shapes are extracted to a sufficient levelthat they can influence the allocation of attention. In the current experiment if one assumes that the objects are not extracted during the presentation phase (500ms in Exp1, 300ms in Exp2), one would have to argue that the extraction of the objects, and the distribution of attention that that causes would have to take place in 225ms (175ms cue, 50 gap in Exp 1) and 175 (75ms cue 100ms gap in Exp 2). Both these times are shorter than the 240ms (single objects presented for 120, followed by a 120ms cue) that proved insufficient in Chen and Cave's experiment, suggesting that there simply is not enough time in the current experiments for an object pair to be extracted following the presentation of the cue, if one assumes there was no pre-attentive processing of that object pair prior to the cue presentation. In the context of Chen and Cave's results therefore the present findings suggest that a significant degree of scene segmentation must be achieved pre-attentively to provide a sufficient representational level to influence the later allocation of attention. It is however important to highlight that Chen and Cave used an endogenous cue, whereas exogenous cues were employed in the current experiment. It is not possible, with the current data available, to rule out the possibility that an exogenous cue might prompt a more rapid segmentation of the objects proximal to it, in which case the current experiment might over-estimate the extent to which objects can be pre-attentively extracted.

Another potential interpretation that needs to be addressed is the possibility that the number of objects to be segmented was still too small to reveal any limitation. In particular one might argue that even when 12 rectangles are presented on the screen there are in essence only 6 objects on the screen (because each rectangle is predictably paired with another) and 6 is still very close to the limitations seen in other paradigms. For the within between effect to manifest however we would argue that each rectangle in each pair has to be parsed as a separate unit of attention, if each pair were treated as one object, it is not clear why a within between object advantage would emerge.

If one accepts the arguments against the two caveats raised above, the present results suggest that the segmentation that drives the Egly et al within-object advantage can be achieved independently of the number of units extracted (at least up to 12). This finding is consistent with results from visual search suggesting that there is an early parallel process that provides representations for further processing that proceeds in a manner that is not constrained by the amount of information to be processed. This notion is congruent with Driver et al's (2001) conceptualisation of the mechanisms that drive object based attention as an early process of scene segmentation. Recent research on the neural basis of object based attention suggest that it is dependent upon the later extraction of visual shape in an area of the ventral stream (de-Wit, Kentridge, Milner 2009a, Martinez et al. 2006) rather than the early registration of basic form properties in the primary visual cortex. Taken together these results suggest that there is a feedforward sweep between the primary visual cortex and LO that allows for the parallel extraction of shape representations across the visual scene. This picture is consistent with the role LO plays in extracting modal completed figures (de-Wit, Kentridge, Milner 2009b, Murray et al. 2002) and the fact that modally completion can also be computed in a parallel scene wide fashion (Davis and Driver 1997).

Is this notion of a seemingly limitless pre-attentive stage of proto-object extraction at odds with the limitations seen in other object based paradigms? For example in Multiple Object Tracking (MOT) participants seem to be limited to something in the order of four objects which can be effectively tracked (Scholl & Pylyshyn 1999). Why should participants be limited to about 4 objects in certain paradigms, whilst it was not possible to find a limit using the current paradigm? In fact it could be argued that more recent evidence from MOT in fact supports the notion of an unlimited parallel or pre-attentive stage of segmentation. Scholl et al 2001, for instance, demonstrate that although participants can only track around four items, those items are selected from a level of representation in which potential proto objects have already been constructed across the visual scene. More specifically Scholl et al. found that even if instructed to only attend to one element of a set of grouped shapes, participants could not help but attend to the object as a whole, suggesting that the individual parts of an object had already been put together before that object could be selected as a unit of attentional tracking. Thus whilst participants might only be able to track 4 objects in MOT, this selection is probably made upon a much larger number of potential object units that are segmented in parallel across the visual scene. The notion of pre-attentive parsing in MOT is consistent with recent data from our lab which demonstrates that the units of MOT are determined on the basis of ‘lower’ level grouping factors, rather than higher level factors such as their biological status (de-Wit, Lefevre, Kentridge, Rees, Ayse, under revision). Thus the pre-attentive parsing apparent in the current study, and the proto object construction apparent before the selection of 4 items in MOT, could in fact be seen as complementary rather than contradictory findings.

One could argue that a fundamental difference between the two paradigms is not the level of parallel scene processing that is required, but the number of units provided by that process that have to be selected or indexed. In the MOT paradigm multiple objects have to be selected, whereas here although multiple potential units have to be processed only one of those has to be selected. Framing the Egly et al. and MOT object based paradigms in this manner brings to light a new research question regarding whether or not the within-between cuing advantage would still occur when more than one object is cued at a time. If one extrapolates from the current argument, it is possible that one could not only extract multiple objects, but that attention could be cued to, and simultaneously select for up to four objects at one time. Clearly this possibility deserves further empirical research.

This account of the current findings, as reflecting an early pre-attentive parsing process that provides units of selection for further processing, is consistent with the fact that this effect is manifest equally for both hemifields. If the process studied here is critical in providing the units for further perception and cognition it is only logical that those units should be extracted equally for both hemispheres. It is important to note that our results do not necessarily contradict the previous findings regarding the lateralisation of the critical neural substrates for engaging in object based attention (Egly, Driver and Rafal 1994, Egly, Rafal, Driver and Starrveveld 1994). It does however show that any potential lateralisation does not impact performance in healthy observers. In the context of potential lateralisation effects in object based attention it is interesting to note that Alvarez and Cavanagh 2005 (see also Chakravarthi & Cavanagh 2009) have recently demonstrated that resources for MOT seem to be to some extent separate for stimuli presented in separate hemifields, such that tracking four objects in one hemifield is more demanding than tracking two on the left and two on the right. If, as suggested above, we could find evidence that Egly et al. within object advantage is able to develop in multiple objects it would clearly be of interest to explore whether such a multi-focal cuing effect would be more effective for stimuli presented in separate hemispheres.

5. Conclusions

What do the current results allow us to conclude about the functional status of the representational process tapped by the within vs between object advantage reported by Egly et al (1994)? Although this process of selection does not seem to play an important role in the guidance of manual actions (de-Wit et al. 2009a), it could be argued to play an important role for establishing a ‘first parse’ representation of the visual scene that allows for the selection, tracking, and further processing of the units established by that first parse (Driver et al. 2001). Both the trivial nature of the representational demands of the typical Egly et al. within object advantage, and the increasing recognition of the interactive dynamic between attention and representation led to the question of whether the representations tapped could really play a genuine role in aiding visual perception and cognition in our daily life. In order to address this concern the Egly et al. paradigm was adapted in two experiments such that a larger number of objects (2-6 in experiment 1 and 4-12 in experiment 2) were presented in random and unpredictable locations. The Egly et al. within vs between cuing effect remained equally present across these different conditions. This result provides a first step in establishing that the Egly et al. paradigm does indeed tap into a process of visual scene extraction that could play a genuine role in our daily perception. However whilst the object extraction must to some extent occur in parallel our results do not imply that it is totally independent of attention, given that Goldsmith and Yeari (2003) have shown that when attention is focused away from the objects they may not be extracted to a sufficient level to influence the later allocation of attention. Nevertheless with attention diffusely distributed across the scene it was not possible to find a limit on the number of objects that could be extracted. At first glance this effect might seem incongruent with the limitations seen in other paradigms, in which the number of objects that can be simultaneously extracted seems to be around 4. It could be argued however that even though there is an approximate limit of four objects that can be selected for multiple object tracking, that selection in fact occurs on the basis of units that are provided via a parallel sweep analogous to that apparent here. This means of reconciling the MOT and Egly et al. paradigms opens the possibility that the within object advantage could manifest when more than one object is cued. Indeed one might predict that in exploring the number of objects that can simultaneously influence the allocation of attention, one might find a similar limitation to that observed in other paradigms. Finally it appears that the potential lateralisation of the neural resources important in object based attention does not influence the manifestation of this phenomenon for healthy participants, and that objects are extracted and attended with equal ease across both hemifields.

Chapter 4 - Object Based Attention and Awareness

Abstract

Is conscious first person experience automatically evoked by certain forms of information processing? The plausible assumption, that the allocation of attention might lead to the representations being selected becoming conscious, has recently been questioned. This paper sets out to explore whether the construction of objects, and the selection of those objects as units of attention would automatically evoke conscious representations of those objects. The results failed to find any evidence that a certain form of object construction, operationalized in terms of the construction of illusory Kanizsa figures, could take place in the absence of conscious experience. This result is interpreted in terms of the necessary role of neural resources in the ventral stream, influenced by the manipulation of conscious experience in this experiment, in the construction of illusory Kanizsa figures.

Introduction

Does consciousness have a functional role in cognitive life? As neuroscientists search for anatomical structures or neural processes that are critically associated with consciousness (Crick 1994), is it possible for cognitive scientists to identify computational processes that are intimately associated with consciousness? Indeed is it possible to identify aspects of cognition that cannot proceed without conscious awareness? Can we, for example, represent a goal, or hold information in working memory, without necessarily being conscious of that information?

One of the cognitive processes most consistently linked with awareness is attention. In some senses attention could be thought of a general process of selection, common to any moving organism, in which important aspects of its environment have to be selected either to approach or avoid. For organisms like humans however, whose decisions to act are derived from complex internal models of the world, rather than simple stimulus response mapping contingencies, this process of selection has to be internalised. At the same time as the brain builds a complex model of the external world, so too it has to bias this model to focus on only those aspects of information processing that are most critical to determining its next response.

It could be argued that it is not the complex nature of information processing that dictates the nature of conscious awareness, but rather it is the manner in which this internal processing is biased and coordinated by attention that determines what we are conscious of. These complex co-ordination processes can be studied at many levels, most often however psychologists explore attention in terms of the biasing of sensory signals, in the auditory or visual domains. These constrained and empirically tractable aspects of cognition enable us to gain some insights into the nature of attentional selection.

What have these early levels of internal selection revealed about the relationship between attention and awareness? Do these early stages of selection play a special role in defining what we become conscious of? Is attention a special cognitive operation that requires awareness or that plays a special role in the generation of conscious experience? In one sense the answer to the last question must be no. Taken together recent neuropsychology (Kentridge, Heywood, Weiskrantz 1999), cognitive neuroscience (Wyart & Tallon-Baudry 2008) and cognitive science (Kentridge, Nijboer & Heywood 2008) has highlighted that we can bias the processing of information via purely internal means without becoming conscious of either the cues that elicit attentional effects or stimuli whose representational status is enhanced.

This exploration of the role of attention in relation to consciousness very much complements the development of our understanding of attention per se, which has progressed over the years from the concept of an all or nothing gate-keeper to higher level information processing (Broadbent 1958), to a process of biasing the competition for processing resources at multiple levels within the visual system (Desimone & Duncan 1995). Indeed framed within Desimone and Duncan’s model for understanding attention, the concept of a one to one relationship between attention and awareness is hard to conceptualise, given that attention is viewed not as a single mechanism or process, but as a reflection of a biasing of information processing that occurs at many levels of information processing. This more nuanced notion however leaves open the possibility that although one can find levels at which information processing can be biased without requiring or leading to awareness, there may be other levels of selection, or other ways in which information processing can be biased, which would necessitate either that that information is conscious before selection, or that it must become conscious after it is selected.

In this context it is interesting to note that Mole (2008) has argued that previous attempts to dissociate attention and awareness have fallen short exactly because they only investigate the biasing of the representation of an area of space. They do not (Mole argues) explore what happens when the human visual system selectively attends to an *object*. Although developed from a philosophical standpoint Mole's arguments clearly feed into a distinction within cognitive visual science regarding the relationship between object-based and spatially based attention.

One could envisage a purely spatial form of visual selection, such that one could enhance information processing by directing a 'spotlight' to the maps of visual information present at different levels of the human visual system. It has become clear however that the biasing of information processing is not applied to a uniform area of space but is able alter the representation of one object or another, even when those objects are overlapping at the same spatial location (Duncan 1984).

There is a vast array of phenomena that are thought to reflect an object based form of selection. One of the most commonly employed phenomena is the Egly, Driver and Rafal (1994) cueing paradigm, in which a cue-target pairing on one object leads to faster target detection compared to a cue-target pairing on separate objects. In a recent, and compelling example of this form of information selection Mitchell, Stoner & Reynolds (2004) showed that in situations of binocular rivalry (in which incompatible stimuli are presented separately to each eye, such that only one of them can become conscious) the selection can act upon objects. In all of these object based phenomena however the objects to be selected are clearly conscious to the participant. Can objects be selected as a unit that biases information processing even when one is not conscious of them?

There are many methods in cognitive science for reducing the conscious access to aspects of visual stimuli (Kim and Blake 2005), including crowding, inattention blindness, and binocular rivalry. One particular form of binocular rivalry, called Continuous Flash Suppression, has proved to be extremely effective in rendering stimuli invisible for extended periods of time (Tsuchiya & Koch 2005). Continuous Flash Suppression, henceforth CFS, involves presenting incongruous stimuli to each eye, but biasing the competition that ensues by presenting a continuously changing Mondrian mask stimulus to one eye. This bias can result in information in the ‘suppressed’ eye never reaching conscious awareness; even if information presented to the other eye does not become conscious; however, it is clearly registered within the visual system, and leads to measurable behavioural and neural consequences (Tsuchiya & Koch 2005). Yang, Zald, & Blake (2007) have recently used CFS to explore the pre-conscious representation of emotion. More specifically Yang et. al. showed that when the strength of the CFS masking was adjusted such that participants could sometimes become aware of the ‘suppressed’ stimuli, they found that faces with fearful emotions were detected consciously more rapidly than neutral or happy faces. This suggests that this ecologically important property was extracted, and that it attracted some form of enhanced information processing, even when it was invisible.

The current project seeks to use an analogous methodology to ask, a) if objects are extracted and reach consciousness more often than non-objects and b) will these unconsciously extracted objects have a different effect on attention?

Up until this point however this discussion has somewhat side stepped the question of what constitutes a coherent object for the human visual system. Whilst highlighting this issue, we will not attempt to fundamentally answer this question, but will employ a well studied means of operationalizing an object, evident in the distinction between coherent vs incoherent Kanizsa stimuli, that is, stimuli which either do or do not cohere to form an illusory shape. This distinction not only provides a perceptually compelling distinction between a coherent and incoherent shape but also clearly has a significant impact on a range neural responses (Tallon, Bertrand, Bouchet & Pernier 1995, Seghier & Vuilleumier 2006). More-over neuro-imaging and neuropsychology suggests that this form of object extraction relies upon areas in the human ventral visual stream (Murray et al. 2002, de-Wit, Kentridge & Milner, 2009b) that are also critically involved in the influence of object representations upon attention (Martinez et al. 2006, de-Wit, Kentridge & Milner 2009a).

This research project therefore uses coherent vs incoherent Kanizsa stimuli, to explore whether coherent objects accrue some form of enhanced bias in the competition for representation induced by binocular rivalry. Our expectation here is that, like fearful faces in Yang et al’s study, coherent Kanizsa patterns should be detected consciously more often than incoherent Kanizsa patterns. In and of itself however this would merely demonstrate that coherent Kanizsa are extracted pre-consciously, and that this pre-conscious representation results in a higher representational status within the visual system. Demonstrating this would of course be valuable in establishing that the cognitive operations involved in extracting an object do not require consciousness. The focus of this experiment however is to establish whether the selection involved in object based attention is dependent upon the conscious representation of those objects. In order to explore this possibility this project will also explore what happens when attention capturing cues are placed in the centre of suppressed coherent and incoherent Kanizsa patterns; if object based attention requires awareness of the objects then any attentional advantage of the cue would be equally present for coherent vs incoherent Kanizsa stimuli. If however object based selection can operate before one is conscious of the object then one would expect that coherent Kanizsa patterns would show a larger cuing effect and become conscious more often.

Methods

Participants

Eight observers (2 men, including the author) were recruited. All had corrected or corrected to normal vision. All (except the author) were naïve to the purpose of the study.

Procedure

Using a custom made mirror stereoscope participants were presented with a different image to each eye. Before the experiment started participants adjusted the mirrors to enable the efficient fusion of the images presented to each eye. Throughout the experiment each eye was presented with a central fixation cross and fusion aiding checker boards against a grey background. One eye was presented with a CFS mask, consisting of 1600 blue or yellow rectangles of varying hue and luminance, ranging between 1 and 0.1 degrees of visual angle, presented within an 8.5 by 8.5 degrees area centred on the fixation cross. This mask changed every 100ms, and was present from the start to the end of each trial. After 400ms of CFS to one eye, a coherent or incoherent[9] Kanizsa pattern was presented to the other. The onset of this Kanizsa stimuli however occurred gradually over 800ms, with the luminance increasing

---

[9] There are many ways one might construct an incoherent Kanizsa, such that Pac-Men are presented so that they do not form an illusory square. The configuration used for this experiment is presented in figure 1. Although many incoherent configurations are possible, this one was used on every trial to eliminate any potential confounds in terms of the number of Kanizsa types in the coherent and incoherent conditions.

by 12.5% every 100ms. The centre of each Kanizsa pattern appeared 2 degrees to the left or right of fixation. The four pac-men forming each Kanizsa pattern had a radius of 0.3 degrees, and were separated by 1.5 degrees of visual angle. A 0.4 by 0.4 square red cue was presented to the same eye as the Kanizsa stimuli for 200ms. This cue was either valid, appearing in the centre of the four pac-men forming each Kanizsa pattern, or invalid, appearing 2 degrees to the other side of fixation. Both the Kanizsa stimuli and the red cue were removed with an abrupt offset at the same time. Depending upon the task the participant then had make a forced choice response based on a feature of the Kanizsa pattern. After each response there was a 600ms pause before the start of the next trial. During this pause the participant saw only the fixation cross and the fuse aiding checker boards. The eye receiving CFS or Kanizsa stimuli changed randomly from trial to trial. Likewise whether or not the Kanizsa pattern was presented to the left or right of fixation, whether the cue was valid, and whether the Kanizsa pattern was coherent or incoherent was determined randomly for each trial with a 50% likelihood. The luminance of the Kanizsa figure was determined individually for each participant before starting the primary experiment such that the participant would become aware of the Kanizsa stimuli but not such that they could score higher than 80% in a two option forced choice discrimination in the same vs different colour task described below. This 80% threshold was employed to ensure that the participant did not become conscious of the Pac-men too frequently (avoiding a ceiling effect). What is of critical interest in this analysis is whether or not coherent Kanizsa, and more specifically attended coherent Kanizsa, can be detected consciously more often than other stimulus types. The ideal result in this experiment would be that coherent and incoherent Kanizsa reach awareness on only a small subset of trials when not cued (for example leading to 55-60% accuracy), and that coherent (but not incoherent Kanizsa) prove to be much more capable of reaching consciousness when attended (for example 80% for cued coherent, 55-60% for cued incoherent Kanizsa). This result would prove that an otherwise unconscious stimulus when unattended (as indicated by a 55-60% detection rate) would be much more likely to reach consciousness when attended than an equivalent incoherent figure, thus suggesting a role for attentional selection that is differentially sensitive to the object status of the figure, but that operates on a representation that would otherwise have remained unconscious.

Every participant completed two slightly different tasks, the order of which was counterbalanced across participants. In task 1 the four pac men forming each Kanizsa stimuli were either all the same colour (blue or green) or one of the pac-men was formed of a different colour (ie 3 green pac-men and one blue, or visa versa). The participant was instructed to report whether the four pac-men were the same or different using their index and middle finger. Fifty percent of the pac-men were all the same, and fifty percent included a differently coloured pac-man, with a random allocation on each trial. In task 2 the four pac men could either all be blue, or one of the pac men could be green. In this experiment all the participant had to do was report (again using index and middle fingers) if there was a green pac-man present. Each participant completed 6 blocks of 48 trials for each task.

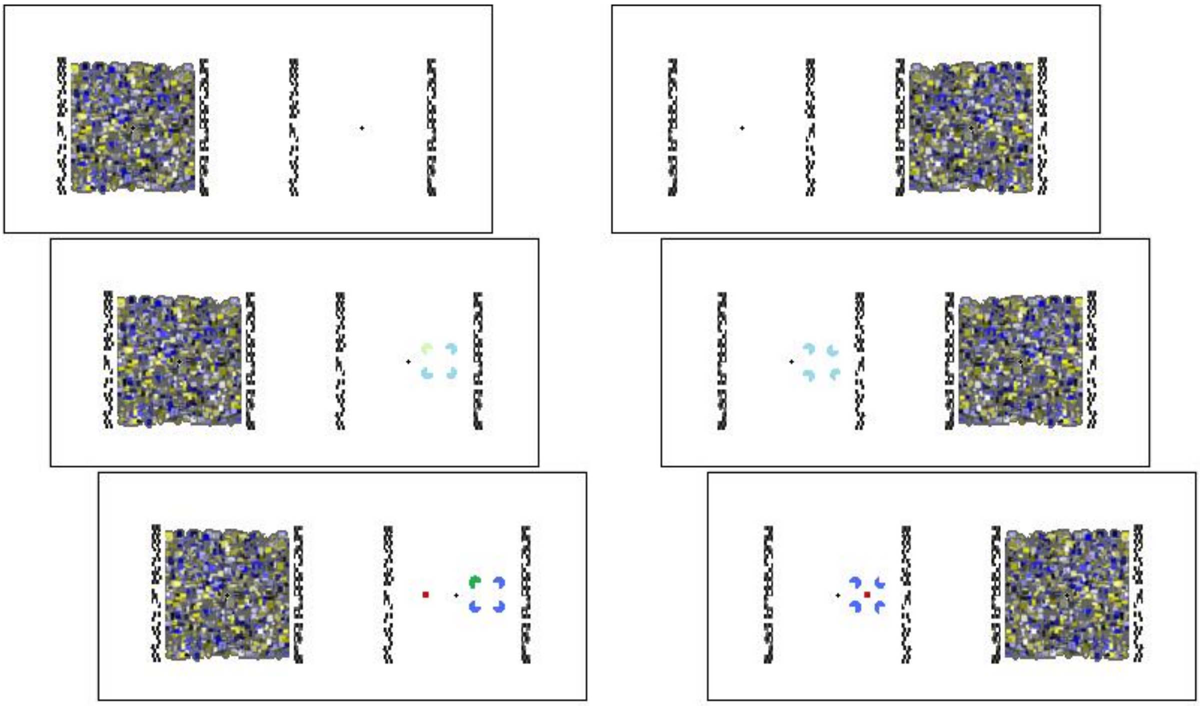


Figure 1. On the left is a coherent invalidly cued Kanizsa stimulus. On the Right is a cued incoherent Kanizsa stimulus. For the first 400ms the participant saw nothing but the CFS mask, randomly presented either to the left or right eye. Then for 800ms the Kanizsa pattern gradually onset. For 200ms a cue was then presented, either on the same or the opposite side of the fixation cross to the Kanizsa pattern. In task one, the participant would be presented with 3 blue pac-men and one green (as illustrated on the left), 3 green pac-men and one blue, or all green or all blue (see the stimulus on the right), and the participant had to report if the pac-men making up the Kanizsa figure were all the same colour or not. In task two, the participant would be presented either with all blue pac-men or with one green pac man, and the participant simply had to report if a green pac-man was present.

During both tasks the participant was instructed to focus on responding accurately rather than rapidly. The participant was also instructed that the red cue (although otherwise task irrelevant) should be used a cue as to when to make a forced choice response, and that if the participant had not seen a Kanizsa by the time the red square offset then they should guess at a response because the Kanizsa will have already offset. This instruction was given, both for the pragmatic reason explained to the participant, but also to ensure that the participant looked out for and made use of the cue, in the expectation that this might enhance the cues effectiveness in attracting the participant's attention.

Results

Given the similarity between the two tasks, the results were first analysed using a 2 by 2 by 2 factor within subject ANOVA, with Task entered a factor. There was however no main effect of Task (F(1,7)=2.52, p=0.156) on the accuracy data, and Task type did not interact with the other factors (all F values below 1). The results presented in figure 2 are therefore collapsed across the two tasks. A within subjects ANOVA on the factors Cue Type (valid vs invalid) and Coherence (coherent vs incoherent Kanizsa patterns), revealed a clear cuing advantage (F1,7)=19.25, p=0.003), a near-significant effect of Coherence (F(1,7)=5.45, p=0.052), and a clear interaction between the two (F(1,7)=33.97, p=0.001). As is clear from the graph below the near-significant effect of Coherence in fact reflects an unexpected advantage for the Incoherent Kanizsa stimuli, it is also clear however that this advantage manifests only in the cued condition.

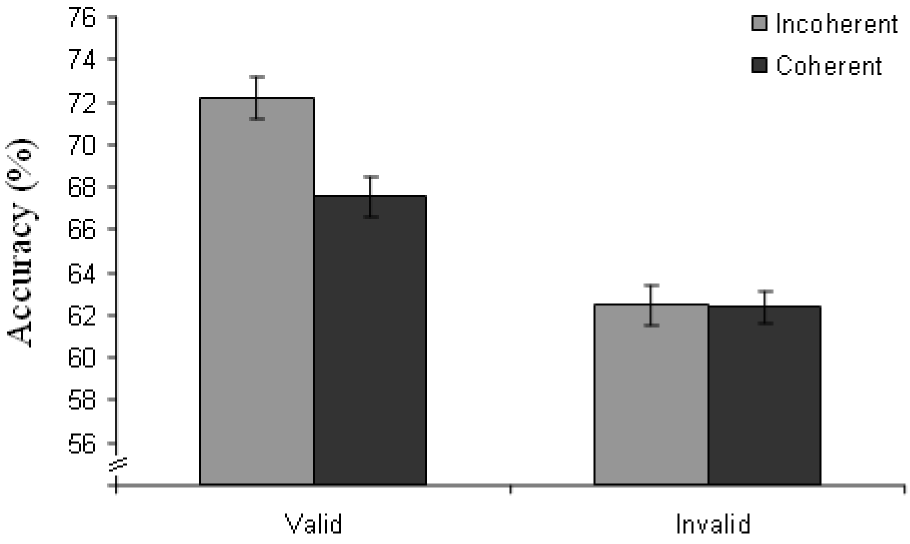


Figure 2, Response accuracy for valid and invalidly cued coherent and incoherent Kanizsa stimuli.

Post-hoc t-tests reveal no difference in the invalid condition (t(15)=0.054, p=0.957), and a clear advantage for incoherent Kanizsa patterns in the valid condition (t(15)=2.71, p=0.016). Looked at in terms of the strength of the cuing effect, it was clear that cuing was effective for both types of Kanizsa stimuli, but was more effective for the Incoherent (t(15)=4.73, p<0.001) compared with the coherent Kanizsa patterns (t(15)=2.68, p=0.017).

The participant was instructed to focus on responding accurately rather than quickly, and an analysis of the reaction time data reveals no significant effects, with one exception. Responses in the “Spot the Green” pac-man task were faster than those in the “Same vs Different” task (1230ms vs 1048ms, F(1,7)=19.93, p=0.003). This difference probably reflects the fact that the “spot the green” task is slightly simpler than the same vs different task.

Discussion

This experiment sought to investigate whether coherent objects could compete more effectively against a Continuous Flash Suppression mask. Furthermore it sought to explore whether coherent objects (even before they had become conscious) could be used as a more effective basis for attentional selection, therefore suggesting that ‘object based attention’ could operate upon object representations that are not present in conscious awareness. The results in fact proved quite counter-intuitive. There was no difference between the coherent and incoherent Kanizsa stimuli when invalidly cued. However whilst attentional cuing biased the representational status of both the coherent and incoherent Kanizsa patterns, cuing had a larger effect on the incoherent Kanizsa stimuli, leading to an advantage for the Incoherent Kanizsa patterns when validly cued. The lack of an advantage for coherent Kanizsa stimuli (cued or uncued) is highly surprising because coherent Kanizsa are known to pop-out (Davis and Driver 1989), thus if they are constructed, they should attract attention. The attraction of attention in a binocular rivalry context should lead to them competing more effectively against a continuous flash suppression mask, and there result in all of the information in the otherwise 'suppressed' eye being more readily accessible to consciousness. This discussion will first address why there was no advantage for coherent Kanizsa stimuli in the invalid condition, and will then turn to address why there was an advantage for validly cued Kanizsa patterns.

This experiment started with the intention to understand the ability of objects to act as a unit of attentional selection when those objects are not represented consciously. Given the challenges faced in trying to define what exactly counts as an object for the visual system, an object was operationalized in terms of the difference between Coherent and Incoherent Kanizsa stimuli. This choice was not arbitrary, and reflected an extensive literature exploring the effects of this distinction on the human visual system. Coherent Kanizsa stimuli have been consistently reported to reveal higher levels of activation at various levels of the ventral visual stream (Seghier, & Vuilleumier 2006). Furthermore coherent Kanizsa figures can be extracted in parallel across a visual scene (Davis & Driver 1998) implying that they do not require serial attentional processing. In fact more recently Senkowski, Rottger, Grimm, Foxe and Herrman (2005) have shown not only that Kanizsa figures can be extracted in parallel, but that they can automatically capture attention. Studies with neglect patients have also revealed that some degree of perceptual completion involved in the construction of a coherent Kanizsa figures can be achieved in the absence of both attention and awareness. Mattingley, Davis and Driver (1997) presented two pac-man to the neglected and two pac-men to the intact visual field of a neglect patient and showed that pac-men on the neglected side were detected more often when they grouped to form a coherent Kanizsa pattern with pac-men presented in the preserved side of visual space (see also Conci et al. 2009, for a replication with a larger sample). It should be noted however that the mode of presentation in these neglect studies was very different to that employed here, because two of the pac-men were presented in clear conscious view in the preserved side of visual space. It could be argued that the construction of the coherent Kanizsa was in some sense kick-started by the construction of an illusory shape in the preserved visual field and that this shape could then connect with implicitly processed grouping cues in the neglected field. In this context it could be very interesting to explore whether pac-men masked by CFS would be more likely to break through suppression when grouped with pac-men viewed consciously in the dominant eye.

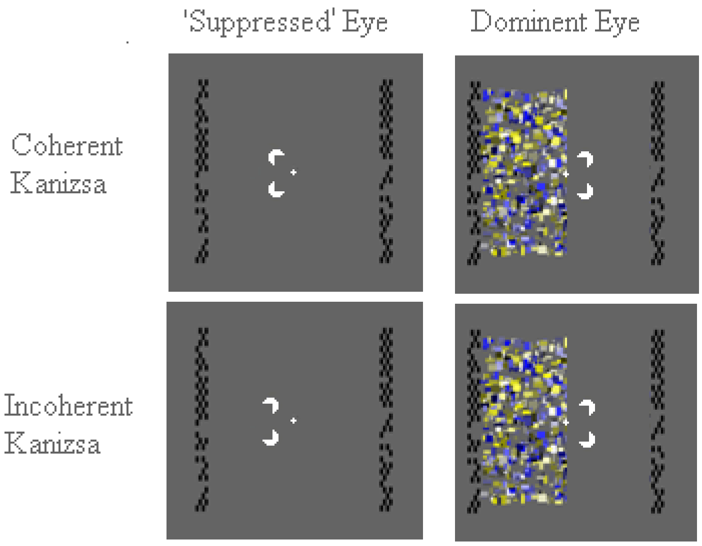


Figure 3, illustrating how Mattingley and Driver's (1997) finding with a neglect patient could be explored using CFS.

Mattingley and Driver (1997) argued that the extraction of grouping cues in the neglected visual field was consistent with neurophysiological evidence that illusory contours are represented at an early stage of information processing in the primary visual cortex (von der Heydt, Peterhans & Baumgartner 1984). More recent neuro-imaging and neuropsychology however has suggested however that the construction of coherent Kanizsa is dependent upon the shape processing area LO. Indeed Murray et al. (2002) argue that LO is critical to the extraction of a coherent kanizsa, and that the activation in the primary visual cortex must reflect feedback from this higher processing stage. This argument is consistent with neuropsychological data highlighting that although a motion illusion induced by coherent Kanizsa can elicit activation in the primary visual cortex (Seghier et al. 2000), this illusion cannot be perceived when LO is damaged (de-Wit, Kentridge and Milner, 2009b). This position is consistent with comparative research in monkeys highlighting that illusory shapes elicited from coherent Kanizsa patterns cannot be constructed when IT (thought to be the homologue to human LO) is damaged (Huxlin, Saunders, Marchionini, Pham & Merigan 2000).

Perhaps therefore the absence of an advantage for coherent Kanizsa patterns in the present study reflects their relatively late extraction within the visual system. Indeed it is clear that binocular rivalry can manifest at early stages of the visual system (Haynes and Rees 2005), so it is possible therefore that suppressed information cannot reach higher areas of visual perception. Fang and He (2005) have in fact argued that binocular rivalry restricts the flow of information from primary visual areas to the ventral stream (like LO) but still enables some degree of information processing in object processing areas of the dorsal stream. Perhaps this restriction of information processing to area LO, and the critical involvement of LO in the construction of illusory shapes from coherent Kanizsa provides a full account of the current findings.

Several studies using continuous flash suppression have however highlighted that high level stimulus properties can be extracted and affect performance. Jiang, Costello, Fan, Huang and He (2006) have for example found that gender and sexual orientation specific images in the suppressed eye can influence the allocation of attention. Similarly Yang, Zald and Blake (2007) have highlighted that faces showing fearful emotions will emerge from suppression before faces with happy or neutral expressions. How could these complex stimulus differences influence perception, when a coherent Kanizsa pattern cannot? There are two potential answers to this. One depends upon the experimental methods employed to explore the representation strength of the suppressed stimuli, and the other centres on the neural structures required to process each type of stimuli. Focusing first on the task differences, in the current study the representational status of the suppressed stimulus was assessed in terms of its ability to enter consciousness, whereas both Jiang et al, and Yang et al. used slightly different behavioural methods. In Yang et al. they did not explore whether a fearful or neutral face *could* enter consciousness, but by gradually reducing the strength of the mask they explored *when* this stimulus entered consciousness. This reaction time based approach could simply prove more sensitive in picking up the subtle differences between suppressed stimuli. The other potential explanation for these behavioural effects for complex stimuli may lie in the fact that the extra attention attracted to these stimuli reflects the fact they can be extracted by structures outside of the ventral stream. In fact this answer is entirely plausible given the flow of visual information directly into the amygdala, that plays an important role in the rapid detection of ecologically relevant stimuli (Phelps & Ledoux, 2005; Zald, 2003). The sole dependence upon an amygdala pathway is however questioned by the fact that neuroimaging studies have revealed residual processing of face information within the ventral visual stream (Yiang et al. 2009; Sterzer, Jalkanen, Rees 2009). Furthermore Jiang, Costello and He (2007) have demonstrated not only an advantage for upright faces in competing against a CFS mask, but also for complex form information like words, which surely must depend upon processing in higher level areas within the ventral visual stream. Indeed although most single cell responses in monkeys highlight that responses in the ventral visual stream respond more clearly only to the stimulus consciously perceived, there is a small proportion of cells that still respond to suppressed stimuli in binocular rivalry (Sheinberg and Logothetis 1997). Taken together these studies imply that a certain amount of information processing is still achieved for suppressed stimuli in higher areas of the ventral stream.

Given that information can reach other areas in the ventral stream, one must either contend that the behavioural paradigm employed here was not sensitive enough to exploit this information, or that LO requires a certain threshold of information to leak through in order for illusory shapes to be extracted. Given the known neural signatures (in both EEG and fRMI) for the representation of coherent Kanizsa figures it could be wise to assess whether these signals still arise for otherwise suppressed stimuli. In parallel it could also be informative to investigate the distinction between coherent and incoherent Kanizsa patterns using exactly the same paradigm employed by Yiang et al. (2009) and Jiang, Costello and He (2007). By allowing the stimuli to eventually reach awareness, and assessing whether coherent or incoherent Kanizsa patterns reach awareness more rapidly one might be able to gain a more sensitive measure of the representation status of Kanizsa figures.

The lack of an advantage for coherent Kanizsa patterns is therefore reconcilable with the effects of CFS on the ventral stream (Fang and He 2005) and the critical structures involved in the perception of coherent Kanizsa stimuli (Murray et al. 2001). The advantage for the incoherent Kanizsa patterns in the cuing condition is however some-what bizarre. The original motivation for this cuing condition was that it might bring to light whether an object (epitomised by the coherent Kanizsa pattern) could be selected as a unit of attention when one would not otherwise be conscious of that object. Any potential interaction between coherence and cue type was therefore expected to reveal an advantage for the coherent Kanizsa pattern in the cued condition. The direction of the interaction found is therefore exactly counter to that originally predicted.

It is at least clear that the cue successfully attracted the participants' attention, and led to a higher rate of detection for both the coherent and incoherent Kanizsa patterns. We can only offer speculation as to why this cuing effect would be stronger for the incoherent Kanizsa. Given that the coherent Kanizsa did not reveal a representation advantage in the invalidly cued condition, it is likely that any explanation for the difference does not lie in their differential status as objects. Indeed if one steps back to think about how attention selects when uninfluenced by objects, it is clear that distance is critical. In this context it is apparent that when directly compared the distance between the cue and the target pac-men are marginally smaller for the incoherent Kanizsa patterns.

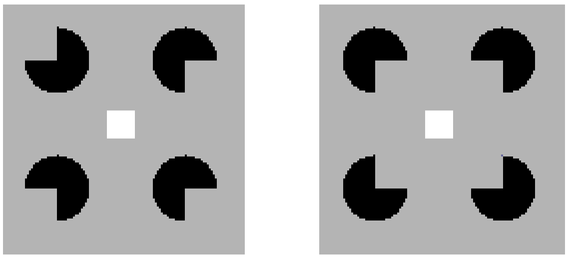


Figure 4, Illustrating a Cued Incoherent and Coherent Kanizsa. Note there is slightly more empty space between the cue and the pac-men for the Coherent Kanizsa. The colours used in the actual experiment are represented here in black and white.

Perhaps this marginal difference between the distance from the cue for the coherent and incoherent Kanizsa patterns results in a significant advantage for the incoherent stimuli. This is an entirely post-hoc explanation, but could be tested by exploring how the strength of the cuing effect changes with systematic changes in the distance between the cue and the targets. As this explanation stands however it is certainly consistent with the idea that when attention is attracted to an aspect of the visual image that was previously unconscious the only form of selection that occurs at that location is spatial. However in the present context this 'spatial only' form of selection is just as likely to reflect the fact that the Kanizsa has not been extracted per se, as that attention cannot select for objects when they are not represented consciously.

In summary, this project began with the intention of exploring whether an object can be selected by attention when one is not conscious of that object. This aim was explored by testing the extent to which cued coherent vs cued incoherent Kanizsa patterns would compete against a Continuous Flash Suppression Mask. Somewhat unexpectedly it was found that cued incoherent Kanizsa stimuli were perceived more often than cued Kanizsa stimuli, possibly reflecting the closer spatial proximity of the pac-man forming the incoherent Kanizsa and the attention capturing cue. As it stands this result highlights only a role for spatial attention when shapes influenced by attention are not consciously perceived. However this purely spatial role of attention may reflect the fact that, when masked via CFS, the Kanizsa was simply not constructed per se. Whether or not an illusory Kanizsa figure can be extracted when masked by CFS needs to be tested with potentially more sensitive behavioural techniques.

Section Two: Object Based Attention, a Tool to Explore the Nature of Perceptual Organisation

The first section of this thesis sought to explore object based attention as a phenomenon in its own right. Section two seeks to use object based attention as a tool to explore the nature of perceptual organisation more generally. This section is nevertheless guided by our developing understanding of object based attention. In particular chapter 5 is motivated by the possibility, highlighted in the introduction and tested in chapter 8, that object based attention is dependent on the extraction of visual form information to the level of the Lateral Occipital area. This area is not only thought to play an important role in the explicit perception of visual form (James et al. 2003, Malach et al., 1995, Kourtzi & Kanwisher, 2001), but is also activated by the tactile exploration of visual shapes (Amedi, Malach, Hendler, Peled & Zohary, 2001, James, Humphrey, Gati, Servos, Menon & Goodale, 2002). If the representations that drive attention are indeed extracted in LO, and LO is also influenced by tactile exploration of a given object, it is possible that objects that have only been explored tactually will be able to influence the visual allocation of attention. Chapter 5 explicitly tests this possibility.

Chapter 6 and Chapter 7 to some extent have similar goals, to explore whether the ecological importance of different stimulus types alters their representational status within ‘mid’ level vision. In Chapter 6 this possibility is explored by examining whether surfaces have a reduced representational status when they can be interpreted as shadows. Chapter 7 in contrast explores whether or not stimuli that induce perceptions of biological motion have a higher representational status within the visual system.

Chapter 5 - Cross-modal object-based attention

Abstract

The present study tested whether object information derived purely from touch could provide structural information sufficient to mediate object-based attentional effects. This aim was met by constructing a 3-dimensional version of the object based attention paradigm introduced by Egly et al. (1994), and testing whether a within-object advantage could be elicited for cues and targets appearing on objects, whose shape was known either on the basis of purely tactile or visual stimulation. Although smaller than the effects found with visual presentation, purely tactile knowledge of an objects shape was sufficient to drive the visual allocation of attention. This behavioral result suggests that the multi-modal activations found previously in visual system using fMRI have a functional status in terms of their ability to influence visual attention.

Introduction

A significant wave of research in cognitive science has been motivated by the possibility that specific computational challenges can be solved most efficiently using algorithms that are individually tailored to that purpose. Developed within vision (Marr 1982), this logic has extended to provide a general framework for understanding human cognition (Pinker 1997). Combined with the notion that these algorithms can be localized to particular areas of the brain, this modular approach to human cognition has motivated (and is indeed mutually motivated by) cognitive neuropsychology and neuroimaging, in which double dissociations between different cognitive functions have been sought in contrasting areas of brain damage or sources of neural activity.

This modular approach to cognition appeared, at first glance, to be congruent with the basic neuroanatomy of the human central nervous system, in which signals from different sense modalities are predominantly channeled to different areas of the brain. Over the last decade, however, although neuroimaging has highlighted how one area of the brain may seem to be associated with a particular cognitive task, it has also begun to reveal that areas of the brain traditionally thought to process only one modality of information are in fact activated by stimulus processing derived from other sense modalities. This cross-modal or multi-sensory processing requires some form of reconciliation with the modular paradigm of cognitive research. One potential resolution is offered by a framework whereby the initial processing of information is unimodal, but this unimodal processing is influenced by feedback from higher levels of processing where multi-modal information converges. More specifically Macaluso and Driver (2001) have argued that spatial information is combined in a high level system that also controls the allocation of attention. This higher level resource for directing attention, localized in the parietal lobe, results in shifts in information processing required for one modality automatically affecting the allocation of resources in all modalities (Macaluso, Frith, Driver 2002). In this way spatial maps for each modality may be computed separately, but influence each other via feedback from a higher level at which information is combined cross-modally.

The feedback effects from an integrated system of spatial attention (Macaluso & Driver, 2001) cannot however explain all multisensory effects. For example the ‘visual’ shape processing area LO (James, Culham, Humphrey, Milner & Goodale, 2003, Malach et al. 2005, Kourtzi & Kanwisher 2001) is responsive to the tactile exploration of objects (James et al 2002, Amedi, Malach, Hendler, Peled & Zohary 2001). Moreover the pattern of neural activation elicited from tactile exploration suggests that this information is processed in a manner that respects the same category distinctions elicited by purely visual stimulation (Pietrini et al. 2004). Furthermore activation derived from tactile stimulation in dorsal LO still manifested in a patient with brain damage resulting in an inability to recognize shapes through vision (Allen & Humphreys 2009). LO is typically not activated by auditory stimulation, even if those sound stimuli evoke perceptions indicative of a particular object (e.g., a car moving past). Recent research using sensory substitution, however, has revealed that if visual images are recorded and transformed into a ‘soundscape’ that maintains visual information, then these “auditory shapes” are able to activate LO (Amedi et al. 2007).

Can the multisensory nature of shape processing in LO be reconciled within a modular framework of human information processing? Yes, but only if modularity is defined, not in terms of the types of sensory input received, but in terms of the functions performed by that area upon its input. In this sense LO could be thought of as a generic shape processing resource. Indeed the specificity of LO’s responses to shape information across different modalities, if anything, highlights its specialization: it is not, for instance, activated by other forms of perception, such as the discrimination of textures, which can also in principle be resolved using information from vision or touch (Cavina-Pratesi, Kentridge, Heywood, Milner, 2009).

If LO is principally concerned with extracting shape representations, in a modality-independent fashion, then one would expect that effects elicited via the visual stimulation of LO could also be elicited via the tactile stimulation of LO. Recent neuropsychological research has suggested that area LO plays a critical role in the phenomenon of object-based attention (de-Wit, Kentridge & Milner 2009a). In this phenomenon the movement of attention within one visual 'object' is more efficient than the movement of attention between two objects. Given the critical role that LO plays in establishing shape/object representations that influence the allocation of attention, the present project set out to explore whether purely tactile knowledge of an object could influence the allocation of visual attention. If the shape representations in LO are developed in a manner that is to some extent independent of the sense modality from which those representations are derived, then purely tactile stimulation should be sufficient to elicit a visual object-based attention effect.

In order to test this hypothesis we adapted Egly, Driver and Rafal's (1994) object based attention paradigm. In this paradigm participants are presented with two outline rectangles, and are required to detect targets appearing at one of the four ends of these rectangles. Before a target appears, however, one of its potential locations is 'cued'. This cue can either appear at the location in which the target will subsequently appear (a valid cue) or at a different location (an invalid cue). Invalid cues come in two types: either the invalid cue is located on the other end of the same rectangle in which the target will appear (an invalid within-object cue) or it appears in an equally distant location but on the other rectangle (an invalid between-object cue). Targets following invalid cues within the same object are processed more efficiently than targets following invalid cues presented on the other object. This 'within-object' cuing advantage exemplifies how the allocation of attention can be influenced by visual shape representations. In order to explore the potential role of shape representations derived from tactile stimulation we replaced the computer generated display traditionally used in this paradigm with an equivalent set of solid objects, with embedded LEDs to act as cues and targets. The LEDs were embedded in such a way that the cues and targets could become visible without revealing the shape of the object in which the LEDs were located. With this set-up we were able to ask: if participants are grasping objects in the dark, can the tactile stimulation thereby elicited lead to a shape representation that is able to influence the allocation of their visual attention?

A consistent challenge in cross-modal research involves the possibility that any representations developed might reflect a high-level, but modality-specific, imagery-based feedback effect. In the current research, one could argue that knowledge about the target objects, acquired through any modality, might be sufficient to drive an effect. To rule out this possibility we tested whether simply knowing about the objects structure in the display could, in the absence of direct tactile or visual stimulation, drive an object-based attention effect.

Methods

Participants

Sixty participants (19 male) gave informed consent to take part in this experiment in exchange for course credits or money. The mean age was 20.36 (SD=2.05), and all participants had corrected or corrected to normal vision.

Design

The study employed a within-subject design, where participants performed a modified version of the Egly, Driver and Rafal (1994) object-based cuing paradigm in a number of different presentation conditions. In our modification of this paradigm, participants were presented with an exogenous cue followed by a simple color (red/green) target discrimination (see figure 1). The cue consisted of a white LED, which was followed by a small group of either red or green LEDs lighting up either at the same location (valid trials) or at another location (invalid trials). On invalid trials the target color could appear either at the other location on the same object as the cue (invalid within) or at the equally distant location on the other object (invalid between). The cues and targets appeared at any one of four locations depending on the condition. These four LEDs were embedded in two rectangular objects, which were either oriented vertically or horizontally (counterbalanced across participants) with a red fixation LED in the center.

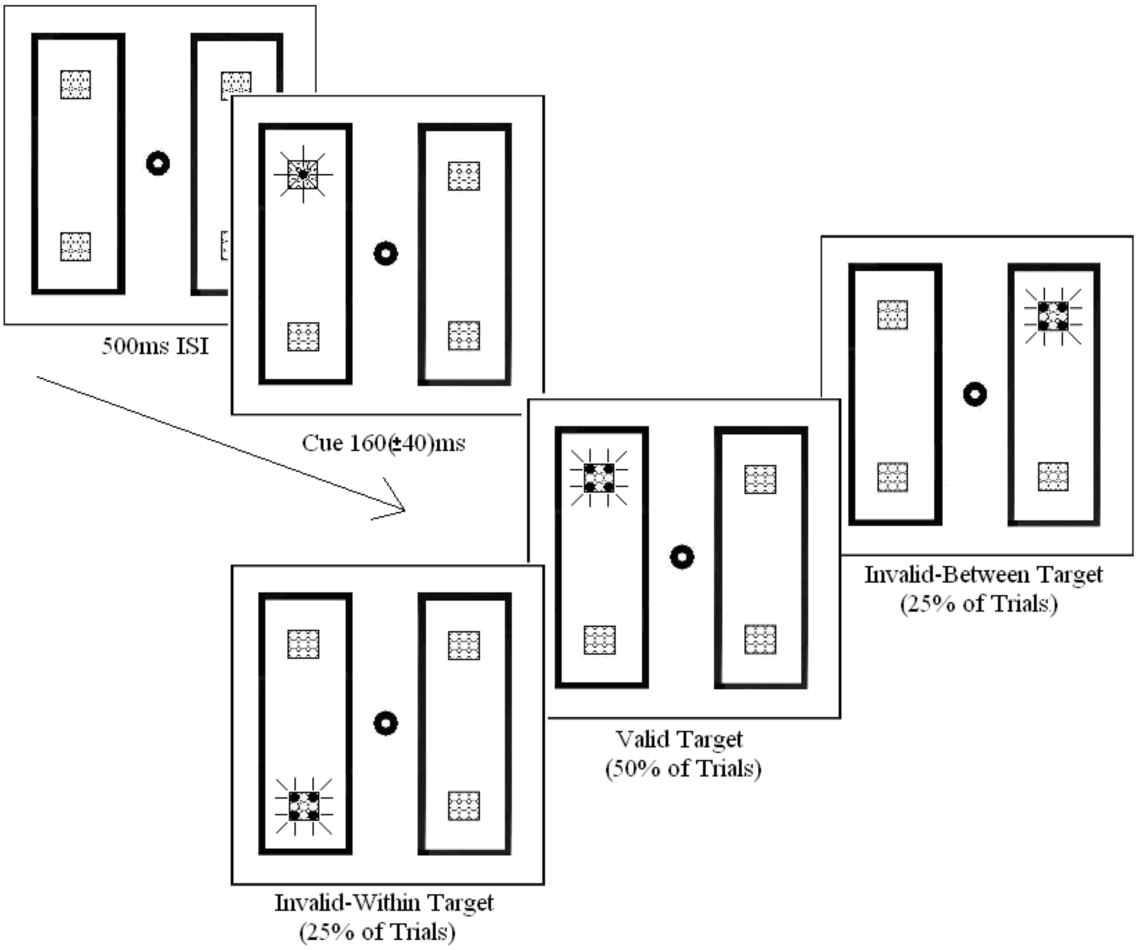


Figure 1, Illustrating the Basic Egly et al. (1994) object based cuing paradigm. The targets illustrated here were in fact either red or green, see text for details.

All participants completed three versions of this paradigm.

Condition 1. In the ‘baseline’ condition, participants performed the cued discrimination task in the dark, having never seen the two rectangles. As a precaution at the end of this block participants were asked if they could identify the objects on which the cues and targets were presented, and none of the participants were able to do so.

Condition 2. Prior to starting the ‘tactile’ condition, participants wore a pair of blacked-out goggles and were instructed to explore the target locations in front of them with both hands. Once they had found the four target locations and the location of the central fixation, the participants were requested to name the shapes, and identify whether the two rectangles were the same. This was done to encourage the participant to tactually explore the objects. After exploring the rectangles the participant was asked to complete the cued discrimination task while grasping the rectangles. The type of grip used varied depending upon whether the objects were oriented vertically or horizontally. For vertical objects, participants maintained an opposing finger-thumb grip at the bottom of each of the rectangles. For the horizontal objects, the participants gripped either end of the bottom object (see figure 2). These grip types where selected to provide a balance between ensuring that the participant had as much on-line tactile stimulation from the object as possible and that they did not become uncomfortable whilst holding the objects. The cued color discrimination task was completed in complete darkness with no other light sources. To avoid any distracting reflections from their hands or arms, the participants wore a pair of black gloves and had a black cloth placed over their arms.

Condition 3. In the ‘visual’ condition, an additional source of light was provided such that the objects were in clear view. This condition therefore provided the closest replication of the original Egly et al (1994) paradigm.

Forty of the 60 participants completed an additional two conditions immediately following the original 3. These conditions were in practical terms identical to conditions 1 and 2.

Condition 4. Here the participant again performed the cued color discrimination task in the absence of any current tactile or visual stimulation from the objects; but this time after having felt and seen the objects.

Condition 5. The participants again held onto the objects (using the same grip types as described above).

Conditions 1-3 (‘baseline’, ‘tactile’, ‘visual’) were completed in that order. Conditions 4 and 5 (‘post vision – no stimulation’, ‘post vision – tactile stimulation only’) were counterbalanced across participants.

Stimuli

LEDs were presented in five locations. The red (fixation) LED (5mm diameter) was embedded in a small tube (12mm diameter) in the center of the display. The four cue/target locations consisted of a 3x3 grid of LEDs, 15mm square. The central LED in this 3 x 3 grid was white, the four corner LEDs were green and the remaining four LEDs were red. The vertical/horizontal separation between the centers of the four LED locations was 81mm. The four cue/target LED locations were embedded in two black plastic rectangular blocks measuring 48 x 134mm.

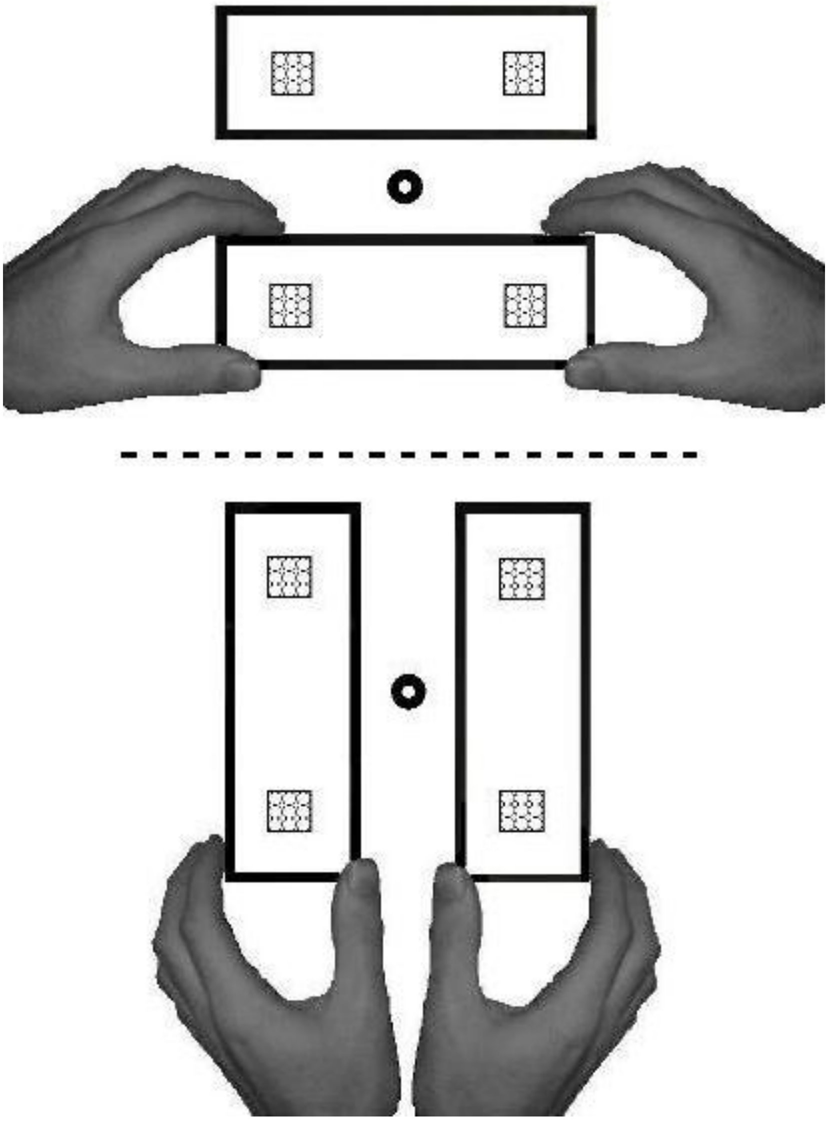


*Figure 2. Illustrating the objects and grip types with horizontal and vertical rectangles. In the center of each display there was a fixation spot, and on each rectangle there are two cue/target locations each of which is composed of a 3*3 array of LED's.*

Procedure

Fifty percent of the trials used valid cues, 25 percent invalid within object, and 25 percent invalid between object. The high percentage of invalid trials (double that employed by Egly et al., 1995) was used to provide more reliable mean invalid RTs without increasing the overall number of trials. In each block there were 192 trials preceded by 4 practise trials. The order of valid, invalid within and invalid between trials was randomized independently for each block. Before each trial there was a 500ms ISI, in which only the red fixation dot was presented. This dot remained on throughout each trial, and throughout each block. A white LED in the center of one of the four target locations (the cue) then appeared for an average of 160ms (randomly ±40ms on each trial). At the white LED offset, the participant was presented with a red or green color cue (consisting of four LEDs) at one of three locations relative to the cue, depending on the cuing condition. The participant was instructed to identify the target (as red or green) as fast and accurately as possible. The participant responded using a pair of foot-pedals, pressing the left pedal for red and the right for green. Trials with reaction times above 1000ms, below 100ms or outside 2 standard deviations for that participant for that condition, were removed prior to the calculation of the means reported in the Results. The participant sat approximately 57cm from the display, depending upon the length of the participant's arms.

Results

The results were analyzed with respect to 2 factors, presentation condition (with 5 levels: 'baseline', 'tactile only', 'vision', 'post vision – no stimulation', 'post vision – tactile stimulation only') and cue type (valid, invalid within, invalid between). The graphs present conditions 1-3 separately from conditions 4 and 5, because all 60 participants completed the first three while only 40 completed the last two. A 5 (Condition) by 3 (Cue) repeated measures ANOVA on the reaction time data revealed main effects of both Condition (F(4,152)=41.42, p<0.001) and Cue (F(1.69,64.3)=96.55, p<0.001), and a significant interaction (F(6.05,230.08)=3.42, p=0.003)[10]. The overall effect of Presentation Condition undoubtedly reflects an order effect whereby participants became faster in using the foot pedals as the experiment progressed. For example the average RT in the first 'baseline'

---

[10] The reported values were given Greenhouse-Geisser corrections because of a potential sphericity violation for the effect of Cue Type (M, p=0.024) and the interaction (M, p=0.017).

condition (collapsing across cue type) is 547ms compared with 484ms in the otherwise identical, ‘post vision – no stimulation’, second baseline ($t(38)=7.5$, $p<0.001$). The effect of Cue Type reflects a combination of both faster responses on valid trials (across the five conditions the overall valid mean is 483ms while the combined (within and between trials) invalid mean is 512ms, ($F(1,38)=137$, $p<0.001$)) and faster responses on within vs between object trials ($F(1,38)=5.52$, $p=0.024$).

The significant interaction, however, means that the cues had different effects in the different presentation conditions. In order to pin down this effect, and to test our hypothesis regarding the potential for tactile stimulation to influence the visual allocation of attention, we performed paired t-tests on the invalid-within vs invalid-between means for each presentation condition. In the first baseline condition (in which the objects had never been seen or touched) there was unsurprisingly no difference in reaction time between within or between-object invalid cues ($t(58)=0.278$, $p=0.782$). In the ‘tactile only’ condition 2, participants took significantly longer to respond following invalid between object cues (5.7ms, $t(59)=2.19$, $p=0.033$)[11]. In the ‘visual’ condition 3 we replicated the standard influence of visual objects upon attention, revealing a significant advantage for within vs between object invalid cues (10.48ms, $t(59)=3.61$, $p=0.001$). In condition 4 (‘post vision – tactile stimulation only’) there was a similar trend to that seen in the first tactile condition, though it did not reach significance (4.58ms, $t(39)=1.35$, $p=0.186$). In condition 5 (‘post vision – no stimulation’), despite the objects having been seen and touched, there was no influence of the objects upon invalid reaction times ($t(39)=0.244$, $p=0.81$).

---

[11] If this analysis is run as a mixed 2 x 2 ANOVA, with the horizontal or vertical orientation of the objects included as a between-subject factor, cue type does not interact significantly with orientation type ($F(1,58)=0.083$, $p=0.774$). This suggests that the tactile object based effect is independent of object orientation, and of the different grip types necessarily associated with each orientation.

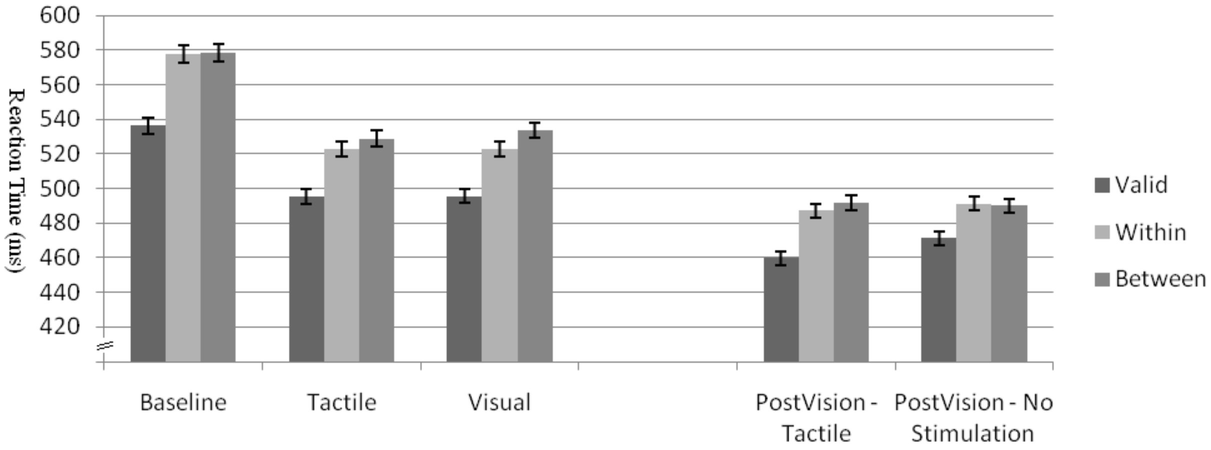


Figure 3. Reaction times (in ms with standard error bars) for the valid, invalid within and between cue types in the different presentation conditions: 'baseline', 'tactile', 'visual', 'post vision tactile' and 'post vision no stimulation'.

In the accuracy data there were again main effects of Presentation Condition (F(4,152)=7.76,p<0.001) and Cue Type (F(2,76)=3.45,p=0.037), but no interaction (F(5.52,209.68)=1.571, p=0.163)[12]. The small Cue effect does not reflect a difference between within or between object trials (F(1,38)=1.04, p=0.31) but in fact a slight reduction in accuracy on *valid* trials (94.5% on valid, 95.2% on invalid, F(1,38)=7.11, p=0.011), counter to the faster reaction times on valid trials. The effect of Presentation Condition also runs counter to that seen in the reaction times, in that participants became less accurate as the experiment progressed. For example the accuracy (averaged across cue type) for the first 'baseline' (condition 1) is 96.46% whilst the accuracy in the second baseline (condition 5, 'post vision – no stimulation') is 94.41% (t(38)=3.5, p=0.001). Thus it seems that both main effects seen in the reaction time data are counterbalanced by speed-accuracy trade-offs. This finding is not critical to the current research question, but it nevertheless raises the question whether the within vs between-object cuing effects in the RT data might themselves reflect differential speed-accuracy trade-offs. To rule out this possibility we performed the same five paired sampled t-tests conducted on the reaction time data on the accuracy scores. Four of the presentation conditions revealed no sensitivity to within vs between-object cuing (all t values < 0.59), the one condition that did show some sensitivity was condition 4 ('post vision –

---

[12] The statistic for the interaction is reported after a Greenhous-Geisser correction, because Mauchly's test was significant (p=0.001).

tactile stimulation only'). The direction of the effect here was however in the same direction as the small trend seen in the reaction time data for this condition, in that between object trials were less accurate than within object trials (1.76%, t(39)=2.10, p=0.043). One would not normally predict the Egly et al. (1994) cuing effect to reveal itself in accuracy data, but this result indicates that any speed-accuracy trade-off acted only to reduce the strength of the RT effect. That is, the accuracy data provide further evidence that tactile stimulation can influence the visual allocation of attention.

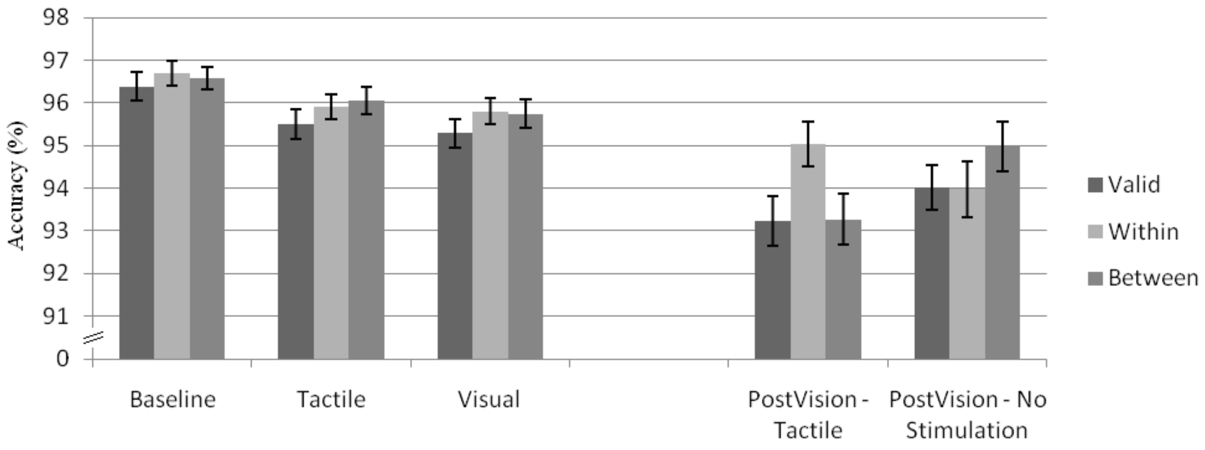


Figure 4. Accuracies (and standard error bars) for the valid, invalid within and between cue types in the 'baseline', 'tactile', 'visual', 'post vision tactile' and 'post vision no stimulation'.

Discussion

Shape representations extracted within area LO have been argued to play a critical role in influencing the allocation of attention (de-Wit, Kentidge & Milner, 2009a). Moreover this 'visual' shape processing area has been shown to become active during the tactile perception of shape (Amedi, et al. 2001, James et al. 2002, Pietrini et al. 2004). We set out to test whether tactile information supplied to LO could evoke a shape representation with the same attentional influence as one elicited by purely visual information. In order to test this possibility we adapted a well studied visual object-based attention paradigm (Egly et al. 1994), such that participants could hold the objects that they could traditionally only see. We found that purely tactile stimulation from an object was sufficient to influence the allocation of visual attention in a manner congruent with that elicited by purely visual stimulation.

We provided participants with two conditions in which they had direct tactile stimulation from a pair of objects, these two conditions occurring before or after having seen the objects. In the first ‘tactile only’ condition, participants showed a significant advantage for cue-target pairings appearing within the same object (within-object cuing advantage of 5.7ms). In the second ‘post vision – tactile stimulation’ condition, participants showed a similar trend (4.6ms), and although this reaction time difference was not significant, it was associated with a significant reduction in the accuracy of responses following between-object cues. The within-object cuing advantage following tactile stimulation (overall around 5ms) was smaller than that seen with purely visual presentation (around 10ms), which may reflect the slightly lower activation of LO following tactile stimulation (Beauchamp, 2005), or a less complete sensory experience of the objects with tactile stimulation. What is critical for the present investigation, however, is that tactile stimulation led to the same pattern of advantage for within-object cues as purely visual stimulation. It is also notable that the within-object cuing advantage found here with real three dimensional objects is of the same magnitude as that elicited with 2D computer generated outline figures (typically 10-15ms). This consistency between the influence of 3D and 2D figures upon attention is consistent with the similar responses seen in LO to the presentation of 3D and 2D shapes (James et al. 2003). It also provides a welcome reassurance that the 2D figures normally used in attentional studies really are treated by the brain as ‘objects’.

The interpretability of many cross-modal findings, particularly in visual cortex, is hampered by the possibility that they might reflect high-level feedback effects derived from the recoding of nonvisual information into a visual format. Several papers investigating the influence of heteromodal stimulation on LO have however excluded imagery as an explanation, such that parsimony would suggest little or no role for such mediation (Amedi et al. 2007, Peitrini et al. 2004). Peitro and colleagues, for example, found tactile-elicited activations in LO even in blind patients who reported no visual memories. This concern is nevertheless relevant in the current experiment, especially in the context of previous work within vision showing that prior experience of an object can influence how that object influences attention (Zemel, Behrmann, Mozer & Bavelier 2002). In order to rule out the possibility that high level knowledge of an object, independent of direct sensory experience, could influence the allocation of attention, the object-based cuing task was repeated with only memory-based object information (arising from the objects having been previously seen and held). Critically, there was absolutely no influence of object structure upon the allocation of attention in this condition. In conjunction with previous research suggesting little or no role for visual imagery in mediating the tactile activations in LO, we would argue that this result indicates that the effect of tactile stimulation we have reported here reflects a truly bottom-up integration of information from this modality.

One of the clear advantages of Egly et al’s (1994) object-based cuing paradigm is that the rectangles presented have no explicit task demands. Although cues and targets are presented upon the rectangles, the participant never has to report anything about the objects to perform the task. This suggests that the extraction of the shapes, and their influence upon attention, occurs automatically. This in turn suggests in the present context that tactile shape representations may be integrated with visual information processing in an equally automatic fashion. That said however, it should be noted that the tactile stimulation in this experiment was not completely passive. Before holding the rectangles, participants had to identify them on the basis of tactile stimulation, and moreover they did not simply touch the objects but had to adopt an opposing finger-thumb grip in order to grasp them.

The original discovery of cross-modal activation in traditionally unimodal areas originally appeared to challenge the modular approach to cognition, according to which dissociable systems process different forms of information. If anything, however, the emerging picture of LO offers a more refined and interesting version of modularity. The activation of LO following tactile, and even auditory, information demonstrates that the specificity of this area needs to be defined in terms of the functions that it performs rather than the sensory source from which its inputs are derived. As far as area LO is concerned information regarding shape can be abstracted from multiple sources: 2-D and 3-D visual stimulation, tactile stimulation, or auditory ‘soundscapes’ that preserve shape information. Thus LO seems to have one very specific function, reconstructing shape, and will exploit information from any modality in solving this challenge. The present results provide evidence consistent with the idea that the cross-modal activations evident in recent fMRI results derive from a pattern of representation that is able to exert a functional influence on the allocation of visual attention.

Chapter 6 - Are Shadows accessible as units of object based attention?

Abstract

It is unclear how shadows are processed in the visual system. Whilst shadows are clearly used as an important cue to localise the objects that cast them, there is mixed evidence regarding the extent to which shadows are exploited to aid the recognition of those objects. Furthermore previous work exploring the perception of shadows per se has provided evidence that the visual system has less efficient access to the form of a stimulus if it is interpreted as being a shadow (Rensink and Cavanagh 2004). The current study sought to clarify our understanding of the extent to which shadows are discounted by the visual system by exploring how they influence attention in two different object based attention paradigms. The results provide consistent evidence that cues to interpret a surface as a shadow do not reduce the salience of that surface as a target for attentional selection. We attempt to reconcile our data with Rensink and Cavanagh's by framing the representational status of shadows and the nature of object based attention within a reverse hierarchy framework.

Introduction

The pattern of stimulation that falls upon the retina is not equally informative. The limited capacity of the visual system can be employed more effectively if it is focused on those aspects of visual structure that are most informative. Recent evidence has highlighted that important stimuli attract preferential processing even before they are consciously recognised (Jiang, Costello, Fang, Huang, & He 2006). In parallel to the active enhancement of important information, does the human visual system actively seek to suppress less informative aspects of the environment? This paper will explore this question by examining the representational status of shadows within the human visual system. Shadows pose an interesting representational challenge because on the one hand, they do not represent inherent structure in the environment, whilst on the other they can potentially provide useful information about the objects that cast them. Perhaps unsurprisingly attempts to understand the status of shadows within the human visual system have provided a somewhat mixed picture, in which shadow properties are hard to identify (Ostrovsky, Cavanagh & Sinha 2005) and in some contexts have no measurable influence on object recognition (Braje, Legge and Kersten 2000; Bonfiglioli, Pavani and Castiello 2004), whilst in other contexts shadows clearly aid and/or interfere with object recognition (Tarr, Kersten, & Bulthoff, 1998). Indeed whilst shadows often seem to be employed in computing the location and movement profile of the objects that cast them (Kersten, Knill, Mamassian & Bulthoff 1996), it is also clear that in certain visual search contexts that explicit access to the form of shadows is actively reduced (Rensink and Cavanaugh 2004). This research project seeks to further refine our understanding of the status of shadows within the visual system by exploring how they influence the allocation of attention in two commonly employed object based attention paradigms.

The common acceptance of the role of objects in influencing the allocation of attention has led a number of researchers to use object-based attention paradigms as tools to explore the nature of object representations per se (see Scholl 2001 for a review). For example Ben-Shahar, Zucker and Scholl (2007) have recently used object based attention as a tool to explore how the visual system uses textures to segment the visual environment into distinct surfaces. The current study seeks to extend this logic by exploring whether or not stimuli that are very likely to be interpreted as shadows still reveal classical object based attention effects. In employing this logic however it should be noted that our understanding of the level at which objects influence attention is still developing. Nevertheless it is possible to argue that the type of visual structure that influences attention does not depend upon the lowest levels of edge extraction (Ben-Shahar et al.) or our higher representations of recognisable/nameable objects but seems to reflect the mid-level stages between edge extraction and recognition in which the visual system seeks to employ heuristics to parse the visual scene into candidate proto-objects (Driver, Davis, Russell, Turatto & Freeman 2001). Recent neuroimaging and neuropsychology has suggested that this ‘mid-level’ extraction of shape information that influences attention is contingent upon the shape processing area LO (Martinez et al. 2006, de-Wit, Kentridge, Milner 2009a). The current study should therefore enable us to address whether shadows are still extracted to an equivalent representational status up to the level of LO.

Returning to the previous literature on shadows in more detail, the clearest indication that shadows have a special status within the human visual system was provided by Rensink and Cavanagh (2004), who demonstrated that when a surface could be interpreted as a shadow its form became less accessible for perceptual discriminations. As the visual system moves from the limited 2 dimensional image on the retina to our 3 dimensional perception of the world the visual system must ensure that the contrast boundaries cast by shadows are not misinterpreted as the boundaries between real objects. Rensink and Cavanagh argue that this is ensured by a rapid mechanism that acts in parallel across our visual representation to first identify and then reduce the access to form information contained in shadows.

In Rensink and Cavanagh (2004) participants were required to search for a target that had a different orientation to other items in a display. Rensink and Cavanagh found that this visual search, or oddity detection task, was in fact harder to perform when the items in the display could be interpreted as shadows. When the stimuli were rendered uninterpretable as shadows the task became easier, indicating that the visual system actively discounted form information related to shadows. The interpretability of the stimuli as shadows was manipulated in a number of ways; one method (also employed in the current set of experiments) involved removing the object that cast the shadow. In Rensink and Cavanagh's study the removal of the object casting a shadow disrupted the interpretation of the shadow, and in fact enabled the visual system to have greater access to the form information of that shadow-like shape.

This active discounting of shadows is predicted by models of illumination-invariant recognition, in which spurious features (like shadows) have to be discounted by the visual system. The idea that shadows are in some sense discounted by the visual system is also borne out by a number of studies that explore the role of shadows in identifying the objects that cast them. For example Braje, Legge and Kersten (2000) found that even when stimuli were transformed (grey scaled and blurry) so as to make the discrimination between genuine objects and shadows more difficult, the recognition of natural objects was highly invariant to the presence or absence of shadows. This invariance to the information inherent within shadows is further supported by Bonfiglioli, Pavani and Castiello's (2004) finding that the recognition of real objects was not influenced by the presentation of incongruent shadows. However whilst Bonfiglioli et al. found that shadows did not influence the recognition of the objects that cast them they did find that manual responses were biased by the presentation of shadows. At first glance it might seem appropriate to frame this finding within Goodale and Milner's (1992) division between the dorsal 'vision for action' and ventral 'vision for perception' pathways. It is however somewhat incongruent with the Milner and Goodale (1995) framework that 'context' (created by the shadow) influenced manual action but not recognition, in apparent contradiction to the notion that context effects (which often cause visual illusions) often have a greater influence on perception than action (Goodale & Westwood 2004). Reflecting on Bonfiglioli et al's study, however, it seems unlikely that the manual vs recognition responses they collected map neatly onto the Milner and Goodale conception of dorsal and ventral stream. For instance the manual action was generated 'at the participant's leisure', introducing a delay that would allow allocentric representations in the ventral stream to influence manual actions (Milner and Goodale 1995, Westwood and Goodale). More critically the spatial components of the action and perception tasks in this study were not closely matched; whilst the manual action involved a degree of localisation (where to direct the hands) the perceptual task could be solved without any concern for the location of the target object. Thus it is probable that the dissociation found in Bonfiglioli et al can be more appropriately framed in terms of shadows influencing a localisation task, but not a recognition task.

Indeed this idea that shadows can influence 'where' components of perception is supported by a range of studies showing that the visual system actively utilizes shadows in order to aid the localisation of the objects that cast them. In a rather elegant study, Kersten, Knill, Mamassain and Bulthoff (1996) found that they could induce an illusory perception of motion in depth by altering the size and location of the object's shadow. Indeed work by Imura et al (2006) and Yonas and Granrud (2006) has established that this ability to utilize cast shadows as a cue to depth emerges (in line with other pictorial depth cues) at around 6 months. This active utilization of shadows to aid 'where' judgements, could be seen as contradictory to Rensink and Cavanagh's findings that shadows are discounted in the context of visual search. Rensink and Cavanagh's conclusions however were derived from an oddity detection (or what) task. Perhaps therefore it is possible to build an argument that whilst shadows are actively discounted in the context of form or shape discriminations (Rensink and Cavanagh, Braje et al. 2000, Bonfiglioli et al 2004) they might still be actively employed in the context of localisation tasks (Bonfiglioli et al, 2004; Kersten et al., 1996; Imura et al., 2006; Yonas and Granrud, 2006). This picture however, of shadows influencing 'where' but not 'what' components of perception, cannot be completely accurate. Tarr et al. (1998), whilst noting that traditional models of edge/contour based recognition assume that spurious features such as illumination should be discounted (Marr & Nishihara 1978), in fact found that the recognition of a set of novel shapes was influenced by the presentation of shadows. Thus whilst there is clear evidence that shadows influence 'where' tasks, involving localisation

(Bonfiglioli et al) or motion and depth perception (Kersten et al., Imura et al. and Yonas & Granrud Y), the contrast between Tarr et al's and Braje et al. and Bonfiglioli's results suggest a more mixed picture regarding the extent to which shadows are ignored in the context of 'what' or recognition tasks.

A possible resolution to this disagreement lies in the fact that as, Rensink and Cavanagh note, the process of actively ignoring shadows in their study was not absolute; the form of the shadow was still accessible to the visual system, it was simply less accessible than the form of identical shapes not interpreted as shadows. Indeed they argue that this partial process of ignoring the form of shadows might explain why their presence can sometimes influence recognition (i.e. in Tarr et al., 1998). Even if one accepts this position, however, it is still unclear how the access to form information contained in shadows is altered (if only to an extent) in 'what tasks' on the one hand and actively used as a cue to localisation on the other. It is in this context that the present study attempts to shed further light on the representation of shadows within the human visual system, by exploring their status within two different object based attention paradigms, in which cue-target pairings on the same object lead to faster responses (Egly, Driver and Rafal, 1994 paradigm, adapted for experiment 1) and the comparison of two targets is facilitated when presented on the same object (paradigm used by Ben-Sahar et al. 2007, adapted for experiment 2). Are these 'within-object' advantages still evident when the 'objects' in question are interpreted as shadows?

Experiment 1

Methods

Participants: Twenty-five participants were recruited in exchange for course credits. All participants were first year psychology undergraduates, with normal or corrected to normal vision. Participants were naive to the aims of the study. Participants ages ranged from 18 to 38 (mean 20). There were 8 males and 17 females.

Materials: The experiment was programmed using C++, with the use of DirectX to ensure accurate timings. Stimuli were constructed off-line before the experiment using the 3-D rendering package 'POV-Ray'. This program allows for accurate ray tracing, thereby enabling an effective simulation of shadows. The stimuli were presented on 17 inch PC monitors with a screen resolution of 1024 by 768 pixels. Responses were recorded using ‘direct input’ to collect key presses from a keyboard.

Design: The experiment employed a 3 by 2 within-participants design. The first factor, Cue Type, had three levels: valid (cue appears in same location as the target), invalid-within (cue appears at a different location but on the same shape), and invalid-between (cue appears at a different location on a different shape), as shown in Figure 1. There were equal numbers of each cue type and these were randomly distributed across trials. The second factor, Shadow Type, had two levels: either the shapes could or could not be interpreted as shadows. In line with Rensink and Cavanagh’s (2004) study this was achieved by removing the object that cast the shadow. Shadow or Non-shadow-like stimuli were presented in separate blocks, the order of which was counterbalanced across participants.

One of the critical factors in the original Egly et al. (1994) paradigm pertains to the distance between the cues and the targets. In order to establish the effect of objects upon attention it was critical to control for the distance between and within objects, such that any differences could not be explained in spatial terms. The Egly et al. result, and the many replications that have followed have clearly highlighted that the influence of object structure upon attention cannot be explained in terms of the spatial distances between cues and targets. The fact that our stimuli are rendered in a Three-Dimensional perspective (in order to enhance the perception of our stimuli as shadows) complicates this aspect of control. This complication arises because it is evident that attention moves in perceived rather than retinal space (Robertson & Kim, 1999). It is therefore more appropriate to align the distances between the stimuli in terms of the 3D environment rather than the 2 dimensional distances that will hit the retina. Given that stimuli naturally get closer as one moves further in depth the two horizontal distances are not identical on the screen (front = 7, back = 4.4) and the vertical distances between the target locations was 4 degrees of visual angle. Given the previous evidence for the alignment of size in line with size constancy it seemed more appropriate to ensure that the targets fell in line with the perspective of the surfaces on which they were presented, rather than in terms of the absolute 2 dimensional distances.

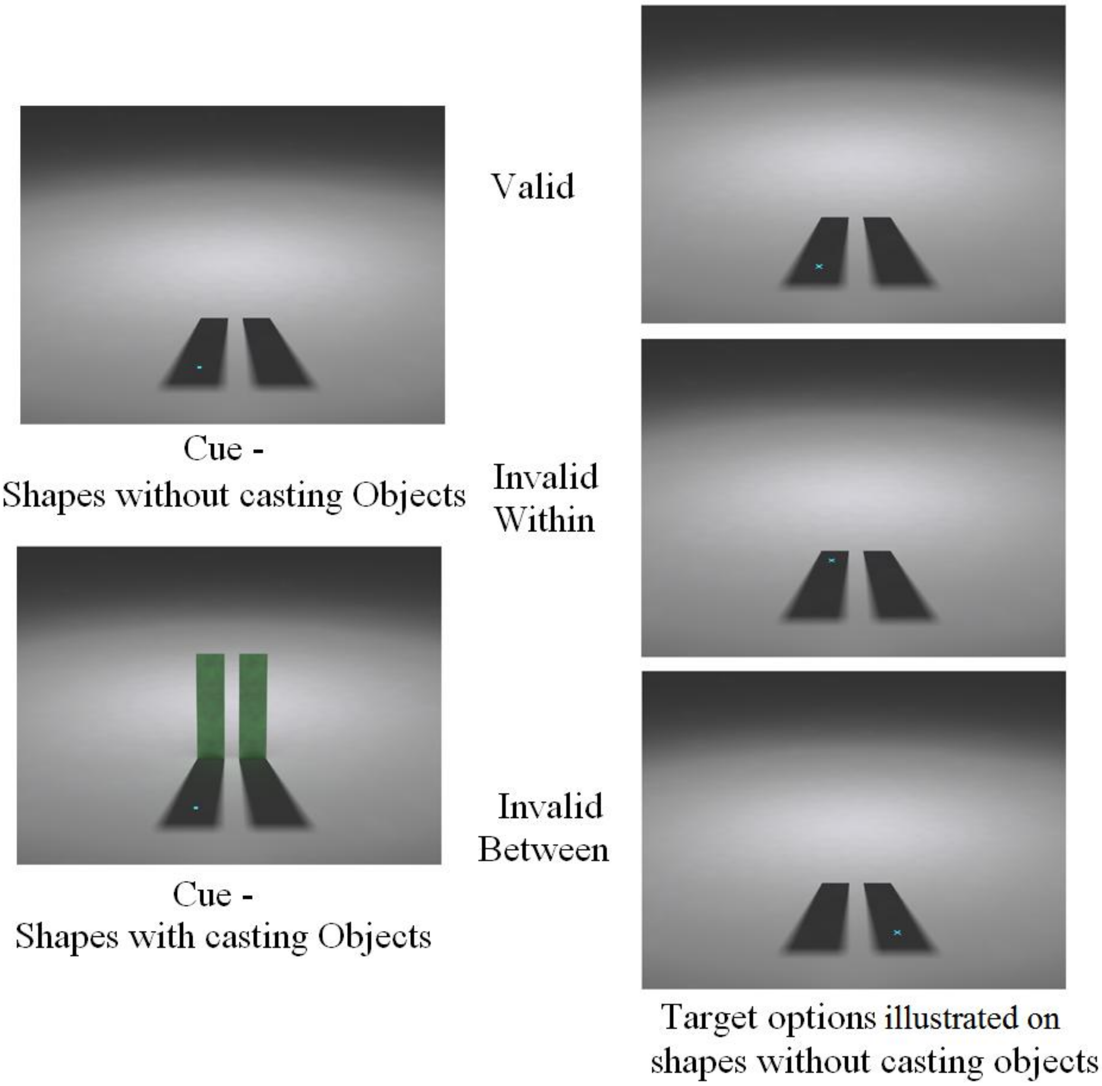


Figure One, on the left one can see the surfaces presented with (bottom) and without (top) a casting object, on the right one can see the location of the target stimulus in the various cuing conditions relative to the cue on the left.

Procedure – Each participant completed two blocks of 32 practise and then 192 actual trials. Each block contained equal numbers of valid, invalid-within and invalid-between trials[13]. On each trial the participant would be presented with a cue for 250 ms. The cue was a small square appearing in one of the four possible target locations. The use of this cue type was important because Rensink and Cavanagh showed that contrast outlines around the edge of a surface (like those typically used as cues in the Egly et al paradigm) can disrupt the

---

[13] This ratio is different to that used in the original Egly et al paradigm (75%, 12.5%, 12.5%), however pilot work showed that cues are still effective with this ratio, and the larger number of invalid trials with this ratio allows for the calculation of a more stable and reliable mean for the invalid trials.

interpretation of that surface as a shadow. After the cue there was a 200ms gap before the participants would be presented with one of two targets (an X or an N). The participant simply had to report (using the x and n keys) which target had been presented. There was then a 500ms pause before the next trial started. The participant was not required to maintain fixation[14].

Results

The results were analysed using a two factor repeated measures ANOVA. The reaction time data showed a clear effect of Cue Type (F(2,24)=70.728, p<0.0001). Comparing only the within-vs-between object Cue Types there is still a significant effect (F(1,24)=16.99, p<0.001) highlighting that targets presented on the same object as the cue were detected more rapidly. This within-vs-between object advantage did not interact with Shadow Type (F(1,24)=0.026, p=0.874), and shadow Type on its own had no influence on the reaction time (F(1,24)=0.164, p=0.689).

The accuracy data revealed no significant main effect of Cue Type (F<1), Shadow Type (F<1) or Cue x Shadow interaction (F(1,24)=1.326, p=0.261). The accuracies for the ‘valid’, ‘invalid within’ and ‘invalid between are 94.6%, 94% and 95.1% for the non-shadow and 93.6%, 94% and 94.6% for the shadow condition, and their associated standard errors were 0.89, 0.84, 0.64, 0.99, 0.88 and 0.85 respectively.

---

[14] It should be noted that a number of neural areas show less object based attention related activation during free viewing (Fink, Dolan, Halligan, Marshall and Frith (1997). This is not critical in the current study however because we are not focusing on the overall strength of the object based attention effect but whether this effect varies when the stimuli can or cannot be interpreted as shadows.

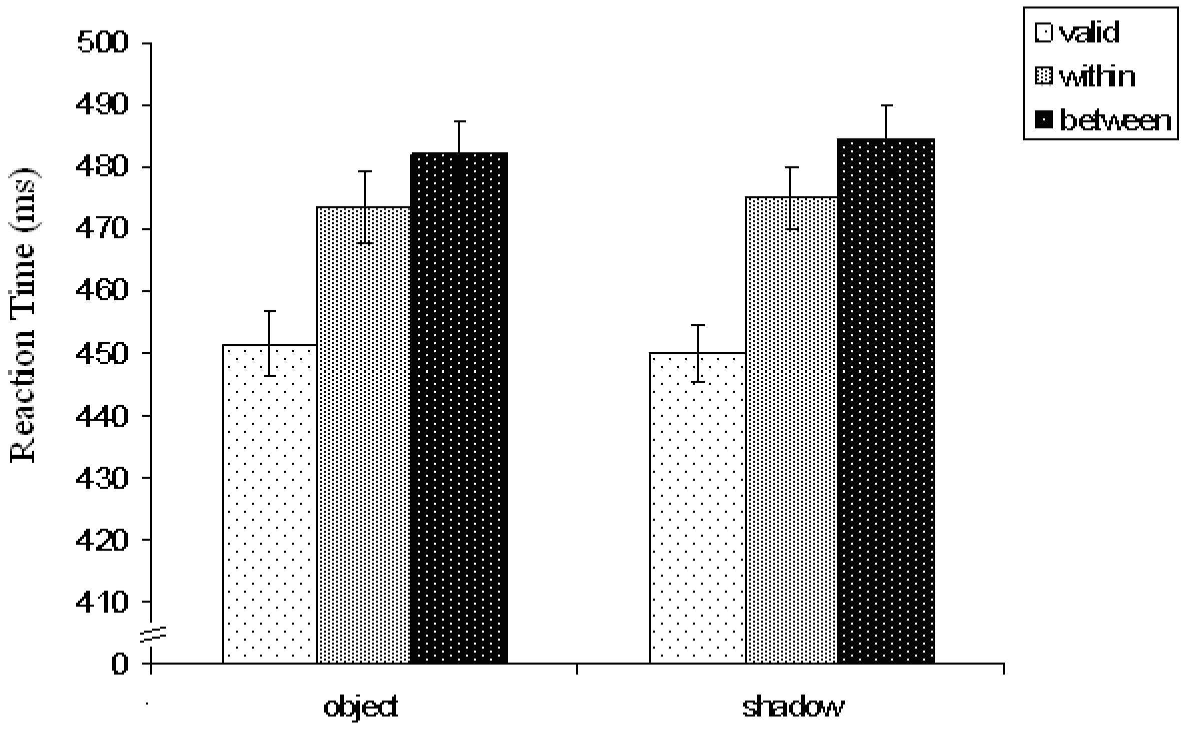


Figure 2, Reaction time data for the valid, invalid within and invalid-between conditions in the shapes with casters ('shadows') and without casters ('object').

Interim Discussion

The results of Experiment 1 revealed a clear replication of Egly et al's (1994) demonstration of the influence of objects upon attention. The results provide no evidence that shadows are treated as less salient aspects of visual structure with respect to object-based attention. The Egly et al paradigm is, however, just one of a range of tasks that can be employed to illustrate the influence of visual structure upon attention (Duncan 1984, Driver and Baylis 1989, see Scholl 2001 for a review). Ben-Shadar, Scholl and Zucker (2007) have for instance employed a two item comparison task in which the participant has to report whether two items (which can either be located on the same or on different objects) are the same or different. They found that when targets were presented on different 'objects', participants were less accurate in reporting whether they were the same or different. Experiment 2 therefore seeks to explore whether the effects found here with the Egly et al cuing paradigm will be replicated with this 'divided attention' object based paradigm. Finally shadows come in two different forms, which could potentially be interpreted differently: shadows can either be attached to the objects that cast them (like those employed in Experiment 1) or they can lie on a separate surface. In order to broaden the scope of the findings in Experiment 1, Experiment 2 employed unattached shadows.

Experiment 2

Methods

Participants: Thirty participants were recruited in exchange for payment or course credits. All participants had normal or corrected to normal vision and were naive to the aims of the study. Participants ages ranged from 18 to 57 (mean 22: 29 of the participants were aged 18 to 27). There were 12 males and 18 females.

Materials: Identical to Experiment 1.

Design: The experiment employed a 2 by 2 within-participants design. The first factor, Object Type had two levels, within vs between object. On within-object trials the two to-be-compared letters appeared on the same shape, whereas on between-object trials the two letters appeared on different shapes. There were equal numbers at each level and these were randomly distributed across trials. The second factor, Shadow Type, had two levels: either the shapes could or could not be readily interpreted as shadows. In line with Rensink and Cavanagh (2004), this was achieved by removing the object that cast the shadow. In contrast to Experiment 1, however, the shadows were cast from an object to which they were not attached, and the light source for the scene was visible (see Figure 4). In the non-shadow-like condition, as well as removing the objects responsible for casting the shadow, all other shadows in the scene were removed, to further reduce the likelihood of the shapes on which the targets were presented would be interpreted as shadows. Shadow or non-shadow-like stimuli were presented in separate blocks, the order of which was counterbalanced across participants.

In the context of previous data highlighting that attention moves in perceived rather than retinal space (Robertson and Kim 1999) one again had to decide how to control for distance. In Experiment Two a more stringent criterion was adopted such that the within object distances were longer than the between object distances. Given the potential ambiguity over how to control for the distance between the target locations this manipulation enables one to be much more confident in arguing that any within vs between object based differences reflects object based effects. The vertical (within object) separation between the targets was therefore fixed at 5.75 degrees whilst the horizontal distances were 3.8 deg at the top and 4.9 deg at the bottom.

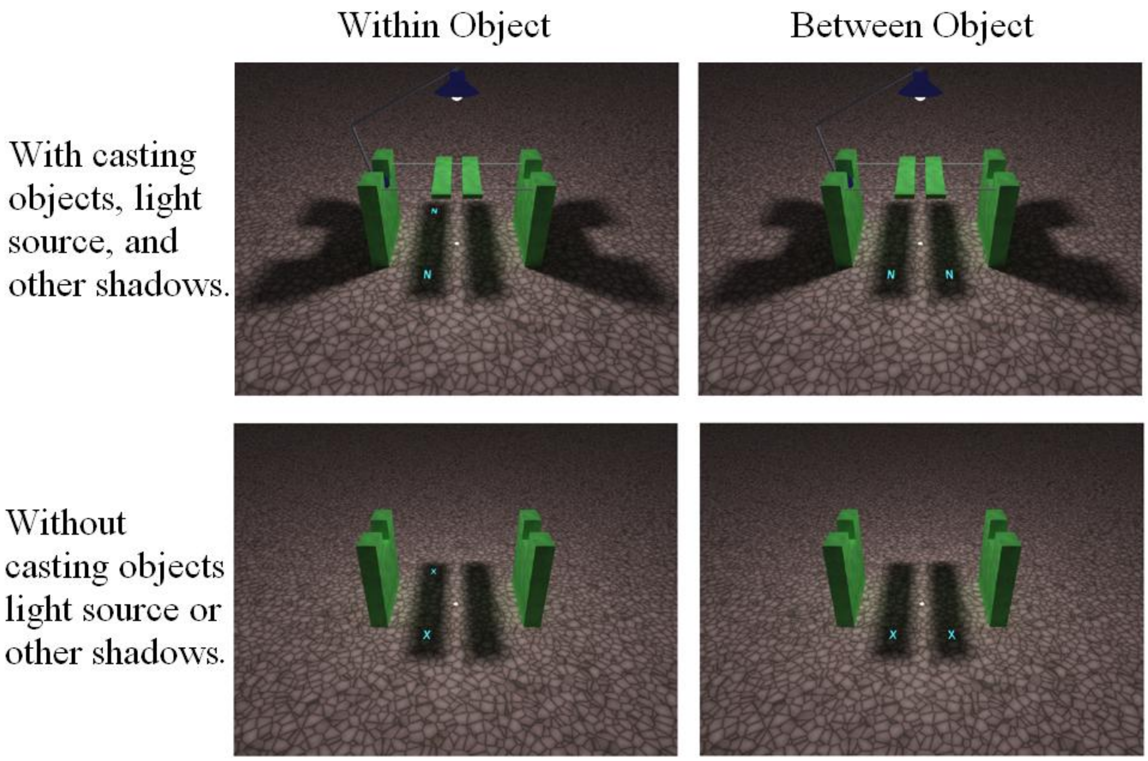


Figure 3. Illustrating the means by which the identification of the surfaces as shadows was manipulated and the relationship between the targets in the within and between object conditions.

Procedure: Each participant completed two blocks with 4 practise and then 192 actual trials. Each block contained equal numbers of within and between-object trials. On each trial the participant would be presented with two letters (‘x x’, ‘n x’, or ‘n n’) that could either both be the same or different. The participant had to report as quickly and accurately as possible whether the letters were the same or different by pressing the ‘s’ key if they were the same and the ‘k’ key if they were different. The two letters would either both be on the same or on different shapes. There were equal numbers of same vs different and within vs between object trials and these were randomised independently. Participants maintained fixation on a small white dot in the centre of the display.

Results

A 2 by 2 repeated measures ANOVA on the accuracy data revealed a significant effect of Object Type (F(1,29)=4.56,p=0.041)[15]. Object Type, however, did not interact with Shadow Type (F=0.851). There was also no effect of Shadow Type on the accuracy data (F<1). The Reaction Time data revealed no significant effects, (Shadow Type, F<1; Object Type, F<1). The Shadow Type x Object Type reaction time interaction (F(1,29)=1.037, p=0.317) was not only non-significant but in fact showed a trend in the opposite direction to that expected if shadows were ignored. The reaction times in the within and between conditions were 609ms and 601ms in the non shadow and 604ms and 605ms in the shadow condition, their associated standard errors were 15ms, 15ms, 16ms and 14ms.

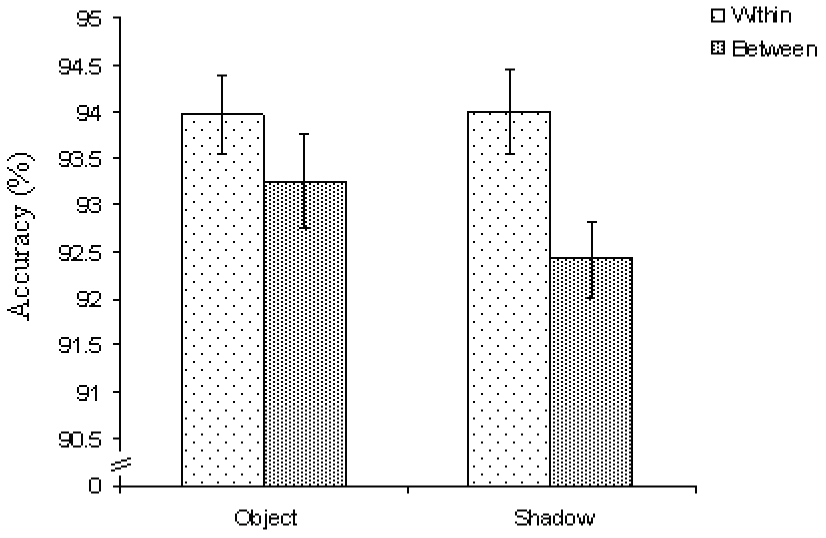


Figure 4, Accuracy data for within and between object target pairs presented on shapes with casters ('shadows') and without casters (object').

---

[15] It is apparent that the magnitude of the accuracy difference reported here is smaller than that previously reported. For instance Ben-Shahar, Zucker and Scholl (2007, exp.1) found effect sizes in the order of 3-4 percent, whereas the difference across the two conditions here is only 1.15%. This difference could well reflect the fact that the spatial distances within the objects were made larger than those between to ensure that any within vs between differences genuinely reflected object based effects. This difference however may also reflect the fact that the targets in Ben-Shahar et al. were presented for very brief periods of time, followed by a mask, leading to much lower accuracies in general (86-76%, depending on their condition). This added difficulty in processing the targets may well bring into clearer light the advantages accrued to target pairings on the same objects and explain why the effect size here is so much smaller. A post target mask was not employed here simply because it is uncertain how this might effect the interpretation of the surfaces as shadows.

Discussion

The representational status of shadows was explored using two object based attention paradigms. The results provide consistent evidence that increasing the extent to which a surface could be interpreted as a shadow had no effect on that surfaces ability to be selected as an 'object' of attention. This study was motivated by the mixed picture regarding the role of shadows in the visual system, which are often actively utilized, in the context of localising the objects that cast them, but often play surprisingly little role in influencing the recognition of the objects casting them. More specifically this study was motivated by visual search data highlighting that access to form information is less efficient when that form can be interpreted as a shadow (Rensink and Cavanagh 2004).

At first glance the current result and that of Rensink and Cavanagh (2004) could appear in contradiction, it could be argued that their study shows that shadows have a lower representational status within the visual system, whilst the current result suggests that the interpretation of a stimulus as a shadow does not have any influence upon the allocation of attention. A possible resolution to this is however brought to light by Lovell, Gilchrist, Tolhurst and Troscianko (2009). In their attempt to replicate Rensink and Cavanagh's study they found that in fact the access to shadow related form information was only less efficient when the perceptual discriminations required a fine scale of comparison. With more coarse discriminations there was no clear difference between shadow and non-shadow stimuli. Lovell et al. employ this data to argue that access to form information is not suppressed per se but that shadows are represented within a functionally distinct, spatially coarse, mechanism.

It is also possible however to frame all three results within one functional framework, namely the Reverse hierarchy model. Lamme and Roelfsema (2000) and Ahissar and Hochstein (2004) have argued that there are distinct feed-forward and feedback modes of visual processing. In both models an initial parse of the visual scene is enabled via a feed-forward parse that establishes salient and/or meaningful aspects of the scene. Access to the fine details of visual stimulation however requires a process of feedback. Lamme and Roelsema draw a parallel between these modes of feed-forward and feed-back processing and the more traditional notions of attentive and pre-attentive forms of visual processing. Recent research from our lab has provided data consistent with the notion that the representations that drive attention are extracted in a parallel, potentially pre-attentive/ feed-forward, visual sweep

(Chapter 3). In this context the current the influence of shadows upon attention suggests that their extraction in this feed-forward sweep must progress in the same manner as visual form information that is not interpreted as belonging to a shadow.

Research using neuropsychology (de-Wit, Kentridge, Milner 2009a) and fMRI/EEG (Martinez 2006) suggests that form information has to be fed up to LO before it can influence the allocation of attention. Although LO has a recognised role in explicit form perception (Malach et al 1995, Kourtzi & Kanwisher, 2001, James, Culham, Humphrey, Milner & Goodale 2003), the feeding-forward of shadow information to LO does not imply that the visual system will have full access to the details of visual form at this level of representation. In fact, quite the contrary. Stanley and Rubin (2003) have shown that the representations in LO seem to be somewhat coarse and reflect the segmentation of an area of space rather than a representation of the precise details of the contour that defines that figure ("the LO... response to salient regions may be the result of fast but crude region-based segmentation processes, which are useful for selecting parts of cluttered images for more detailed, computationally intensive processing" p. 323). The access to more detailed form information related to shadows might therefore require a feedback process to extract this information from earlier levels of representations (for example from the primary visual system). It could be exactly at this stage, whereby the feedback works to extract more detailed information about a shape that has been parsed at a coarse scale within LO, that the visual system seeks reduce the accessibility of information recognised as belonging to a shadow. This explanation is appealing because it not only enables the processing and representation of shadows as a segmented area of space that could be used to inform judgements related to the localisation of the casting object, but at the same time could explain why shadows seem to play little role in identifying the objects that cast them. Indeed previous work demonstrating that incongruent shadows have little impact on the recognition of the objects casting them could be explained if the access to detailed shadow information is reduced such that this incongruence is not extracted.

The recognition of objects is however not always invariant with respect to the shadows that those objects cast. Whilst Braje et al. 2000 and Bonfiglioli et al. 2004 found that the recognition of **natural objects** was highly invariant, the recognition of **novel objects** appears to be influenced by the presence of shadows (Tarr et al. 1998). A possible resolution to this discrepancy could lie in the inability of a feed forward sweep to segment shadows from the objects that cast them for novel objects. It could be that experience in processing a given object is required before it can automatically be parsed as separate from any shadow that it casts. Indeed one might predict both that one's level of experience with an object and the ease with which it can be segmented in any given context will influence the extent to which incongruent shadows are able to interfere with the recognition of that object. The fact that objects were able to elicit an object based attention effect in the present experiments is clearly congruent with the idea that the shadow can be segmented as a separate area of the visual scene. In order to truly test this hypothesis however one would need to explore whether or not an attentional cue on an object automatically led to the facilitated processing of the shadow it is attached to. If the interpretation outlined here is correct, one would argue that when shadows are interpreted as shadows they should be segmented as a distinct area such that cues presented on the object would not lead to an attentional enhancement of the shadow cast by that object. At the same time however one would predict that when the distinction between a shadow and the object casting it is less clear (as would be the case for novel shapes) cues presented on the casting object would also lead to attentional enhancements on the shadows that object casts.

To review this potential interpretation: shadows are able to act as units of attentional selection in the current study because they are extracted and represented as a distinct shape or region in the visual area LO. Indeed this forward pass that delineates a shadow region from other objects in the scene is a critical first step in shadow processing. Other results which suggest that our access to shadow related form information is reduced can be explained by assuming that once a shadow region is extracted via a feed forward sweep to LO access to the precise form information of that shadow would require a feedback sweep of information processing that is inhibited for shadows. This interpretation is not necessarily singled out by the data (indeed many others have to be considered), nevertheless it is offered to highlight how the seeming indifference to shadows in the current study and their reduced accessibility in others can be explained within a reverse hierarchy framework.

The interpretation outlined so far, is however not the only one available. Essentially the results 'prove' a null hypothesis and one could argue that there are reasons to suspect this might reflect a type 2 error. It could be argued that the simple attentional demands of the current paradigms do not provide a stringent test of the representational status of shadows. One could for instance argue that under normal circumstances shadows would be extracted to a lower representational status even in a feed-forward sweep, but that in the present context the simple and consistently repeated visual scenes rendered the shadows as salient targets for attention. Indeed to take this argument to an extreme, one could argue that attention is such a flexible system that it is able to latch onto and select any feature of the visual scene, independent of its general representational status. The clearest counter-argument to this infinitely flexible model of attention comes from the 'chopstick illusion' (Anstis 1990). In this illusion the intersection of two lines is easy to track when those lines are viewed through an aperture that occludes the ends of those lines, but the same intersection becomes almost impossible to track when the full length of the lines is visible. Thus the same low level feature (the intersection between two lines) is sometimes accessible and sometimes inaccessible as a unit of selection depending upon the visual context. Thus although in this experiment there was no evidence that the context in which contrast boundaries can be interpreted as shadows could influence their salience as units of selection, such contextual effects are possible. In the chopstick illusion however the visual system is challenged by a dynamic moving stimulus, and one might argue that potential differential nature of shadow processing might only be brought to light by similar scenes in which the items to be segmented and selected are constantly moving or changing.

In the domain of object based attention, for example, Multiple Object Tracking (Pylyshyn & Storm 1988) might offer a more dynamic object based attention phenomenon to bring to light the nature of shadow processing within the visual system. In the MOT paradigm participants are asked to simultaneously track a set of randomly and independently moving targets amongst identical distracters. With the MOT paradigm one could explore not only the extent to which shadows can be selected as units of attention in a more dynamic scene, but also the manner in which shadows can be ignored as distracters (Pylyshyn 2006). It could be the case that the role of shadows in object recognition might not reflect a reduction in the ease with which they can be selected as targets (as suggested by the current results), but the ease with which shadows can be inhibited as distracters.

Before concluding, one final possibility that needs to be considered is that the manipulation of the status of the shadows and non-shadows in this study was simply not effective. For example it could be that the inclusion of casting objects was not sufficient to cause participants to perceive the shapes as shadows. It is however unclear why this manipulation would be effective in Lovell et al. (2009) and Rensink and Cavanagh (2004) but not in the present context. Nevertheless this possibility could be ruled out by including an independent measure of the extent to which the manipulation of the shapes as shadows is effective. One could for example test whether the manipulation used in this study would lead to a similar pattern of visual search performance as seen in Rensink and Cavanagh (2004).

To conclude, this paper started with the recognition that whilst shadows seem to play some role in localising the objects that cast them, they often play surprisingly little role in recognising those objects. Indeed research on the perception of shadows per se had suggested that explicit access, at least to a fine scale of representation, of shadows was in some manner suppressed by the visual system. Following previous work using object based attention as a tool to explore the representation status of different stimuli, two object-based attention paradigms were employed to try to clarify the status of shadows within the human visual system. The results provide clear evidence that cues to interpret a surface as a shadow did not reduce their ability to be extracted to a level that would influence the allocation of attention. This result could reflect an unaltered feed-forward sweep of information processing to the level of the Lateral Occipital area, at which representations are maintained somewhat coarsely. If one assumes more detailed form discriminations require some form of feedback to earlier areas of visual representation one could argue that the previous ‘discounting’ effects of shadows in fact reflect a reduced ability to apply feedback processing to stimuli interpreted as shadows. This pattern of an intact feed-forward sweep and a restricted feedback sweep may explain both the ability of shadows to be fed on and influence localisation tasks whilst at the same time have relatively little impact on the recognition of the objects that cast them. Some caution should be applied in interpreting the current null result, however, and it could be wise to explore the extent to which shadows can be selected as targets and suppressed as distracters with more dynamic object based attention phenomena such as Multiple Object Tracking.

Chapter 7 - Multiple Object Tracking (MOT): A specialised system for multiple people tracking or a result of indexing early visual ‘proto-objects’?

Abstract

This paper explores whether Multiple Object Tracking (MOT) is specialised for “multiple people tracking”. Tracking performance is examined for both biological and non-biological stimuli composed of point lights (point-light biological motion and letters), in upright inverted and scrambled conditions. Inverted and scrambled stimuli were tracked less successfully, but this effect was found for both biological and non-biological targets. Thus the results provide no evidence for selective mechanisms for “multiple people tracking”. Nevertheless the results suggest that information regarding the nature of the stimulus does have a role in MOT performance. This finding is discussed in relation to hypotheses regarding the encapsulated nature of proto-object formation, and is interpreted to demonstrate that gestalt principles, especially common fate, play an important in proto-object formation.

1. Introduction

Humans are able to detect and track moving objects in the presence of multiple distractors. In real life, the objects we need to follow are often individuated because of their biological significance. If one thinks of a football player mindful of the location of their team mates, or a stressed mother trying to keep track of her children at the January sales on Oxford Street, it is evident that the tracking of biological targets is both commonplace and important. Indeed, biological targets have specific properties that make multiple object tracking a particularly critical skill. Unlike non-biological objects, biological targets can move in independent, unpredictable and rapidly changing ways. It is not clear, however, whether the cognitive processes underpinning the ability to track moving objects are in some manner specialised for the tracking of biological stimuli, or instead provide a more general-purpose mechanism.

The multiple object tracking or MOT paradigm has been extensively used to study this ability (Pylyshyn & Storm 1988). To date, a major focus has been on understanding what performance in this task can reveal about the operation of the early human visual system. For example, Pylyshyn (2001) operationalizes the notion of a ‘primitive visual object’ (or proto-object) exactly in terms of the visual units that enable individuation and tracking in the

Multiple Object Tracking (MOT) paradigm (Pylyshyn & Storm 1988). Pylyshyn argues that tracking is enabled using a limited number of visual indexes that point to, or reference, these proto-objects.

Nevertheless, what constitutes an object in multiple object tracking remains poorly understood (Scholl, Pylyshyn, Feldman 2001). Pylyshyn (2001) has argued that the objects (in fact the 'proto-objects') that underpin MOT are computed on the basis of pre-conceptual mechanisms encapsulated in the early visual system. According to this view, although proto-objects can be used as a kind of scaffold on which to frame conceptual knowledge, this knowledge cannot be used in the formation of these objects per se. Within this conceptualization it seems unlikely that the higher-level properties of an object, such as its biological status or "meaning", could play a role in MOT. However, the biological status of an object can influence what would traditionally be considered as early orienting responses in other paradigms. For example, upright faces have a much more disruptive influence on involuntary saccade programming compared with inverted faces that are matched in low-level properties (Gilchrist and Proske, 2006). If the biological status of a stimulus can rapidly influence the computations underpinning eye movements, perhaps it can influence other aspects of visual cognition involved in multiple object tracking?

A number of studies highlight factors that suggest that MOT could play a useful role in our daily life: it can operate with wide fields of view (Franconeri, Lin, Pylyshyn, Fisher, Enns 2008), in complex scenes (Zelinsky & Neider 2008), and is more effective in 3D scenes, with surfaces at different depths (Viswanathan and Mingolla 2002). Indeed Zelinksy & Neider (2008) have validated MOT using biological targets (sharks). However, no previous study has explicitly compared MOT for biological and non-biological targets in order to determine whether MOT has a biological specialisation.

Here, the potential importance of biological information in MOT was explored by studying the ability of human observers to track multiple objects comprising point-light biological motion displays. Such objects consist of a few moving points, yet evoke a clear percept of a body in motion (Johansson, 1973). Prior work has established procedures for creating control stimuli that preserve the local information in these animations whilst disrupting the global form or gestalt, namely, inverting or scrambling the point-lights (see Blake & Shiffrar, 2007 for a review). Indeed, inversion or scrambling can influence perception, even when all other aspects of the stimuli relevant to the task remain unaltered (for evidence from motion coherence and audiovisual temporal perception domains see Saygin, Driver & deSa, 2008; Tadin, Lappin, Blake, & Grossman, 2002).

In a pilot study, the ability of participants to track point light biological motion animations was impaired when these figures were inverted[16]. This inversion effect suggested that the biological status of the target may play a role in MOT. This pilot data motivated our main experiment in which the specificity of this inversion effect is tested by using biological as well as non-biological stimuli (a point light letter R). The small size of the inversion effect in our pilot data also prompted the inclusion of a scrambled condition, in which the dots composing each stimulus were shifted in their starting position in order to further eliminate the perception of biological motion. This contrast between biological and non-biological targets under upright, inverted and scrambled conditions enabled us to both explore the potential importance of biological information in tracking, and the formation of visual proto-objects more generally.

2. Methods

2.1 Participants

Thirty adults (18 females, mean age = 21 years, SD =6.4) gave written informed consent to participate in the experiment. All participants had normal or corrected-to-normal vision. The experiment was approved by the local ethics committee.

2.2 Stimuli

Stimuli were presented using Matlab (Mathworks, Natick, MA, USA) and the Psychophysics Toolbox (Brainard, 1997). On each trial, eight individual object stimuli were presented simultaneously. These comprised either all biological or all non-biological point-light stimuli. On each trial all eight stimuli were presented either upright, inverted or scrambled. All eight stimuli were identical (see figure one).

---

[16] Pilot data were collected with the same methods described below employing only the upright and inverted conditions. Average tracking performance for 12 participants was 70% in the upright and 66.6% in the inverted conditions ($t(11)=2.56$, $p=0.026$).

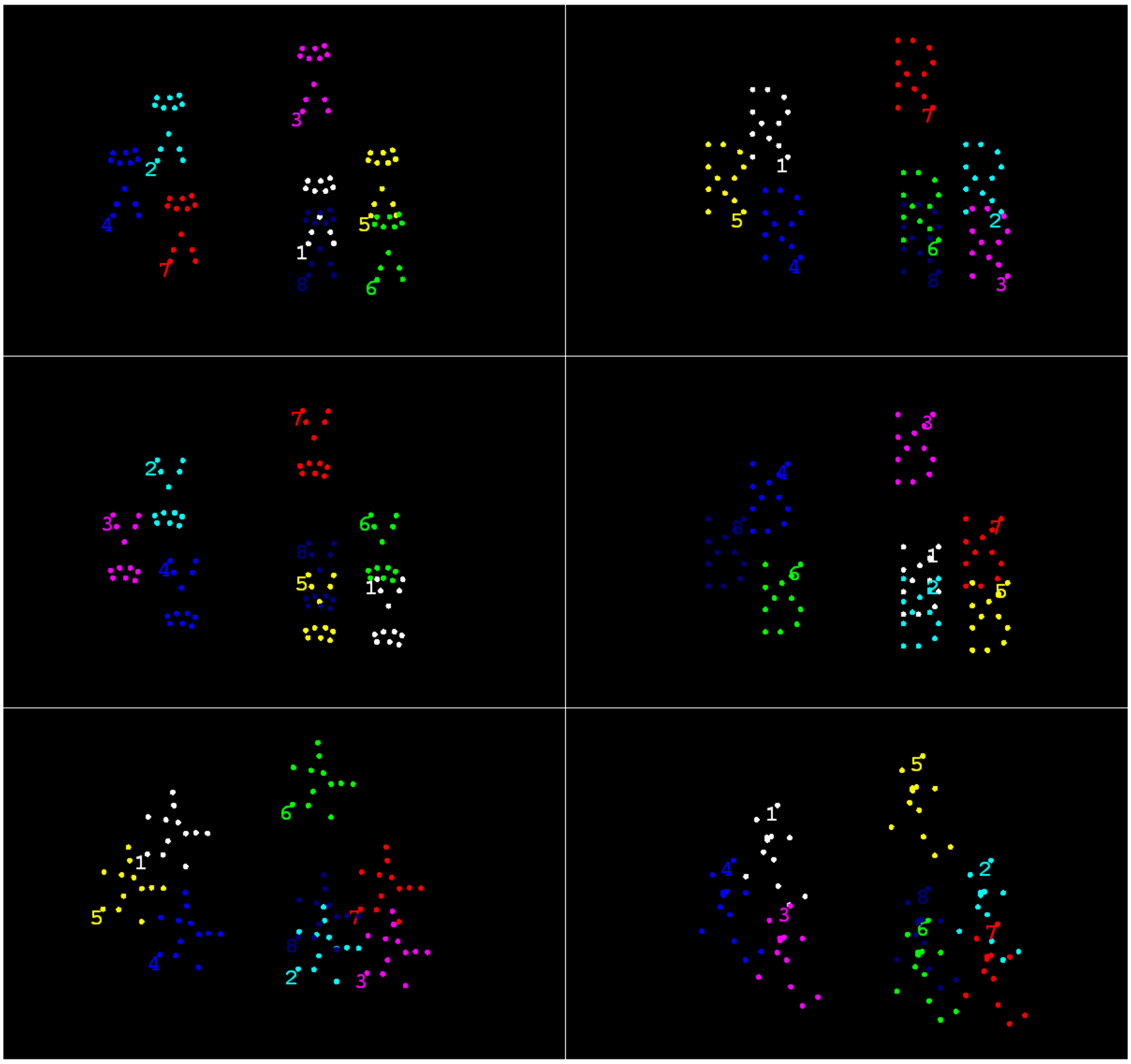


Figure One

Frames from the biological (left panel) and non-biological (right panel) stimuli in the upright, inverted and scrambled conditions. The participants were presented with exactly these displays at the end of each trial and were prompted to report the numbers corresponding to the four targets they had been tracking. During the trial, when the targets and distracters were moving around the screen, all points appeared in the same colour (white), apart from the first 30 frames in which the four targets flashed in red. Colours are added to this figure merely to aid the separation between each figure which is easier to see when each figure is in motion.

Each individual biological motion stimulus comprised an animation that was created by videotaping an actor and then encoding the joint positions in the digitized videos (Ahlstrom et al., 1997). We selected one specific animation depicting a ‘star jump’ (or jumping jacks) as this action does not have an obvious implied direction of left/right motion. The joints were represented by twelve small white dots each subtending ~0.015 degrees of visual angle (participants viewed the screen at approximately 57 cm) against a black background. The height of each figure subtended 2 deg of visual angle, the width varied (with the motion of the arm and leg joints) from 1.5 to 0.9 deg. Each ‘star jump’ consisted of 20 frames, which looped continuously throughout the trial.

Non-biological control stimuli were constructed using 12 points of light to form an uppercase letter R that had a height of 2 deg and width of 1.2 deg of visual angle. The internal structure of the non-biological stimulus, unlike the biological ‘star jump’, did not change throughout the trial.

On each trial participants were presented with eight point-light stimuli simultaneously. Each moved independently around a 10 by 8 degree area of the screen, following independent random paths. These paths were constructed off-line prior to the experiment using linked Bezier curves. The rate of change in these paths was constrained to avoid any sharp changes in direction. Each stimulus moved through 80 points defined along a curve. These points were constrained so as to not be separated by more than 8 pixels. A trial included 80 frames, and each frame was presented for 50ms. For the first 30 frames, four of the stimuli flashed in red on alternate frames to indicate that they should be tracked as the targets while the other four should be ignored. All targets and distractors then appeared white as they moved around the screen for 50 frames. At the end of a trial each individual stimulus changed to a different colour, and a number was presented next to each stimulus in the same colour (Fig. 1). Participants were required to report the numbers associated with the four target stimuli.

On each trial all eight stimuli were presented, either in the upright, inverted, or scrambled format. The inverted stimuli were generated by vertically rotating each stimulus around its centre. Scrambled biological motion animations were constructed by randomizing the starting positions of the points while keeping the motion trajectories of each individual dot intact. The starting positions were chosen randomly within a region such that the total area encompassed by each figure was similar to that of the upright figures. The inverted and scrambled animations therefore contained the same local motion cues but did not have the same global form as the upright animation. The non-biological (letter) stimuli were transformed into inverted and scrambled versions in the same manner as the biological motion figures.

2.3 Design and Procedure

Participants were presented with one block of biological and one block of non-biological targets, the order of which was counterbalanced across participants. Each block contained 60 trials with an equal number of upright, inverted and scrambled figures, presented in a random order.

3. Results

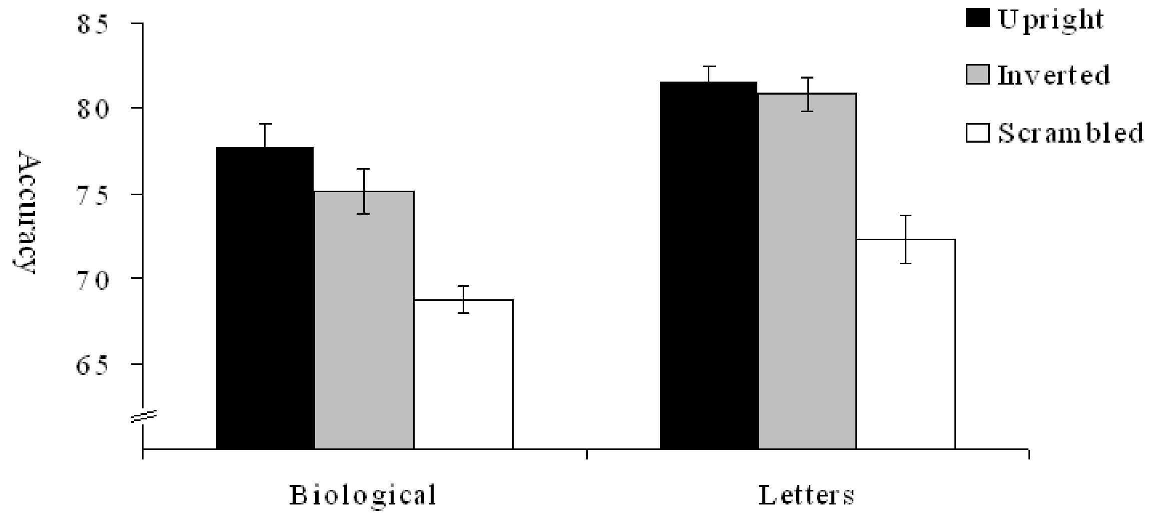


Figure Two, average percentage accuracy and associated standard error for tracking the upright, inverted and scrambled presentation of biological and letter targets.

The data were analysed using a repeated measures ANOVA, with two factors, stimulus type (biological, non-biological) and presentation type (upright, inverted, scrambled). The results, in terms of percentage correct responses, are presented in Figure 2. There was a main effect of stimulus type, such that the non-biological letter targets were in fact easier to track (F(1,29)=62.74, p<0.001). There was also a main effect of presentation type (F(2,58)=85.84, p<0.001). Scrambling the stimuli (relative to upright stimuli) had a substantial effect on tracking with significantly lower accuracy for scrambled stimuli (F(1,29)=140.7, p<0.001). This effect of scrambling did not interact with target type (F=0.11, p=0.916). Critically, the absence of an interaction between target type and scrambling shows that the advantage for non-biological letter targets in contrast to biological targets does not alter when scrambling renders these stimuli un-interpretable as biological motion, or as letters.

The effect of inversion (relative to upright targets) was smaller but significant (F(1,29)=5.197, p=0.03). The interaction between the upright and inverted conditions for the different target types was not significant (F(29)=1.38,p=0.25), although as seen in Figure 2, the main effect of inversion was primarily driven by the biological motion condition (paired samples t-tests for up-right vs inverted targets in with biological stimuli reveals a significant difference (t(29)=2.58, p=0.031) whereas the same is not the case for letter stimuli (t(29)=0.71, p=0.49)).

4 Discussion

This study sought to investigate whether multiple object tracking (MOT) mechanisms are specialised for tracking biological stimuli. We employed stimuli composed of point lights depicting biological or non-biological objects (a person performing a star jump, and the letter R, respectively), under different presentation conditions.

Participants were better at tracking letters in comparison to biological motion animations. Critically, however, this advantage for non-biological stimuli was maintained in the scrambled condition, suggesting that the difference between these target types cannot be framed in terms of an advantage for letters per se. This absence of an interaction between scrambling and stimulus type suggests that the difference between tracking accuracy for these stimuli in the upright condition reflects some property of these stimuli that is maintained by scrambling. Such a property is clearly evident in the local motion profile of the dots making up each figure. For the biological targets, the individual dots move relative to each other; but for the non-biological targets, the relationship between the dots making up each letter R remains constant throughout the tracking period. The main effect of target type can therefore be framed in terms of the Gestalt principle of “common fate”: the dots that make up the letters all share common fate, whilst those making up the biological targets move relative to each other. Common fate might therefore be an important factor in the formation of the early visual ‘proto-objects’ that are thought to underpin tracking performance.

There was a clear and significant decrease in performance with scrambling for both non-biological and biological stimuli. This indicates that there was information in the coherent upright figure that facilitated tracking. Our failure to find a significant interaction between scrambling and stimulus type suggests however that biological information receives no special treatment in MOT, at least for the stimuli used here.

The present results replicated the inversion effect of biological motion on MOT that we found in our pilot experiment (see Introduction). The inversion effect was only significant for the biological targets. In the absence of a statistical interaction of inversion and stimulus type however, we cannot reliably take this as evidence for a special role for biological information in MOT. In any case we would argue that any potential difference between the inversion effect for biological and non-biological targets can be most meaningfully framed in terms of the common fate properties of each stimulus type: Inversion is classically thought to undermine the ability to utilise previous processing of a stimuli that has a canonical orientation, this disruption of the 'top-down' facilitation of information processing is clearly likely to have a greater impact when bottom up cues to grouping (such as common fate) are less salient. The strong bottom up cues to grouping via common fate for the letter targets could therefore render the role of top down facilitation (that would be disrupted by inversion) less significant.

The role of common fate in our results resonates with Pylyshyn's arguments that pre-conceptual principles play an important role in parcelling out a visual scene into candidate proto-objects. Pylyshyn (2001) additionally notes however that whilst proto-objects can be used as a scaffold upon which conceptual knowledge can be assembled, they are constructed in a manner that is encapsulated from (and therefore unaffected by) conceptual knowledge. To the extent that proto-objects have been viewed as synonymous with the units of selection in MOT, our data suggest instead that 'knowledge' of a candidate object can play some role in the formation of proto-objects. However, the scrambling manipulation, where we have our strongest effects, not only disrupts the high-level nature of the stimulus but also alters low level cues to grouping (for instance the straight 'back' on the upright letter R). The role of higher level knowledge in MOT is more evident in the inversion effect. This inversion effectively leaves the low level properties of the stimulus intact, but disrupts the facilitation of processing a familiar stimulus that has a canonical orientation. This suggests that 'knowledge' of the stimulus is used in forming a proto-object, thereby leading to an improvement in tracking.

Of course, it is possible that this ‘higher-level’ information may have become embedded in the process of early object formation (Pylyshyn 1999). Thus, rather than the visual system making a high level interpretation of the stimulus as biological or meaningful, which then affects the process of proto-object representation, it may be that the statistics or regularities associated with the stimuli (such as the dynamics and kinematics of biological motion) have become embedded in early visual processing, in a similar manner to other non-conceptual gestalt cues. Indeed, primary visual cortex, where receptive fields are not large enough to “see” the stimuli in its entirety, responds retinotopically to the contrast of biological motion to scrambled controls, even in the absence of attention (Saygin & Sereno, 2008). On the other hand neuropsychological studies have revealed higher-level areas in the frontal and superior temporal lobes as the critical neural substrates for biological motion perception (Saygin, 2007), however such data are not inconsistent with the possibility that biological motion might also receive processing in early visual cortices, even if these are not in themselves critical to perceiving the stimuli per se. It is probably not possible to identify the precise level at which stimulus properties influence MOT from the current data. Future work with neuroimaging and neuropsychological patients could shed light on whether the selection of these targets as ‘objects’ is determined by computations performed solely within the early visual system, or rely on higher brain areas and processes.

5. Conclusions

In summary, the present study compared the ability to track point light biological motion and letter stimuli in order to assess the potential role of biological information in multiple object tracking. The results provide no evidence for a special role for biological information in MOT. Thus whilst MOT might well aid activities in our daily life, it does not appear to be specifically adapted for “multiple people tracking”. Our data suggest the importance of the gestalt principles (e.g., gestalt law of common fate) in MOT. This idea continues to build a bridge between gestalt psychology and more recent concepts of proto-objects (Marino and Scholl 2005), the formation of which are arguably critical to our ability to parse a visual scene prior to the application of conceptual knowledge. The results also suggest that previous experience regarding the nature of a given stimulus is utilized in MOT. It will be interesting to explore whether these effects reflect computations embedded in the statistics of the early visual system, or the ability of higher-level representations to influence early visual proto-object construction, which has previously been hypothesized to be an encapsulated process.

Section Three: Object Based Attention, Perceptual Organisation and Shape Processing Area LO.

The previous two sections have attempted to advance our understanding of object based attention and to use object based attention to ask questions about the nature of perceptual organisation more generally. In this final section the nature of perceptual organisation and object based attention are explored in a patient with visual form agnosia. This patient, DF, has most predominantly been studied in the context of a profound dissociation between her preserved ability to manually act upon objects that she is otherwise unable to perceive (Goodale & Milner 2002, Milner & Goodale, 1995). Combined with a prior understanding of anatomical divisions in the primate visual system Milner and Goodale argued that this patient proved that vision-for-action and vision-for-perception could in principle be computed by distinct neural resources. Milner and Goodale argued that the preserved vision-for-action in this patient was mediated by a geniculo-striate pathway that continued to pass information to areas of the dorsal stream in the parietal lobe. In contrast the inability to develop perceptual representations of visual form information was thought to reflect a lesion to occipito-temporal areas within the ‘ventral stream’. This interpretation has been borne out in more recent neuroimaging data (James et al. 2003), which has revealed that the vision-for-action computations required to grasp objects in this patient are reflected in a preserved pattern of activation in the primary visual cortex and areas of the parietal lobe. In contrast, DF’s lesion in the ventral stream seems to centre on the area LO which is typically activated during the perceptual recognition of shape (Malach et al. 1995), indeed these areas reveal absolutely no functional activation in DF.

In the introduction it was noted that neuro-imaging data was consistent with the idea that this area might play a critical role in constructing the representations that influence attention (Martinez et al. 2006). Indeed this possibility was further reinforced by the data collected in Chapter 5, showing that just as this area is able to construct shape representations on the basis of tactile information, so too object based attention is influenced by objects that can be held manually, even when those objects have never been seen. Patient DF therefore offers an opportune chance to test the whether or not this area is truly critical to generating the representations that influence attention. Furthermore patient DF also offers an opportunity to contribute to the behavioural debate laid out in the introduction, regarding the possibility that simple edges (Avhrami 1999) vs the higher level interpretations applied to those edges in terms of their figure ground relationships (Ben-Shahar et al. 2007, Albrecht et al. 2008), are required in order to influence the allocation of attention. DF’s pattern of behavioural performance is also informative in this context because oriented edges are clearly extracted at early stages of the visual system that not only enable vision-for-action when passed to the dorsal stream (James et al. 2003), but also give rise to an orientation specific after image (or illusion) known as the McCollough effect (Humphrey, Goodale & Gurnsey 1991). If DF shows an object based attention effect it would suggest that this early registration of edge information is sufficient to drive the allocation of attention. If however DF’s ability to allocate processing resources is not sensitive to the classical Egly et al. rectangular objects this would suggest that higher level processing resources are required in order to interpret what edges might mean in terms of figure ground segmentation before they can influence attention. Chapter 8 therefore explores the existence of both spatial, and object based forms of attention in DF, in order to clarify the role of basic edge information, and the shape processing area LO in generating object based attention effects.

Chapter 9 employs a similar logic to explore the resources required to generate a perception of illusory contours or Kanizsa figures. Kanizsa figures were employed in Chapter 4 as a means of operationalizing the construction of a ‘mid level’ object within the visual system. As noted in that chapter there is some debate regarding whether or not this phenomenon reflects a process of edge completion in the primary visual cortex, or requires a genuine interpretation of shape based on resources in the lateral occipital cortex. Single unit recordings and fMRI data have provided a mixed picture on the possible role of many areas of the visual system in the construction of this percept. Chapter 9 exploits DF’s preserved motion perception to explore whether or not her visual system is implicitly able to construct the Kanizsa figures.

Combined with previous evidence regarding DF’s ability to exploit ‘gestalt cues’ to visual organisation (Goodale et al. 1994), chapters 8 and 9 highlight that visual area LO plays a critical role in many aspects of perceptual organisation. All these previous examples however explore this patient’s ability to organise information on the basis of shape cues. Chapter 10 therefore explores whether all forms of perceptual organisation are dependent upon this area, or whether groupings between aspects of perception that are somewhat preserved for DF (colour and motion) are still organised to a level where they can influence her perceptual performance.

Chapter 8 - Object based attention and visual area LO[17]

Abstract

We investigated the neural basis of so-called "object-based attention" by examining patient D.F., who has visual form agnosia caused by bilateral damage to the lateral occipital area (LO) of the ventral visual stream. We tested D.F.'s object-based attention in two ways. In the first experiment, we used a spatial cueing procedure to compare the costs associated with shifting attention within versus between two separate outline figures. D.F. did not show the normal advantage of within-object over between-object attention shifts. In the second experiment, we used a complementary paradigm in which two separate stimuli, presented either on the same or on different objects, have to be identified as the same or different, We found no evidence for the normal pattern of superior performance for within versus between figure comparisons. In a third experiment, we checked that D.F. showed normal shift costs for invalid as opposed to valid cueing in a standard Posner spatial attention task. In a final experiment, we compared horizontal versus vertical attention shifting in group of healthy controls without the presence of outline rectangles, and found that their pattern of shift costs was indistinguishable from that seen in D.F. when the rectangles were present (Experiment 1). We conclude that whilst D.F. has a normal spatial orienting system this is completely uninfluenced by object structure, we suggest that area LO may mediate form processing precisely at the stage where visual representations normally influence the spread of attention.

1. Introduction

Ungerleider and Mishkin (1982) first proposed that the cortical processing of visual information was divided into two functionally and anatomically distinct streams. Within their interpretive framework the extraction and representation of objects was viewed as the preserve of the ventral stream. It is now clear from both primate and human research, however, that shape selectivity is a feature of both the dorsal and ventral streams (Milner & Goodale, 1995, 2008). These two forms of shape encoding are not mere replicas of each other but show distinct patterns of selectivity and generalization (James et al., 2002; Lehky and Sereno, 2007). Furthermore Milner and Goodale (1995) have argued that the shape representations in each stream differ not only in their nature, but more importantly in their

---

[17] This chapter has been published in Neuropsychologia, de-Wit, Kentridge and Milner (2009a)

function. Shape representations in the ventral stream are used to inform perception, while shape representations in the dorsal stream are used to guide actions.

Visual form, however, not only drives recognition and action but also influences the allocation of attention: it is easier to shift attention within a single object form than between two separate forms (Duncan, 1984; Egly, Driver & Rafal, 1994). This influence is thought to play a role in fine tuning the construction of perceptual representations, by focusing visual processing on those elements of a scene that group together (Driver, Davis, Russell, Turatto & Freeman 2001). It is unclear whether the same or similar influences of visual structure might also operate on attentional selection within the shape representations that guide manual actions to those objects. In terms of Milner and Goodale's (1995, 2008) two visual-systems framework, we would predict that the classical perceptual "object-based" attention effects would be mediated within the ventral stream, and therefore not play a direct role in guiding manual actions. The well studied patient D.F. has a bilateral brain lesion that is centred on ventral-stream area LO, which is known through neuroimaging research to play a crucial role in shape perception (Kourtzi & Kanwisher, 2001; James, Culham, Humphrey, Milner, & Goodale, 2003). D.F. therefore provides a strong test of the potential role of area LO not just in perception, but also in object-based attention. D.F. is, however, clearly able to segregate candidate objects for manual action, and to use that visual information to calibrate her movements towards such objects. Thus if object-based attention operates not only within the perceptual domain, but also during the processing of candidate objects for action, it should be intact in D.F., at least within the dorsal stream. An absence of classical perceptual object-based attention effects in D.F. would suggest that this influence of visual structure upon attention is either not important for guiding manual actions towards visual objects, or that it is organized quite separately within the two visual streams.

DF's impairment in shape perception may also be informative in terms of the debate within the object-based attention literature regarding the level in the ventral stream at which visual structure influences attention. At present, although the ability of visual structure to influence attention is uncontroversial (Scholl, 2001), it is still unclear at exactly what level of representation object shape exerts its influence on attention. Indeed Driver and colleagues (2001) have persuasively argued that referring to the influence of visual structure upon attention as "object-based" is strictly inappropriate (though we will continue to use this questionable nomenclature within the present paper). They argue that it is not objects, in the sense of the visual units that provide the mental furniture for visual recognition and visual cognition, that influence our allocation of attention. Rather, they argue, attention interacts with earlier visual information at the stage of scene segmentation, thereby serving to structure the visual input to enable genuine object representations to be extracted later in the system.

Early segmentation mechanisms and later object representations certainly appear to have quite different neural bases: there is evidence that edge extraction and figure-ground assignment take place in the primary visual cortex (Lamme, 1995), whereas shape as a property of objects is associated with activity in the lateral occipital cortex (Malach et al., 1995; Kourtzi & Kanwisher, 2001). The neural basis of object-based attention, however, still remains unclear. For example whilst Müller and Kleinschmidt (2003) have provided neuroimaging evidence that the enhanced activation of locations within objects manifests itself within the primary visual cortex, they accept that this could well reflect feedback from higher ventral- or dorsal-stream areas. Evidence supporting the putative role of the lateral occipital area (LO) comes from a recent study by Martinez and colleagues (2006), who investigated object-based attention using a combination of EEG and fMRI. They found that object-based attention was clearly associated with the enhancement of the N1 component of the ERP waveform. They go on to argue that this component principally reflects activation in the lateral occipital complex, while noting that other areas too can contribute to the N1 component (Di Russo, Martinez & Hillyard, 2003).

The distinction drawn by Driver et al. between early scene-parsing mechanisms and later object-based processing can also be examined at a behavioural level. Again, however, the picture remains as yet unclear. On the one hand object-based attention can be demonstrated when attention simply has to move across some horizontal lines (Avrahami, 1999), suggesting that such simple segmentation cues are able to modulate attention. On the other hand, several researchers have demonstrated that the influence of visual structure upon attention is modulated by high-level factors such as the perceptual interpretation of the nature of the shape (Chen 1998) and previous perceptual experience of that shape (Zemel, Behrmann, Mozer & Bavelier, 2002).

The present study seeks to shed light on this issue by examining object-based attention in patient D.F. Her pattern of impairment, both in terms of its neural locus and its manifestation as visual form agnosia, provides a strong test for the level at which visual structure influences attention. DF’s lesion has resulted in a complete bilateral disruption of processing in area LO whilst leaving functionally intact processing in the primary visual cortex, which plays a pivotal role in feeding information to her largely-intact dorsal stream to guide manual actions (James et al., 2003). Despite her severe impairment in visual form perception (Milner et al., 1991) D.F. is still able to perform basic scene segmentation tasks. For example, if asked to count the number of disks in a display DF is able to segment these from their background and report with near-perfect accuracy (Carey, Dijkerman, Murphy, Goodale, & Milner, 2006). By examining the nature of object-based attention effects in D.F. we can therefore test the level at which visual structure in the ventral stream influences attention. We have done this by using two complementary experimental paradigms. In Experiment 1, we used the paradigm of Egly, Driver and Rafal (1994), in which a Posner-style spatial cueing procedure is used to compare attention shifting made either within or between two separate outline figures. Despite the similar distance over which attention has to shift from an invalid cue in either case, healthy subjects show a reliable reaction-time advantage of within-object over between-object attention shifts. In the second experiment, we used the complementary paradigm of Marino and Scholl (2005), in which two stimulus elements have to be compared either within the same outline figure or between two separate figures. Healthy subjects are facilitated in making same/different judgements for targets that fall within the same figure.

2. Participants

DF was aged 53 at the time of testing. Her clinical background has been described in detail elsewhere (Milner, et al. 1991; Goodale & Milner, 2004) and MRI scans of her lesions have been presented by James et al. (2003). Experiments 1, 2 and 3 additionally used eight age-matched control subjects (mean age = 51; SD= 4, for experiments 1 and 2, mean age = 52: SD = 4 for experiment 3). Experiment 4 used 6 healthy control subjects (mean age = 30; SD= 8). All control subjects had normal or corrected to normal vision and were students or staff at Durham University.

3. Experiment 1

3.1. Methods

Experiment 1 used an adaptation of the cueing paradigm devised by Egly and colleagues (1994). Participants were presented with two white outline rectangles subtending 17 deg by 3.6 deg of visual angle oriented vertically 6.5 degrees either side of the centre of the screen. The rectangles had a white outline, 8 pixels thick, presented on a black background. This outline display was present throughout each experimental session. At the start of a trial one of four possible target locations at the end of one of the rectangles would be cued with a bright grey square that would appear for 200ms (see Figure 1, top right). The participant was then presented with a red or green target of similar size, which they were asked to discriminate by pressing one of two buttons as quickly and accurately as possible, using the left index finger for one button and the right index finger for the other. The target could appear in one of three locations relative to the cue: either in the identical location (valid cue trial), at the other location on the same object (invalid within-object trial), or at the homologous location on the other rectangle (invalid between-object trial). These three trial types are illustrated in Figure 1 (bottom three panels). There was 1500ms inter-trial interval during which only the two rectangles remained present on the screen. Participants were tested in 6 blocks of 96 trials. Fifty percent of the trials were validly cued, 25% were invalidly-cued within-object and 25% invalidly cued between-objects. The order of trial types was randomly counterbalanced within each block. Trials with inaccurate responses and those with reaction times likely to be anticipations (less than 150ms) or mis-keyings (more than 1500ms) were removed prior to analysis. Any remaining outliers (responses that were two standard deviations above or below the mean for each subject in each condition) were also removed. This procedure removed 6% of trials as outliers.

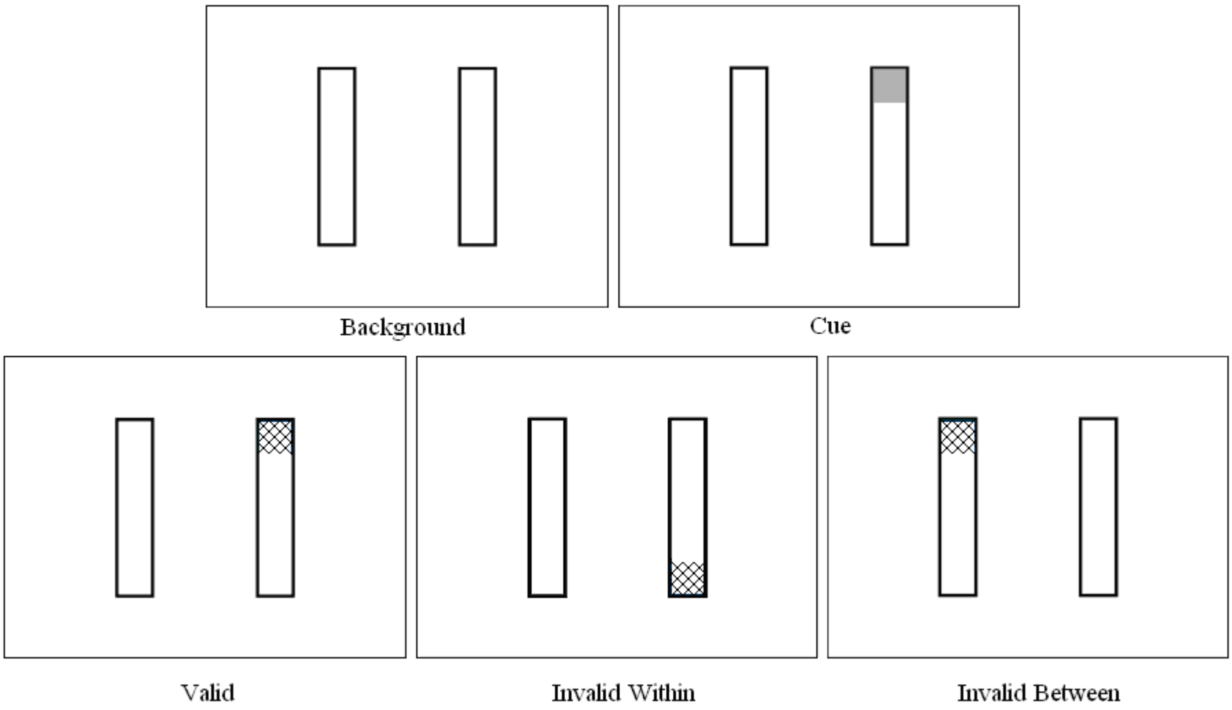


Fig 1 *- Experiment 1.*

- Illustrative stimuli for the object-based cueing paradigm. The targets (cross hatched) were green or red, the outline rectangles were white on a black background.

The stimuli in Experiments 1, 2 and 4 were displayed on an Eizo F55 17-inch (600 x 800 pixels) colour monitor, driven by a Cambridge Research Systems VSG 2/5 graphics system at 100 Hz. Responses in all experiments were recorded using a Cambridge Research Systems CT3 button-box.

3.2. Results

Here and throughout the paper, all statistics are reported as two-tailed. The control group clearly demonstrated the object-based attention effect reported by Egly et al. (1994), in that invalidly cued targets appearing on a different object took longer to recognize than invalidly cued targets appearing on the same object [t(7) = 2.35, p = 0.051, within object mean RT 468 ms, between object mean RT 483 ms]. D.F., however, showed a quite different pattern of performance. While she was faster to respond to validly cued than to invalidly cued targets (t(498)=7.993, p<0.001), she did not show the normal advantage for within-object invalid trials over between-object invalid trials [valid mean RT 726 ms, invalid within-object mean RT 882 ms, invalid between-object RT 839 ms].

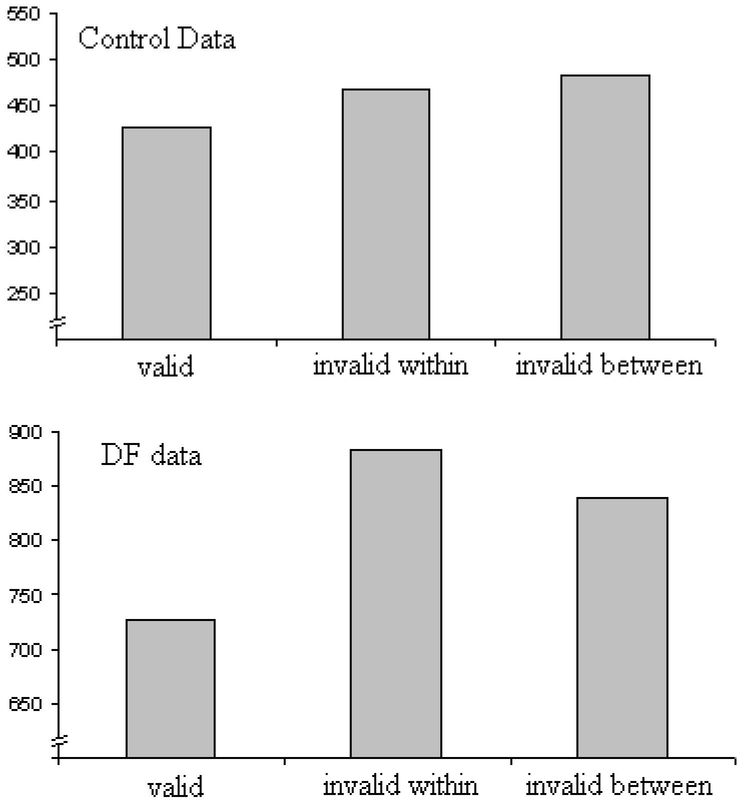


Fig 2 *- Experiment 1.*

- Reaction times (ms) for target discriminations across the three cueing conditions for a healthy control group and for patient D.F... Note the y axis uses a different intercept for D.F. and the control sample.

It is quite apparent that D.F.'s performance is in general slower than that of controls. However, what is most critical in interpreting our studies is determining whether the difference in D.F.'s reaction times across within- and between-objects conditions is what one might expect of a normal observer or whether her behaviour differs significantly from that of our control sample. It is common to treat the means and variances of data obtained from control *samples* as if they were *population* parameters rather than sample statistics. Doing so when comparing difference scores with relatively small control sample sizes of the order we use here can lead to highly inflated Type-I error rates (see e.g. Crawford & Garthwaite, 2005). We have therefore adopted an approach suggested by Crawford & Garthwaite (2002, 2007) in which the mean and variance of the scores obtained from the control sample are treated correctly as statistics in the course of assessing the likelihood that the patient's scores were drawn from the same population as control sample's. We have used Crawford & Garthwaite's

(2007) Bayesian calculation for within subject control scores as this allows us to preserve raw, unstandardised, reaction-time measures in our calculations.

Comparing D.F.'s difference score (between-object invalid vs within-object invalid RTs) with those of the sample indicates that D.F.'s RT difference lies at the extreme of normal observers' performance. The Bayesian estimate of the percentage of the control population estimated to show a more extreme RT difference is 2% (i.e. p=0.02). Accuracy was near ceiling for the control subjects in both between and within object trials and these did not differ statistically [t(7)=1.37, p=0.21]. D.F. has good colour discrimination (Milner & Heywood 1989) and indeed responded correctly on every trial. Her accuracy did not differ from that of the controls on Crawford and Garthwaite's test [p=0.66].

3.3. Interim discussion

According to our use of a variation of Egly et al.'s (1995) paradigm it is clear that D.F. did not show the normal influence of objects upon attention. If anything, she showed an apparent reversal of the normal trend. It is possible, however, that this trend merely reflects the differential difficulty of making a vertical as opposed to a horizontal shift within this task, since we did not balance our experimental design to rule out this possibility. We examined this interpretation in Experiment 4. But first we sought in Experiment 2 to establish whether the absence of an object-based attention effect in D.F. could be replicated using a quite different experimental paradigm.

4. Experiment 2

4.1. Methods

Experiment 2 employed a modified version of the two-item comparison paradigm employed by Marino and Scholl (2005), in which participants are asked to report whether two items, either located on the same object or on two different objects, are the same or different. The display in this experiment contained the same pair of rectangles as used in Experiment 1, which again formed the background that was present throughout each experimental session. On each trial the participant was presented with two square patches of colour (equi-frequently red or green), which could either occupy opposite ends of the same rectangle, or homologous ends of the two different rectangles.

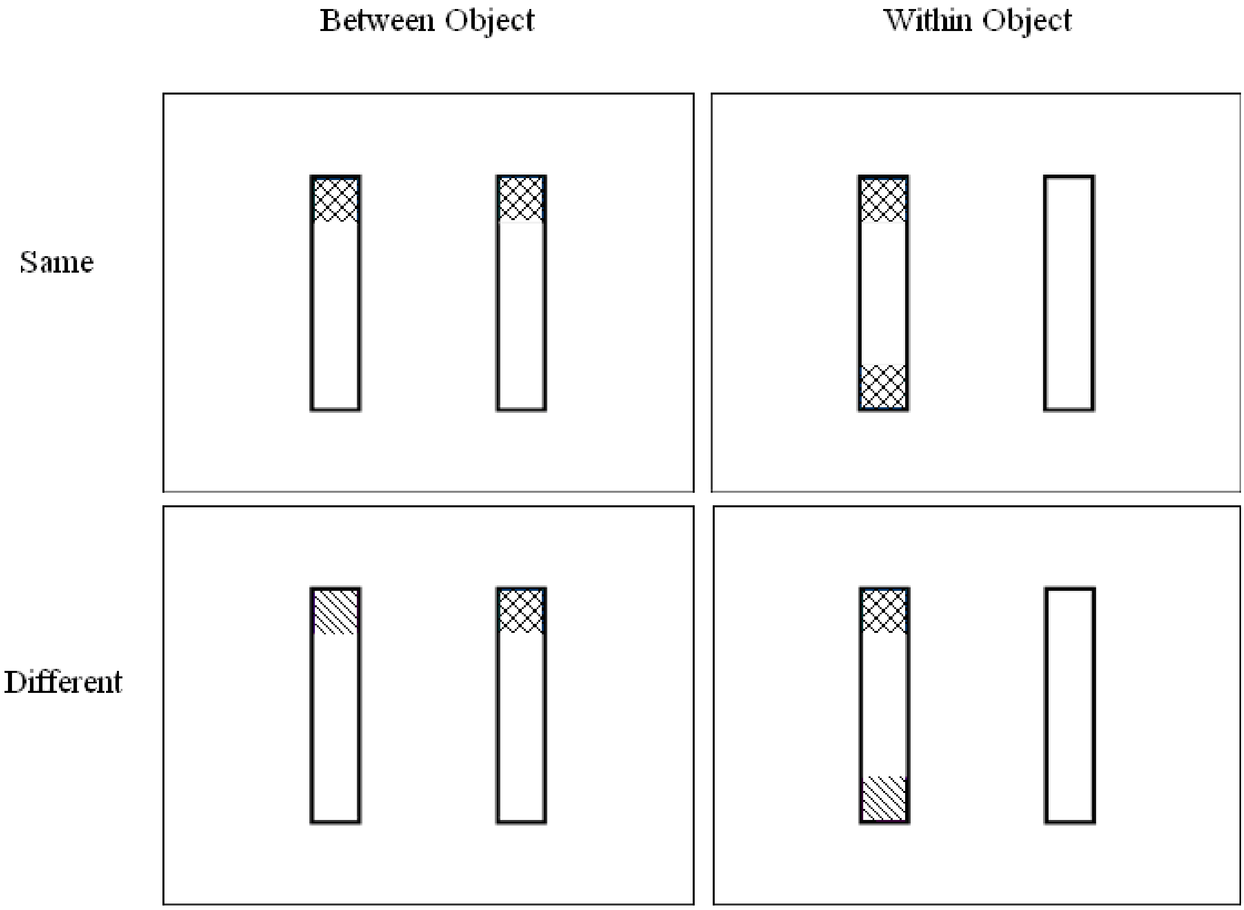


***Fig. 3** – Experiment 2*

- Illustrative stimuli used in the object-based two-item comparison paradigm (Experiment 2). The targets were red (shown as cross-hatched) or green (shown as dashed). The outline rectangles were white on a black background.

The subject’s task was to report whether the colours were the same or different using one of two buttons, again using the left index finger for one button and the right index finger for the other. The targets remained present on the screen until a response was made. There was a 1700ms inter-trial interval during which the rectangles remained visible. Each participant completed 6 blocks of 96 trials. Each block contained equal numbers of same and different pairings and of within- and between-object trials. The presentation order was randomly counterbalanced within each block. Reaction times were trimmed for outliers and errors in the manner described in Experiment 1. This procedure removed 10% of trials as outliers.

4.2. Results

The control subjects took reliably longer to make discriminative responses when the targets were both located on different objects than when they were on the same object [t(7) = 3.7, p <0.01; within-object mean RT 499 ms, between-object mean RT 511 ms][18]. Thus their results were clearly consistent with the operation of object-based attention in this paradigm. Again D.F.'s results showed a quite different pattern. She not only showed no cost for between-, as opposed to within-, object comparisons, but in fact took longer to make within object comparisons [D.F. within object RT 1287, between object RT 1224]. Crucially, Crawford and Garthwaite's (2007) Bayesian inferential test showed that D.F.'s mean difference score (within-object minus between-object) differed significantly from that of the controls [p<0.001].

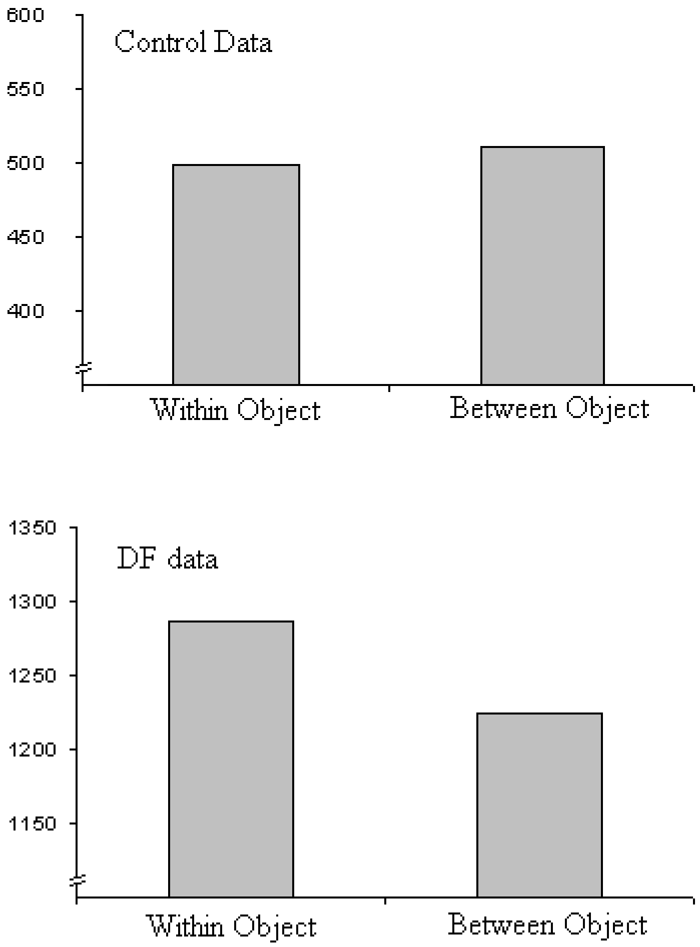


***Fig. 4** – Experiment 2*

- Data for a healthy control group and for patient D.F.. Reaction times (ms) for same/different comparisons for pairs of stimuli appearing within the same object or between the two objects. Note the y axis uses a different intercept for D.F. and the control sample.

---

[18] The first block of 96 trials for one of the control participants had to be removed because performance fell below 50% (i.e. below chance) in one of the conditions, suggesting that they had not learnt the correct key responses.

Control subjects did not differ in accuracy on the between- and within-object trials (97.9% correct on both within and between trials), and Crawford and Garthwaite's inferential test on the difference scores shows that this pattern is not different from that shown by D.F. (p=0.219, within 97.2%, between 95.5%).

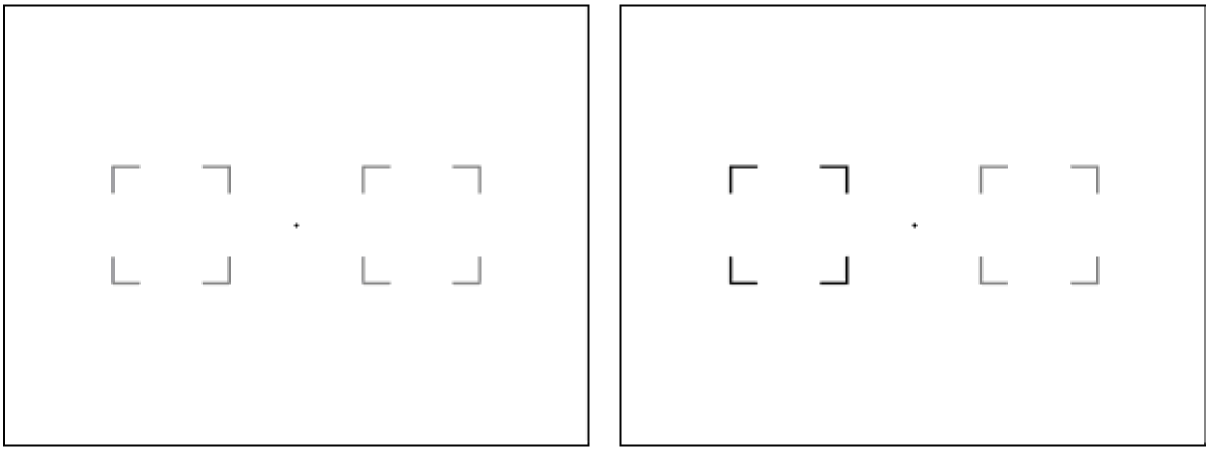
4.3. Interim discussion

The results from this experiment again demonstrate that D.F.'s allocation of attention is not influenced by objects in the usual manner. As was seen in Experiment 1, D.F. responded if anything faster in the between-object condition, a pattern opposite to that shown by the controls. Taken together, the results from Experiments 1 and 2 demonstrate that D.F. does not exhibit the normal effects of object-based attention. Before we can draw any meaningful conclusion from these findings, however, it is vital to establish that D.F.'s results do not simply reflect a failure of attention *per se*, but are specific to object-based attention. Experiment 3 therefore examines simple covert spatial orienting in D.F., using the classic Posner spatial attention paradigm.

5. Experiment 3

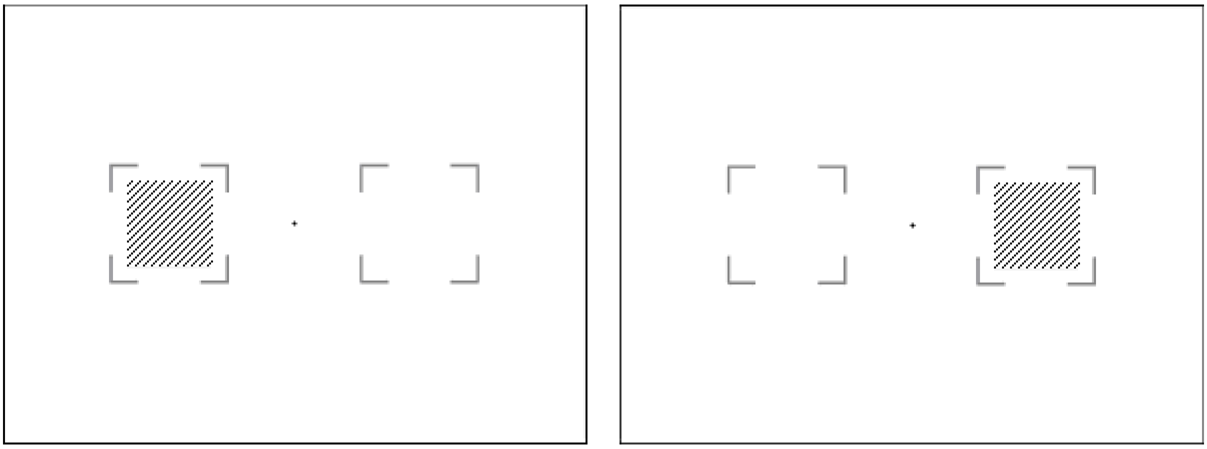
5.1. Methods

Experiment 3 employed a variant of the Posner's (1980) peripheral spatial cueing paradigm. The participant was presented with a background against which two low contrast placeholders (each represented by the corners of a square) were displayed. These placeholders subtended a visual angle of 10.68 deg, and were located from centre to centre 15.4 deg either side of fixation. The lines making up the placeholders were 0.15 deg thick. They remained present on the screen throughout each experimental session. At the onset of each trial, one of the placeholders would brighten for 30 ms as a spatial cue (see Fig. 5, top right). On cue-offset there was an interval of either 90, 120 or 150ms before target presentation. The target was a red or green square subtending 7.84 deg, which appeared either on the same side as the cue (valid target) or on the opposite side to the cue (invalid target). There were equal numbers of valid and invalid trials.



Backgound

Cue



Valid Target

Invalid Target

Fig 5 *Experiment 3*

- Illustrative stimuli for the Posner spatial cueing paradigm. The placeholder frames and fixation cross were white on a black background. The dashed pattern is used to illustrate the target, which could be either green or red.

The participant identified the target's colour as quickly and accurately as possible using one of two buttons, again using the left index finger for one button and the right index finger for the other. Following the response there was a 1500ms inter-trial interval during which the placeholders remained visible. Participants were instructed to maintain fixation on the cross in the centre of the screen throughout each experimental trial. The stability of fixation was monitored using an infra-red camera. Testing was conducted in 6 blocks of 96 trials, the order of stimulus type (red, green) cue type (valid, invalid) and cue-target interval (90, 120 or 150ms) were independently randomized within each block. The reaction times were trimmed for outliers and errors in the manner described for Experiment 1. This procedure removed 6% of trials as outliers.

5.2. Results

D.F. showed a normal pattern of spatial orienting, with validly-cued targets (i.e. appearing on the same side as the cue) being identified significantly faster than targets appearing on the opposite side to the cue [t(470)=2.38, p=0.018, valid mean RT 752ms, invalid mean RT 793ms]. A group analysis revealed the same effect for the control subjects [t(7) = 5.51, p<0.001, mean RT for validly cued targets 421 ms, mean RT for invalidly cued targets 448 ms]. Crawford and Garthwaite’s (2007) Bayesian inferential test revealed that D.F.’s shift costs (invalid minus valid mean RTs), although quantitatively greater, did not differ significantly from those of the controls [p=0.335]. Indeed, allowing for her longer reaction times, the proportionate shift costs were very similar to those of controls.

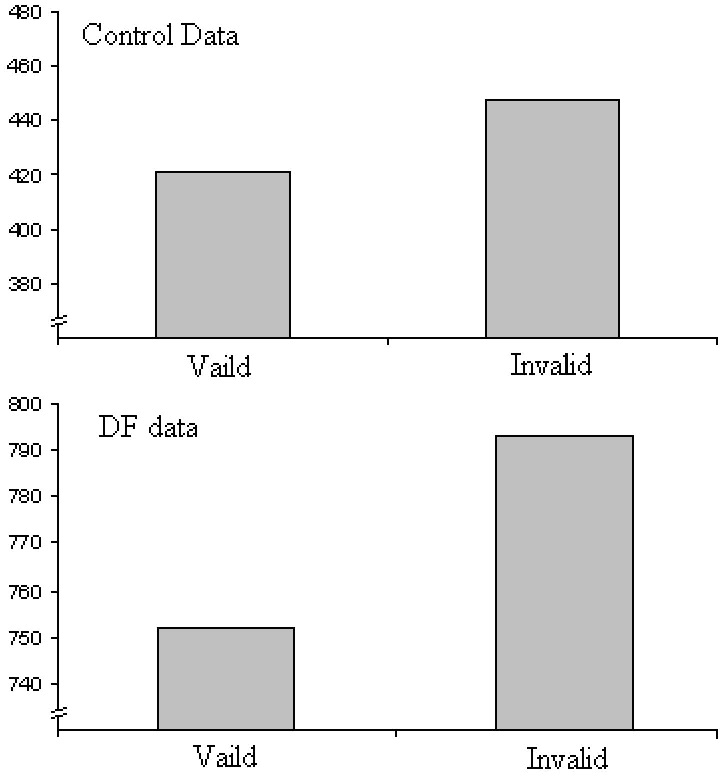


Fig 6 *Experiment 3*

-Reaction times (ms) for target discriminations following valid and invalid spatial cues. Note the y axis uses a different scale for D.F. and the control sample.

Cue validity did not affect accuracy for control subjects [t(7)=0.113, valid mean 98.1%; invalid mean 98.2%] and Crawford and Garthwaite's inferential test showed no difference between D.F. and the controls in terms of accuracy [p=0.669, valid 99.6%, invalid 99.2%].

6. Experiment 4

In Experiments 1 and 2, D.F. not only showed no cost for shifting attention between objects but, if anything, appeared to shift her attention more quickly between objects than within them. A possible explanation for those findings, however, is that in both of those experiments, the rectangles were always presented in a vertical orientation, so that between-object shifts were made horizontally whereas within-object shifts were made vertically. In another experiment currently in progress in our laboratory, we recently observed that healthy participants respond faster to targets following a horizontal shift in a variation of the Egly et al (1994) paradigm in which the cue and targets were visible but the objects were not. It might be, therefore, that D.F.'s contrary pattern of performance in our object-based attention paradigms emerged totally independently of the presence of the rectangles on the screen. That is, D.F.'s performance is completely explicable if horizontal shifts can be achieved more rapidly in this task when the outline rectangles do not exert any influence. In Experiments 1 and 2 any such effect will have been masked for normal observers because of the influence of the objects. In Experiment 4 we sought to test this idea directly by retesting healthy controls using the procedures of Experiment 1, but in the absence of the rectangular shapes. With the rectangles absent we predicted that normal observers would show the same pattern of performance as D.F. showed, when the rectangles actually were present.

6.1. Methods

The procedure for Experiment 4 was identical to Experiment 1, except that the participants (healthy controls only) completed only one block of 96 trials. The cue and target stimuli were identical to Experiment 1, but they were presented on a blank screen, without the rectangular shapes. Responses were made and recorded in the same way as before. The results were trimmed for outliers and errors in the same manner as described for Experiment 1. This procedure removed 5% of trials as outliers.

6.2 Results

When the rectangles were absent from the screen, our healthy control participants showed the same pattern of performance as D.F. had done in Experiment 1. Participants showed a general cueing effect, with responses to validly cued targets being significantly faster than those to invalidly cued targets (t(5)=-7.59, p=0.001, valid mean RT 382.5, invalid mean RT 426. More critically their RTs to invalidly-cued targets requiring a horizontal shift of attention (corresponding to 'between-object' trials in Experiment 1: mean RT 416 ms) were significantly faster than those to targets requiring a vertical shift of attention (corresponding to 'within-object' trials: mean RT 436 ms). Notably, while this difference across 'between' and 'within' object trials is significant [t(5) = 3.4, $p<0.05$], it does not differ, in terms of Crawford and Garthwaite's (2007) Bayesian inferential test, from the performance of D.F. in Experiment 1 (p=0.238).

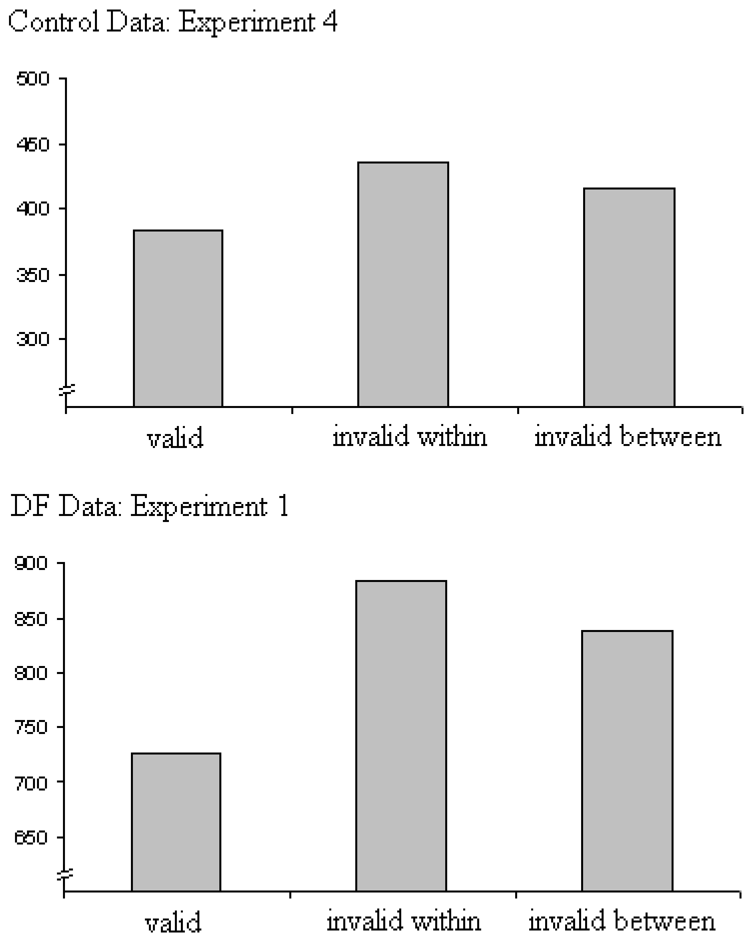


***Fig.** 7 - Experiment 4*

Reaction time data (ms) for control subjects on a target discrimination task in the equivalent of the three cueing conditions used in Experiment 1, but without the outline shapes. Thus the "invalid-within" data refers to vertical attention shifts, while the "invalid-between" data refers to horizontal shifts. D.F.'s data is reproduced from Experiment 1. Note the y axis uses a different intercept for D.F. and the control.

There were no differences in accuracy across ‘between’ and ‘within’ object trials in the subjects tested [t(5)=0.963, invalid ‘within’ mean 98.5%, invalid ‘between’ mean 97%], and again Crawford and Garthwaite’s inferential test shows that this pattern is not different from that of D.F. in Experiment 1 [p=0.989].

7. Discussion

Despite D.F. clearly showing normal covert spatial attention effects (as reported in Experiment 3) we have found, using two different object-based attention paradigms (Experiments 1 and 2), that there was no evidence that the deployment of her attention was sensitive to the presentation of objects. In fact her performance was so insensitive to the presence of objects in the display that it matched that produced by healthy participants performing the task when no objects were present at all (Experiment 4). In their recent paper Martinez et al. (2006) argue that object-based attention effects emerge because object representations are highlighted by spatial attention. In these terms, our data indicate that in D.F., despite normal spatial attention, there are no ‘objects’ to highlight.

Given D.F.’s severe form perception deficit, the inability of visual form representations in the ventral stream to influence attention could be argued to give little cause for surprise. Recent theorizing about object-based attention (Driver et al., 2001) has, however, emphasized that object representations *per se* might not be influencing the allocation of attention. Rather, the spread of attention might be modulated by the segmentation processes that parse the visual scene prior to the construction of object representations. Basic segmentation processes operating within in the primary visual cortex are likely to be still intact in D.F., given her ability to carry out simple visually guided acts of object grasping quite normally, and given that these acts are accompanied by normal patterns of dorsal-stream activation in the object-grasping area AIP (James et al., 2003). Indeed D.F. is able to differentiate figures from their background and give accurate perceptual reports of the number of items presented on a display (Carey, et al., 2006). Our results show, however, that these early stages of figure-ground segmentation are not able to influence the spread of attention. At the same time, our results do not demonstrate that attention is influenced solely by ‘high level’ object representations. Whilst D.F. is clearly impaired in recognizing objects on the basis of shape or form, this does not imply that the representations underpinning perceptual recognition are the locus at which shape influences attention. Although D.F. can perform some perceptual tasks not requiring object recognition (figure-ground segmentation) there are other aspects of visual processing, also not requiring object recognition, that are disrupted in D.F. She is, for example, insensitive to a number of gestalt cues to perceptual grouping that normally influence perception (Goodale, Jakobson, Milner, Perrett, Benson & Hietanen, 1994). It may be the case that processes such as these are intrinsically linked to the genesis of object-based attention.

The findings of James and colleagues (2003) allow us to relate the clear bilateral disruption to the object-selective area LO to D.F.'s visual form agnosia. First, her lesions overlap almost completely with area LO as determined in healthy controls through subtracting fMRI activations caused by viewing line drawings from those caused by viewing scrambled versions of those drawings. More tellingly still, a similar subtraction in D.F. yielded no net ventral stream activation at all (James et al., 2003). In the context of this disruption of form processing in area LO, the lack of any influence of visual structure upon D.F.'s attention is clearly consistent with the suggestion of Martinez and colleagues (2006) that this area plays a critical role in object-based attention. Again, however, it would be premature to conclude that our data prove that LO is essential for object-based attention, as the damage to DF's ventral stream is not solely restricted to LO. We are currently attempting to test the specificity of LO involvement in object-based attention by studying the effects of transcranial magnetic stimulation over LO in healthy participants.

Viewed in the context of D.F.'s established abilities in calibrating manual actions to selected visual objects, the present data demonstrate that classic perceptual object-based attention effects are not required for guiding manual actions. This finding is consistent with the theoretical framework of Milner and Goodale (1995) and reinforces the idea that not only is form processing in each stream computed separately but that it serves different functions. Our results, however, can only address the influence of one stream upon the other in one direction: specifically, they argue against ventral-stream object-based attention effects being required for the visual guidance of manual actions. Humphreys and Riddoch (2007) have suggested that computations in one stream might influence the other in the opposite direction. Specifically, they argue that the "objects" that influence attention are not solely perceptual entities: their nature and attentional status is also influenced by the manner in which one acts upon those objects. They argue that this influence manifests itself even when specific actions are only planned, and have not yet been executed. The results of the current study demonstrate that ventral-stream form representations are not able to influence the spread of attention in D.F., thus providing an ideal opportunity to test whether object representations localized solely within the dorsal stream might be able to influence attention. We therefore plan to explore the possibility that when D.F. is acting upon an object the representations subserving that action might be able to influence the perceptual allocation of attention within and between object structure (cf. Schenk & Milner, 2006).

In conclusion, although D.F. is able to show perfectly intact covert orienting of visuospatial attention, her allocation of attention seems to operate in a manner that is independent of the typical influence of visual structure.. That is, her attentional shifts appear to depend solely on the spatial constraints of the task, and not at all on any configural constraints. This result shows that the basic figure-ground segmentation that is intact in D.F. is not able to influence the spread of her visual attention. The absence of object-based attention in D.F., together with her clear bilateral lesion to area LO, is consistent with the idea that LO plays a key role in object-based attention (Martinez et al., 2006). Viewed in the context of D.F.'s ability to guide manual actions with respect to selected objects, this finding suggests that classic perceptual object-based attention effects are not required when acting upon visual objects, and strengthens the view that form processing in the two streams is not only distinct but serves different functions.

Chapter 9 - Shape processing area LO and illusory contours[19]

Abstract:

Recent functional MRI has highlighted that illusory contours can activate the primary visual cortex. Our investigation sought to demonstrate whether this correlation reflects computations performed in the primary visual cortex or feedback effects from shape processing area LO. We explored this in a patient who has a bilateral lesion to LO, but a functionally spared V1. Our data indicate that illusory contours are unable to influence behaviour without visual area LO. Whilst we would not claim that our data provide evidence for the 'cognitive' nature of illusory contours, they certainly suggest that illusory contours are dependent upon the computations involved in extracting shape representations in LO. Our data highlight the importance of neuropsychological data in interpreting the role of feedforward and feedback effects in the generation of visual illusions.

Experiment:

Typical illusory contours result when the visual system modally completes a surface whose presence is implied by the removal of corners from real objects (see Figure 1 for the example of the stimuli used in the current study). It is unclear whether these illusions result from local computations involved in the visual system's attempt to extract edges from the visual scene or from a higher-level interpretation of the possible shapes that could cause that pattern of visual input ("cognitive contours": Gregory 1972). Recent neuroimaging studies, using stimuli identical to those employed in this paper, have shifted the debate on illusory contours by demonstrating that they activate early visual areas (Seghier et. al., 2000). Of course this result does not preclude the possibility that these activations develop from feedback arising at higher levels of the visual system.

---

[19] This chapter has been published in the journal Perception, de-Wit, Kentridge, Milner (2009b)

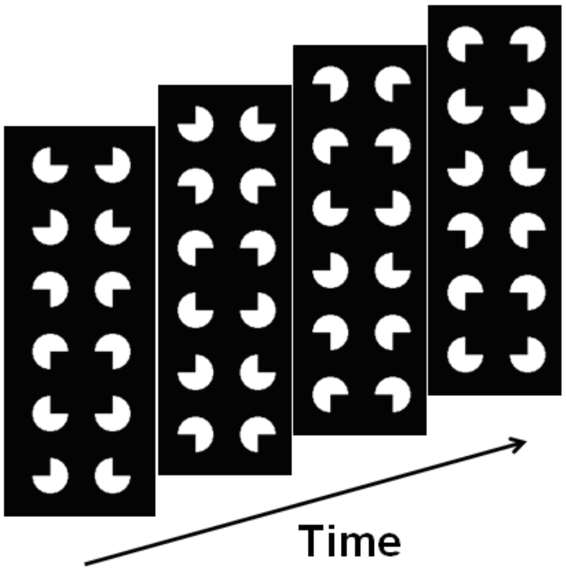


Figure 1: Moving Kanizsa Stimulus.

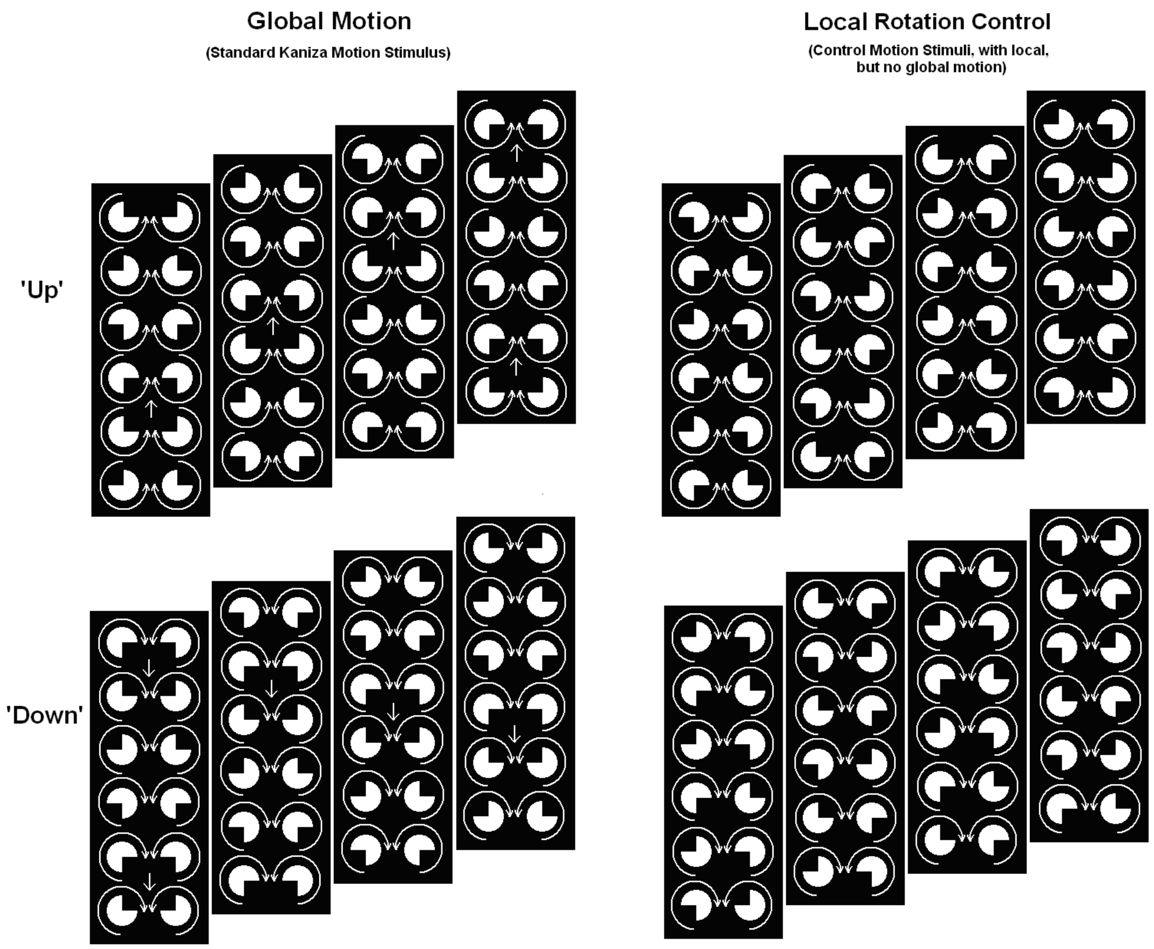
The stimuli were displayed on a Taxan Ergovision 8855LR (600 x 800 pixels) monitor, driven by a Cambridge Research Systems VSG 2/5 graphics system at 100 Hz. Each frame was presented in the centre of a black screen for 300ms. Each "pac-man" subtended 1.5 degrees of visual angle.

The neurological patient D.F. provides an opportunity to test the role of different areas in computing illusory contours. D.F. has a visual form agnosia resulting from bilateral damage to the putative shape processing area LO (James et al. 2003). D.F.'s ability to perceive visual form is extremely limited, and her perceptual system is unable to make even very basic form discriminations. On the other hand it is also clear that D.F.'s primary visual cortex is functionally spared, and is able to perform a number of important computations: it can supply information to guide manual actions (Milner and Goodale 1995), enables the perception of motion direction, the segmentation of figure-ground relationships (Carey et al., 2006) and the integration of chromatic and orientation information (Humphrey et al., 1991).

This pattern of impaired and preserved function allows us to ask whether computations performed in the primary visual cortex are sufficient for creating illusory contours. Indeed D.F.'s preserved motion perception allows us to use a variation of the illusory contour phenomenon in which the contours give rise to a percept of motion (see Figure 1). If orientation and segmentation processing in D.F.'s primary visual cortex are sufficient to generate this illusory stimulus then they should evoke, or at least influence, her motion perception.

Before exploring D.F.'s ability to perceive the motion of an illusory square, it was necessary to establish that she can perceive the motion of a real black square (presented without the 'pac-men' inducers, at the same speed and size as the illusory square). D.F. was not perfect at this, but did achieve an accuracy of 90% on 2 blocks of 24 trials. She was then presented with 4 blocks of 24 trials of illusory motion. D.F. certainly did not experience the compelling perception of motion experienced by normal observers – she felt she was guessing. Her guesses were, however, above chance, averaging 60.4%. On a two-tailed binomial test this performance is almost significant (N=96, p=0.082). Thus D.F.'s pattern of performance could be taken to suggest that despite the complete bilateral disruption to LO, she is still able to extract sufficient information to represent illusory contours and for those contours to bias her motion discrimination.

In the next series of observations we had two aims: the first was to establish the reliability of the trend reported above. The second aim was to establish whether or not this near-significant performance could be explained by any means other than the extraction of illusory contours.



<u>Figure 2: The local rotation confound in the motion stimulus of Figure 1.</u>

The control stimuli on the right contain a local rotation cue but no subjective square.

On viewing Figure 2 it becomes apparent that when the illusory square moves up the screen all the "pac-men" on the left rotate anti-clockwise while those on the right rotate clockwise, and vice-versa when the illusory square is seen to move down the screen. We therefore constructed a control stimulus in which this local rotation confound remains present, but there is no global illusory motion. If D.F.'s performance on the global motion stimulus is in fact based on this local rotational cue than she should also be above-chance in her judgement of the local rotation control stimulus.

D.F. was presented with 11 blocks of 24 trials using the original illusory motion stimulus. Her performance was slightly lower than previously (58%), but with the larger sample a two tailed binomial test revealed that this reached significance (N=246, p=0.011).

However when presented with 19 blocks[20] of 24 trials employing the local rotation cue alone D.F. was also able to perform significantly above chance (57%, N=432, p=0.004). Critically a Fisher's exact test between the performance associated with each stimulus type revealed that her performance on the two tasks did not differ (N=696, p=0.813).

Thus whilst D.F. has a functioning primary visual cortex that enables her to perceive figure-ground relationships and enables an implicit processing of orientation, she is unable to perceive this normally compelling motion illusion generated by illusory contours. Indeed D.F. is not only unable to perceive this illusion she is also unable to guess at the direction of motion, when an appropriate stimulus is used as a control baseline. In reviewing fMRI studies of the neural substrates activated by illusory contours, Seghier and Vuilleumier (2006) highlight a range of different areas, including V1, area LO, and the right fusiform gyrus. Although informative, these neuroimaging data cannot tell us where the critical computations responsible for the illusion take place. D.F. shows functional activations in both V1 (James et al. 2003) and right fusiform gyrus (Steeves et al., 2006) but has a clear anatomical disruption to, and indeed shows no functional activation in, area LO (James et al. 2003). The current findings therefore suggest that area LO is critical to the generation of illusory contours. The orientation processing and segmentation mechanisms evidently supported by D.F.'s primary visual cortex are insufficient in themselves to support the formation of illusory contours. Thus although Seghier and colleagues (2000) have demonstrated that moving illusory contours activate the primary visual cortex, our data suggest that this activation may reflect feedback from the results of computations performed at higher stages of information processing. Indeed this conclusion, based on human neuropsychology, is complementary to lesion work with monkeys (Huxlin, et al. 2000) suggesting that the inferior temporal cortex (considered the monkey homologue to human LO) is required for the monkey's visual system to generate illusory contours. We contend that this work highlights the value of neuropsychological research for teasing out the role of feed-forward and feedback effects in the human visual system.

---

[20] The uneven number of blocks (11 for the original and 19 for this control) occurred because of time constraints in testing the patient.

Chapter 10 - Perceptual organisation in visual form agnosia

Abstract

Previous research with the visual form agnosia patient DF has suggested that following damage to the visual area LO this patient is unable to exploit cues to perceptual grouping, either for the basis of guiding action (Goodale et al. 1994), attention (de-Wit, Kentridge, Milner 2009a), or motion perception (de-Wit, Kentridge & Milner 2009b). All of these examples have however explored this patient’s ability to group on the basis of shape. The current experiment seeks to explore whether or not DF can group information via a common colour-motion conjunction. The results show that although this grouping does not lead to a compelling explicit perception, DF’s guesses are biased in a manner that suggests that her visual system is able to achieve this form of grouping. This above-chance performance is consistent with a form of ‘base-grouping’ (Roelsfema 2006) encoded in the primary visual cortex that could have developed as an internalisation of the nature of natural scene statistics.

Introduction

One of the earliest research programmes in vision developed from the recognition that the human visual system does not just detect physical stimulation, it automatically seeks to organise that stimulation. This research programme led to a catalogue of organisational principles known as the Gestalt “Laws” of perception. Over the decades the perceptual organisation generated according to these “Laws” has been shown to play an important role in the detection and recognition of visual objects (Pelli et al. 2009).

Although a matter of debate for some time, the automatic nature of gestalt grouping principles has been verified by demonstrations that the resulting computations are performed even when the stimuli being interpreted are not consciously seen or attended (Moore & Egeth 1997, Lo & Yeh 2008). The ‘automatic’ implementation of these principles however does not mean they are fundamentally inflexible: far from it. It is arguable that these principles emerge via a registration of the statistical regularities experienced in our lifelong interactions with the world. Clear evidence for the development of grouping principles via statistical co-occurrence has recently been provided by Vickery and Jiang (2009) who found that the detection of two commonly coloured shapes was facilitated when those shapes had previously been experienced as an associated pair. The previous co-occurrence of a pair of shapes in their study was found to be able to guide the rapid integration of those shapes, such that they could more readily be organised and compared as a single group.

The flexible, experience dependent, nature of perceptual organisation brings into focus another critical debate that has unfolded since these principles were first described; at what level of the visual system are these organisational principles implemented? Moore and Egeth's demonstration that perceptual organisation does not require attention was thought to indicate the 'early' (pre-attentive) nature of these principles. More recent research has however demonstrated that a plethora of 'higher' level perceptual distinctions are registered in the absence of both attention and awareness (for a recent example see Jiang, Costello & He, 2007). Thus it is not possible to draw any clear conclusions about the level at which a process occurs in the visual system on the basis of its independence from attention. Somewhat more direct evidence for the early manifestation of grouping principles could be argued to be provided by von der Heydt, Peterhans, Baumgartner's (1984) finding that illusory contours (arguably derived from the gestalt principle of good continuation, although see Conci et al. 2009) appear to be evident at very early stages of cortical visual information processing. More recent research however, using fMRI and EEG in humans (Murray et al. 2002), and neuropsychology in monkeys (Huxlin, Saunders, Marchionini, Pham & Merigan 2000) and humans (de-Wit, Kentridge, Milner 2009b), has suggested that the activations seen in early visual areas reflect the products of perceptual organisational processing at higher levels of the visual system. Indeed more generally the dividing lines between what counts as high or low level within the human visual system have become blurred, for example recent fMRI results have demonstrated that illusions that manifest on the basis of 'high level' interpretations of the depth relationships in a scene are able to influence activation at the first stages of cortical information processing in the primary visual cortex (Murray, Boyaci & Kersten 2006).

Within this interactive framework it is possible to argue that grouping may be a ubiquitous aspect of perception, which operates "for each level of representation, rather than as a single stage that can be definitively localized" (Palmer, Brooks & Nelson 2003). That said however there do appear to be certain visual areas that play a more specialised role in aspects of perceptual organisation. Neuropsychological research with the visual form agnosia patient DF, for instance, highlights that whilst the processing of visual form information in early visual areas (Humphrey, Goodale & Gurnsey 1991) and in higher areas of the dorsal stream (James et al. 2003) is preserved, damage to LO in the ventral stream renders this patient unable to exploit cues to perceptual grouping in a number of different contexts. For example there is evidence that patient DF is unable to exploit cues to perceptual grouping not only in terms of her allocation of attention (de-Wit, Kentridge, Milner 2009a) and her motion perception (de-Wit, Kentridge, Milner 2009a) but also in terms of her guidance of manual actions (Goodale, Jakobson, Milner, Perrett, Benson & Hietanen 1994). These results could be argued to suggest that higher areas of the ventral stream play a definitive role in perceptual grouping. These demonstrations with DF however are limited to perceptual grouping via the property of shape. DF is known to have a selective preservation in the processing of colour and motion information, and is able to consciously perceive and accurately report the presentation of these stimulus attributes (Milner et al. 1991). Is this residual processing of colour information sufficient to guide perceptual grouping, or are the same resources that are required for grouping by shape also required for grouping by other properties such as colour? Framed more generally: is grouping intrinsic to the processes that mediate the conscious representation of a particular stimulus attributes (such as colour or motion), or are additional resources required?

In order to investigate this question the current research project sets out to test whether grouping by colour could influence DF's motion perception. Like her colour perception, motion processing is to some extent spared. Critically also, despite the original claims of independence between colour and motion processing, it is clear that salient colour groupings can influence and aid the perception of motion (Martinovic, Meyer, Muller and Wueger 2009, Croner & Albright 1997, although see Li & Kingdom 2001) and that these properties are in fact not necessarily encoded separately in early stages of the visual system (Sincich & Horton, 2005). Croner and Albright (1997), for example, demonstrated that the coherent motion of 10% of dots in a random dot kinematogram was more easily detected when those dots shared a common colour.

Previous experiments with DF have explored whether grouping via form information could bias DF's preserved motion perception (de-Wit et al. 2009b). The current project seeks to explore whether grouping via a common colour-motion property could bias DF's motion perception. To meet this aim a novel motion display was created; in this novel stimulus one of three directions of motion becomes dominant when all the dots moving in that direction share a common colour.

One could speculate that this increased dominance, derived from a common colour/motion conjunction, probably reflects an internalisation of the probability that two identically coloured surfaces (even if separated by an occlusion) moving in the same direction will belong to the same object in the real world. Pilot testing suggested that this increased salience did not necessarily lead to the target direction ‘popping out’, but certainly the observer would ‘settle on’ a perception of this direction of motion after a few seconds of viewing. If this common grouping in the target direction can bias DF’s selection of the most salient direction of motion then we would have evidence that her visual system is able to organise the information it receives without any contribution from higher level form processing areas of the ventral stream.

In addition to DF’s relatively preserved colour perception there is evidence that the early registration of basic form information is preserved. Humphrey et al. (1991) have shown that DF is sensitive to the McCollough effect, an orientation-contingent colour after image, believed to depend on the joint encoding of wavelength and orientation in area V1 neurons. Indeed DF’s visuo-motor responses are clearly sensitive to oriented edges, although her manual responses appear to be insensitive to even the most basic of shape based grouping cues (Goodale, et al. 1994). This project seeks to extend the evidence derived from DF’s visuo-motor responses to assess whether her motion perception can be biased by the common conjunction of a common motion and simple shape (an array of edges of the same orientation). In addition to exploring the ability of DF’s visual system to group on the basis of this simple form property, this project also examined whether grouping could be achieved on the basis a more complex form.

To summarise, the means by which gestalt organisational principles are implemented remains under debate. It is unclear whether there are specific resources for the grouping of visual information, or whether the grouping of different stimulus properties is in some sense an intrinsic aspect of the extraction of those properties per se. If grouping is intrinsic to the process of representing a given stimulus, what level of extraction is required before grouping principles become engaged? This project seeks to shed light on these issues by exploring the ability of DF, who has a severe lesion to higher areas of the ventral stream, to group on the basis of three properties:

(i) Colour, which is extracted to the level of conscious recognition,

(ii) basic form (oriented edge) information, that is implicitly processed but unable to inform perception, and

(iii) complex form stimuli (2D shapes) that are unlikely to be extracted at any level within DF's ventral stream.

Methods

Participants:

DF suffers from a visual form agnosia resulting from carbon monoxide poisoning. Her profile has previously been described in detail (see Milner et al., 1991 for her behavioural and structural profile, and James et al. 2003 for more recent structural and functional MRI). At testing (April 2009) DF was 55 years old. Two aged matched control subjects were also tested, however unlike DF they performed essentially at ceiling (above 95% accuracy) in all conditions, and so their results will not be discussed further.

Stimuli and Procedure:

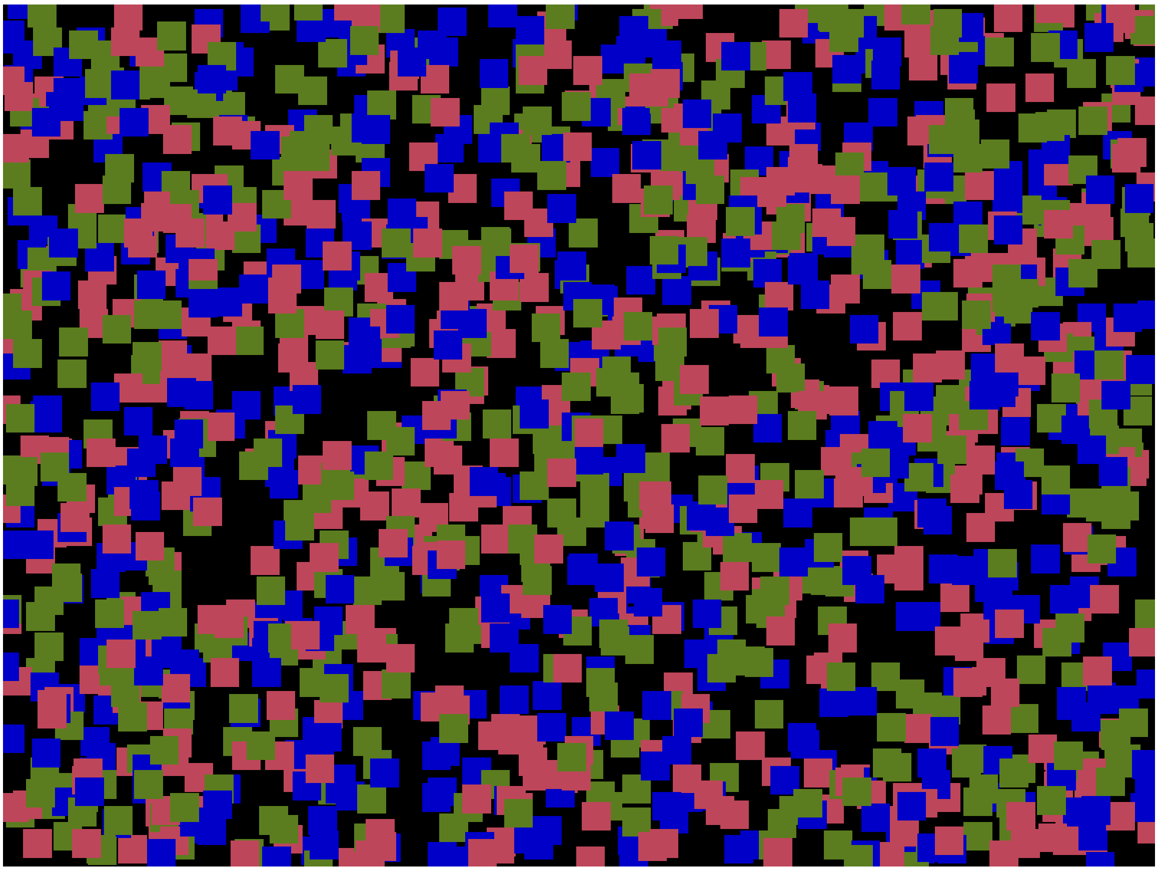
DF was presented with three stimulus types. Each stimulus was composed of hundreds of dots moving in one of three directions (downwards, bottom left to top right, and bottom right to top left). The nature of each dot differed for the three stimulus types. Coloured dots were either red, green or blue squares[21], the 'simple form' dots were edges (oriented vertically, or tilted from top left to bottom right or top right to bottom left) whilst the 'complex form' dots were one of three shapes (diamond, hexagon or triangle). In the colour and 'complex form' stimuli 750 dots were presented, while in the 'simple shape' stimulus 1500 oriented edges were presented. The increased number for this condition reflected the fact that the grouping in this condition was more salient when more edges were present. In every display type each

---

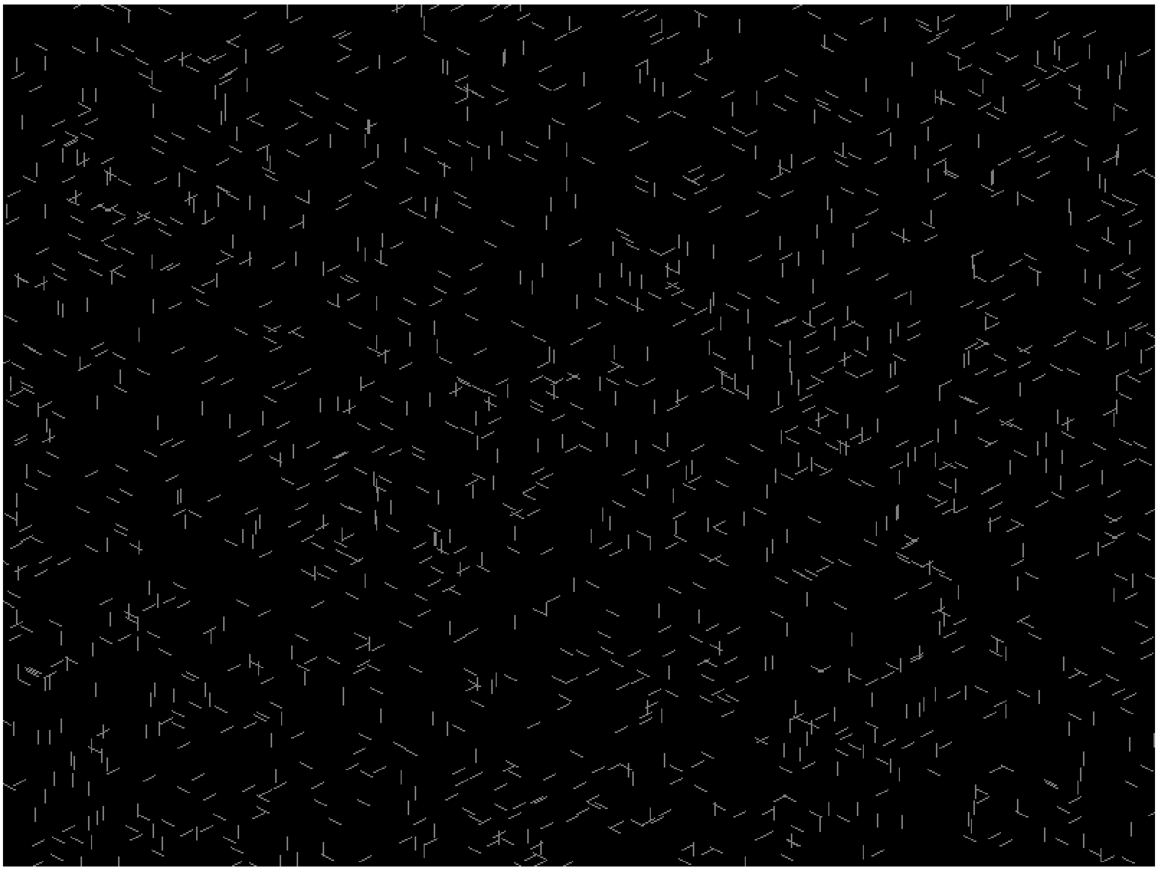
[21] Prior to testing we had equated the three colours for luminance and discriminability (in L,u',v' space), however when presented with these colours DF reported that she could not distinguish the blue colour. The value of this was therefore changed. Any effects in this condition might therefore result from luminance rather than purely colour differences. It should be noted that given this distortion in DF's colour space any attempts to match the luminance or discriminability of the three colours would require a lengthy calibration with DF that was not feasible during this testing session.

dot moved 4.5 pixels per frame, and a new frame was presented every 33ms. Stimuli were presented on a 800*600 PC Monitor. The key features of the stimuli employed in this experiment are only evident when the stimulus is set in motion; nevertheless a static frame from each stimulus type is presented in figure 1.

Colour (Figure 1a)



Simple Form (Edges – Figure 1b)



Complex From (simple shapes – Figure 1c)

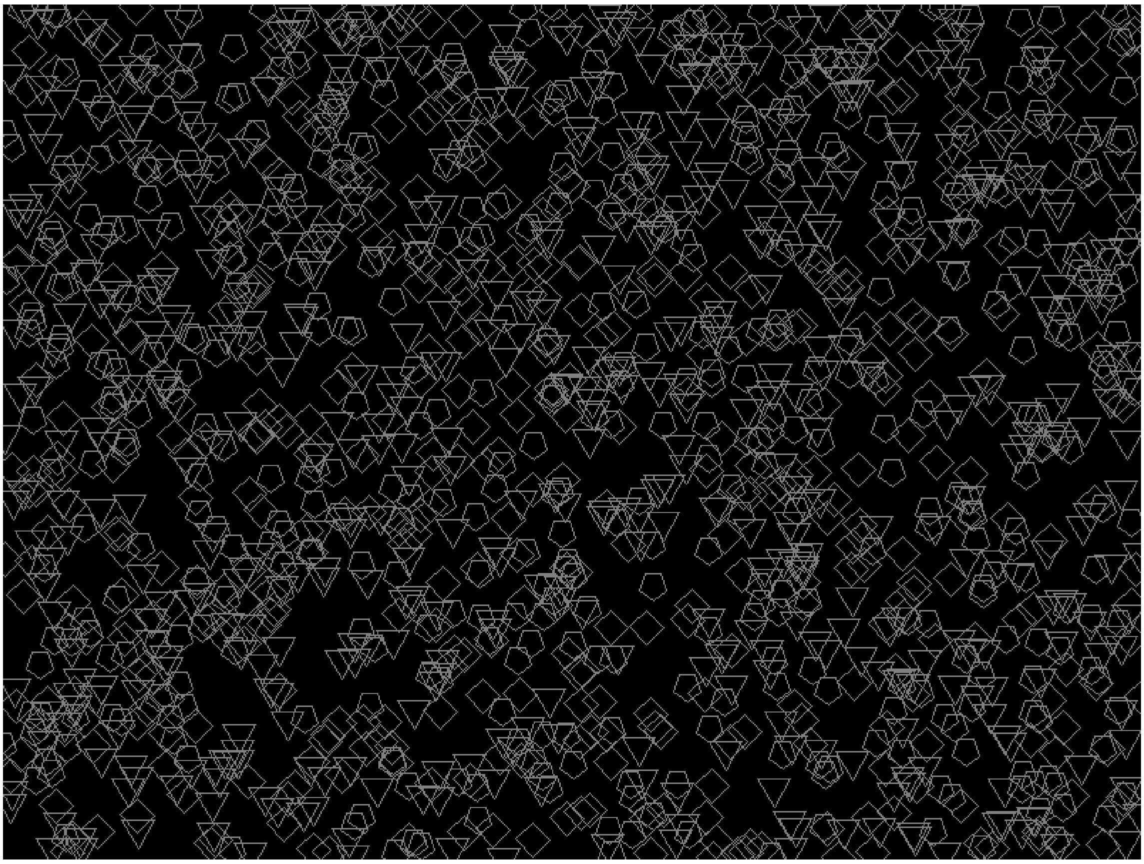


Figure 1, Colour, simple, and complex form stimuli.

In each display two of the properties are equally distributed across two of the directions of motion, whilst the ‘target’ direction of motion is composed of only one stimulus type. In the colour version for example, on one trial the ‘bottom-left to top-right’ and the ‘bottom-right to top-left’ moving dots might be either blue or green, whilst the dots moving downwards would all be red. The direction of the ‘target’ motion and the single feature (colour, simple or complex shape) associated with it was determined randomly for each trial. The three stimulus types were presented in separate blocks. For the colour stimuli DF could report both the ‘target’ direction of motion, and the colour associated with it, 5 blocks of 30 trials were therefore collected for each of these response modes. DF completed 5 blocks reporting the colour associated with the ‘target’ motion direction and then 5 blocks reporting the motion direction itself. For the simple and complex form stimuli DF could only report the direction of motion, for which 5 blocks of 30 trials were collected for each stimulus type in that order.

Results

DF did not report a clear perception of a dominant direction of motion for the colour stimuli, and her responses in the colour condition were therefore essentially guesses. For the coloured displays however DF correctly reported the ‘target’ direction of motion on 40% of trials (60/150: chance 33%) and the colour associated with the ‘target’ direction on 38% of trials (57/150). The overall performance with coloured stimuli is significantly better than chance (117/300 = 39%, binomial p=0.017) as is her performance when making direction responses (chance 33%, binomial p=0.036) although her performance when making colour responses does not reach significance (binomial p=0.10). There is, however, no statistical difference between her colour and direction naming performance (chi-squared(1)=0.056 with continuity correction, n.s.). DF reported some confidence in her ability to pick out a salient direction of motion for the simple and complex form stimuli, nevertheless her performance was clearly at chance in both the simple form (46/150=31%) and complex form (50/150=33%).

Her mean reaction times (of the order of 7 seconds, is only approximate because the experimenter made key presses corresponding to her verbal responses) did not differ significantly between response modes or between correct and incorrect responses, or in the interaction between those factors.

Discussion

If two identical surfaces separated by an occluding shape are moving in the same direction it is probably that those surfaces belong to the same object. The human visual system appears to be well adapted to internalise and exploit such statistic regularities to organise the information that it receives (Vickery & Jiang 2009). The present experiment has established that an organisational principle can still operate and exert a behavioural influence for a colour[22]/motion grouping in a visual form agnosia patient with extensive damage to areas of the visual ventral stream that are critical to grouping via form (James, et al. 2003).

When presented with displays containing hundreds of dots moving in three different directions the visual form agnosia patient DF was able to guess at the ‘target’ direction of motion defined by a common conjunction of colour and motion. For normal observers the even distribution of two colour types over two directions of motion, results in the motion

---

[22] We refer to this as a colour property, however it could be a combination of colour and luminance (see methods).

direction composed of only one colour becoming perceptually more salient. This common motion colour grouping did not lead to a clear salient perception for DF, nevertheless her ability to guess at this direction of motion shows that some form of grouping is still able to exert a weak influence on her selection criterion. How exactly could this grouping effect come about?

The common conjunction of colour and motion must have been encoded in order for that conjunction to lead to a slightly stronger motion signal. Roelfsema (2006) contrasts two distinct means by which perceptual groupings such as this could be achieved. Perceptual groupings can reflect either the common coding of two features, a ‘base grouping’, or can be bound using a more flexible process in which horizontal and feedback connections are used to simultaneously direct attention to grouped features, a process referred to as ‘incremental grouping’. ‘Base grouping’ is dependent on the common coding of different properties by the same neurones. Previous models of perceptual organisation (Treisman & Gelade 1980) and early neural physiology (Hubel & Livingstone 1987) assumed that basic stimulus features (such as colour and motion) were encoded separately. However more recent reviews (see Sincich & Horton 2005) have emphasized the fact that neurones in early visual areas can code not just for a single features (such as orientation or colour) but the combination of different features (not just a vertical bar, but a red vertical bar). Indeed it has long been known that neurones in V1 can code simultaneously for multiple stimulus dimensions (Michael 1978). More recently Gegenfurtner, Kiper and Fenstemaker (1996) have argued that the encoding of information along more than one stimulus dimension could enable V2 to play an important role in the integration of information about colour, motion and form. This process of base conjunctions could provide a very parsimonious explanation for the current findings. If there are neurones that code not only for a specific colour or motion direction but the combination of these, one only needs to assume that there are stronger excitatory lateral interactions between these neurones (which could have developed simply on the basis of the principle that “neurones that fire together wire together”) for a common colour and motion grouping to become more active.

Roelfsema (2006) argues that in addition to these ‘base’ conjunctions the visual system must also be able to flexibly recognise the common occurrence of novel feature conjunctions that might not have been experienced before, let alone share common neural encoding. In Roelfsema’s model, flexible or incremental grouping is enabled via an attentional enhancement directed only to the features that belong to the same object. This sort of flexible conjunction is probably supported by areas of the parietal lobe (Robertson 2003) that are likely to be spared in DF (James et al. 2003). This form of grouping however relies upon a top down selection of one of the three motions, and it is not clear why this process would selectively pick out the common motion-colour direction, unless that grouping had been suggested by some form of bottom up, or base conjunction.In this case however bottom up grouping could provide a more parsimonious explanation of the slight bias in behavioural performance without invoking more complicated flexible binding processes.

If some form of base grouping in the primary visual cortex does account for the present findings our results are in line with the hypothesis that perceptual organisation is not dependent upon specific or separately localisable neural resources, but rather is a process that can emerge at many levels of the visual system as an intrinsic feature of the processing of that information. In other words the early visual system does not simply seek to recognise different stimulus features; the sensitivity of cortical wiring at this level of the cortex allows it to play a role in organising that information. This ‘bottom-up’ explanation of the grouping observed in this experiment is arguably consistent with the fact that it led to a small bias in the participants’ guesses, and was not associated with a compelling or conscious percept of a dominant direction of motion. It is tempting to conclude that our results in fact indicate some form of non-conscious grouping in the visual system which is able to influence representations that are conscious, without itself becoming conscious. In the present context however this argument would rely on an interpretation of DF’s subjective description that may not provide an accurate reflection of her actually phenomenological experience. This possibility is however entirely plausible given previous demonstrations of perceptually bound, yet unseen stimuli (Melcher & Vidnyanszky 2006).

This bottom-up model however leaves open the question of why a similar bias did not appear for the common conjunction of orientation and motion. It is of little surprise that DF was not biased by the conjunction of complex form and motion, because DF should not have the perceptual resources to construct the relevant representation. Simple form information however is implicitly processed by DF’s visual system. Oriented edges are for example able to guide visuo-motor responses, and support the generation of an orientation specific after-image, the McCollough effect, which is thought to be dependent upon a common coding of orientation and colour in the primary visual cortex (Humphrey et al., 1991). If DF’s visual system is able to encode edges why were her motion responses not biased in the same manner seen for colour? One possibility is that the orientation task is simply fundamentally more difficult. This difference is not necessarily apparent in the control data, however as noted before the control data may well be at ceiling, and one could therefore argue that an adapted procedure would be required in order to truly equate the difficulty involved in recognising the motion direction of the two stimulus types with controls. Another possible answer to this lies in the fact that the level of processing of colour and form in DF is very different; whilst form processing in DF stops at the primary visual cortex, the processing of colour continues within the ventral stream to the level of V8 (James et al., Hadjikhani, Liu, Dale, Cavanagh & Tootell 1998). This cortical difference probably accounts for DF's general performance with colour in which she is able to consciously report an experience of colour, whilst she is unable to even make educated guesses at the orientation of different edges. The evidence for grouping without conscious experience (Melcher & Vidnyanszky 2006) however demonstrates that the differential access to consciousness is unlikely to be causally critical to the differential grouping of these features in DF. Nevertheless whilst the base grouping of colour in the primary visual cortex provides a compelling and parsimonious explanation for the current data, the inability of basic form information to elicit the same effect could be taken to imply that orientation processes do not overlap with motion encoding to the same extent in the primary visual cortex as do orientation and colour (base) conjunctions. There is however no a priori reason to believe that the association between colour and motion in primary visual areas is any more pronounced than the association between oriented edges and motion (Sincich and Horton 2005). This would suggest that some higher level factor, possibly reflecting some form of feedback loop enabled by the higher processing of colour was essential for the colour motion grouping to manifest.

Further research is probably required both to ensure the reliability of the effects found with colour here and to make certain that no grouping can be achieved on the basis of common oriented edges. Indeed it could be more productive to employ a different paradigm in which the participant has to identify the predominant direction of motion in a ‘random dot motion field’, in which all but a small proportion of dots are moving in a random direction. Previous research has shown that when a small number of dots moving in a common direction (among an array of dots moving randomly) share a common feature (even S-cone isolated colour signals) their direction of motion can be more easily detected (Martinovic et al. 2009). Indeed future work could also extend the current research to explore other properties that one might expect to be encoded in the primary visual cortex, for example one could employ the same logic to ask if DF would be able to group on the basis of a common stereo-defined depth plane.

To summarise, DF's visual system can exploit a grouping between colour and motion, such that she can report the dominant direction of motion evoked in normal observers with above chance accuracy. This above chance performance is consistent with a form of 'base-grouping' (Roelsfema 2006) encoded in the primary visual cortex that could have developed as an internalisation of the nature of natural scene statistics. This bottom-up perceptual organisation provided an above chance bias, rather than a compelling conscious perception, consistent with previous demonstrations of grouping in the absence of consciousness. This result suggests that perceptual organisation occurs as an intrinsic feature of information processing throughout the visual system. As expected DF was not able to exploit a grouping based on a complex shape, which her visual system should not be able to represent. DF however was also not biased when the grouping was provided by a simple form distinction (oriented edges) that according to previous data would receive some implicit processing in her visual system.

General Discussion

This thesis was fundamentally motivated by the Kantian perspective that, in order for us to understand and cognise about the sensory input we receive, that input has to be organised. More specifically Kant (1781) argued that sensory information has to be interpreted in terms of certain a-priori principles, which do not reflect inherent properties in the world, but provide a crucial framework for the mind to conceptualise the world. Early research in psychology revealed that two of these principles, namely causality and object-hood, have interesting parallels in the seemingly automatic manner in which visual information is organised (Koffka 1935, Michotte 1963). Understanding the extent to which these early processes of perceptual organisation in themselves shape and organise later cognition is undoubtedly an important question (Pylyshyn and Carey & Xu 2001). This thesis however did not seek to address this question head-on, but was motivated by it to explore our understanding of 'object-hood'. More specifically this thesis focused on one of the most basic level at which 'object-hood' influences subsequent cognition; that is in terms of its influence upon the allocation of processing resources, a process commonly referred to as 'object based attention'.

As reviewed in the Introduction however, scratching the surface of the neat, simple and commonly used expression, 'object based attention' reveals a bundle of different phenomena whose relationship both to each other and to our general understanding of 'object-hood' is far from clear. Rather than trying to address our understanding of 'object based attention' in general, this thesis focused largely (although not exclusively) on the Egly, Driver and Rafal (1994) within-vs-between-object advantage. In focusing on the Egly et al cuing paradigm this thesis has attempted to advance our understanding of how 'objects' influence the allocation of processing resources and how those 'objects' are extracted and represented.

The Egly et al. paradigm is one of the most cited and replicated examples of 'object based attention', it is therefore perhaps somewhat surprising that many authors (reviewed in Chapter 2) have argued that the within-vs-between-object advantage does not reflect a fundamentally object based mode of selection. In fact a number of authors have argued that the within-vs-between-object advantage reflects the way in which the spatial selection of an area is guided by object structure (Martinez et al. 2006, Davis et al. 2000, Richard, et al. 2008). This assumption was so prominent in the literature that Chapter 2 originally set out to explore how this form of 'objects biasing spatial attention' would manifest over space and time. The results therefore proved quite counter intuitive: there was no evidence consistent with the idea that the 'within-vs-between-object' advantage reflected a spatial spread of attention guided by object structure. Furthermore the data from Chapter 2 brought to light the fact that a significant part of the 'within-vs-between-object' advantage in fact does not reflect a 'within-object' advantage but a between-object switch cost. Indeed this switch cost may have been differentially sensitive to the figure/ground status of new target locations, suggesting that switching attention between two figures incurred a greater cost in comparison to switching from one figure to an area of background. This potential sensitivity to the figure ground status of the units of selection in Chapter 2 is consistent with recent behavioural work showing that whether or not a rectangular set of edges are used as a unit of selection depends upon whether those edges can be interpreted as a gap, or as a hole, that allows one to see the ground behind the edges rather than a figure standing on top of that ground (Albrecht et al. 2008). The figure ground sensitivity suggested in Chapter 2, and evidenced in Albrecht et al. suggests that the within-vs-between-object advantage in the Egly et al. paradigm reflects a form of selection that, like that seen in many other contexts (Duncan 1984, O'Craven et al. 2001), is 'object based'.

If the Egly et al. within-vs-between-object advantage reflects an object based mode of selection, what are these objects, and where are they extracted in the visual system? As reviewed in the Introduction, there are not only two potential streams of information processing within the visual system that might extract the 'objects' that influence attention, there is also a hierarchy of different levels of representation within both of those streams. In terms of the larger motivation of this thesis, one could argue that, if object based attention is to play a meaningful role in shaping later perception (let alone cognition), the representations that influence attention would need to be extracted early within these hierarchies. The notion of an 'early' stage of information processing has, in the past, been operationalized in terms of whether or not that process happens in a 'pre-attentive' fashion (Treisman & Gelade, 1980). In this context some authors have argued that if 'object based attention' is to play a meaningful role in guiding later perception those 'objects' have to themselves be extracted in a pre-attentive fashion (Egly et al., Driver et al. 2001). Indeed the fact that objects influence attention has been taken as evidence, in itself, that these representations are extracted 'pre-attentively' (Egly et al.). As reviewed in Chapter 3 however there are good reasons to question the extent to which representations can influence attention if they have not themselves received some degree of attentive processing (Goldsmith and Yeari 2001).

The fact that even simple outline shapes require some degree of attention before they achieve a sufficient representational status to influence the subsequent allocation of attention (Goldsmith and Yeari, 2001), raises a concern as to whether the within-vs-between-object advantage in the simplified Egly et al. paradigm would be manifest in the more complicated and dynamic scenes we face in our daily interaction with the world. This concern is brought into even clearer focus when one considers the long presentation times and predictable arrangement of the simple outline 'objects' typically used in this paradigm. Chapter 3 sought to address some of these concerns by systematically increasing the number of objects, randomising their location, and reducing their presentation time. Reassuringly the results of Chapter 3 showed that the within-vs-between-object advantage was completely robust to these complications to the basic paradigm. Thus whilst Goldsmith and Yeari have demonstrated that some degree of attention must be directed towards the space in which 'objects' are located before those objects are organised (to the extent such that they can be selected as a unit of attention), Chapter 3 provides evidence that with a diffuse focus of attention this process of perceptual organisation can be achieved in parallel across the entire visual scene. Indeed Chapter 3 provides a first step towards establishing whether the within-vs-between-object advantage can still manifest itself in situations that come closer to approximating our real world visual interactions.

The parallel extraction of these representations would be taken as evidence in traditional cognitive models (Broadbent 1958) to imply that this extraction was an early, non-computationally demanding process. Another means of framing the level at which a process occurs can be operationalized in terms of whether or not that process occurs in the absence of conscious awareness. Chapter 4 set out to test this possibility in the context of a paradigm known as Continuous Flash Suppression (CFS); originally chapter 4 sought to examine whether the Egly et al. within-vs-between-object advantage would still manifest when the 'objects' that are selected in this paradigm are not consciously perceived. Pilot work aiming to combine the Egly et al. and CFS paradigms suggested however that it might be wise to employ a simplified paradigm to explore, first if 'objects' are represented when they are rendered non-conscious via CFS, and second whether attentional selection acts differently upon coherent vs incoherent objects when those objects are not consciously perceived. In order to address this question Chapter 4 operationalized the notion of a coherent vs incoherent object in terms of the well studied perception of illusory Kanizsa figures. The results of Chapter 4 suggested that when suppressed via CFS the visual system was unable to extract these illusory figures, which normally emerge effortlessly and in parallel when consciously perceived (Driver & Davis 1998). Indeed the allocation of attention to Kanizsa figures was only influenced by their spatial arrangement, not their status as objects. Under conditions of suppression via CFS therefore it seems that the visual system is unable to construct this form of ‘object-hood’ and consequently this form object-hood is unable to influence attention.

Does this result prove that consciousness is required for the perception of objecthood? No, in all likelihood it says more about the limited nature of visual information processing in the context of the binocular rivalry induced by CFS. Fang and He (2005) for example provided fMRI data suggesting that although object information is able to ‘leak’ into the dorsal stream during CFS (either via V1 or sub-cortical routes) it is not able to influence object selective regions of the ventral stream. Thus although Chapter 4 does not allow us to make any claims about the relationship between awareness and object based attention, it does provide data to suggest that higher level resources within the ventral stream are required in order to extract the form of object-hood reflected in Kanizsa figures. This idea is further explored in the context of neurological damage to higher level areas of the ventral stream in Chapter 9.

In the context of the established dissociation between basic spatial attention mechanisms and awareness (Kentridge et al. 1999) it would clearly be valuable for further research to clarify whether the Egly et al. within-vs-between-object advantage still manifests itself when those objects are rendered unconscious. Chapter 4 however suggests that the extreme suppression induced in further ventral stream processing means that CFS is not the optimal paradigm with which to test this hypothesis. There are however a range of different paradigms within which one can manipulate the conscious status of a given object (Kim & Blake, 2005) and future research is planned using other paradigms, such as peripheral ‘crowding’ to explore whether the Egly et al. rectangles can influence the allocation of attention when one cannot consciously identify whether those rectangles are oriented horizontally or vertically[23].

Taken together the first set of experiments suggest that the within-vs-between-object advantage in the Egly et al. paradigm manifests not because ‘objects’ influence spatial

---

[23] I’d like to acknowledge Patrick Cavanagh for suggesting this option.

attention, but because some form of 'object-hood', extracted in parallel, is itself selected in order to bias information processing, in a manner that respects the distinction between figures and ground. Although further research is needed to pin down the relationship between attention and awareness, Chapter 4 opens a discussion about the role of higher ventral stream areas in constructing illusory or Kanizsa figures that will be further taken up in Experiment 8.

In principle the first section of this thesis aimed to advance our understanding of object based attention per se. The second set of chapters sought to use object based attention to explore the nature of perceptual organisation. As outlined by John Duncan (1984) however studies of perceptual organisation and attention need to progress hand-in-hand, and the next three chapters therefore attempted to shed light simultaneously on the nature of perceptual organisation and its influence upon attention.

Chapter 5 was motivated by the fact that accumulating evidence had suggested that areas in the ventral stream involved in the visual representation of shape also become active (Amedi, Malach, Hendler, Peled & Zohary 2001, James et al. 2002, Pietrini et al. 2004), and indeed may be critically required for (James, James, Humphrey & Goodale 2006), the tactile representation of shape. Indeed more recent research has suggested that when visual information is converted into a 'soundscape', auditory information can also activate what was traditionally considered a visual shape representation area (Amedi et al. 2007). Taken together this research suggests that this area of the brain, namely the Lateral Occipital area, should not be conceptualised in terms of the sensory modality from which it receives information, but rather should be defined in terms of the shape-extracting computations that it performs with this information. If this definition is correct however it should be the case that effects elicited via the visual stimulation of this area should also manifest on the basis of purely tactile stimulation. Motivated therefore by the potential role this area might play (highlighted in the Introduction, and further evidenced in Chapter 8) in the development of the representations that lead to the Egly et al. within-vs-between-object effect, Chapter 5 sought to establish whether or not this effect would manifest when participants only knew about the 'objects' via tactile stimulation. In more general terms Chapter 5 sought to use object based attention as a tool to explore whether the construction of objects, or shapes more specifically, was a genuinely cross-modal one.

To this end participants were tested using a 3-dimensional version of the Egly et al. paradigm, where the effect of purely tactile knowledge of an object could be assessed in terms of its influence on cues and targets that simultaneously appeared on those objects. Chapter 5 found consistent evidence that even when the pair of rectangles had never been seen, purely tactile knowledge of their shape was sufficient to influence the manner in which cues and targets were processed, leading to a classical within-vs-between-object advantage. Although the within-vs-between-object advantage was smaller following purely tactile stimulation, Chapter 5 also critically revealed that this small effect was not mediated simply by conceptual knowledge of the object configuration, because a follow up control condition proved that in the absence of direct sensory stimulation, previous experience (tactile and visual) of those rectangles was not sufficient to develop a representation that could influence the allocation of attention.

This result is quiet literally counter-intuitive in terms of the completely separate phenomenological experience of visual and tactile information. Even if shape information derived from touch can evoke a similar influence on attention, it must somehow be tagged in a manner that insures that it has a very different influence upon conscious perception. The juxtaposition between the equivalent status in terms of attention and yet strikingly different conscious experience, may relate to the activation of different shape processing areas within the larger shape processing Lateral Occipital Complex. For example it may be the case that multi-modal representations are only evoked in the more posterior aspect of LOC (LO), and indeed Haushofer, Livingstone and Kanwisher (2008) have argued that more subjective aspects of shape perception are mediated by the more anterior aspect of LOC, the posterior Fusiform sulcus area (pFs).

The convergent influence of visual and tactile shape information is in itself interesting to frame in the context of Carey and Xu's (2001) argument that the 'object based attention' effects seen in adulthood may reflect exactly the same underlying mechanisms that are responsible for the rudimentary understanding of objecthood evident in very early infant behavior. Reflecting upon such putative cross-modal convergence at this point in human cognitive development is particularly pertinent because, probably more so than at any other stage of human development, an infant's exploration of the world is a largely cross-modal process. Infants are not content to simply look at the world, they want to pull, hold and feel the objects in front of them. It is therefore probably very important at this stage of development that infants come to understand the relationship between the things they explore tactually and the things that they can see. If Carey and Xu are correct and the effects seen in adulthood, in terms of object based attention, reflect underlying mechanisms present in infancy, it is interesting to speculate on the possibility that the cross-modal object based effect seen here could play a role in biasing infants' perception such that they visually attend to the features of objects that they are simultaneously exploring tactually, thereby ensuring that those features are processed together. The integration of shape information, and the influence of that information upon attention could in principle therefore play a role in aiding the development of a coherent, bound model of the many different stimulus features coming from different modalities.

Chapters 6 and 7 set out to explore whether the differential ecological importance of different stimulus types would differentially alter their status as units for attentional selection. Chapter 6 sought to explore whether shadows, which seem to be discounted in some contexts (Rensink and Cavanagh 2004) are less accessible targets for attention selection using an object based two item comparison paradigm and the Egly et al. within-vs-between-object paradigm. Chapter 7 sought to explore whether biological targets would be more readily accessible as targets for attentional selection. Rather than using the Egly et al. cuing paradigm, Chapter 7 made use of the Multiple Object Tracking paradigm (henceforth MOT). It is important to acknowledge that, as highlighted in the Introduction, one cannot necessarily assume that different 'object based attention' paradigms all tap the same level of representation. Despite the general focus on the Egly et al. paradigm in this thesis however there were two principal advantages to using the MOT paradigm. The first of these was pragmatic. The perception of biologically relevant targets was operationalized in this study via the employment of point-light biological motion stimuli (Johansson, 1973), which were relatively easy to integrate into the more general MOT paradigm. The second was more theoretical. As reviewed in Chapter 7, biological targets represent a particularly important class of stimuli to track, both because of their potential significance as predators or prey, but also because their self-locomotion means that they can rapidly and unpredictably change their rate and direction of motion at any time.

Despite these predictions however Chapter 6 suggested that shadows could be selected as units of attention just as easily as 'objects' not interpretable as shadows. At first glance the results of Chapter 7 appeared to suggest that stimuli moving in a biological manner were less accessible as units to track. An analysis of the scrambled control conditions in this experiment however revealed that MOT was indifferent to the biological status of the motion stimuli per se but was generally less effective within stimuli that did not have a coherent (fixed or constant) internal motion profile. Thus whilst Chapter 7 set out to test whether the biological status of a stimulus affected the ease with which it could be tracked, the results in fact emphasized the importance of lower level grouping principles like common fate, in the formation of the representations that underpin tracking.

Although again one should acknowledge that it is not possible to argue that MOT and the Egly et al. within-vs-between-object advantage reflect selection at the same level of representation, the parallel between these two chapters is consistent with an argument developed by Pylyshyn (1999) that, in order for mid-level perceptual organization to provide a useful scaffold for further cognition, that organization has to be in-and-of-itself inaccessible to the products of later cognition. Thus whilst shadows or biological motion might have a differential importance in terms of later cognition, in order for that cognition to work effectively, Pylyshyn would argue, perception has to provide it with the basic building blocks upon which to apply its computations without affecting those building blocks per se. In future work we plan to further explore the potential for the higher level nature of a given representation to influence its selection in MOT by exploring whether the implied direction of motion of a biological “walker” influences the ability with which it can be tracked. Pylyshyn for example would argue that whether the direction of motion of a walker is congruent with the direction implied by their biological motion profile would not impact the manner in which this figure is parsed as a potential proto-object in early vision, and would not affect the utility with which it can be tracked. Our future work should therefore further illuminate the extent to which the implied meaning of a given stimulus can affect its status as a unit of attentional selection.

Taken together this section appears to reveal a level of perceptual organization that is at once cross modal, but at the same time seemingly insensitive to the higher-level status of visual objects. The cross modal construction of objects that influenced the visual allocation of attention was interpreted in terms of the known cross-modal sensitivities of the ‘visual’ ventral stream area LO. The role of LO in object based attention was more explicitly tested in the final section of this thesis in which the nature of object based attention and perceptual organization more generally were explored in a patient with visual form agnosia.

This case of visual form agnosia (patient DF) provided an opportune test of many of the issues developed in the introduction of this thesis regarding the nature of the ‘objects’ that are selected in object based attention effects. Most of the work focusing on this patient has explored the fact that she shows a dissociation between the object representations required for generating explicit perceptions and the object representations required for guiding manual actions (Milner & Goodale 1995). Neuro-imaging has provided data consistent with the interpretation that the guidance of manual actions in this patient is dependent upon the extraction of object representations within the dorsal stream, and that the inability to develop a perceptual representation of objects reflects a bilateral lesion to the ventral stream area LO (James et al. 2003). This pattern of preserved dorsal, but impaired ventral, object representations provides an opportune chance to test the role of these representations in developing the structural descriptions required to generate the within-vs-between-object advantage in the Egly et al paradigm. Indeed in addition to this, the patient offers some insight onto the debate raised in the Introduction regarding the stage in the hierarchy of visual form extraction at which attention can select. A number of studies reviewed in the introduction highlighted that although edges could sometimes in and of themselves affect the allocation of attention (Avhrami 1999), what seemed to be critical was the manner in which those edges were interpreted in terms of figure ground relationships (Ben-Shahar et al. 2007, Albrecht et al. 2008). Patient DF shows a preserved processing of edges within the primary visual cortex that enables both the development of representations in the dorsal stream required to guide manual actions, and also causes her to be sensitive to an orientation specific after image called the McCollough effect (Humphrey, Goodale & Gurnsey 1991). An exploration of the potential influence of object structure upon attention in this patient should therefore allow us to address whether edges extracted in the primary visual cortex are in and of themselves sufficient to drive the allocation of attention, or whether those edges have to be interpreted in terms of their figure ground relationships by areas like LO, before they can generate the units required for attentional selection.

Chapter 8 revealed that despite clear behavioral markers of an intact spatial selection, this patient’s movement of attention was (in both the Egly et al, and a two-item comparison paradigm) completely unaffected by object structure. Chapter 9 employed a very similar logic to address whether activations seen in fMRI in the primary visual cortex following the perception of a motion illusion induced by a moving version of the Kanizsa illusion (Seghier et al. 2000), reflected computations performed within the primary visual cortex, or feedback based on higher levels of activation. The results from Chapter 9 showed that the performance of this patient was in no way biased by this generally compelling visual illusion. Thus suggesting that computations at levels higher than the primary visual cortex are required in order to extract illusory Kanizsa figures at least in this version of the illusion. This finding is consistent with neuroimaging evidence suggesting that shape representations in the lateral occipital cortex are central to the generation of illusory Kanizsa figures (Stanley & Rubin 2003, Murray et al. 2002). Across this thesis therefore there is evidence that LO plays an important role in the construction of different forms of object-hood (Chapter 9, and potentially Chapter 4) and the capacity for those proto-objects to influence the allocation of attention (Chapter 5 & 8). Indeed the potential importance of LO in generating the representations required for further selection is interesting to frame in terms of Stanley and Rubin's (2003, p. 323) conceptualization of LO function:

"the LOC[24] *response to salient regions may be the result of fast but crude region-based segmentation processes, which are useful for selecting parts of cluttered images for more detailed, computationally intensive processing".*

This conceptualisation of LO/LOC (see footnote 25) accords with Driver et al's (2001) argument that object based attention effects reflect a 'mid' level of representation that seeks to provide a rapid, first-pass decomposition of a visual scene in order to provide an initial structure for the later processes of perception and cognition.

This thesis did not set out to explore the role of LO in perceptual organisation, however the study of perceptual organisation as it manifests in the context of object based attention has led to a cluster of results that, although with differing levels of directness, point to a critical role for LO. This consistent role for LO could lead one to question whether this area is in some

---

[24] This paper refers more generally to area LOC, which can be sub-divided into the posterior region LO and anterior region pFs (Haushofer, Livingstone and Kanwisher 2008). Both of these sub-regions are active during the perception of coherent shapes, and might therefore both potentially play a role in developing the representations that influence attention. The more specific reference to LO in the current thesis is motivated by the fact that DF's lesion in the ventral stream focuses on the posterior aspect of LOC (James et al. 2003). Furthermore the cross modal activations in LOC focus on LO, or as it has been called LOtv (James et al 2002, Amedi, et al., 2001, Pietrini et al., 2004). This thesis has therefore referred more specifically to area LO, rather than LOC, however the extent to which representations in pFs can influence the allocation of attention should be tested in future research. For example Haushofer et al. argue that pFs and LO are differentially sensitivity to the perceived vs the physical appearance of different shapes respectively, a finding that could be adapted to explore whether the within-vs-between-object advantage, and object based attention effects more generally, are influenced more by the perceived or by the physical nature of visual shape representations.

senses the neural seat of early perceptual organisation. This possibility led us to question whether one could find evidence of perceptual organisation that was not contingent upon area LO. This question could most effectively be addressed by exploring perceptual organisation in the visual agnosia patient tested in Chapters 8 and 9.

Chapters 8 and 9 had sought to explore whether some form of perceptual organisation could influence this patient's allocation of attention or perception of motion. In both these chapters perceptual organisation was explored with respect to one visual property, shape. Chapter 10 therefore sought to explore whether or not some form of perceptual organisation would be evident on the basis of a combination of non-shape cues, namely colour and motion. The results of Chapter 10 revealed a weak bias indicating that this patient was indeed able to group via a conjunction of colour and motion cues. The weak nature of the bias in Chapter 10 means this effect needs to be replicated, and possible extensions to this finding are suggested in that chapter. Nevertheless as it stands, it suggests that although LO appears to play a key role in providing a rapid segmentation of the visual scene that provides that units for 'object based' attentional selection, LO is not a general resource underpinning all forms of perceptual organisation.

Concluding Comments

This thesis started with the motivation that early forms of perceptual organisation might play a critical role in shaping the nature of higher level cognition. Rather than address this potential role directly however the thesis sought to examine one of the most basic levels at which perceptual organisation can be seen to influence later perception, that is in terms of its influence upon attention. Given the wide variety of phenomena that come under the umbrella term 'object based attention' the thesis did not seek to explore object based attention more generally, but focused primarily on the Egly et al. within-vs-between-object cuing advantage and to a lesser extent the formation of illusory Kanizsa figures and Multiple Object Tracking. After retreating to this more empirically stable ground, can we look back over the field and see any important advances that this thesis has made, or at the least new fronts for future research?

Each Chapter has hopefully added its own little brick to our understanding of perceptual organisation and object based attention. To briefly review each chapter again, Chapter 2 provided evidence that the within-vs-between-object advantage in the Egly et al. paradigm does not reflect a spatial spreading of attention, but a switch cost that may be sensitive to the distinction between figures and grounds. Chapter 3 provided a first step towards establishing whether the Egly et al. within-vs-between-object advantage still manifests with more complicated scene dynamics, and indeed suggests that the representations required for attentional selection could be extracted in parallel across the visual scene. Chapter 4 provided evidence consistent with the notion that the construction of Kanizsa stimuli is not possible when access to higher level ventral resources such as LO are blocked via binocular rivalry induced by Continuous Flash Suppression. Chapter 5 demonstrated that the representations that can be selected in order to elicit the within-vs-between-object advantage can be established purely on the basis of tactile stimulation. Moreover this finding was argued to be consistent with the role of area LO in developing cross modal representations of shape. Chapters 6 and 7 provide evidence consistent with the notion that object based attention effects reflect a 'mid' level of representation that extracts units of selection on the basis of their visual grouping properties, not their higher-level or ecological significance. Chapter 8 showed that despite evidence for perfectly normal spatial attentional selection a patient who could extract form information both to enable implicit visual illusions and for guiding manual actions, did not reveal the classical influence of object structure upon attention, following a bilateral lesion to the ventral shape processing area LO. Chapter 9 extended the logic of Chapter 8 to demonstrate that this patient's motion perception was completely uninfluenced by what is otherwise a compelling motion illusion induced by the construction of illusory Kanizsa shapes. This thesis concluded by demonstrating that LO is not required for all forms of perceptual organisation, and that some degree of grouping between colour and motion can still be achieved following a bilateral lesion to LO.

Probably the most consistent picture that has emerged throughout this research has revolved around the potential role of LO in generating the representations required to elicit the Egly et al. within-vs-between-object advantage. The potential role of LO in determining this within-vs-between-object advantage could provide a useful framework for exploring the extent to which the plethora of different phenomena that come under the heading of 'object based attention', reviewed in the Introduction, actually tap into the same level of representation.

Furthermore the role of LO in object-based attention could provide a useful framework for exploring the infant's developing concept of object-hood in terms of the functional development of this area of the brain. Indeed the results from Chapter 5 could also be explored in a developmental context to examine the extent to which multi-modal representations of shape and their influence upon attention might provide some form of bridge between the different stimulus dimensions extracted by each sensory modality. Finally the results of Chapter 3, indicating that object based attention effects can manifest themselves with more complicated scene dynamics needs to be tested further. If object based attention effects really do reflect a stage of representation that provides a crucial scaffold for later perception and cognition, then we need to go further in proving that these effects operate under ecologically valid settings.

Ahissar, M., & Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends in Cognitive Sciences, 8*(10), 457-464.

Ahlstrom, V., Blake, R., & Ahlstrom, U. (1997). Perception of biological motion. *Perception, 26*(12), 1539-1548.

Albrecht, A. R., List, A., & Robertson, L. C. (2008). Attentional selection and the representation of holes and objects. *Journal of Vision, 8*(13), 10.

Allen, H. A., & Humphreys, G. W. (2009). Direct tactile stimulation of dorsal occipito-temporal cortex in a visual agnosic. *Current Biology, 19*(12), 1044-1049.

Amedi, A., Malach, R., Hendler, T., Peled, S., & Zohary, E. (2001). Visuo-haptic object-related activation in the ventral visual pathway. *Nature Neuroscience, 4*(3), 324-330.

Amedi, A., Stern, W. M., Camprodon, J. A., Bermpohl, F., Merabet, L., Rotman, S., et al. (2007). Shape conveyed by visual-to-auditory sensory substitution activates the lateral occipital complex. *Nature Neuroscience, 10*(6), 687-689.

Anstis, S. (1990) Imperceptible Intersections: The Chopstick Illusion. In: Blake, A. and Troscianko, T. (Eds.) AI and the Eye. John Wiley & Sons. Chichester.

Avrahami, N. (1999). Objects of attention, objects of perception. *Perception & Psychophysics, 61*(8), 1604-1612.

Beauchamp, M. S. (2005). See me, hear me, touch me: multisensory integration in lateral occipital-temporal cortex. *Current Opinion in Neurobiology, 15*(2), 145-153.

Ben-Shahar, O., Scholl, B. J., & Zucker, S. W. (2007). Attention, segregation, and textons: Bridging the gap between object-based attention and texton-based segregation. *Vision Research, 47*(6), 845-860.

Blaser, E., Pylyshyn, Z. W., & Holcombe, A. O. (2000). Tracking an object through feature space. *Nature*, 408(6809), 196-199.

Bonfiglioli, C., Pavani, F., & Castiello, U. (2004). Differential effects of cast shadows on perception and action. *Perception, 33*(11), 1291-1304.

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

Braje, W. L., Legge, G. E., & Kersten, D. (2000). Invariant recognition of natural objects in the presence of shadows. *Perception, 29*(4), 383-398.

Broadbent, D. (1958). Perception and Communication. London: Pergamon Press.

Cavina-Pratesi C, Kentridge RW, Heywood CA, & Milner AD (2009). Separate processing of texture and form in the ventral stream: Evidence from fMRI and visual agnosia, *Cereb Cortex*, published online May 28 2009.

Carey, S., & Xu, F. (2001). Infants' knowledge of objects: beyond object files and object tracking. *Cognition, 80*(1-2), 179-213.

Carey, D. P., Dijkerman, H. C., Murphy, K. J., Goodale, M. A., & Milner, A. D. (2006). Pointing to places and spaces in a patient with visual form agnosia. *Neuropsychologia, 44*, 1584-1594.

Chen, Z. (1998). Switching attention within and between objects: The role of subjective organization. *Canadian Journal of Experimental Psychology, 52*, 7-17.

Chen, Z., & Cave, K. R. (2008). Object-based attention with endogenous cueing and positional certainty. *Perception & Psychophysics*, *70*, 1435-1443.

Conci, M., Bobel, E., Matthias, E., Keller, I., Muller, H. J., & Finke, K. (2009). Preattentive surface and contour grouping in Kanizsa figures: Evidence from parietal extinction. *Neuropsychologia, 47*(3), 726-732.

Cole, G. G., Kentridge, R. W., & Heywood, C. A. (2004). Visual salience in the change detection paradigm: The special role of object onset. *Journal of Experimental Psychology-Human Perception and Performance, 30*(3), 464-477.

Crawford, J. R., & Garthwaite, P. H. (2002). Investigation of the single case in neuropsychology: confidence limits on the abnormality of test scores and test score differences. *Neuropsychologia, 40*(8), 1196-1208.

Crawford, J. R., & Garthwaite, P. H. (2005). Testing for suspected impairments and dissociations in single-case studies in neuropsychology: Evaluation of alternatives using Monte Carlo simulations and revised tests for dissociations. *Neuropsychology, 19*(3), 318-331.

Crawford, J. R., & Garthwaite, P. H. (2007). Comparison of a single case to a control or normative sample in neuropsychology: Development of a Bayesian approach. *Cognitive Neuropsychology, 24*(4), 343-372.

Crick, F. (1994). The Astonishing Hypothesis: The Scientific Seach For The Soul. London Touchstone.

Croner, L. J., & Albright, T. D. (1997). Image segmentation enhances discrimination of motion in visual noise. *Vision Research, 37*(11), 1415-1427.

Davis, G., & Driver, J. (1997). Spreading of visual attention to modally versus amodally completed regions. *Psychological Science, 8*(4), 275-281.

Davis, G., & Driver, J. (1998). Kanizsa subjective figures can act as occluding surfaces at parallel stages of visual search. *Journal of Experimental Psychology-Human Perception and Performance, 24*(1), 169-184.

Davis, G., Driver, J., Pavani, F., & Shepard, A. (2000). Reappraising the apparent costs of attending to two separate visual objects. *Vision Research*, 40(10-12), 1323-1332.

Desimone, R., & Duncan, J. (1995). Neural Mechanisms of Selective Visual-Attention. *Annual Review of Neuroscience, 18*, 193-222.

de-Wit, L. H., Kentridge, R. W., & Milner, A. D. (2009a). Object-based attention and visual area LO. *Neuropsychologia, 47*(6), 1483-1490.

de-Wit, L. H., Kentridge, R. W., & Milner, A. D. (2009b). Shape processing area LO and illusory contours. *Perception, 38*, 1260-1263.

Di Russo, F., Martinez, A., & Hillyard, S. A. (2003). Source analysis of event-related cortical activity during visuo-spatial attention. *Cerebral Cortex, 13*, 486-499.

Driver, J., & Baylis, G. C. (1989). Movement and visual attention - the spotlight metaphor breaks down. *Journal of Experimental Psychology–Human Perception and Performance,* 15(3), 448-456.

Driver, J., Davis, G., Russell, C., Turatto, M., & Freeman, E. (2001). Segmentation, attention and phenomenal visual objects. *Cognition, 80*, 61-95.

Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology–General*, 113(4), 501-517.

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.

Egly, R., Rafal, R., Driver, J., Starrveveld, Y. (1994). Covert orienting in the split-brain reveals hemispheric specialization for object-based attention. *Psychological Science*, 5, 380-383.

Eriksen, B. A., & Eriksen, C. W. (1974). Effect of noise letters upon identification of a target in a nonsearch task. *Perception & Psychophysics, 16*(1), 143-149.

Fang, F., & He, S. (2005). Cortical responses to invisible objects in the human dorsal and ventral pathways. *Nature Neuroscience, 8*(10), 1380-1385.

Feldman, J. (2003). What is a visual object? *Trends in Cognitive Sciences, 7*(6), 252-256.

Fink, G. R., Dolan, R. J., Halligan, P. W., Marshall, J. C., & Frith, C. D. (1997). Space-based and object-based visual attention: shared and specific neural domains. *Brain, 120*, 2013-2028.

Franconeri, S. L., Lin, J. Y., Pylyshyn, Z. W., Fisher, B., & Enns, J. T. (2008). Evidence against a speed limit in multiple-object tracking. *Psychonomic Bulletin & Review, 15*(4), 802-808.

Freeman, E., Sagi, D., & Driver, J. (2001). Lateral interactions between targets and flankers in low-level vision depend on attention to the flankers. *Nature Neuroscience, 4*(10), 1032-1036.

Gegenfurtner, K. R., Kiper, D. C., & Fenstemaker, S. B. (1996). Processing of color, form, and motion in macaque area V2. *Visual Neuroscience, 13*(1), 161-172.

Gilchrist, I. D., & Proske, H. (2006). Anti-saccades away from faces: evidence for an influence of high-level visual processes on saccade programming. *Experimental Brain Research, 173*(4), 708-712.

Goldsmith, M., & Yeari, M. (2003). Modulation of object-based attention by spatial focus under endogenous and exogenous orienting. *Journal of Experimental Psychology-Human Perception and Performance, 29*(5), 897-918.

Goodale, M. A., & Milner, A. D. (1992). Separate Visual Pathways for Perception and Action. *Trends in Neurosciences, 15*(1), 20-25.

Goodale, M. A. & Milner, A. D. 2004. *Sight Unseen: An Exploration of Conscious and Unconscious Vision*. Oxford: Oxford University Press.

Goodale, M. A., Jakobson, L. S., Milner, A. D., Perrett, D. I., Benson, P. J., & Hietanen, J. K. (1994). The nature and limits of orientation and pattern processing supporting visuomotor control in a visual form agnosic. *Journal of Cognitive Neuroscience, 6*, 46-56.

Goodale, M. A., & Westwood, D. A. (2004). An evolving view of duplex vision: separate but interacting cortical pathways for perception and action 2. *Current Opinion in Neurobiology, 14*(2), 203-211.

Gregory, R. (1972). Cognitive contours. *Nature, 238*,, 51-52.

Hadjikhani, N., Liu, A. K., Dale, A. M., Cavanagh, P., & Tootell, R. B. H. (1998). Retinotopy and color sensitivity in human visual cortical area V8. *Nature Neuroscience, 1*(3), 235-241.

Haynes, J. D., & Rees, G. (2005). Predicting the stream of consciousness from activity in human visual cortex. *Current Biology, 15*(14), 1301-1307.

Humphrey, G. K., Goodale, M. A., & Gurnsey, R. (1991). Orientation discrimination in a visual form agnosic – evidence from the McCollough effect. *Psychological Science, 2,* 331-335.

Humphreys, G. W., & Riddoch, J. (2007). How to define an object: Evidence from the effects of action on perception and attention. *Mind & Language, 22*, 534-547.

Hubel, D. H., & Livingstone, M. S. (1987). Segregation of Form, Color, and Stereopsis in Primate Area-18. *Journal of Neuroscience, 7*(11), 3378-3415.

Huxlin, K. R., Saunders, R. C., Marchionini, D., Pham, H. A., & Merigan, W. H. (2000). Perceptual deficits after lesions of inferotemporal cortex in macaques. *Cerebral Cortex, 10*(7), 671-683.

Imura, T., Yamaguchi, M. K., Kanazawa, S., Shirai, N., Otsuka, Y., Tomonaga, M., et al. (2006). Perception of motion trajectory of object from the moving cast shadow in infants. *Vision Research, 46*(5), 652-657.

James, T. W., Culham, J., Humphrey, G. K., Milner, A. D., & Goodale, M. A. (2003). Ventral occipital lesions impair object recognition but not object-directed grasping: an fMRI study. *Brain, 126*, 2463-2475.

James, T. W., Humphrey, G. K., Gati, J. S., Servos, P., Menon, R. S., & Goodale, M. A. (2002). Haptic study of three-dimensional objects activates extrastriate visual areas. *Neuropsychologia, 40*(10), 1706-1714.

Jiang, Y., Costello, P., Fang, F., Huang, M., & He, S. (2006). A gender- and sexual orientation-dependent spatial attentional effect of invisible images. *Proceedings of the National Academy of Sciences of the United States of America, 103*(45), 17048-17052.

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The Reviewing of Object Files - Object-Specific Integration of Information. *Cognitive Psychology, 24*(2), 175-219.

Kant, E. (1781). Critique of Pure Reason, trans. Norman Kemp Smith (N.Y: St

Martins,1965).

Kentridge, R. W., Heywood, C. A., & Weiskrantz, L. (1999). Attention without awareness in blindsight. *Proceedings of the Royal Society of London Series B-Biological Sciences, 266*(1430), 1805-1811.

Kentridge, R. W., Heywood, C. A., & Weiskrantz, L. (2004). Spatial attention speeds discrimination without awareness in blindsight. *Neuropsychologia, 42*(6), 831-835.

Kentridge, R. W., Nijboer, T. C. W., & Heywood, C. A. (2008). Attended but unseen: Visual attention is not sufficient for visual awareness. *Neuropsychologia, 46*(3), 864-869.

Kim, C. Y., & Blake, R. (2005). Psychophysical magic: rendering the visible 'invisible'. *Trends in Cognitive Sciences, 9*(8), 381-388.

Klein, R. (1988). Inhibitory Tagging System Facilitates Visual-Search. *Nature, 334*(6181), 430-431.

Koffka K (1935) Principles of Gestalt Psychology. London: Lund Humphreys.

Kourtzi, Z., & Kanwisher, N. (2001). Representation of perceived object shape by the human lateral occipital complex. *Science,* 293(5534), 1506-1509.

Kravitz, D. J., & Behrmann, M. (2008). The space of an object: Object attention alters the spatial gradient in the surround. *Journal of Experimental Psychology–Human Perception and Performance*, 34(2), 298-309.

Lamme, V. A. F., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences, 23*(11), 571-579.

Lehky, S. R., & Sereno, A. B. (2007). Comparison of shape encoding in primate dorsal and ventral visual pathways. *Journal of Neurophysiology, 97*, 307-319.

Li, H. C. O., & Kingdom, F. A. A. (2001). Segregation by color/luminance does not necessarily facilitate motion discrimination in the presence of motion distractors. *Perception & Psychophysics, 63*(4), 660-675.

Lo, S. Y., & Yeh, S. L. (2008). Dissociation of processing time and awareness by the inattentional blindness paradigm. *Consciousness and Cognition, 17*(4), 1169-1180.

Lovell, P. G., Gilchrist, I. D., Tolhurst, D. J., & Troscianko, T. (2009). Search for gross illumination discrepancies in images of natural objects. *Journal of Vision, 9*(1), 14.

Macaluso, E., & Driver, J. (2001). Spatial attention and crossmodal interactions between vision and touch. *Neuropsychologia, 39*(12), 1304-1316.

Macaluso, E., Frith, C. D., & Driver, J. (2002). Directing attention to locations and to sensory modalities: Multiple levels of selective processing revealed with PET. *Cerebral Cortex, 12*(4), 357-368.

Macquistan, A. D. (1997). Object-based allocation of visual attention in response to exogenous, but not endogenous, spatial precues. *Psychonomic Bulletin & Review*, 4(4), 512-515.

Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, H., Kennedy, W. A., et al. (1995). Object-Related Activity Revealed by Functional Magnetic-Resonance-Imaging in Human Occipital Cortex. *Proceedings of the National Academy of Sciences of the United States of America, 92*(18), 8135-8139.

Marino, A. C., & Scholl, B. J. (2005). The role of closure in defining the "objects" of object-based attention. *Perception & Psychophysics, 67*, 1140-1149.

Martinez, A., Teder-Salejarvi, W., Vazquez, M., Molholm, S., Foxe, J. J., Javitt, D. C., et al. (2006). Objects are highlighted by spatial attention. *Journal of Cognitive Neuroscience, 18*(2), 298-310.

Martinovic, J., Meyer, G., Muller, M. M., & Wuerger, S. M. (2009). S-cone signals invisible to the motion system can improve motion extraction via grouping by color. *Visual Neuroscience, 26*(2), 237-248.

Marr, D. (1982). Vision. H. Freeman and Co, New York.

Mattingley, J. B., Davis, G., & Driver, J. (1997). Preattentive filling-in of visual surfaces in parietal extinction. *Science, 275*(5300), 671-674.

Melcher, D., & Vidnyanszky, Z. (2006). Subthreshold features of visual objects: Unseen

but not unbound. *Vision Research, 46*(12), 1863-1867.

Michotte, A. (1963). The perception of causality. (T.R. Miles & E. Miles, Trans.).

London: Methuen. (English translation of Michotte, 1954).

Mitchell, J. F., Stoner, G. R., & Reynolds, J. H. (2004). Object-based attention determines dominance in binocular rivalry. *Nature,* 429(6990), 410-413.

Milner, A. D., & Goodale, M. A. (1995) *The Visual Brain in Action*. Oxford: Oxford University Press.

Milner, A. D., & Goodale, M. A. (2004) *The Visual Brain in Action, Second Edition*. Oxford: Oxford University Press.

Milner, A. D., & Goodale, M. A. (2008) Two visual systems re-viewed. *Neuropsychologia, 46,* 774-785.

Milner, A. D., & Heywood, C. A. (1989). A disorder of lightness discrimination in a case of visual form agnosia. *Cortex, 25*, 489-494.

Milner, A. D., Perrett, D. I., Johnston, R. S., Benson, P. J., Jordan, T. R., Heeley, D. W., et al. (1991). Perception and action in visual form agnosia. *Brain, 114*, 405-428.

Mole, C. (2008). Attention and consciousness. *Journal of Consciousness Studies, 15*(4), 86-104.

Moore, C. M., & Egeth, H. (1997). Perception without attention: Evidence of grouping

under conditions of inattention. Journal of Experimental Psychology-Human Perception and Performance, 23(2), 339-352.

Moore, C. M., Yantis, S., & Vaughan, B. (1998). Object-based visual selection: Evidence from perceptual completion. *Psychological Science, 9*(2), 104-110.

Müller, N. G., & Kleinschmidt, A. (2003). Dynamic interaction of object- and space-based attention in retinotopic visual areas. *Journal of Neuroscience, 23*, 9812-9816.

Murray, S. O., Boyaci, H., & Kersten, D. (2006). The representation of perceived angular size in human primary visual cortex. *Nature Neuroscience, 9*(3), 429-434.

Murray, M. M., Wylie, G. R., Higgins, B. A., Javitt, D. C., Schroeder, C. E., & Foxe, J. J. (2002). The spatiotemporal dynamics of illusory contour processing: Combined high-density electrical mapping, source analysis, and functional magnetic resonance imaging. *Journal of Neuroscience, 22*(12), 5055-5073.

O'Craven, K. M., Downing, P. E., & Kanwisher, N. (1999). fMRI evidence for objects as the units of attentional selection. *Nature, 401*(6753), 584-587.

Ostrovsky, Y., Cavanagh, P., & Sinha, P. (2005). Perceiving illumination inconsistencies in scenes. *Perception, 34*(11), 1301-1314.

Palmer, S. E., Brooks, J. L., & Nelson, R. (2003). When does grouping happen? *Acta Psychologica, 114*(3), 311-330.

Phelps, E. A., & LeDoux, J. E. (2005). Contributions of the amygdala to emotion processing: From animal models to human behavior. *Neuron, 48*(2), 175-187.

Pietrini, P., Furey, M. L., Ricciardi, E., Gobbini, M. I., Wu, W. H. C., Cohen, L., et al. (2004). Beyond sensory images: Object-based representation in the human ventral pathway. *Proceedings of the National Academy of Sciences of the United States of America, 101*(15), 5658-5663.

Pinker, S. (1997). How the Mind Works. Penguin, London.

Posner, M. I. (1980). Orienting of Attention. *Quarterly Journal of Experimental Psychology, 32*, 3-25.

Pylyshyn, Z. (1999). Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *Behavioral and Brain Sciences, 22*(3), 341-+.

Pylyshyn, Z. W. (2001). Visual indexes, preconceptual objects, and situated vision. *Cognition, 80*(1-2), 127-158.

Pylyshyn, Z.W., & Storm, R.W. (1988). Tracking multiple independent targets: evidence for a parallel tracking mechanism. Spatial Vision, 3(3), 1-19.

Richard, A. M., Lee, H., & Vecera, S. P. (2008). Attentional spreading in object-based attention. *Journal of Experimental Psychology-Human Perception and Performance,* 34(4), 842-853.

Robertson, L. C. (2003). Binding, spatial attention and perceptual awareness. *Nature Reviews Neuroscience, 4*(2), 93-102.

Roelfsema, P. R. (2006). Cortical algorithms for perceptual grouping. *Annual Review of Neuroscience, 29*, 203-227.

Saygin, A. P., Driver, J., & de Sa, V. R. (2008). In the footsteps of biological motion and multisensory perception: Judgments of audiovisual temporal relations are enhanced for upright walkers. *Psychological Science, 19*(5), 469-475.

Saygin, A. P. (2007). Superior temporal and premotor brain areas necessary for biological motion perception. *Brain, 130*, 2452-2461.

Saygin, A.P. & Sereno, M. I. (2008) Retinotopy and attention in human occipital, temporal, parietal and frontal cortex. *Cerebral Cortex*, 18(9): 2158-68.

Schenk, T. & Milner, A. D. (2006). Concurrent visuomotor behaviour improves form discrimination in a patient with visual form agnosia. *European Journal of Neuroscience* 24: 1495-1503.

Schoenfeld, M. A., Tempelmann, C., Martinez, A., Hopf, J. M., Sattler, C., Heinze, H. J., et al. (2003). Dynamics of feature binding during object-selective attention. *Proceedings of the National Academy of Sciences of the United States of America,* 100(20), 11806-11811.

Scholl, B. J. (2001). Objects and attention: the state of the art. *Cognition, 80*(1-2), 1-46.

Scholl, B. J., Pylyshyn, Z. W. (1999). Tracking multiple items through occlusion: Clues to visual Objecthood. *Cognitive Psychology*, 38(2), 259-290.

Scholl, B. J., Pylyshyn, Z. W., & Feldman, J. (2001). What is a visual object? Evidence from target merging in multiple object tracking. *Cognition, 80*(1-2), 159-177.

Seghier, M., Dojat, M., Delon-Martin, C., Rubin, C., Warnking, J., Segebarth, C., et al. (2000). Moving illusory contours activate primary visual cortex: an fMRI study. *Cerebral Cortex, 10*(7), 663-670.

Seghier, M. L., & Vuilleumier, P. (2006). Functional neuroimaging findings on the human perception of illusory contours. *Neuroscience and Biobehavioral Reviews, 30*(5), 595-612.

Senkowski, D., Rottger, S., Grimm, S., Foxe, J. J., & Herrmann, C. S. (2005). Kanizsa subjective figures capture visual spatial attention: evidence from electrophysiological and behavioral data. *Neuropsychologia, 43*(6), 872-886.

Shenberg, D. L., & Logothetis, N. K. (1997). The role of temporal cortical areas in perceptual organization. *Proceedings of the National Academy of Sciences of the United States of America, 94*(7), 3408-3413.

Sincich, L. C., & Horton, J. C. (2005). The circuitry of V1 and V2: Integration of color, form, and motion. *Annual Review of Neuroscience, 28*, 303-326.

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

Steeves, J. K. E., Culham, J. C., Duchaine, B. C., Pratesi, C. C., Valyear, K. F., Schindler,

I., Humprhey, G. K., Milner, A. D. & Goodale, M. A. (2006). The fusiform face area is not sufficient for face recognition: Evidence from a patient with dense prosopagnosia and no occipital face area. *Neuropsychologia*, *44*, 594-609.

Sterzer, P., Jalkanen, L., & Rees, G. (2009). Electromagnetic responses to invisible face stimuli during binocular suppression. *Neuroimage, 46*(3), 803-808.

Tarr, M. J., Kersten, D., & Bulthoff, H. H. (1998). Why the visual recognition system might encode the effects of illumination. *Vision Research, 38*(15-16), 2259-2275.

Tallon, C., Bertrand, O., Bouchet, P., & Pernier, J. (1995). Gamma-range activit evoked by coherent visual-stimuli in humans. *European Journal of Neuroscience, 7*(6), 1285-1291.

Tipper, S. P., Driver, J., & Weaver, B. (1991). Object-centered inhibition of return of visual attention. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology,* 43(2), 289-298.

Treisman, A. M., & Gelade, G. (1980). Feature-integration theory of attention. *Cognitive Psychology*, 12(1), 97-136.

Tsuchiya, N., & Koch, C. (2005). Continuous flash suppression reduces negative afterimages. *Nature Neuroscience, 8*(8), 1096-1101.

Ungerleider, L. G., & Mishkin, M., (1982) Two cortical visual systems. In: Analysis of Visual Behavior (Ingle DJ, Goodale MA, Mansfield RJW, ed), pp549-586. *Cambridge MA: MIT Press.*

Vecera, S. P. (1994). Grouped locations and object-based attention – comment. *Journal of Experimental Psychology-General*, 123(3), 316-320.

Vecera, S. P., & Farah, M. J. (1994). Does visual attention select objects or locations? *Journal of Experimental Psychology-General*, 123(2), 146-160.

Vickery, T. J., & Jiang, Y. V. (2009). Associative grouping: Perceptual grouping of shapes by association. *Attention Perception & Psychophysics, 71*(4), 896-909.

Vinberg, J., & Grill-Spector, K. (2008). Representation of shapes, edges, and surfaces across multiple cues in the human visual cortex. *Journal of Neurophysiology,* 99(3), 1380-1393.

Viswanathan, L., & Mingolla, E. (2002). Dynamics of attention in depth: Evidence from multi-element tracking. *Perception, 31*(12), 1415-1437.

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

Watson, S. E., & Kramer, A. F. (1999). Object-based visual selective attention and perceptual organization. *Perception & Psychophysics, 61*(1), 31-49.

Wolfe, J. M., Oliva, A., Horowitz, T. S., Butcher, S. J., & Bompas, A. (2002). Segmentation of objects from backgrounds in visual search tasks. *Vision Research, 42*(28), 2985-3004.

Wolfe, J. M. (2003). Moving towards solutions to some enduring controversies in visual search. *Trends in Cognitive Sciences, 7*(2), 70-76.

Wyart, V., & Tallon-Baudry, C. (2008). Neural dissociation between visual awareness and spatial attention. *Journal of Neuroscience, 28*(10), 2667-2679.

Yang, E., Zald, D. H., & Blake, R. (2007). Fearful expressions gain preferential access to awareness during continuous flash suppression. *Emotion, 7*(4), 882-886.

Yonas, A., & Granrud, C. E. (2006). Infants' perception of depth from cast shadows. *Perception & Psychophysics, 68*(1), 154-160.

Zald, D. H. (2003). The human amygdala and the emotional evaluation of sensory stimuli. *Brain Research Reviews, 41*(1), 88-123.

Zelinsky, G. J., & Neider, M. B. (2008). An eye movement analysis of multiple object tracking in a realistic environment. *Visual Cognition, 16*(5), 553-566.

Zemel, R. S., Behrmann, M., Mozer, M. C., & Bavelier, D. (2002). Experience-dependent perceptual grouping and object-based attention. *Journal of Experimental Psychology-Human Perception and Performance, 28*(1), 202-217.

Appendix: Replications of the Egly, Driver and Rafal (1994) object based attention paradigm. Papers are listed in order of publication, starting with the most recent.

de-Wit, L. H., Kentridge, R. W., & Milner, A. D. (2009). Object-based attention and visual area LO. *Neuropsychologia, 47*(6), 1483-1490.

Schindler, I., McIntosh, R. D., Cassidy, T. P., Birchall, D., Benson, V., Ietswaart, M., et al. (2008). The disengage deficit in hemispatial neglect is restricted to between-object shifts and is abolished by prism adaptation. Experimental Brain Research, 192 (3), 499-510.

He, X., Humphreys, G., Fan, S., Chen, L., & Han, S. (2008). Differentiating spatial and object-based effects on attention: An event-related brain potential study-with peripheral cueing. *Brain Research, 1245*, 116-125.

Albrecht, A. R., List, A., & Robertson, L. C. (2008). Attentional selection and the representation of holes and objects. *Journal of Vision, 8*(13), 10.

Li, X. S., & Logan, G. D. (2008). Object-based attention in Chinese readers of Chinese words: Beyond Gestalt principles. *Psychonomic Bulletin & Review, 15*(5), 945-949.

Chen, Z., & Cave, K. R. (2008). Object-based attention with endogenous cuing and positional certainty. *Perception & Psychophysics, 70*(8), 1435-1443.

Kravitz, D. J., & Behrmann, M. (2008). The space of an object: Object attention alters the spatial gradient in the surround. *Journal of Experimental Psychology-Human Perception and Performance, 34*(2), 298-309.

Hecht, L. N., & Vecera, S. P. (2007). Attentional selection of complex objects: Joint effects of surface uniformity and part structure. *Psychonomic Bulletin & Review, 14*(6), 1205-1211.

List, A., & Robertson, L. C. (2007). Inhibition of return and object-based attentional selection. *Journal of Experimental Psychology-Human Perception and Performance, 33*(6), 1322-1334.

Brown, J. M., & Denney, H. I. (2007). Shifting attention into and out of objects: Evaluating the processes underlying the object advantage. *Perception & Psychophysics, 69*(4), 606-618.

Ariga, A., Yokosawa, K., & Ogawa, H. (2007). Object-based attentional selection and awareness of objects. *Visual Cognition, 15*(6), 685-709.

Ben-Shahar, O., Scholl, B. J., & Zucker, S. W. (2007). Attention, segregation, and textons: Bridging the gap between object-based attention and texton-based segregation. *Vision Research, 47*(6), 845-860.

Martinez, A., Ramanathan, D. S., Foxe, J. J., Javitt, D. C., & Hillyard, S. A. (2007). The role of spatial attention in the selection of real and illusory objects. *Journal of Neuroscience, 27*(30), 7963-7973.

Shomstein, S., & Behrmann, M. (2006). Cortical systems mediating visual attention to both objects and spatial locations. *Proceedings of the National Academy of Sciences of the United States of America, 103*(30), 11387-11392.

Avrahami, J. (2006). Selective attention as tuning: The case of stroke weight. *Perception & Psychophysics, 68*(2), 208-215.

Martinez, A., Teder-Salejarvi, W., Vazquez, M., Molholm, S., Foxe, J. J., Javitt, D. C., et al. (2006). Objects are highlighted by spatial attention. *Journal of Cognitive Neuroscience, 18*(2), 298-310.

Marino, A. C., & Scholl, B. J. (2005). The role of closure in defining the "objects" of object-based attention. *Perception & Psychophysics, 67*(7), 1140-1149.

Lee, H., & Vecera, S. P. (2005). Visual cognition influences early vision - The role of visual short-term memory in amodal completion. *Psychological Science, 16*(10), 763-768.

Moore, C. M., & Fulton, C. (2005). The spread of attention to hidden portions of occluded surfaces. *Psychonomic Bulletin & Review, 12*(2), 301-306.

Linnell, K. J., Humphreys, G. W., McIntyre, D. B., Laitinen, S., & Wing, A. M. (2005). Action modulates object-based selection. *Vision Research, 45*(17), 2268-2286.

Buchholz, J., & Davies, A. A. (2005). Adults with dyslexia demonstrate space-based and object-based covert attention deficits: Shifting attention to the periphery and shifting attention between objects in the left visual field. *Brain and Cognition, 57*(1), 30-34.

Bekkering, H., & Pratt, J. (2004). Object-based processes in the planning of goal-directed hand movements. *Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology, 57*(8), 1345-1368.

Fischer, M. H., & Hoellen, N. (2004). Space- and object-based attention depend on motor intention. *Journal of General Psychology, 131*(4), 365-377.

He, X., Fan, S. L., Zhou, K., & Chen, L. (2004). Cue validity and object-based attention. *Journal of Cognitive Neuroscience, 16*(6), 1085-1097.

Shomstein, S. & Yantis, S. (2004). Configural and contextual prioritization in object-based attention. Psychonomic Bulletin & Review, 11 (2), 247-253.

Marrara, M. T., & Moore, C. M. (2003). Object-based selection in the two-rectangles method is not an artifact of the three-sided directional cue. *Perception & Psychophysics, 65*(7), 1103-1109.

Goldsmith, M., & Yeari, M. (2003). Modulation of object-based attention by spatial focus under endogenous and exogenous orienting. *Journal of Experimental Psychology-Human Perception and Performance, 29*(5), 897-918.

Muller, N. G., & Kleinschmidt, A. (2003). Dynamic interaction of object- and space-based attention in retinotopic visual areas. *Journal of Neuroscience, 23*(30), 9812-9816.

Mortier, K., Donk, M., & Theeuwes, J. (2003). Attentional capture within and between objects. *Acta Psychologica, 113*(2), 133-145.

McCarley, J. S., Kramer, A. F., & Peterson, M. S. (2002). Overt and covert object-based attention. *Psychonomic Bulletin & Review, 9*(4), 751-758.

Shomstein, S., & Yantis, S. (2002). Object-based attention: Sensory modulation or priority setting? *Perception & Psychophysics, 64*(1), 41-51.

Lamy, D., & Egeth, H. (2002). Object-based selection: The role of attentional shifts. *Perception & Psychophysics, 64*(1), 52-66.

Abrams, R. A., & Law, M. B. (2002). Random visual noise impairs object-based attention. *Experimental Brain Research, 142*(3), 349-353.

Pratt, J., & Sekuler, A. B. (2001). The effects of occlusion and past experience on the allocation of object-based attention. *Psychonomic Bulletin & Review, 8*(4), 721-727.

Iani, C., Nicoletti, R., Rubichi, S., & Umilta, C. (2001). Shifting attention between objects. *Cognitive Brain Research, 11*(1), 157-164.

Arrington, C. M., Carr, T. H., Mayer, A. R., & Rao, S. M. (2000). Neural mechanisms of visual attention: Object-based selection of a region in space. *Journal of Cognitive Neuroscience, 12*, 106-117.

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 and Performance, 26*(4), 1387-1400.

Abrams, R. A., & Law, M. B. (2000). Object-based visual attention with endogenous orienting. *Perception & Psychophysics, 62*(4), 818-833.

Avrahami, N. (1999). Objects of attention, objects of perception. *Perception & Psychophysics, 61*(8), 1604-1612.

Robertson, L. C., & Kim, M. S. (1999). Effects of perceived space on spatial attention. *Psychological Science, 10*(1), 76-79.

Moore, C. M., Yantis, S., & Vaughan, B. (1998). Object-based visual selection: Evidence from perceptual completion. *Psychological Science, 9*(2), 104-110.

Johnson, M. H., & Gilmore, R. O. (1998). Object-centered attention in 8-month-old infants. *Developmental Science, 1*(2), 221-225.

Chen, Z. (1998). Switching attention within and between objects: The role of subjective organization. *Canadian Journal of Experimental Psychology-Revue Canadienne De Psychologie Experimentale, 52*(1), 7-17.

Macquistan, A. D. (1997). Object-based allocation of visual attention in response to exogenous, but not endogenous, spatial precues. *Psychonomic Bulletin & Review, 4*(4), 512-515.

Buck, B. H., Black, S. E., Behrmann, M., Caldwell, C., & Bronskill, M. J. (1997). Spatial- and object-based attentional deficits in Alzheimer's disease - Relationship to HMPAO-SPECT measures of parietal perfusion. *Brain, 120*, 1229-1244.

Vecera, S. P. (1994). Grouped Locations and Object-Based Attention - Comment. *Journal of Experimental Psychology-General, 123*(3), 316-320.