Visual Selection as an Object-oriented Mechanism: An Ecological Perspective towards the Primacy of Objects over Space

Atanaska Nikolova

Cardiff University

School of Psychology

Doctor of Philosophy
2016

Summary

The visual world consists of objects. Planning or performing actions requires some form of engagement with an object. This requirement has shaped our perceptual systems to be highly tuned to ‘objecthood’ and construct objects from minimal available information. This project aimed to explore to what extent the importance of objects influences visual selection: the mechanism that prioritises the necessary information subsets in order to perform an action, and investigate on what basis this information is prioritised. Current visual selection theories argue prioritisation is accomplished as a combination between space-based and object-based mechanisms, with space having a prime role in how information is selected from the environment. This project proposes an alternative view, suggesting selection is a fully object-oriented mechanism and space-based effects are a consequence of object-based selection. This possibility was tested in three empirical chapters with the use of cueing paradigms, in the context of immediate perceptual decisions (luminance change identification), and colour change detection involving visuo-spatial short term memory.

The key premise is that there is an intrinsic link between the spatial separation of any two points and the likelihood they belong to the same object. If these points are perceived to be within the same object, visual selection is not affected by the distance between them and they are equally prioritised for action. Prioritisation level decreases with increasing distance only when this likelihood of object-belongingness is low, because points closer together have a higher probability to originate from the same object. The current work tested this premise by varying independently object-belongingness and spatial proximity of cue-target stimuli pairs. Results indicated that visual selection is fully object-oriented and can be distance-independent. It is proposed that the perceptual system assesses the probability that information is integrated into potential objects, and then prioritises selection based on this object-belongingness probability.

Acknowledgements

This work would not have been possible without the support and positive encouragement from a number of people.

First of all, I would like to thank my supervisor, Bill Macken, who always stayed optimistic and encouraged me to pursue results further when I was in doubt or feeling discouraged. I always left meetings feeling more motivated and excited to explore novel hypotheses. Bill always supported me to aim higher and put a lot of faith in me, even when I was lacking faith in myself. I appreciate the independence he granted me during the process of this PhD, as well as my Master's thesis prior to that. It has been a long journey, and I have certainly learnt and developed a lot.

Special gratitude also goes to the ESRC for funding this work and making it possible to attend national and international conferences for dissemination and networking.

Another person who helped me immensely and deserves a special mention is Geoffrey Megardon. I would like to thank him for being so patient in supporting my first steps in learning Matlab during those late nights in the lab, and for being next to me through every happy or tough moment – both for this PhD and in general.

I would also like to thank all the incredible friends I have made during this journey. These are friends for life, no matter where life takes us. We all went through this together, building a friendship based on experiences of joy and sorrow that nobody else would understand – and this is what makes these friendships special.

Finally, I would like to thank all my family for always believing in me and encouraging me to develop. Especially my mum, Tanya, for her unconditional love and support in absolutely everything I do.

Contents

Chapter 1

General Introduction

Visual selection and its functional purpose

The survival of living organisms necessitates interaction with the external environment, and therefore sensitivity and reactivity towards the properties of this environment. These properties are not random, but exhibit statistical regularities and predictable structures, which have consequently played a key role in the development and evolution of the perceptual system of organisms (Simoncelli & Olshausen, 2001). Therefore, as a realistic and logical consequence, it can be proposed that the mechanisms of the perceptual system are optimised for the environment within which the organism functions, and the information from the sensory organs must be well integrated in a way that allows adaptive functioning (Gibson, 1966).

One critical source of sensory information is visual perception. Visual perception is more than the visual experience itself, i.e. the act of seeing, but it is also relatable to the processes that influence behaviour and cognition as a result of the visual input (Cavanagh, 2011). An important part of these processes is *visual selection*. As the term suggests, visual selection serves to prioritise for processing a subset of the visual information from all the available input (Duncan, 2013), and it is often referred to as visual attention. It can have an overt form, so what is selected is also foveated, or it can operate covertly, when the target of selection does not coincide with the direction of gaze (Fuller & Carrasco, 2006). Since selection plays a central part in the functioning of organisms, it is important to understand and explain its mechanism of operation. Thus, the principal aim of the current project is to contribute towards establishing a parsimonious answer to the question regarding what is the basis of visual selection.

A critical question to pose in order to address this issue relates to what is the purpose or function of visual selection. In other words, why there is a need to prioritise. Asking questions about functional purposes is crucial to understanding how any system operates, as approaching the issue from an adaptive perspective addresses the very basics of the concept, relating to its usefulness to the organism. In simpler terms, finding out how something works may be greatly facilitated by understanding why it exists and what its purpose is. Therefore, it is important to address these issues with regards to visual selection, or attention, as they would be inevitably linked to the questions addressed throughout this work.

One of the key accounts regarding the purpose of visual selection is that it is necessary to cope with capacity limitations: the brain has limited processing resources and cannot cope with all the available information, so selectivity is needed to economise these resources (e.g. Awh, Dhaliwal, Christensen, & Matsukura, 2001; Awh, Belopolsky, & Theeuwes, 2012; Carrasco, 2011; Emmanouil & Magen, 2014; Fiebelkorn, Saalmann, & Kastner, 2013; Pestilli & Carrasco, 2005, to name a few). Consequently, a lot of research effort is focused on trying to quantify the attentional capacity and characterise the purported limited cognitive resource (Cowan et al., 2005; Franconeri, Alvarez, & Cavanagh, 2013; Holcombe & Chen, 2013; Dale Lee, Koch, & Braun, 1999; McAvinue et al., 2012; Sperling & Hsu, 2014). However, it is often the case that capacity as such cannot be strictly quantified, and no unitary resource can be found (e.g. Duncan, 2006; Huang & Pashler, 2005; Marois & Ivanoff, 2005; Navon, 1984).

To illustrate this point, it was originally established that when presented with a set of identical and randomly moving items (e.g. a field of circles or crosses), a person can simultaneously keep track (i.e. attend to, or select) of up to four or five such items from a total of ten (Pylyshyn & Storm, 1988). However, subsequent research indicated this limit can vary between one and eight items depending on factors such as speed of motion (Alvarez & Franconeri, 2007), as well as the extent to which the target items can be grouped into a spatial configuration while in motion (Yantis, 1992; Zhao et al., 2014). That is, more items can be simultaneously tracked if they can be perceptually grouped into a single moving shape. Similarly, in the domain of visuo-spatial short term memory, the perceptual organisation of the to-be-remembered items can have a profound impact on memory capacity estimates, e.g. varying between four and sixteen features (Anderson, Vogel, & Awh, 2013; Luck & Vogel, 1997). Therefore, it is very challenging to quantify the limits of visual selection, which in turn questions the proposal that selection results from limited processing resources.

Essentially, limits to performance may vary greatly depending on task demands, local stimulus properties, as well as global properties emerging from the interaction between all items on the visual scene (Davis, Welch, Holmes, & Shepherd, 2001; Macken, Taylor, & Jones, 2015). Nevertheless, the idea that selection has evolved due to the apparently elusive concept of limited resources is an *a priori* assumption for the majority of research in this domain (Neumann, 1996). Indeed, to some extent it may seem reasonable that selection is necessary to deal with limited capacity, because if there were no processing limits, then selectivity would be redundant (Mesulam, 1985). However, this is not necessarily the case. As Neumann (1996) argues, even if a hypothetical organism enjoying unlimited processing capacity exists, it will still need to choose an appropriate action and the most relevant target for this action, relative to its current needs and circumstances.

The ‘limited capacity’ assumption has been critically evaluated and questioned in detail by Neumann (1987, 1996) who proposes the inverse possibility: selection is not the result of limited capacity, but capacity limitations are the side effect of dealing with selection problems. The purpose of selection, therefore, is to ensure the appropriate behavioural output, a notion termed *selection for action* (Allport, 1987). According to this perspective, there are no capacity limitations on the senses per se, or a limited internal resource that needs to be sparingly allocated. Instead, at any one point in time a subset of the information needs to be acted upon and the organism has to resolve the problem of selecting the appropriate afforded action. As a result, performance is constrained (and capacity limitations inferred) by the level of compatibility and integrity between possible actions, and the extent to which the organism can coordinate their coherent execution (Allport, 1987, 1989, 1992).

According to the *selection for action* account, the function of visual selection is not to cope with limited processing capacity, but to ensure adaptive existence by responding to the most relevant aspect of the environment in the most optimal manner. One consequence of this process is that performance may vary depending on how the task is structured, i.e. is it in a way that affords most efficient use of the available apparatus (effector system), or in a way that introduces a conflict in the system (Neumann, 1987). For example, attending to two objects or events simultaneously may be more demanding than attending to a single event, so if the same amount of information that constitutes the two events is restructured to form a single perception, this may result in substantial improvement of performance. Indeed, such effects are clearly observable in divided attention tasks, where judging two properties (e.g. shape and orientation) is slower and less accurate when these attributes belong to two different objects, compared to when they are incorporated within the same perceptual object (e.g. Duncan, 1984; Lavie & Driver, 1996; Matsukura & Vecera, 2011; Vecera & Farah, 1994; Watson & Kramer, 1999).

Similarly, a limitation may arise for motor action execution if two separate actions require the use of the same mechanism, but if the two actions are joined into a single sequence, then this conflict can be eliminated (Klapp & Jagacinski, 2011). For example, the difficulty of simultaneously performing two different tapping patterns with each hand can be eliminated when the patterns are merged into a single sequence or rhythm (Klapp, Nelson, & Jagacinski, 1998). Therefore, the limitation was not due to resource depletion per se, but rather it resulted from the problem of selection conflict. A consequence of the non-capacity model (selection for action) is that visual selection is not dependent on a unitary resource, but arises via the interaction of multiple brain regions and neural pathways, not all of which have a functional specialisation and thus can be flexibly adjusted in accordance with the current demands (Duncan, 2006, 2013; Hannus, Cornelissen, Lindemann, & Bekkering, 2005; Woolgar, Williams, & Rich, 2015).

Going back to the original question asked – what is the basis or ‘unit’ of selection - if the functional purpose of selection is action-oriented, then it is the current target object itself which is the basis of selection. Actions are directed towards objects for the purpose of interacting with the external environment, so the perceptual system needs to select the relevant object and take its properties into account for the correct execution of the current behavioural goal. This mechanism is in line with an adaptive functionality of visual selection, and it suggests that selection is object-based. Nevertheless, research on visual selection globally proposes that the mechanism behind it is not fully object-based, but rather it is directly dependent on the spatial distribution of the visual information on the scene, that is, visual selection is space-based (e.g. Vecera, 1994).

The main implication from this account is that the strength of selection decreases with increasing distance from the attention focus. The role of objects is also recognised, e.g. there is evidence for the automatic processing of all properties of a selected object, even if some of these properties are task-irrelevant (Kahneman & Henik, 1981). Also, given equal spatial distance between a target and a cue, a target is processed faster and more accurately when it is part of the same object as the immediately preceding cue, compared to when it is a part of a different object to the one that contained the cue (Egly, Driver, & Rafal, 1994). However, the object basis is often seen as also modulated by space, i.e. factors such as spatial separation between selection targets or objects (Hollingworth, Maxcey-Richard, & Vecera, 2012). The current consensus in the literature is therefore that space-based and object-based selection coexist (Chen, 2012; Egeth & Yantis, 1997).

The main argument behind the current empirical work is that the coexistence of space-based and object-based selection is not in line with an adaptive mechanism, and instead a more parsimonious account of visual selection can be proposed, one which regards selection as purely object-oriented. There are a number of challenges which arise as a result. First of all, to successfully support the notion that selection is fully object-based, it is necessary to account for the findings which postulate that space is a key factor in selection. This involves introducing an alternative explanation of existing results, an explanation which poses objects as the primary and only factor in visual selection. In relation to this, the second challenge is to define what constitutes an object for the perceptual system, as this is also partially where inconsistencies and limitations in previous research arise, and it is a critical point for understanding object-based effects. Finally, another challenge is to try to overcome previous limitations by utilising stimuli which address confounding factors and have the potential to provide a genuine test for the object-based proposal. The present work aims to tackle these challenges, and as a starting point the purported evidence for space-based selection is discussed, in order to identify its limitations and how they can be addressed by a pure object-oriented account.

The basis of visual selection: The role of space

Space is often described as an ‘indispensable’ and primary factor in visual perception, just as pitch is for audition (Kubovy, 1981). Visual space is the medium which all tangible things, or potential selection targets, are situated in, and spatial separation is necessary in order to enumerate objects (Kubovy & Van Valkenburg, 2001). In other words, two objects cannot occupy the same space at the same time, so space is considered as a critical factor for the visual dimension of perception.

Given this proposition, a large proportion of research is focused on the role of space in visual selection, and a standard methodology for assessing the influence of space is the cueing paradigm (Posner, 1980; Posner & Cohen, 1984). The task involves maintaining fixation at a central point, while responding with a manual key press to the onset of a peripheral target, such as a dot or a square, i.e. selection is executed covertly. Importantly, the target is preceded by a visual cue, which is typically a brief luminance increment (e.g. 50 ms) occupying one of the potential target locations. The cue can have a predictive value, e.g. it indicates that the target is 80% likely to appear at the cued location, or it can be neutral by not being correlated with the target location. Faster reaction times to the subsequent target are observed when its location matches the location indicated by the predictive, relative to the neutral cue, and also a cost is observed, i.e. increased reaction time, if the target appears at a location other than the cued one. Therefore, visual selection can be spatially allocated, i.e. constricted to a specific spatial location (Kiefer & Siple, 1987; McCormick & Klein, 1990). This has led to conceptualising selection as a spatial spotlight mechanism, whereby any visual information falling within the boundaries of the spotlight enjoys privileged processing.

A spotlight mechanism, however, can be too simplistic to encompass the influence the complexity of the visual environment can have on selection. Additional research has led to the conclusion that the shape and range of the spotlight can be adjusted as required by the task, shifting the spotlight metaphor into a zoomlens (Eriksen & James, 1986; Eriksen & Yeh, 1985; Hoffman & Nelson, 1981; Laberge & Brown, 1986). This model proposes that attention can be widely or narrowly distributed across space depending on the difficulty of the task, such that a more demanding task (e.g. identifying a target letter in a field of various non-target letters or other type of visual clutter) can lead to constraining the attentional zoomlens to a smaller spatial region, leading to less distraction from irrelevant stimuli (Forster & Lavie, 2008). Consequently, an easier task (e.g. with lower perceptual or cognitive load) can lead to the zoomlens encompassing a larger spatial area of the display, but an increased spread also results in decreased resolution (Castiello & Umiltà, 1990).

This reasoning is in line with the idea of limited resources – there is only so much attentional resource to be allocated, so it can either be highly concentrated or diluted across a wider area. Furthermore, it is not clearly established whether the attentional focus can be ‘split’, i.e. selecting more than one discrete location at a time, or there is a single zoomlens adjusted to include a broader spatial region in order to incorporate distant targets with the space in-between also selected (Eriksen & Yeh, 1985; Jans, Peters, & De Weerd, 2010; Kiefer & Siple, 1987; McCormick & Klein, 1990). Although a zoomlens model provides more flexibility for selection compared to a rigid spotlight, it still suggests that all information within the selected area is equally processed.

Further research into the role of space in selection has developed the zoomlens model into a Gaussian gradient-type distribution. For example, in a spatial cueing task where a visual cue occurring at one of ten horizontally arranged square placeholders (five on each side of fixation) is followed by a target (luminance increment) at one of these locations results in fastest detection reaction times when the target matched the cued location, and gradually slower reaction times with increasing cue-target distance (Downing & Pinker, 1985). This is taken as evidence that visual selection operates not just as an all-or-nothing spotlight region, but the selected locus is surrounded with a spatial gradient of gradually decreasing facilitation for the processing of visual stimuli, and the level of processing also decreases with visual eccentricity (Downing, 1988; Shulman, Wilson, & Sheehy, 1985; Shulman, Sheehy, & Wilson, 1986). Even when eccentricity is controlled, e.g. using targets centred on an imaginary circle around fixation and thus enjoying approximately equal visual acuity, cue-target distance affects both accuracy and reaction time. For example, cueing one of eight possible locations on the imaginary circle leads to superior performance for identifying a subsequent letter (X or O) when it matched the cued location (cueing effect), and a gradual drop in performance as cue-target distance increases (Henderson, 1991; Henderson & Macquistan, 1993).

Spatially graded selection is present not only in facilitation effects, but also in interference from distractors. Discriminating the identity of a central target is substantially impaired if the target is surrounded (flanked) with peripheral distractors (flankers) on each side, and these distractors have an identity which introduces a response conflict, e.g. they coincide with the alternative response possibility, such as if the target is the letter “A” (the alternative being “B”), while the distractors are the letter “B” (Eriksen & Eriksen, 1974). However, the spacing between the target and distractor letters can modulate the response, as more interference (slower response times) was observed by proximal compared to distant distractors. There is also evidence that the spatial profile of visual selection can follow a ‘Mexican hat’ shape, whereby interference gradually decreases with target-distractor distance, and then starts to increase again (the bottom of the function, i.e. the point where interference starts to increase again, may depend on cognitive and perceptual load factors) (Caparos & Linnell, 2010; Linnell & Caparos, 2011). In any case, whether the performance function is linear or not, the implication is that the facility with which covert visual selection operates is directly influenced by spatial separation between stimuli, and this is also corroborated with physiological indices. For example, the amplitude of visual event related potentials (ERPs), specifically components P135 and N190, is found to progressively decrease as distance between the primary focus of covert selection and the location of a visual target increases (Mangun & Hillyard, 1987, 1988).

Apart from directly influencing the strength of visual selection, the role of space is also seen as a primary dimension for selecting a visual stimulus, more influential than other stimulus properties such as colour or shape. This is evidenced in studies demonstrating that selection is mediated by the spatial location of the stimulus. For example, Tsal & Lavie (1988) presented participants with a circular array of six letters which varied in colour or shape (e.g. angular or curved). The primary task was to identify a letter of a pre-specified colour or shape, and as a secondary task to report any other letters from the display. Letters spatially proximal to the key target were more frequently reported than letters that shared the same colour or shape with the target. In other words, selection on the basis of spatial proximity was preferred, compared to selection based on other stimulus aspects. This was additionally supported in a later study by Tsal & Lamy (2000) who used a similar methodology, but some of the letters were surrounded by a coloured shape or partially superimposed on it. Participants were initially required to report what is the shape of a given colour, and then report any letter from the display. The preferred choice was consistently for the letter contained within the target shape, rather than a letter sharing its colour. Together, this evidence is taken to suggest that location is a special property, and selection is mediated by space.

Considering the studies reviewed above, the key implication is that visual selection is directly affected by and dependent on space. Spatial proximity between stimuli (e.g. cue and subsequent target, or target and distractor) is a critical constraint on what aspects of the visual environment are selected, and the strength of interference from irrelevant visual events. The research evidence points towards a selection mechanism operating on the basis of space, placing key importance on aspects such as spatial proximity and location. However, to what extent is visual selection on the basis of space appropriate in functional terms? In other words, what is the adaptive value of space-based selection?

The targets for visual selection can be described as being distributed in space, i.e. they occupy specific spatial locations. However, as pointed out earlier, selection is the outcome or a consequence of the process required for executing a specific action plan, and the target for this action is necessarily an object of some type, not space itself (Allport, 1987). It can be argued, of course, that processing space for the purpose of calculating distance is important for action. However, in this case as well, the action target is an object whose distance needs to be calculated in order to, for instance, calibrate a movement towards it. In other words, it is objects and their dimensions that need to be selected and processed in order to react appropriately and optimally to the environment (i.e. ensuring survival). This suggests selection can be a flexible mechanism, adjusting to the current goals and affordances of the environment. Space-based selection may not be efficient in such circumstances, leading to the need for an alternative medium for visual selection, that is, object-based selection.

The mechanisms of visual selection need to be optimised for the way the visual environment is structured and the statistical properties of the natural visual scene (Field, 1987, 1989). Therefore, as mentioned earlier, the visual system and visual cognition in general have evolved, as all senses do, in response to continuous interaction with the environment (Geisler & Diehl, 2003; Geisler, 2008). An observation of the environment can confirm that it is composed of solid objects varying in size and position from the observer, resulting in multiple occlusions and discontinuities. Also, as objects vary in size, the same amount of distance can encompass a few small objects, or a single large one. If space is the primary dimension for visual selection, this may be problematic for the appropriate selection of an action target. A spatially graded mechanism, for example, may result in prioritising aspects of the environment which are not important for the current goal or setting. For instance, this may lead to selecting aspects from an irrelevant (relative to the current task) element just because it is proximal to the current target object. An object-based mechanism is more efficient and accurate in a world populated with objects. For example, selecting an object allows to prioritise its parts which are not visibly connected to it (e.g. due to another object creating an occlusion), but are still part of it and thus behaviourally relevant.

Indeed, evidence from natural scene statistics suggests an object-level relationship between any two randomly chosen points on an image (Ruderman, 1997). Specifically, the analysis of natural images involves calculating a pixel difference function (with respect to the logarithm of each pixel's luminance), obtained by averaging the squares of pixel differences between two points in an image. The analysis reveals that when two points originate from the same object this difference function does not vary with distance (as measured in pixels), while if the two points happen to come from different objects, the difference is larger, but again distance-invariant (Ruderman, 1997).

Ruderman (1997) suggests this distance-invariance exists because if two points belong to different (or the same) objects, then it is not essential how distant these objects are, as the statistical relationship between them will always be the same, i.e. those points will remain (un)correlated to the same extent. Spatial separation is not likely to change this relationship (that is, not considering issues such as lighting, which may be more similar if the objects are proximal to each other). Another point of importance extracted by the natural image analyses is that the probability of a certain point pair belonging to the same object drops with increasing distance, and at the same time natural images are scale-invariant (i.e. statistical relationships do not change with a change in observation scale), as they are composed of statistically independent objects spanning a variety of shapes and sizes (Ruderman, 1994; 1997; Ruderman & Bialek, 1994). These results were corroborated by Baddeley (1997), summarising that within natural images, a correlation statistic which gradually decreases with distance results from the combination of three key factors: high within-object correlation, low between-object correlation, and multiple objects of different size and viewing angle.

In summary, natural scene statistics suggest that it is the objects on the scene that dictate or define the relationship between any two points in space. Spatial factors and location per se are not at all critical, but rather it is object-level factors that lead to the observed statistical relationships in the environment. In relation to this point, there is evidence to suggest that visual perception (and in fact, perception globally) works on the basis of Bayesian inference, i.e. integrating prior knowledge of the visual scene with the current visual input in order to arrive to the most likely perceptual interpretation (Kersten, Mamassian, & Yuille, 2004; Kersten & Yuille, 2003; Nakayama & Shimojo, 1992). This is especially evident in the case of ambiguities and can often result in visual illusions, as the perceptual system ‘fills in’ missing or incoherent information with the most probable outcome (Brown & Friston, 2012). Since the prior knowledge needed to form the visual perception is based on the prolonged interaction of the organism with a visual environment (Gibson, 1966), and in turn, this environment consists of multiple objects and surfaces (Ruderman, 1997), then it follows that visual selection is likely to be specifically adapted for interacting with object structures, rather than space per se.

An object-based mechanism of selection has a higher adaptive value than space-based mode of selection. Moreover, given these circumstances, there is no obvious need for a space-based mechanism at all. Such ecological view of visual cognition in general, where selection is influenced by natural scene statistics and regularities in the visual world, fits well with the selection for action account (e.g. Allport, 1992), as actions are ultimately directed towards objects.

How does the perspective of having object structures as the sole influence on selection fit with the dominant view that it is space that is the primary medium for selection and all other dimensions come second or are qualified by space? As already mentioned, effects of spatial separation are globally interpreted as evidence for the primacy of space, i.e. space-based visual selection (Carrasco, 2011). However, given that in the natural world two points close together are more likely to originate from the same object than two points that are far apart (Ruderman, 1997), effects of spatial separation, i.e. graded facilitation or interference, can in fact be interpreted from an object-oriented perspective. In other words, because spatial proximity and object belongingness are positively correlated, instead of assuming a special role for space and location, it may be the case that the observed effects of spatial proximity have an object-level origin. Proximal points may be prioritised because of their high likelihood of belonging to the same object, all other things being equal.

Given this proximity-objecthood correlation, good performance when cue and target are close to each other, and more interference from proximal flanker distractors, may be due to object-level perceptual organisation factors, rather than space. These factors relate to typical characteristics of objects, such as integrity and coherence. If the stimuli are perceived to be parts of the same object (which is more likely when they are proximal), then they would be equally prioritised for processing, resulting in better performance, or stronger interference, depending on the task. However, given that image measurements for within-object points in natural scenes correlate close to 1, while between-object correlations are close to 0 (Baddeley, 1997), it also follows that altering spatial separation between two different objects, or between stimuli perceived to be parts of the same object, should have little or no effect on visual selection. In other words, selection is guided by the objects in the environment, and effects of spatial separation are not necessarily evidence for a space-based mechanism of selection. Importantly, it is not simply the case that the perceptual system uses spatial information (e.g. proximity) to construct objects, but rather it is the proximity cue itself that is constructed by object-based perception.

Indeed, there is evidence for the correlation between spatial proximity and object formation in the Gestalt literature and perceptual organisation research, but it is nevertheless largely interpreted in a way that gives primacy to space, rather than objects. A particularly relevant point here is that spatial proximity is established as a very strong grouping cue (Claessens & Wagemans, 2005; Kubovy & van den Berg, 2008; Kubovy & Wagemans, 1995). There are various Gestalt laws and cues to the formation of perceptual objects, but proximity is amongst the most potent ones, capable of overriding others (Elder & Goldberg, 2002; Han, Humphreys, & Chen, 1999). What this means in practical terms is that the closer two points are to one another, the more likely they are to be grouped into a single perceptual unit (Kubovy, Holcombe, & Wagemans, 1998; Oyama, 1961; Pomerantz & Schwaitzberg, 1975). And this may be especially true when there are no other cues to objecthood, or if those cues are weak or ambiguous. This can be linked back to the point regarding perception as a Bayesian inference (Kersten et al., 2004), whereby how the scene gets constructed, i.e. perceived, depends on learnt regularities and patterns in the environment.

Although this role of proximity as an object formation cue is strongly related to the findings in natural scene statistics discussed earlier, it is interpreted from a spatial perspective. Specifically, spatial information cues are used in order to establish the perception of objects. Space is the primary dimension and objects can emerge from spatial relationships as a secondary effect, i.e. a consequence. In reality, however, the reason for the potent effect of proximity is most likely rooted in the structure of the natural world. Therefore, it may equally be interpreted as a consequence of object-oriented perception, such that the perception of the available objects on the visual scene determines the extent of spatial separation effects, i.e. these effects are an emergent property of object-based selection.

The basis of visual selection: The role of perceptual objects

Although the role of objects in constraining visual selection is indeed recognised, it is often seen as secondary to the role of space (e.g. Vecera, 1994). However, in order to spell out an alternative interpretation where space is not a factor, it is critical to define what an *object* is with regards to the perceptual system. As already mentioned, providing a clear account of what constitutes an object is a challenge, both for the current empirical work, and for visual cognition and perception research in general. This project argues that this challenge is also one of the reasons for inconsistent interpretations and conflicting results in visual selection research regarding the proposed coexistence of space-based and object-based effects. Defining what an object is may seem trivial at first, as it is a notion that is very familiar, but it can in fact be very complicated and subjective, depending on a range of phenomenological factors (Feldman, 2003; Scholl, 2001). In other words, what is semantically treated as an object may not be the same as what the perceptual system treats as an object. That is, an object is not necessarily a concrete, physically defined and tangible entity (e.g. an apple), but it can be rather abstract and formed on the basis of the current behavioural and visual context – a perceptual object (Figure 1) (Feldman, 2003).

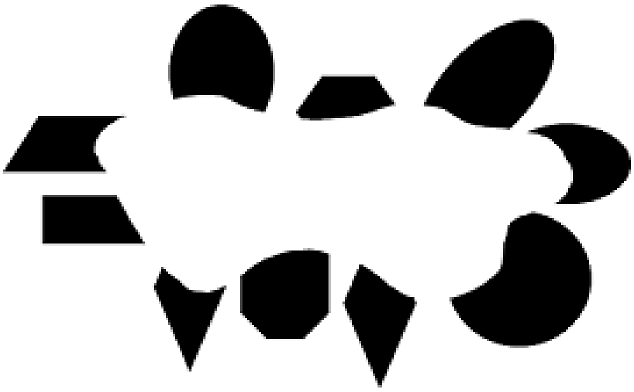

Figure 1: A perceptual object (reproduced from Feldman, 2003, Figure 1). The image induces a perception of a while 'object' situated on top of multiple black objects.

An object can be operationalised, based on principles of the Gestalt school of thought, as a unit which observes a number of ‘laws’ that contribute to a coherent perception of a ‘whole’ (Koffka, 1922; Wagemans, Elder, et al., 2012). Importantly, these laws relate to aspects of non-accidental regularities in the environment (Strother & Kubovy, 2006), such as closure and good continuation (Hess & Field, 1999; Marino & Scholl, 2005), similarity (Kubovy & van den Berg, 2008), proximity (Kubovy et al., 1998), connectedness (Han et al., 1999), common fate (Sekuler & Bennett, 2001), as well as general spatiotemporal continuity, i.e. consistency and integrity over time (Scholl, 2007; Scholl & Pylyshyn, 1999). Therefore, here again it becomes evident that it is the object structures in the natural environment, which have coherent and non-random structure, that have influenced what the perceptual system is likely to class as an object. How perceptions emerge is directly shaped by experience with the regularities in the environment, emphasising the link between natural scene statistics and the mechanisms of visual cognition.

Another important point relating to the formation of perceptual objects is that this formation is a probabilistic process. That is, depending on the presence or absence of certain regularity and coherence cues (cues to objecthood), a different perception may emerge (De Winter & Wagemans, 2006). This has been termed a “degree-of-objecthood measure” (Feldman, 2003). For example, the closer two or more dots are placed together, the more likely they are to be grouped as a perceptual object by virtue of proximity (Oyama, 1961), but one of them can be perceptually re-grouped with a more distant dot, if it is physically connected with it, i.e. the connectedness principle overrides proximity (Palmer & Rock, 1994). Also, the proximity principle can override grouping by similarity, e.g. with respect to colour or shape (Elder & Goldberg, 2002). Given the complexity of these principles and their interactions, studying how objects affect selection and perception in general can be challenging, but nonetheless equally interesting.

The effects of an object sometimes cannot be predicted until after it has been phenomenologically experienced, due to its emergence from the structure of the visual elements that eventually compose it (Feldman, 1999). Thus, an *emergent* object represents a perceptual entity which is formed by the experience of the sum of its parts (which may vary depending on the perceptual organisation and regularities of the scene), and it does not exist as a singular entity prior to this experience (Pomerantz, 2006; Wagemans, Feldman, et al., 2012). The illusory white shape depicted in Figure 1 is an example of an emergent object – it is perceived and exists by virtue of the regularities (in this case, good continuation) formed by the surrounding objects. If these objects were to be oriented in a different way, one which does not afford these regularities, the perception of the white shape would be lost. This phenomenon is frequently demonstrated with various illusory shapes, commonly known as Kanizsa figures (Kanizsa, 1976), corroborating the idea that the perceptual system is predisposed to seek and 'see' objects, and this predisposition is likely to be a result of the evolutionary importance of objects. The key challenge here is that it is sometimes difficult to know in advance how a set of stimuli would interact to give rise to a perceptual object, and its impact on visual selection can only be established a posteriori.

It is also worth mentioning that objects are not reserved only for vision, as the conditions for having a coherent unit expressing perceived regularity and structure can also be found in auditory perception, e.g. tones can be segregated into different objects based on frequency (De Freitas, Liverence, & Scholl, 2014; Turatto, Mazza, & Umiltà, 2005). Objects can also be formed in tactile perception, e.g. discrete vibrations (Gillmeister, Cantarella, Gheorghiu, & Adler, 2013). An object can be even more abstract, as it can constitute a temporal event which is treated by the perceptual system as a unit determined on the basis of semantic boundaries, such as a sequence of related actions (i.e. a behavioural unit) or movie clips (Newston & Engquist, 1976; Swallow, Zacks, & Abrams, 2009). Further complexity is added by the fact that an object can be a dynamic unit, e.g. if it is defined by common motion of similar elements (e.g. a field of dots), it can change or split into multiple objects if some of the elements change their speed or direction (Festman & Braun, 2010; Wegener, Galashan, Aurich, & Kreiter, 2014). The focus of the current work is on visual objects, but recognising that the same principles hold for other perceptual domains emphasises that object-based perception is a supra-modal phenomenon, and thus likely to be the default way of experiencing the world.

Given the importance of objects, they may have a strong influence on visual selection and can determine which part of the visual scene is selected and how attention is distributed. This object-oriented processing has certain benefits over space-based selection, for example, a spatial gradient or a spotlight model does not take into account what the visual space is filled with, while an object-based account typically predicts that selection can be limited by or ‘spread along’ the body of an object (Chen, 2012; Egly et al., 1994). The advantage of an object-oriented account becomes obvious in the case of overlapping objects. Namely, a pure spatial account would suggest that all information within the zone of the spatial beam is selected, while an object-oriented perspective would suggest that a single object can be prioritised, even if the same spatial area is occupied by parts of a different object (Baldauf & Desimone, 2014; Brawn & Snowden, 2000; Cohen & Tong, 2013; Mitchell, Stoner, & Reynolds, 2004; O’Craven, Downing, & Kanwisher, 1999; Simons & Chabris, 1999).

One implication from object-based selection is that when a single feature or part of an object is selected, other parts of the same object will automatically be perceptually enhanced and gain privileged processing by virtue of belonging to the same object (Kahneman & Henik, 1981). Also in terms of executing eye movements, given the same distance between the current locus of gaze and two potential visual targets, a saccade towards a target within the currently fixated object is much more likely (Emberson & Amso, 2012; Theeuwes, Mathôt, & Kingstone, 2010). Since an object can be described as a (subjectively) coherent unit, it is also selected as such – all information associated with it is prioritised relative to other objects on the scene which are not currently relevant. This within-object enhancement and prioritisation of information, whether it is as a result of exogenous cueing or voluntary selection of the object, can be termed an *object-based benefit* when it leads to improved performance. However, it should be noted that object-based selection can also result in performance deterioration due to, for example, increased distraction from response-incongruent elements within the selected object (Banks & Prinzmetal, 1976).

This object-based benefit can be demonstrated in a number of ways. For example, superior performance for processing a set amount of information when it is contained within a single object, compared to when it is distributed between more than one object (Duncan & Nimmo-Smith, 1996). This is illustrated with the divided attention paradigm, which typically requires making two discriminations along different dimensions (e.g. orientation – left or right, and size – large or small), while these dimensions belong either to the same object, or to two separate objects. For example, the two objects used by Duncan (1984) were a line varying in tilt orientation and texture, and an outlined box which could be small or large, with a gap on its left or right side. The two objects were presented superimposed (Figure 2) for a brief amount of time (50-100 ms), and participants had to make either a single judgement concerning one of these dimensions for one of the objects (e.g. only line texture or only box size), or a double judgement. The latter could either involve discriminating two dimensions from the same object (e.g. line orientation and line texture), or one dimension from each object (e.g. box gap position and line texture).

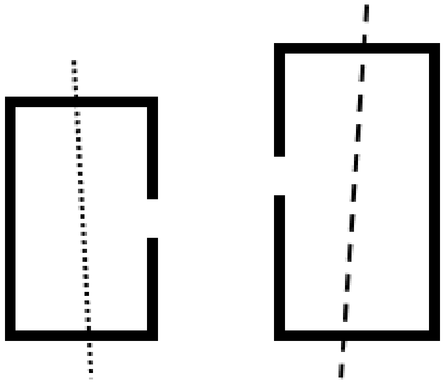

Figure 2: Reproduction of the stimuli used by Duncan (1984). Left: small box with a gap on the right and a dotted line tilted anticlockwise; right: large box with a gap on the left and a dashed line tilted clockwise.

Results indicated that judging two attributes (dimensions) from the same object was as efficient as judging a single attribute (so no cost for double judgement), but performance (proportion correct) dropped when the double judgement involved two attributes from different objects. In other words, a given object is selected in order to determine some property (e.g. its orientation), and this selection involves the facilitated processing of all attributes related to this object, thus not involving any more effort or ‘resources’ to determine its colour or texture, since its characteristics are integrated into a holistic perception. However, if the same amount of information is distributed within different objects, a drop in performance is observed. Furthermore, this impairment is not due to the number or similarity of the judged dimensions, but specifically to the number of objects that contain them (Duncan, 1993). The critical factor is whether the dimensions (e.g. size, shape, orientation) are contained within the same object. Therefore, as far as these examples are concerned, visual selection is based on discrete objects, rather than processing a specific location in space.

Critically, the divided attention task also provides evidence for spatially-invariant object-based selection. Specifically, a variation of the paradigm described above with an additional condition where the line and box were not superimposed, but each appeared 1.91° visual angle laterally from fixation, demonstrated that the magnitude of the object-based benefit did not vary with spatial separation (Vecera & Farah, 1994). More specifically, whether the objects were superimposed (so occupying roughly the same region in space) or spatially separated, made no difference to the cost associated with judging two attributes of separate objects, compared to attributes of a single object. If space-based selection was taking place, then this object-based benefit would be larger in the spatially separated condition, since under these circumstances a larger spatial area needs to be selected, or attention needs to move across space. Since this was not the case, the results indicate that visual selection was not mediated by the spatial distribution of the objects, but was uniquely based on object representations. Therefore, changing spatial separation makes no difference for selection, which is reminiscent of the characteristics of natural scene statistics described earlier (e.g. Ruderman, 1997), and it also poses a challenge for a space-based model of selection.

The space-invariant effect in this type of paradigm has been successfully replicated (Awh et al., 2001; Kramer, Weber, & Watson, 1997; Matsukura & Vecera, 2011), although a follow-up task within the same study questioned the extent of pure object-based selection (Vecera & Farah, 1994). Cueing one of the objects by briefly highlighting it, followed by the presentation of a dot on the surface of one of the objects revealed faster reaction times when the dot appeared on the cued object in the spatially separated, but not in the superimposed condition. The conclusion derived by Vecera & Farah (1994) is that space-based and object-based modes of selection coexist. Which one is employed depends largely on the task demands, such that a task involving decisions about object properties necessitates object-level coding and representation, while simply responding to a visual event (e.g. a dot onset) requires ‘low-level’ location-mediated selection and no need to have a full object representation. In this case it is reasoned that the lack of cueing benefit in the superimposed condition is due to a spotlight-like spatial selection, which encompasses both objects (Vecera & Farah, 1994).

There is, however, an alternative explanation for this finding, which does not resolve to space as a factor. When the objects are superimposed, this may significantly increase the likelihood that they are perceived as a single object – a rectangle with a line. This is especially likely since in the cued version of the task the line object was solid, as opposed to dotted or dashed. Therefore, the two objects were perceptually more similar than in the divided attention task, making it even more probable to be perceived as one when superimposed. The logical consequence then is the lack of cueing effect in that condition, since whether the dot appeared on the line or the rectangle makes no difference to the perceptual system – they are the same object. This possibility also emphasises the challenge outlined earlier: establishing what an object is and dealing with discrepancies between what the perceptual system treats as one object, and what the experimental design assigns as an object.

In the case above, a space-based account of the results can be well accommodated with an alternative, fully object-oriented perspective. However, later research based on a similar task again puts weight on the possibility for space-based selection. Specifically, the same methodology of making attribute judgements for separated versus superimposed objects was used, but on a small proportion of trials there was a secondary task where participants had to detect the onset of a red dot, which appeared after the offset of the objects (Kramer et al., 1997). In the separated condition when both attributes were contained within the same object, e.g. the orientation and texture of the line had to be reported, reaction times were faster when the dot occupied the location of the object that contained both target dimensions (in this example, the line object), compared to when it appeared at the location of the irrelevant object (the box). In other words, there was a location-based facilitation for responding to the dot when it occupied the region where attention was previously focused. Kramer et al. (1997) suggested that selection may be object-based, but it is not fully independent from space, as it is the location of the object(s) that is selected. If selection was confined to the object representation only, then responding to the dot was expected to produce similar reaction times, regardless of its location relative to the objects.

Although the results from the dot post-detection task may appear convincing for the case of space-based selection, if the perceptual system is biased towards immediately subsequent visual events at the location of the behaviourally relevant object, then this can be equally classed as an object-based effect. Furthermore, such events are more likely to be associated with, or be part of this object, and thus be behaviourally relevant as well. The dot appeared immediately after the offset of the objects, so from a neural coding perspective residual activity can still result in response facilitation, even more so since eye movements were not explicitly restricted during the trial (although unlikely to be executed given the short timings) (Moore & Fallah, 2003). Although this may still seem as reinforcing a space-based explanation, as it refers to spatial neural maps and location-based coordinates, it is not necessarily so. Enhanced neural activity which is spatially specific relative to a previously selected object does not necessarily imply space-based selection per se, as it is itself the consequence of object selection.

As the review so far suggests, the interpretations of research evidence generally favour a space-based account for visual selection, even if the object-based benefit is recognised. It appears that the perception of space as a primary dimension is accepted to be a starting or even a default state, which may lead to the dismissal or incomplete exploration of alternative, object-oriented explanations for the results. An example can be observed in another version of the divided attention task, one which requires the comparison of two simultaneously presented targets as ‘same’ or ‘different’, and these targets are properties of one or two objects. Specifically, the targets represent structural changes (appearance of a gap or a dot) along either two dashed lines (red and green) intersected at their midpoint (one horizontally oriented, and the other diagonally), or the target changes are situated along just one of these lines (Figure 3) (Lavie & Driver, 1996). Critically for this task, the spatial separation between the two targets was roughly equal whether they appeared on a single line (named by Lavie & Driver as *object* condition, Figure 3a), or along the two different lines (referred to as *far* condition, Figure 3b). When each target was contained within a different line, the targets could also appear in close proximity, i.e. each target was on the same side of the line it belonged to (*near* condition, Figure 3c).

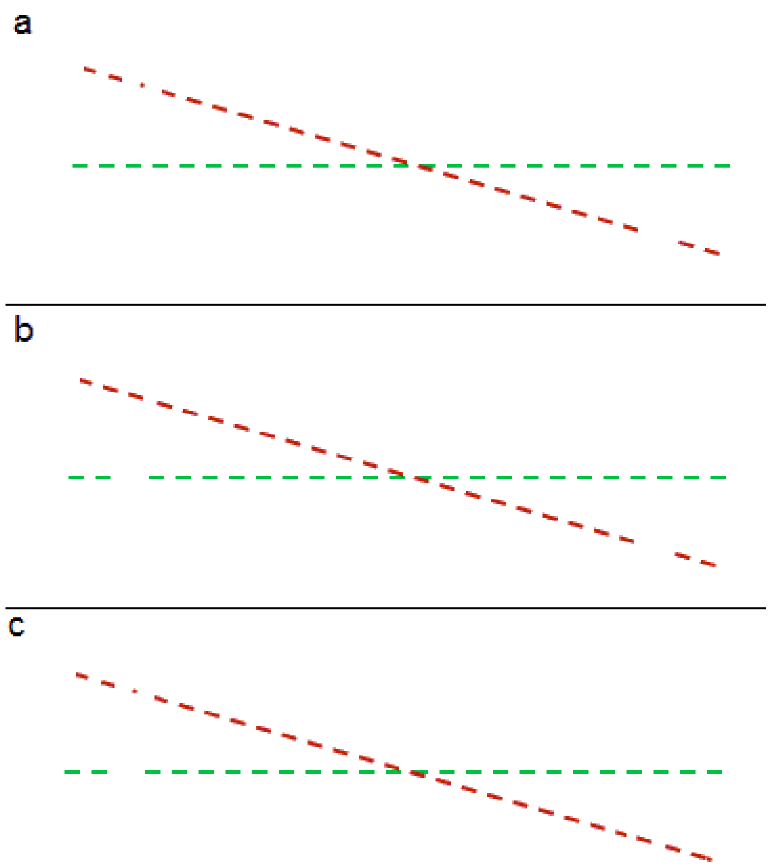

Figure 3: Reproduction of the object and target stimuli in Lavie & Driver (1996); a: *object* condition, different targets; b: *far* condition, same targets; c: *near* condition, different targets.

Despite the equated spatial distance between targets, performance (accuracy and reaction time for classifying the structural changes as ‘same’ or ‘different’) was better when the targets were part of the same line (e.g. Figure 3a) compared to two different lines, suggesting an object-based benefit. In terms of *near* versus *far* conditions results were inconclusive, as no reliable benefit for close compared to distant targets was found. However, a different pattern emerged when one side of the display was cued prior to target occurrence by highlighting the ends of the two lines. In this case, a space-based effect was observed, such that performance was superior when both targets appeared on the cued side (i.e. the *near* condition on the side of the cue), and there was no longer an object-based advantage. The proposed implication is that object-based effects occur when there is a diffused attentional focus, but if attention is constrained by cueing a narrow part of the display, performance is affected by spatial factors (i.e. the distance between targets). However, a later attempt to replicate Lavie & Driver's (1996) results was unsuccessful, as a series of experiments by Lamy (2000) and Law & Abrams (2002) demonstrated object-based selection effects with the same type of stimuli under conditions of both diffused (no cueing) and focused attention (using a cue prior to target presentation). It is likely that the lack of object-based effect in the cued version of the task was due to the nature of the cueing – highlighting the ends of both objects – which results in essentially selecting both objects to an equal extent (Law & Abrams, 2002). The observed space-based benefit is then due to top-down expectancy as the cue predicted the most likely location of the targets, while the lack of a within-object advantage is a consequence of object-based selection because both objects were activated by the cue.

Although introducing cueing seems to generate conflicting results within the divided attention paradigm, in its standard 'uncued' format it provides consistent support for object-oriented selection. For example, the fact that Lavie & Driver (1996) did not find distance effects (i.e. no consistent difference for *near* versus *far* condition) may be due to the fact that in both of these conditions the targets belonged to two different lines, which can be interpreted in terms of a space-invariant object-based selection. In support of this possibility, comparing the reaction time and accuracy data between the *near* and *object* conditions reveals an interesting trend: the *object* condition resulted in an overall better performance even though, in spatial terms, the targets were further apart. Unfortunately, Lavie & Driver (1996) did not report statistical analyses for this comparison. Nevertheless, this trend suggests that the object-based effect can hold true even when it conflicts with spatial proximity.

A possibility is raised, however, that the object-based effect demonstrated in the divided attention task which uses lines as objects is not genuinely due to object-level factors, but it is simply the result of co-linearity cues (Crundall, Cole, & Galpin, 2007). This is potentially due to the fact that within-object targets are always situated on a straight line across each other, while when each target is on a different object (whether it is the *near* or *far* condition), there is no physically visible straight line connecting the two. An alternative design of the study aimed to remove co-linearity as a confounding factor, while also using dashed lines to form objects (Crundall et al., 2007). In this case, the lines were bent in the middle in order to create conditions where co-linearity and object belongingness of the targets can be systematically manipulated. Therefore, there was a condition where different-object targets are collinear, while same-object targets are not (i.e. still belonging to the same line object, but separated by an angle). Performance benefited when the targets were collinear, regardless of whether they were part of the same object, which may question whether the previously observed object-based benefit in this context (e.g. Lavie & Driver, 1996; Lamy, 2000) was really a result of object selection.

It should be noted, however, that co-linearity is in fact a characteristic of the contours of objects in natural scenes (Sigman, Cecchi, Gilbert, & Magnasco, 2001), and as a result it is also considered a natural cue to objecthood according to Gestalt principles (Wagemans, Elder, et al., 2012). As such, it is very likely to influence selection because parts of the same object are likely to be collinear, even if not continuous in space. This may be especially true when the target objects are dashed line contours as in Lavie & Driver (1996) and Crundall et al. (2007), since the integrity of (especially) a dashed line is critically dependent on the co-linearity of the dash segments, even if colour is introduced as an additional cue to object differentiation. Therefore, evidence for co-linearity effects is not a valid argument against object-oriented selection, but in fact one that supports the critical importance of objects and highlights the challenge of defining a perceptual object.

Additional support for the strength of the same-object advantage over and above co-linearity cues in divided attention tasks comes from displays where the two objects are made more distinguishable, i.e. instead of simple outlines or lines, more salient and solid objects are utilised. Object-based effects are observed when two solidly outlined rectangles are overlapped and intersected in the middle so they form an ‘X’ shape (Figure 4a), or when the contours at the intersection are rearranged to form the perception of two oppositely oriented ‘V’ shapes overlapping at the apex (Figure 4b) (Behrmann, Zemel, & Mozer, 1998). The targets in this task are changes in the shape of two of the rectangle ends (e.g.

changing from a straight line into two or three curved ‘bumps’), which may belong either to one and the same object, or to two different objects. Participants have to decide whether both ends changed into the same shape or a different shape.

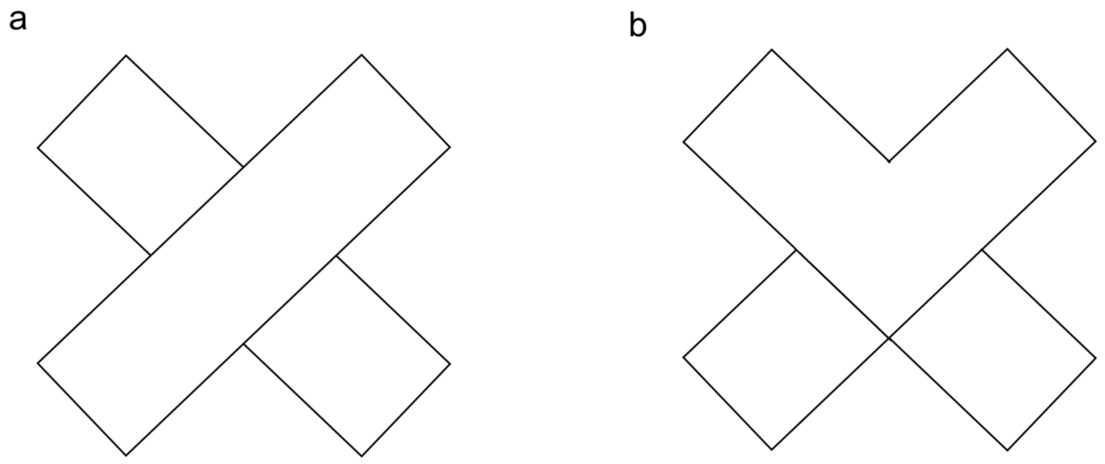

Figure 4: Reproduction of the stimuli used in Behrmann et al. (1998); a: overlapping rectangle objects into an 'X' shape; b: overlapping V-shaped objects (targets not depicted).

In the ‘X’ shape version of the task, the bottom rectangle is occluded by the top one, so targets at its two ends are not visibly connected. Nevertheless, performance is superior for discriminating within-object targets compared to targets at the ends of two different rectangles. It should also be noted that in this case ends belonging to different rectangles are more proximal than the ends of the same rectangle. In addition, the object-based effect is of equal magnitude regardless of whether the targets are integrated within the top (unoccluded) object or the bottom object (when the within-rectangle ends are separated by an occlusion). More importantly, the object-based effect is replicated even when the objects form overlapping ‘V’ shapes, i.e. same-object targets are always situated at an angle from each other, while different-object targets are collinear with respect to the body of the objects.

In addition, in the ‘X’ layout of the objects, when the two ends of the partially occluded rectangle are displaced so that they no longer face each other directly (i.e. co-linearity and symmetry are removed as illustrated on Figure 5a), the same-object advantage for the occluded object is preserved if the displacement is gradually introduced to participants via apparent motion (Behrmann, Zemel, & Mozer, 2000). This advantage is also preserved if a preceding block of trials introduced a novel object shape corresponding to what the displaced rectangle ends would look like if they were part of a single, non-occluded object (i.e. resembling a ‘Z’ shape similar to Figure 5b) (Zemel, Behrmann, Mozer, & Bavelier, 2002). However, in a static presentation and without previous exposure to the implied occluded object, the two displaced rectangle ends are treated as separate objects, so no advantage is observed when targets are integrated within these parts. Given the persistence of the object-based effect in the face of these manipulations, it can be concluded that it is a very robust and adaptive phenomenon, i.e. affected by learning from previous perceptual experiences. Importantly, co-linearity alone or space-based explanations cannot accommodate these results, as performance is modulated by emergent object perception.

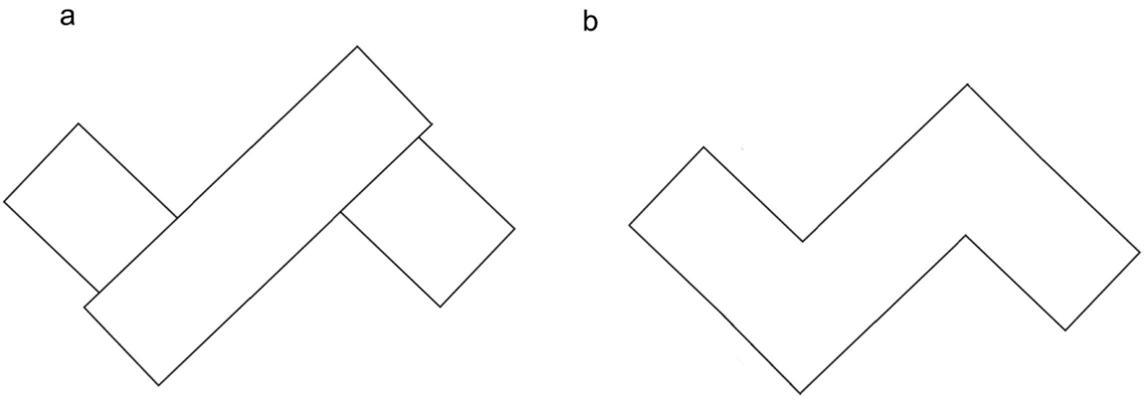

Figure 5: Reproduction of the stimuli used in Zemel et al. (2002): (a) displaced ends of an occluded rectangle and (b) the corresponding integral object implied under the occlusion.

Indeed, objects can have a powerful effect on how information is selected, as the same visual display can produce different selection patterns based on the implied perceptual organisation of the scene. For example, a variant of the divided attention task requires comparing the texture of two out of four lines as ’same’ or ‘different’, while those lines are either perceived as the outlines of a partially occluded diamond shape (so part of the same object, but not collinear), or as independent, unconnected lines (Figure 6) (Naber, Carlson, Verstraten, &

Einhauser, 2011). In both cases the spatial and visual arrangement of all lines is identical. However, prior to two of the lines changing their textures from solid to dotted and/ or dashed (i.e. target appearance), in the single-object condition the four lines move together in the same direction behind the vertical occluders, creating the perception of a diamond shape. In the alternative condition, all four lines move independently from each other, encouraging the perception of separate, unrelated objects. Discrimination performance was better when the lines were perceived as parts of a bound object than when they were perceived as independent entities, even though the visual display at the time of target onset was identical in both cases. Therefore, performance was affected by the emergent object on the scene, which was differently constructed with the same amount of visual information, depending on how the cues to object formation were manipulated.

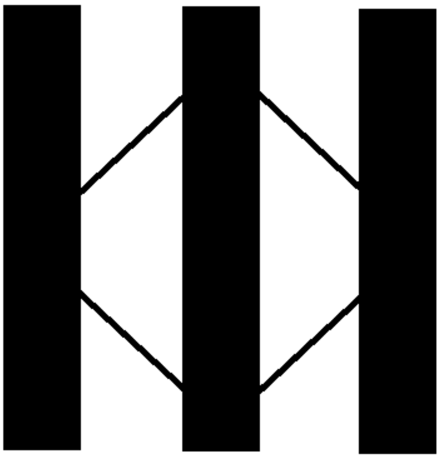

Figure 6: Reproduction of the stimuli used in Naber et al. (2011): a perceived diamond shape. Prior to target appearance the four thin lines move either as an ensemble to reinforce a coherent perception of an occluded object, or in different directions to create the perception of unconnected lines.

In addition, object-oriented benefits for divided attention are found in cases when the objects constitute orientation consistent surfaces (textons) that observe basic object characteristics such as surface coherence and similarity, but do not necessarily resemble a familiar or identifiable object (Ben-Shahar, Scholl, & Zucker, 2007). In this case a benefit is found for comparing two targets that appear on the surface of a single continuous texture, compared to when the targets are at the same distance, but on different textures. The examples provided above demonstrate the perceptual system is predisposed to perceive objects, and these objects have a potent effect on biasing visual selection which cannot be easily accounted for by a space-based perspective.

Although the discussion so far was focused on the manifestation of the object-based benefit under various divided attention settings, another consequence of object-based selection is the difficulty to selectively focus on a subset of information from an object while ignoring other visual information within this object. In other words, the ability to easily incorporate and process all of the features of a selected object, as demonstrated so far, can also have a negative consequence leading to distraction and interference. Such negative effects are typically illustrated with the flanker paradigm, which requires the identification of a target surrounded with distractors. As mentioned earlier, Eriksen & Eriksen (1974) demonstrated spatial separation effects of decreasing response interference as the spacing between the central target letter and incompatible distractor flankers increased, which is taken as evidence for space-based selection.

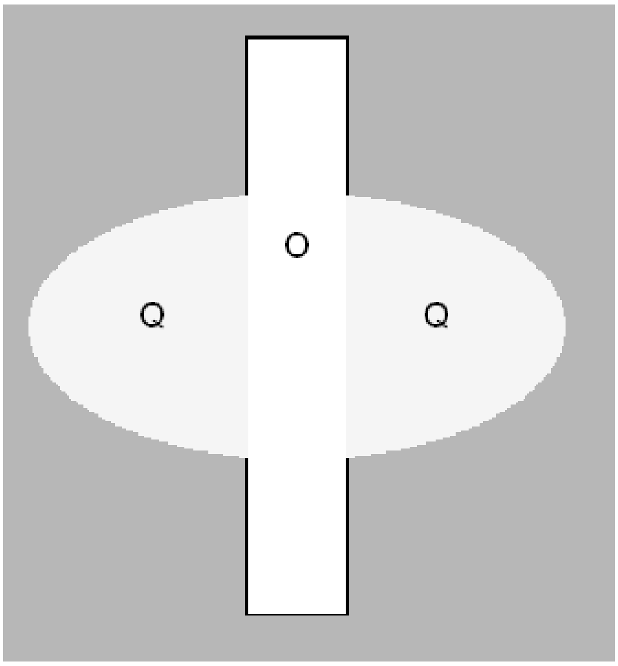

Figure 7: Reproduction of the stimuli used in Davis & Driver (1997). A central target letter "O" and two lateral response incongruent distractor letters "Q", all situated on the illusory surface of an ellipse.

However, effects of perceptual organisation relating to object-level factors can also be observed when target-flanker distance is held constant. For instance, if the central target and two flanking letters are situated within the outlines of the same rectangle (object), interference from incompatible distractors is stronger than if they are situated outside the rectangle containing the target letter (Chen & Cave, 2008; 2006; Cosman & Vecera, 2012; Ho, 2011). As with divided attention tasks, phenomenally completed objects also influence selection in a comparable way to intact objects. Specifically, interference from distractors is stronger when they were perceived to be on the same surface as the central target, even if this surface is a product of illusory contours (Figure 7), i.e. it is an emergent property of the visual display accomplished via modal completion (Davis & Driver, 1997). Therefore, object-based selection is also manifested as within-object distraction and interference.

Although the object-based distraction is well established, it has been suggested that it is not a very robust effect, as it is not manifested under all circumstances. For example, it has been proposed that object-based interference in flanker tasks is only observed if the position of the target is not known in advance, while under positional certainty there is no automatic selection of the entire object surface (Shomstein, 2012; Shomstein & Yantis, 2002). This is reminiscent of the focused versus diffused spread of attention proposed earlier by Lavie & Driver (1996). However, there is evidence to suggest that when the target and distractors are well embedded into the object structure, object-based selection (in this case, evidenced by interference) is robust even under conditions where the target and distractor locations are known with a 100% certainty (Richard, Lee, & Vecera, 2008; Zhao, Kong, & Wang, 2013). Specifically, an important factor to elicit these effects under high predictability of target location is that the targets represent structural changes in the objects (i.e. appearance of ‘bites’ or chips on the surface), as opposed to superimposed letters. This emphasises once again the influence of perceptual organisation, and suggests an important role for good target-object integration. If the influence of an object on selection is to be tested by measuring reactivity to some visual target, then it must be ensured that this target is indeed perceived to be a part of the relevant object. Otherwise the measured performance could reflect the processing of a separate object (the target as an independent entity), rather than the object intended by the experimental design.

In addition to integrating the target and distractors into the bodies of objects, comparable flanker effects can also be achieved on the basis of common characteristics of the target and distractors (e.g. colour). As mentioned earlier, similarity is a natural cue to objecthood and so is likely to increase the probability that the stimuli sharing similar characteristics belong to the same object. For example, in addition to varying target-flanker distance (as in Eriksen & Eriksen, 1974), the colour similarity of the flanker and target letters can also be manipulated. For the same spatial arrangement, this results in larger interference when the target shares the colour of the distractors (Baylis & Driver, 1992). More importantly, in an array of laterally distributed distractor letters, those which were far from the target but had the same colour produced a larger interference effect than near distractors of a different colour. In other words, common characteristics (in this case, colour) between the target and distractor can override the effect of spatial proximity. This reversed proximity effect is also evident when the distractors and target move in the same direction, i.e. they display common motion (Driver & Baylis, 1989). This evidence illustrates the power of perceptual grouping, but also poses challenges for space-based models of visual selection and claims that location *per se* is a special attribute (Tsal & Lamy, 2000). Again, the critical argument here is that such common attributes contribute to the perception of the target and distractor as a single unit, even if it is discontinuous in space.

How can such results be reconciled with the evidence for space-based interference inferred from the classic effect of decreased distraction when flankers are far from the target (Eriksen & Eriksen, 1974)? Considering again the data from natural scene statistics, together with the effect of target-distractor similarity, it is possible to interpret the effects of spatial separation as resulting from degree of objecthood variability, rather than space-based selection. In other words, increasing the distance between a target and flankers results in decreasing the probability they belong to the same object, and thus less interference for responding to the target. However, when there are additional cues that can influence this probability, such as similarity, then spatial separation is not as critical. Importantly, all of these factors – proximity, similarity, or physical integration – exert an influence on selection because they are related to within-object characteristics in the environment. Therefore, effects of spatial separation should not be necessarily taken as evidence for space-based selection, as they can equally be evidence for object-based selection.

Given this reasoning, it also follows that different cues to objecthood can interact to influence selection, which can explain that in some cases effects of spatial separation are present and in other cases they are not. For example, discriminating the texture (dashed or dotted) of a central line, which varies in distance and strength of perceptual integration (by colour similarity and connectedness) from two lateral flanker lines, results in stronger interference when the stimuli are perceptually similar for the same distance (Kramer & Jacobson, 1991). However, even when target and distractors shared colour or were physically connected, interference was stronger when the distractors were closer to the target, suggesting a combination of proximity, similarity, and integrity effects, which is interpreted by Kramer & Jacobson as combined space-based and object-based selection. Nevertheless, Fox (1998) conducted a similar flanker study for letter discrimination, where the central target letter and lateral distractor letters were each centred in a circle outline. Target-flanker distance and whether the target and the two flankers were connected by a horizontal line (i.e. forming an object) was systematically varied. There was no effect of target-distractor distance, as long as the stimuli were all connected with horizontal lines to appear as the same object. Therefore, spatial separation is not always a critical factor, although it is not entirely clear how exactly it interacts with other perceptual cues to objecthood. However, based on the evidence reviewed so far, perceptual organisation pertaining to typical characteristics of natural objects has a key influence on visual selection, suggesting that space-based effects may not be truly *space*-based, but emerging from the perceptual organisation of the scene.

There is also evidence to suggest that spatial separation effects may be modulated by the object structure of the stimuli. For example, in a flanker task, Eriksen, Pan, & Botella (1993) required participants to discriminate the orientation of an oblique line situated within a rectangular object. The shape of the edges of the rectangle were used to indicate whether this was a *go* or *no-go* trial, in order to encourage participants to spread their attention within the full area of the object. The distractor flankers (lines of either the same or different orientation as the target) were located outside this object on either side. The length of the object and the position of the target line inside were also varied. Interestingly, the results indicated that the distractor interference varied as a function of the distance between the edge of the rectangular object and the flankers. Interference level was independent of the spatial separation between the actual target line (situated inside the object) and the distractors. Eriksen et al. (1993) interpret the results in terms of a zoom lens model, whereby the size of the 'attentional beam' is adjusted according to the attended area due to the need to attend to the rectangle edges in addition to the internal target line. However, this effect can be interpreted in object-based terms, such that the whole rectangle is selected regardless of its length, and the target line situated inside is a feature of this object. Consequently, the proximity of the distractors is correlated with their probability to be part of the rectangle, thus the space-based effect emerges. However, varying the location of the target line within the object itself makes no difference to performance, since it is always part of the critical object. This is interesting evidence implicating the dependence of the so called space-based effects on perceptual organisation factors, which in turn suggests that space-based effects are not primary or special.

It is worth mentioning that spatial separation effects can be modulated by other object-level perceptual organisation factors, e.g. similarity between stimuli, also in phenomena such as visual crowding. This is an event where the identification of visual items in the periphery is substantially impaired when they are flanked by other stimuli (i.e. when visual clutter is introduced), compared to when they are presented at the same spatial location, but on their own (see Whitney & Levi, 2011 for a review). Importantly, the level of interference due to crowding is typically regarded as dependent on spatial factors, affected by the eccentricity of the target and the spacing between the target and the flankers, as well as whether the flanker is on the inner (closer to fixation) or outer side of the target (Bouma, 1970; Pelli & Tillman, 2008). However, there is substantial evidence that if the spatial separation is unchanged while distractors are perceptually segregated from the target, i.e. perceived as separate objects, crowding is significantly decreased and the previously indistinguishable items become identifiable again (Levi & Carney, 2009; Livne & Sagi, 2007; Saarela, Sayim, Westheimer, & Herzog, 2009). This can be achieved by making the target distinct from the surrounding flankers, for example by grouping the distractors into a symmetric shape, e.g. a circle that the target does not form a part of, or changing the shape of the target relative to the flankers to make it more distinct. These manipulations demonstrate that varying the perceptual organisation of a set of stimuli is sufficient to substantially alter the visual experience of the observer.

So far the discussion has emphasised the fact that although visual selection was originally thought to be space-based, as evidenced by spatial separation effects and automatic processing of stimulus location, there is considerable evidence that in fact selection, and visual cognition in general, is heavily influenced by perceptual organisation and the subjective experience of objects in the environment. Moreover, the spatial separation effects (e.g. graded facilitation or inhibition) can also be interpreted as evidence for object-based selection, as proximity is a cue to objecthood (Sigman et al., 2001). However, a strong case for the coexistence of space-based and object-based selection, as well as the hypothesised superiority of space-based selection, comes from spatial cueing paradigms (Posner, 1980). As mentioned earlier, introducing a cueing element in divided attention tasks typically results in emphasising a role for space-based selection (e.g. Lavie & Driver, 1997). In a cueing paradigm, a visual cue indicates a designated region in space, which may subsequently contain the task-relevant target. Therefore, especially in the case of exogenous cueing where the cue is spatially congruent with the potential location of the future target, attention is supposedly confined to the cued location where visual processing is immediately enhanced, resulting in a 'space-based' cueing effect. In other words, this is a methodology which lends itself well to the study of the 'spatial' element of visual selection and affords adjustments in order to assess the influence of perceptual objects.

The role of objects within the context of spatial cueing is typically studied with different variations of the so called *two rectangle paradigm* (Egly et al., 1994). In its original format, this technique involves the presentation of two parallel rectangle outlines, either horizontally or vertically oriented, centred at equal distance from fixation (4.8° in the original study). A brief peripheral cue consisting of brightening one end of one of the rectangles predicts the most likely target location (Figure 8a). A target (e.g. a square or a dot) then appears inside one of the rectangle ends, and the participant has to respond to its onset (detection task) or identify it (discrimination task). The critical aspect of the paradigm is that two of the uncued rectangle ends are equidistant from the cued location, but differ based on whether they belong to the cued rectangle (within-object position) or not (Figure 8b). The typical finding under these circumstances is that in addition to the standard cueing effect, i.e. best performance at the cued location (indicated by number 1 on Figure 8b), detection and discrimination at uncued locations is better for targets within the cued rectangle (number 2 on Figure 8b). Therefore, similarly to flanker tasks, when distance is held constant, object-based selection can be observed, leading to privileged processing of visual information within the cued object structure.

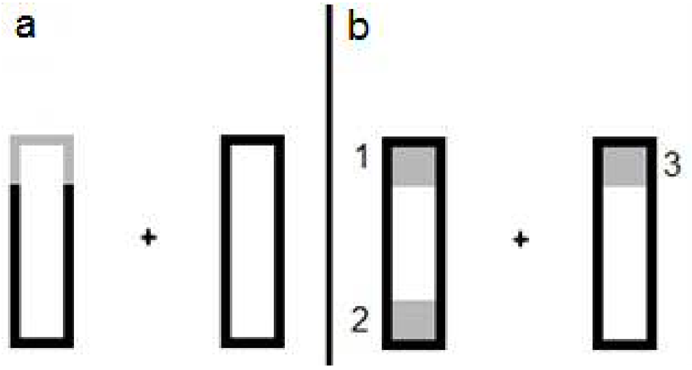

Figure 8: Two rectangle paradigm illustration (Egly et al., 1994); a: cueing by highlighting one end of a rectangle; b: potential target locations illustrated by grey squares, adjacent numbers reflect the type of target: 1 = cued, 2 = uncued same-object, 3 = uncued different-object (only one target is presented per trial).

The object-based advantage under such spatial cueing circumstances is a robust phenomenon, and the objects need not be solid outlines, but can be perceptually completed by illusory contours or an occluding shape superimposed on top of the rectangles (Moore, Yantis, & Vaughan, 1998; Pratt & Sekuler, 2001). Interestingly, the same-object advantage can be observed even when the target appears on a location within the cued object which was occluded during the time of cueing by a diagonally superimposed rectangle covering the uncued corner of the cued object and the uncued-different object corner (for example on Figure 8b, the corners labelled 2 and 3 would be occluded at the time of cueing, prior to target appearance) (Moore & Fulton, 2005). In this version of the task, after cue offset the occluder is either displaced to reveal both rectangles in full, or remains in place. Consequently, the same-object advantage is preserved for targets appearing on the revealed surface of the cued object, but not when the occluder remains stationary and the target appears on top, even though the two-dimensional target location is identical in both cases. This is compelling evidence that it was the object per se which was selected, and not an absolute location on the display.

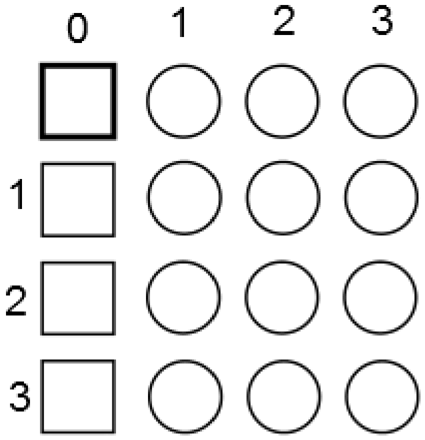

Figure 9: Reproduction of the stimuli used by Dodd & Pratt (2005). The thick outline indicates the cued location, numbers correspond to potential target locations with increasing cue-target distance as illustrated by numbers 0 (cued location) to 3. This layout results in privileged processing for targets within the square outlines compared to targets of equal distance within circle outlines.

The importance of object-level perceptual organisation factors for modulating the same-object advantage is crucial. For example, when instead of two rectangles the display consists of four squares equally distant from fixation (i.e. occupying the space that typically corresponds to the ends of the two rectangles) there is no difference in performance for targets at the locations equidistant from the cue, i.e.

no object-based effect since each square may be perceived as an independent object (Marrara & Moore, 2003). However, there is a critical point to be emphasised here. Specifically, the lack of performance variation in the case of four identical squares can in fact be interpreted as a same-object benefit, since these squares form the corners of a larger single square - a superordinate emergent object. Therefore, this is another example where a potential expression of an object-based effect is not recognised as such, as there is a possible discrepancy between an ‘object’ for the perceptual system, and an ‘object’ for the purpose of the experimental design. However, when perceptual organisation results in more ‘obvious’ objects, it is the case that object-based selection can be demonstrated also with objects formed on the basis of shape similarity, i.e. an array of squares is selected as an object when it is situated among a field of circles (Dodd & Pratt, 2005). In this context, targets (dot onsets) appearing within the perceptually integrated squares are prioritised compared to targets of equal distance from the cue, but occurring in a different superordinate object formed by the field of circles (Figure 9).

Object-based selection following cueing is also demonstrated at a neural level. Indicating the most likely target location (using the two rectangle layout) leads to enhanced retinotopic activity at the uncued end of the cued rectangle compared to the equidistant different-object location, and this is evident before target presentation as indexed by blood oxygenation level-dependent signal (BOLD) using fMRI (Müller & Kleinschmidt, 2003). This is interesting since the cue predicted the most likely target location and the two uncued locations (within or outside the cued object) had the same probability of containing the target. Therefore, in anticipation of the target, neural activity is automatically heightened in spatially distant locations if they are perceptually contained within the cued object.

This object-oriented modulation was replicated in a similar paradigm, but using ERP measurements for better temporal resolution, and also a cue that predicted the target location with a 100% certainty (Martínez et al., 2006). Therefore, object-based selection was evident even for parts of the scene which were task irrelevant, but perceptually integrated with the part of the object that required identification. In addition, enhanced neural activity is detected at the receptive field corresponding to a task irrelevant line object, only when it is perceptually grouped via co-linearity with a target (task-relevant) line (Wannig, Stanisor, & Roelfsema, 2011). When the two lines are orthogonal, i.e. they do not appear as continuations of each other, no enhancement is observed at the receptive field of the irrelevant line. These studies demonstrate that even at an early neural level there is already a bias to select perceptual objects. The nervous system itself is highly tuned to objecthood.

To summarise the discussion so far, research evidence is formally interpreted as suggesting a combination of space-based and object-based selection (Chen, 2012; Reppa, Schmidt, & Leek, 2012). However, the former is inferred primarily from graded spatial separation effects on performance, and such effects may equally be accounted for from an object-based perspective where spatial separation reflects a gradient of object belongingness likelihood. This view can be challenged by emphasising that any effect can be “explained away” with some type of cue to objecthood, or evoking the idea that there may always be a perceptual object not accounted for by the experiment. This is especially so since one of the challenges mentioned earlier is that an emergent object and its impact on selection can sometimes only be established post-factum.

A formal account which advocates the coexistence of space- and object-based selection while attempting to take this challenge into consideration is the so called *grouped array hypothesis* (Kramer et al., 1997; Vecera & Farah, 1994). This view suggests that object-based effects are the result of selecting an array of locations perceptually grouped by Gestalt cues to objecthood, and perceived object structure. Therefore, although the role of objects is recognised, the key implication is that selection is ultimately location-mediated and thus space-based. This accommodates well some of the findings described earlier, such as the spatial cueing effect or the combined effect of perceptual grouping and proximity on flanker interference (Kramer & Jacobson, 1991). However, at its core, this theory still interprets spatial proximity effects as evidence for space-based, or location-mediated selection. Nevertheless, if these space-based effects are viewed as resulting from object-level factors, then a more parsimonious account of visual selection can be established, namely one uniquely centred on object-based representations.

Demonstrations of the purported interaction between space-based and object-based selection, where the object-based effect within a cueing paradigm setting is interpreted as space-mediated, are evident from methodologies that vary the cue-target distance but (supposedly) keep the object structures constant. An example of this is a variation of the two rectangle paradigm with an additional ‘near’ condition where the two rectangles were placed closer together than in the standard equidistant version (Vecera, 1994). This condition results in a display where the uncued different-object target is closer to the cued location than the uncued same-object target (Figure 10). While a same-object advantage for target detection was evident in both the standard and near conditions, reaction time for the different-object target was slower when the two objects were further apart compared to when they were close together. Therefore, processing of the different-object target (labelled number 3 on Figure 10) was more efficient when it was closer to the cued location, i.e. there is a lower cost of switching attention to the uncued object due to its proximity. This interaction between cue-target separation and target type (with reference to the objects) it taken as evidence for location-mediated selection, i.e. selection of an array of locations grouped into an object, as opposed to being space-invariant and fully object-based.

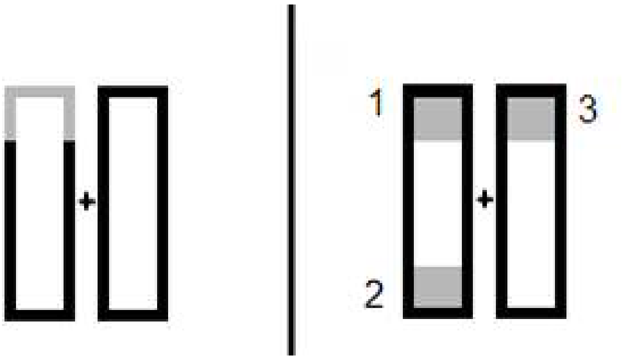

Figure 10: Reproduction of the stimuli used in Vecera (1994), 'near' condition. Left panel: cueing by highlighting one end of a rectangle; right panel: potential target locations illustrated by grey squares, adjacent numbers reflect the type of target: 1 = cued, 2 = uncued same-object, 3 = uncued different-object (only one target is presented per trial). To be compared with the standard equidistant condition illustrated in Figure 8.

However, here again there is a plausible explanation for the observed proximity effect, and this explanation relies on object-level factors and perceptual organisation of the scene. Simply put, when the two rectangles are brought closer together, they are more likely to be perceived as a single object, especially given their identical appearance. Therefore, performance improves for the different-object target not necessarily because of the decreased distance from the cued location, but rather because it is now more likely to be part of the cued object. It is just the case that this probability is cofounded with proximity. In addition, within the 'near' condition the object-based effect was preserved, so performance was still better for the uncued same-object target (labelled number 2 on Figure 10), even though it was more distant from the cued location compared to the uncued different-object target. This is so because the perceptual organisation of the display affords a higher probability for targets at that location to be part of the cued object, but on a hypothetical probabilistic continuum, targets at the different-object location (Figure 10, label 3) are more likely to also belong to the cued object than in the standard equidistant version of the display (refer back to Figure 8). Therefore, the observed effects may not be rooted in the locations of the targets per se, but are just as likely to depend on the perceived relationship of the target with the cued (selected) object.

Object-based interpretation can also be applied to evidence suggesting a spatial gradient of facilitation is observed *within* the body of the cued object. For example, Hollingworth et al., (2012) employed a single object task where the cue and target appeared within a three-dimensional ring shape centred at fixation (Figure 11). This allowed testing performance on a total of eight locations within the circular object, representing five different cue-target distances (Figure 11b). Participants had to identify a transient target, which could be either the letter 'X' or 'O'. Results indicated a smooth gradient of decreasing accuracy and increasing reaction times (for an onset detection version of the task) for up to 8.4° Euclidian cue–target distance (corresponding to two targets on either side of the cue). The fact that performance within the same object varied with cue-target spatial separation is taken to suggest that selection is space-based. This is because if selection was fully object-based, distance should not affect target processing as the target is always within the selected object, i.e. selection would be distance-invariant. However, the results can be explained from an object-oriented perspective. This can be done by proposing the hypothesis that performance varies as a function of the probability that a target is part of the cued object feature, and this probability can decrease with distance.

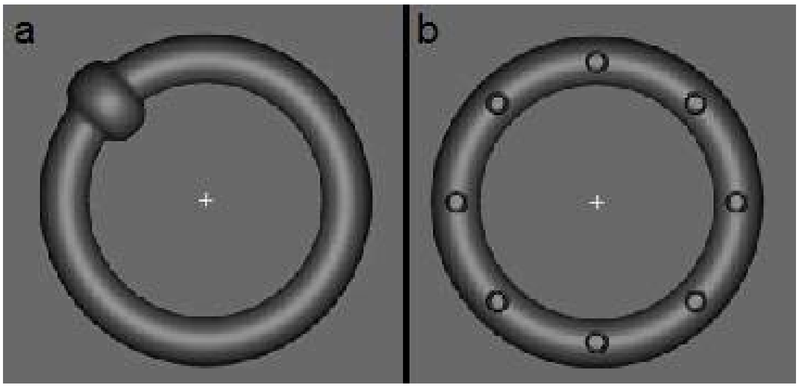

Figure 11: Representation of the stimuli used in Hollingworth et al. (2012), Experiment 2; a: a 'bulge' cue indicating the most likely target location; b: eight potential target locations illustrated by target "O" (only one of these targets is presented per trial).

A key factor here is likely to be the target-object integration. Hollingworth et al. (2012) used a salient three-dimensional object stimulus and two-dimensional superimposed letters as targets. Consequently, there is a high probability that these targets were perceived as independent objects, as opposed to being truly integrated within the cued object. In addition, evidence from flanker tasks suggests that object-based effects are of a larger magnitude when the abrupt onset targets represent structural changes in the objects as opposed to superimposed stimuli (Zhao et al., 2013).

Since the cue in Hollingworth et al. (2012) predicted the most likely target location, top-down control would lead to prioritisation for targets at that location. More importantly, given the potentially poor integration of the stimuli with the circular object, it may be the case that the observed gradient for uncued targets was a result of privileged processing of objects close to the cued location because they were more likely to be part of it. Conversely, if the targets were unambiguously parts of the circular object, rather than superimposed letters, then their distance from the cued location may be less likely to affect performance because they would be equally selected and prioritised. Thus, the observed spatial gradient may be due to probabilistic object-level factors, rather than space-based selection. Interestingly, in support of this possibility, there is evidence that detecting visual targets outside the body of a cued object results in a spatial gradient relative to the centre of mass of the object, rather than the cued location within the object (Kravitz & Behrmann, 2008).

The level of target-object integration and the perceptual organisation of the display can indeed help accommodate a range of evidence for combined space- and object-based selection into a single object-oriented account. For example, a pure object-based account would predict that cueing an object and then changing its location should result in enhanced processing of targets composing parts of this object even after is has changed location, while targets occurring at the cued location previously occupied by the object should not be prioritised. In other words, the cueing benefit for the object should stay with the object and not transfer to the cued location per se. However, there is evidence to suggest that both types of targets are processed to the same extent, thus supporting a combination of space-based and object-based selection (Theeuwes, Mathôt, & Grainger, 2013).

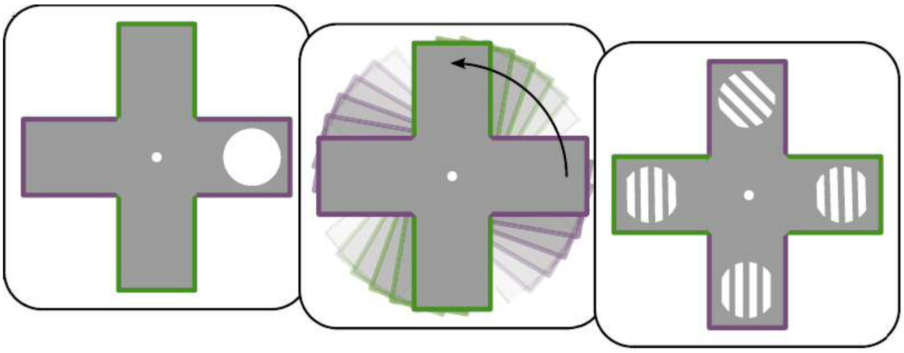

Figure 12: Stimuli used in Theeuwes et al. (2013). The illustration depicts part of the procedure (from left to right): cueing one end of the cross object, object rotation, and target appearance. In this example the target appears at the relative cued location with reference to the object after rotation.

Theeuwes et al. (2013) used a cross-shaped object, whose opposing arms had the same colour outline in order to create the perception of two separate crossed rectangles (Figure 12). Following the cue offset, the cross rotated 90° and participants had to report the orientation of an oblique Gabor patch appearing at the end of one of the cross arms, whereas the remaining three ends were occupied by vertically oriented distractors. Critically, there was a reaction time advantage for identifying the target when it was on the cued location on the same arm after it had changed position (as illustrated on Figure 12), and equally so when it was on the uncued arm end which occupied the recently cued (retinotopic) position. Thus, there were space-based and object-based effects of equal magnitude.

An overlooked issue in the suggested interpretation is that the perceptual structure of the stimuli can be ambiguous. For example, the superimposed targets may not be perceived as well integrated into this object. Consequently, this can affect the probabilistic selection mechanism by prioritising targets at the relative (object-centred) cued location since there is a chance they are indeed parts of the cued object, and also prioritising targets at the retinotopically cued location, reflecting the probability that the targets are independent from the body of the object(s) and thus the rotation is not necessarily relevant for the task. In addition, the methodology assumes that the display is perceived as two crossed rectangles, i.e. two separate objects. However, these objects are only distinguished by a thin colour outline, while the surface inside is of the same grey colour for both rectangles, and there is no visible intersection in the middle (Figure 12). Therefore, it is likely that there is a much stronger visual perception of a single cross object. This single-object perception is potentially reinforced even more after the synchronised motion in the same direction. As a result, whether the target appears on the absolute or relative location of the cue, it may still be within the same object, hence the equal magnitude of the effect. The structure of the stimuli does not allow any strong conclusions about the mechanisms of visual selection, as the perceptual organisation of the display, both in static and dynamic terms, does not afford a clear distinction between space-based and object-based effects.

Ultimately, explaining object-based effects within a space-oriented perspective, such as the grouped array hypothesis (Vecera, 1994), accounts for a lot of results where spatial proximity factors appear to play a key role, while within-object benefits are also observed. The assumption that location is a special and primary attribute of visual objects (Kubovy, 1981) has to a great extent shaped and directed research in the domain (Fernandez-Duque & Johnson, 1999). For example, even when discussed in the context of object-based effects, visual attention is frequently referred to as ‘spatial’ (Stigchel et al., 2009). Therefore, there is already a bias in the way data is interpreted, which leads to conclusions of space-based, or location-mediated object-based selection, even in cases when an alternative, fully object-oriented explanation is also possible. It should be noted that space in itself is often coded and perceived relative to objects, as it is nearly impossible to select or memorise a location without using some object structure as a reference point (Boduroglu & Shah, 2014; Humphreys, 1999; Mutluturk & Boduroglu, 2014). In other words, when remembering a location indicated by a transient visual cue, this location is coded relative to other visual objects available on the scene, i.e. there is a need for an abstract or concrete landmark. Therefore, a more parsimonious explanation for the mechanism of visual selection and cognition may be rooted in the perceptual organisation of visual information which leads to the emergence of space-based effects.

Considering all the evidence above, it is possible that space-based effects of visual selection can be accommodated within an object-based perspective, since spatial proximity is positively correlated with object belongingness. Importantly, however, this correlation only holds true when there is ambiguity regarding the object-level origin of the stimuli in question (e.g. target and cue). An important implication is that the target in the experimental designs needs to be well integrated within the objects that are used to study the selection mechanism, and also these objects need to be well defined for the perceptual system. Given a level of probability whether the target is an intrinsic part of the (cued) object, selection will vary as a function of this probability, which may give rise to spatial separation effects when this likelihood is low. One such example is the case when the level of object-based flanker interference is modulated by whether the flankers and targets represent structural changes in the object(s), as opposed to superimposed elements (Richard et al., 2008; Zhao et al., 2013). Similarly, space-invariant effects in divided attention tasks are found when judging attributes of the underlying objects, as opposed to reacting to sudden-onset, superimposed targets (Kramer et al., 1997; Vecera & Farah, 1994).

In addition, cueing paradigms typically use superimposed dots, shapes or letters on top of the objects, which also can lead to the perceptual segregation of these targets from the objects. In the case of equal cue-target distance for uncued same-object and different-object targets this is unlikely to be a problem, since the belongingness probability correlated with spatial proximity is equal for both type of targets. At the same time, the uncued target situated within the boundaries of the cued object would gain an additional advantage as its probability of being part of this cued object is higher, hence the consistent replications of object-based effects in the standard two rectangle paradigm. However, if cue-target distance is varied either by changing the proximity between the objects (e.g. Vecera, 1994), or introducing various target distances within the same object (e.g. Hollingworth et al., 2012), then target-object integration and more global changes in the perceptual organisation of the visual scene may be critical. Under such circumstances, when the targets are ambiguously integrated with the underlying objects, selection would be additionally affected by differences in proximity giving more weighting to targets close to the cued location/ object feature.

Empirical outline and aims

To this end, the present project aimed to address the potential limitations and confounding factors that may have led to interpreting object-based selection as space-based, and to explore the role of perceptual organisation in the emergence of space-based effects. The project is organised in three empirical chapters which address these issues from a different angle.

Chapter 2 dealt with the challenge to disentangle effects that can genuinely be attributed to a mechanism selecting spatial locations from one selecting information in an object-oriented fashion, given that the probability of object belongingness is correlated with spatial separation (Ruderman, 1997). This was accomplished by using a cueing paradigm and overlapping object stimuli that introduce a conflict between object-oriented and space-oriented selection, as well as targets that are well integrated with the objects they belong to. A series of four experiments introduced manipulations of cue-target timing, variations in scale (absolute distance), variations in colour-based grouping, and target-object integration manipulations. It was demonstrated that visual selection is not affected by spatial separation factors, but by the probability of the selecion target being integrated with the underlying objects.

Chapter 3 focused on the role of perceptual object formation factors in the context of visuo-spatial short-term memory (VSTM). Two experiments using a cued change-detection task demonstrated that VSTM is object-oriented, and influenced by phenomenally completed object structures. Finally, Chapter 4 utilised a similar cueing paradigm to Chapter 2, aiming to explore if object-based effects (as evidenced in the lack of spatial gradients within the same object) can also be observed for the phenomenon of inhibition of return (IOR). Unfortunately, the IOR effect was not obtained, but results nevertheless provided some additional evidence for object-based facilitation under circumstances where the cue does not predict the target location.

Finally, the General Discussion chapter focused on the overall significance of these results, concerning global issues in visual cognition, and important implications for experimental methodologies that aim to manipulate spatial factors while disregarding how such manipulations affect perceptual organisation. Areas for improvement and future directions are also discussed. There are many potential routes for exploring the factors contributing to the powerful influence and significance of objects, and also a lot of implications spanning from aspects of simple experimental stimulus generation and results interpretation, up to the level of reasoning about the functions of visual perception.

In addition to the main empirical work presented in the chapters outlined above, there are four appendices with additional experiments. These experiments represent piloting attempts to develop the stimuli and methodological parameters used in the main work, and are referred to at the relevant sections. Importantly, the work presented in these appendices is another reminder of the challenges associated with defining an object. Specifically, the appendices illustrate in practice a lot of the issues discussed earlier in relation to the problem of divergence between ‘an object’ from the experimenter’s point of view, and ‘an object’ when it comes to the reality of performing a task and making a perceptual decision. The fact that an object is ‘in the eye of the beholder’ will always be the main challenge and the root of most limitations of empirical work in this domain, and incidentally, one of the most intriguing aspects.

A note on statistics: Bayesian analysis

Bayesian statistics have become increasingly popular as they target a lot of the limitations and problems associated with null hypothesis significance testing and sample sizes (Dienes, 2014; Wagenmakers, Wetzels, Borsboom, & van der Maas, 2011). Since a large proportion of the conclusions derived from the current chapters are based on null results, e.g. the lack of a spatial separation gradient or no difference between conditions, additional analyses were conducted to calculate the Bayes Factor (BF) for the relevant hypotheses being tested. In this way, any two predictions can be directly compared against one another, assessing to what extent the data provides evidence for each. A BF is essentially a likelihood ratio of obtaining the observed data under any two hypotheses (Rouder, Morey, Speckman, & Province, 2012). As a general rule, a BF over 3 is considered positive evidence for the tested model versus the alternative (Jeffreys, 1961), although setting up a cut-off point undermines the purpose of Bayesian statistics, and these values should really be regarded as an infinite continuum.

Providing BF values is informative in the case of null results, especially when the lack of variation in performance is taken to be evidence for the experimental hypothesis. Frequentist statistics indicating a *p* value over the cut-off point of, for example, .05 do not provide evidence strength to make a confident conclusion about the lack or presence of an effect (Wagenmakers, 2007), while BFs can help quantify the available evidence. Therefore, the analyses of the following experiments presented a combination of standard and Bayesian statistics. The latter was performed using the BayesFactor package for R (Morey, 2015). The details regarding which predictions are being directly compared are provided at the relevant Results sections.

Chapter 2

The Mechanism of Visual Selection: Objects versus Space

Introduction

Visual selection is necessary for optimal interaction with the environment. In order to perform behaviourally relevant tasks, the target objects for action need to be selected, and the response output appropriately adjusted to accommodate the characteristics of these objects with the available motor and perceptual systems (Allport, 1987). Therefore, it is important to understand how visual selection operates and the factors which influence its efficiency. This chapter is concerned specifically with addressing the possibility that visual selection is purely object-oriented, and that evidence suggesting a critical role for space can be re-evaluated in light of the correlation between object belongingness and spatial proximity (Ruderman, 1997). As outlined in the General Introduction, a large proportion of space-based effects can be hypothetically accommodated within an object-based perspective by drawing on the structure of natural images, and the importance of perceptual organisation. Therefore, the purpose of the following series of experiments was to test this possibility.

One of the key issues to be addressed here is targeting the potential limitations and confounding factors in previous studies that may have led to the interpretation of object-based effects as effects arising as a result of spatial processes. Specifically, of interest is testing the hypothesis that effects of spatial separation, typically taken to suggest space-based selection (e.g. Vecera, 1994), are in fact resulting from a graded probability of object belongingness, i.e. at the root of these effects are object-level factors. However, since there is often a correlation between spatial proximity and object belongingness (Kubovy et al., 1998; Oyama, 1961) (hence, the misinterpretations), creating conditions where these two aspects are disentangled can be challenging. To overcome this challenge, it is necessary to design stimuli and tasks where object belongingness and spatial proximity between two points is varied in a non-monotonic fashion. That is, conditions where two proximal stimuli belong to different objects, while two more distant stimuli are unambiguously integrated within the same object. In addition, preserving the spatial organisation of the stimuli while altering their object-level structure can be informative about the factors that truly affect selection – objects or space. Under such circumstances, it would be possible to test if space can be ruled out as a factor in visual selection.

To spell out the matter more concretely, the basic layout of the main stimuli used throughout the current chapter is illustrated on Figure 13. The critical aspect is that the task-relevant stimuli are always six equally spaced circles arranged at the same eccentricity from fixation. These circles are identical in size and spatial location, and more importantly, they are either perceptually organised to belong to the same star-shaped object (Figure 13a), or to two separate overlapping equilateral triangle objects (Figure 13b). Therefore, in both cases these circles have the same spatial coordinates and thus the spatial factors are kept constant (i.e. location and distance between stimuli), but they are perceived as the features of either the same object, or two different objects. Critically, when these features are integrated into different objects, this is done in a way that any two immediately proximal features always have different object belongingness status, while two more distant features are always within the same triangle object. This two object version of the perceptual organisation creates conditions of non-monotonic relationship between spatial proximity and object belongingness.

A peripheral cueing paradigm with a target discrimination task was adopted for all experiments in the current chapter, as opposed to, for example, a divided attention method. The reason is that cueing a location in visual space directly, as opposed to using a central arrow or another type of implicit indication (e.g. a sound signal), should explicitly influence selection of the indicated spatial region, which has been argued to lead to a ‘narrow focus of attention’ associated with an emphasised space-based effect (e.g. Lavie & Driver, 1996; Posner & Cohen, 1984). Also, as mentioned in the General Introduction, using cueing in divided attention tasks typically leads to the expression of spatial separation effects where otherwise space-invariant object-based selection is observed (e.g. when conditions of superimposed versus separated objects are compared) (Vecera & Farah, 1994). Therefore, a cueing paradigm should be especially sensitive to space-based effects, if such effects exist.

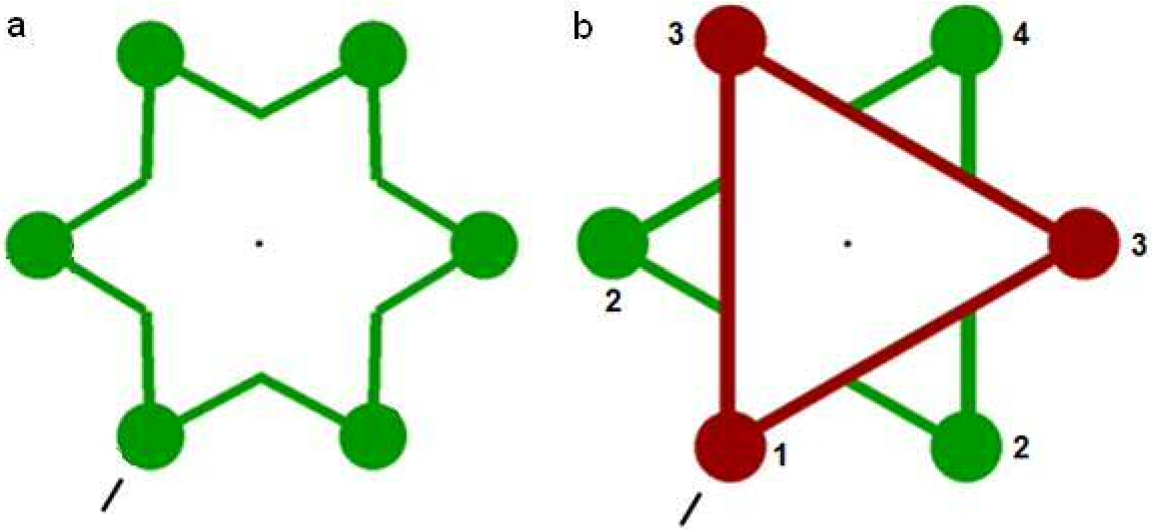

Figure 13: Main object stimuli illustration for Chapter 2; a: single object layout, b: two objects layout. The black line illustrates the cue, and numbers proximal to the object apices indicate the possible cue-target distances. The target was a luminance change of one of the circles situated on the object apex (not illustrated here). The spatial layout of the stimuli in the two objects condition is adapted from Brawn & Snowden (2000).

Adopting these stimuli results in conditions where the typical proximity-object belongingness positive correlation is removed (Figure 13), i.e. two proximal points do not belong to the same object. Therefore, this paradigm has the potential to test for the possibility of pure object-based selection. Specifically, the stimuli used in the current chapter allow testing the effect of four set cue-target distances (illustrated with numbers 1-4 on Figure 13b) under different object belongingness conditions. By cueing one of the object features with the presentation of a transient black line, it can be assessed if all of the features belonging to the same object are selected to the same extent (resulting in a lack of performance variation within the same object), or if selection is affected by the spatial separation between the cued and target features instead.

Other than addressing the challenge of disentangling proximity and object belongingness, another key issue in the current chapter relates to the problem of target-object integration discussed earlier. It is essential that if the effects of object-level factors are to be tested, the stimuli should be composed of salient objects, and the target event should be well incorporated into these objects. This is important because if the target is perceived as a separate object, rather than part of the critical object structures manipulated by the experimenter, then any reasoning based on the effect of the intentionally manipulated objects would not be valid since the perceptual system would not associate the target with these objects.

If a key argument to be tested is that the gradient effects demonstrated in the majority of previous studies are the result of variations in probability of object belongingness of the target and the cued location/feature, then it is critical to ensure that the target has a high probability of belonging to the objects. Therefore, the first three experiments from the current chapter utilised a target designed to be perceived as a change in an integral property of the object(s), i.e. a change in the luminance of an object feature. In contrast, the final experiment adopted a poorly integrated (superimposed) target, similar to the ones used in studies that demonstrate spatial gradient of selection and interpret it as evidence for the primacy of space over objects (e.g. Hollingworth et al., 2012). In other words, the experiments in this chapter also manipulate the probability of object belongingness of the target, in order to test if spatial separation effects, and thus space-based selection, are in fact an emergent property of this probability.

In sum, the current chapter consists of four experiments organised under the common theme of investigating the possibility that visual selection, as measured by target discrimination reaction time, is fully object-oriented, and what is typically considered as space-based selection is an emergent property of object-level perceptual organisation. Experiment 2.1 introduced manipulation of cue-target distance, and length of cue-target onset interval (stimulus onset asynchrony), under conditions where the same critical stimuli correspond to features of one (Figure 13a) or two objects (Figure 13b). The timing manipulation aimed to check for potential coexistence of space-based and object-based selection within a temporal frame, given that space-based selection is considered as an early process, while object-oriented effects may take longer to emerge (Shomstein & Behrmann, 2008). Such coexistence can manifest itself in a spatial gradient of facilitation for short timings, and an object-based pattern of performance for longer cue-target intervals.

Experiment 2.2 used a similar layout of the stimuli, but instead of varying the time interval between cue and target, the scale of the objects was manipulated across three levels. This manipulation aimed to test if performance is affected by increasing the spatial separation between object features while keeping the perceptual organisation into one or two objects constant. An object-based account would predict scale-invariant effects, with equal magnitude of object-based selection at each scale, while a space-based or a combination account would predict effects of cue-target spatial separation, which will increase with increasing scale.

Since the first two experiments supported object-based selection, Experiment 2.3 aimed to clarify if this effect was simply due to grouping on the basis of colour as a common feature of same-object stimuli, often referred to as feature-based selection (Freiwald, 2007). In Experiment 2.3 both objects were of the same colour, in order to eliminate common colouring as a potential confounding factor for selection. Finally, Experiment 2.4 used a single object condition with superimposed instead of integrated targets, in order to test if this manipulation would lead to the emergence of a spatial gradient of selection. Overall, the key aim that underpins the experiments in Chapter 2 was testing the extent to which selection, as measured by reaction time, is directly linked to the *probability* that a target belongs to a selected (cued) object, rather than the spatial separation between cue and target.

Experiment 2.1: Space-invariant Object-based Selection

In order to test the possibility of space-independent, pure object-oriented selection, Experiment 2.1 introduced a systematic manipulation of cue-target distance, time interval between cue and target, and number of objects within which the possible target features were distributed. The basic layout of the object stimuli, together with the cue (black line proximal to one of the object features) is illustrated on Figure 13. Importantly, the six possible locations of the cue and target were always the same in terms of absolute spatial coordinates, but they differed with respect to whether they corresponded to the features of a single star-shaped object – *one object* condition (Figure 13a), or the features of two overlapping equilateral triangles – *two objects* condition (Figure 13b). This critical manipulation of the perceptual organisation of the visual scene creates conditions with a total of four target distances relative to the cued feature.

The aim of the experiment is to test for a potential gradient in performance, given the integrated nature of the target event. Specifically, a transient cue is followed by a positive or negative luminance change of the circular feature at the apex of the object. The distance of this target from the cue, as well as whether it belonged to the cued object or not, is systematically varied. For ease of interpretation, distance 1 always corresponds to target events occurring at the cued feature. Importantly, in the *two objects* condition the four cue-target separations are non-monotonically related to the probability of the target belonging to the cued object, thus unconfounding spatial proximity and object belongingness. Distances 2 and 4 always correspond to a target integrated within the uncued object, while distance 1 and distance 3 are associated with the cued object. Meanwhile, the *one object* condition introduced four levels of cue-target separation within the same object.

The second critical manipulation was the stimulus onset asynchrony (SOA), which is the time interval between cue and target event. Three different time intervals were introduced (100 ms, 200 ms and 300 ms), in order to examine the potential influence of object- and space-oriented mechanisms over time. Specifically, since object-based effects are proposed to be secondary to space-based selection, whereby an object constitutes an array of grouped spatial locations (e.g. O'Grady & Müller, 2000), it may be the case that a short interval between cue and target results in space-based selection, while as this interval increases, the effect of the perceptual organisation (the location-grouping influence) becomes more emphasised. This may also be the case as there is some evidence that object-based selection is a slower process and takes time to be manifested (Avrahami, 1999; Chen & Cave, 2008; Feldman, 2007; Shomstein & Behrmann, 2008).

It is critical to note again that with the use of the current stimuli, space-based selection is to be inferred from spatial gradients in performance. However, in general terms, such gradients may be resulting from a gradient in the probability that a target belongs to the cued feature/object. Given that the current stimuli were specifically designed to try and increase this object belongingness probability to 1 (i.e. certainty that the target is part of the object it appears in), a gradient under these circumstances may indeed be evidence of genuine space-based selection. Therefore, it may be the case that a short SOA results in a spatial gradient of facilitation, followed by a flat performance function within the same object for longer SOAs. Finally, the reason for using SOA intervals within this specific range (100 ms – 300 ms) is because 300 ms is considered as the threshold above which inhibition of return (IOR) can be observed (Klein, 2000). This is a phenomenon where the facilitating effect associated with the cued location (and its hypothesised spatial gradient) starts to reverse into inhibition, resulting in slower reaction times compared to uncued locations (this phenomenon is explored in more detail in Chapter 4).

Given these manipulations, if selection is genuinely space-oriented, a discrimination reaction time gradient centred at the cued feature is expected to occur. However, if selection is exclusively object-oriented, responses should only vary depending on whether the target is part of the cued object, regardless of cue-target distance. In other words, the pattern of performance should be space-invariant and object-dependent. Therefore, a pure object-based effect would be manifested as no variation in performance for the *one object* condition, and a non-monotonic pattern of performance for the *two objects* condition that reflects the object structure of the stimuli, rather than the spatial separation between cued and target features.

If the primary mechanism involves spatial selection subsequently modulated by object formation, then such an interaction might be time-dependent. Since perceptual grouping and object formation are often regarded as late processes and emerging from space (Feldman, 2007; Korjoukov et al., 2012; Shomstein & Behrmann, 2008), while space-oriented selection is considered early and primary, it may be the case that the hypothesised space-invariant effects are not present at short SOA intervals. In this case, a spatial gradient may be observed at early SOA, but an object-based pattern can take place at a later SOA. Alternatively, if the two mechanisms are simultaneously manifested, it may be expected that while a graded performance is observed in the *one object* condition, the opposing mechanisms may cancel out in the *two objects* condition, resulting in a flat performance function. Finally, independent participant samples were used for the two critical conditions (*one object* and *two objects*), in order to avoid confounding effects due to experience with the perceptual organisation of the stimuli (Zemel et al., 2002; Libo Zhao, Cosman, Vatterott, Gupta, & Vecera, 2014).

Method

Participants

Thirty six (mean age 19.97, *SD* = 4.01) undergraduate psychology students from Cardiff University took part for partial course credit in the *one object* condition, and thirty-nine (mean age 20.15, *SD* = 4.18) in the *two objects* condition. The data from three participants in the *one object* condition and four participants in the *two objects* condition were removed from the analysis due to scoring under 50% on accuracy. As a result, the analyses were conducted on a sample of 33 participants (2 male, mean age 20, *SD* = 4.19) for *one object*, and 35 (2 male, mean age 20.23, *SD* = 4.39) for *two objects*.

Participants had normal or corrected-to-normal vision, and normal colour vision. Samples for visual selection experiments tend to vary between 10 and 40 participants. It was aimed for the upper end of the scale to allow room for potential exclusions. Power calculations were not performed a priori due to uncertain effect sizes with these types of stimuli and paradigm.

Stimuli and Apparatus[1]

In the *one object* condition, the six potential target locations corresponded to the integral features of a single six-pointed star-shaped object (outlined in either green or red), while in the *two objects* condition the same absolute locations corresponded to the three apices of each of two superimposed, opaque equilateral triangles (one with a green outline, the other red). The object features were composed of filled circles (with the corresponding object colour) with a diameter of 1.5°, centred at 5° from fixation. The Euclidian distance between the centres of two neighbouring circles was 5°, corresponding to 60° angular separation with reference to the central fixation point. The luminance of the 12-pixel boundaries of the objects was 17 cd/m^2 against a grey background (34.7 cd/m^2). The cue consisted of a dark grey 10-pixel stroke line (6.3 cd/m^2) subtending 0.82°, and was positioned on the axis passing through the central fixation point and the centre of the cued apex, 0.38° from the edge of the cued circle (refer to Figure 13). The target event constituted a ± 50% luminance change of one of the six circular object features.

Stimulus presentation and response recording (via a standard keyboard) were controlled using Matlab R2012a with Psychophysics Toolbox 3 (Kleiner, Brainard, & Pelli, 2007) running on Windows XP operating system and a 23-inch monitor with 1920 x 1080 pixel resolution, 32-bit colour quality, and a 60 Hertz refresh rate.

---

[1] Refer to Appendix 1 for detailed information on pilot studies which contributed to choosing the current stimuli and design characteristics.

Design and Procedure

Each participant completed 432 trials on only one of the object conditions, which conformed to a 4 (cue-target distance) x 3 (SOA: 100 ms, 200 ms, and 300 ms) repeated measures design. Four levels of cued feature-target feature separation were tested in each of the object conditions, corresponding to 0°, ± 60°, ± 120° and 180° of angular separation, which translates into 0°, 5°, 8.5°, and 10° Euclidian distance. For the sake of simplicity, these are referred to as distance 1, distance 2, distance 3, and distance 4, respectively (Figure 13b). For each object condition there were 144 trials per onset asynchrony, with the target event occurring at the cued feature on 44 of these, and at one of the other 5 features for 20 trials each. Therefore, there was a 30.5% chance of the target event appearing at the cued feature, and a 13.9% chance for any uncued feature (i.e. the cue was informative, but only to a limited extent). Within each SOA, the cue appeared an equal number of times on each of the six possible locations. In addition, each combination of cued feature, object colour (red or green for the *one object* condition), object position (top or bottom for the *two objects* condition), and target change polarity occurred an equal number of times.

The study was approved by the Cardiff University School of Psychology Ethics Committee in accordance with the British Psychological Society ethical guidelines. All participants underwent an identical procedure, apart from the differences in the stimuli described above. Participants were tested individually in semi-enclosed booths. Each session began with 10 randomly selected practice trials. Feedback on response accuracy was provided for the practice trials but not for experimental trials. Between each trial, the word ‘Ready’ (for experimental trials) or ‘Correct’/’Incorrect’ feedback (for practice trials) subtending approximately 1.93° was superimposed 2.21° by 0.80° to the bottom left of fixation. It was displayed on top of a dynamic visual mask consisting of different sized red and green circles randomly changing location every 50 ms for duration of 900 ms. The mask subtended 14.47° by 11.8° and was presented at the centre of the screen. Its purpose was to minimise the after-image effect due to prolonged fixation.

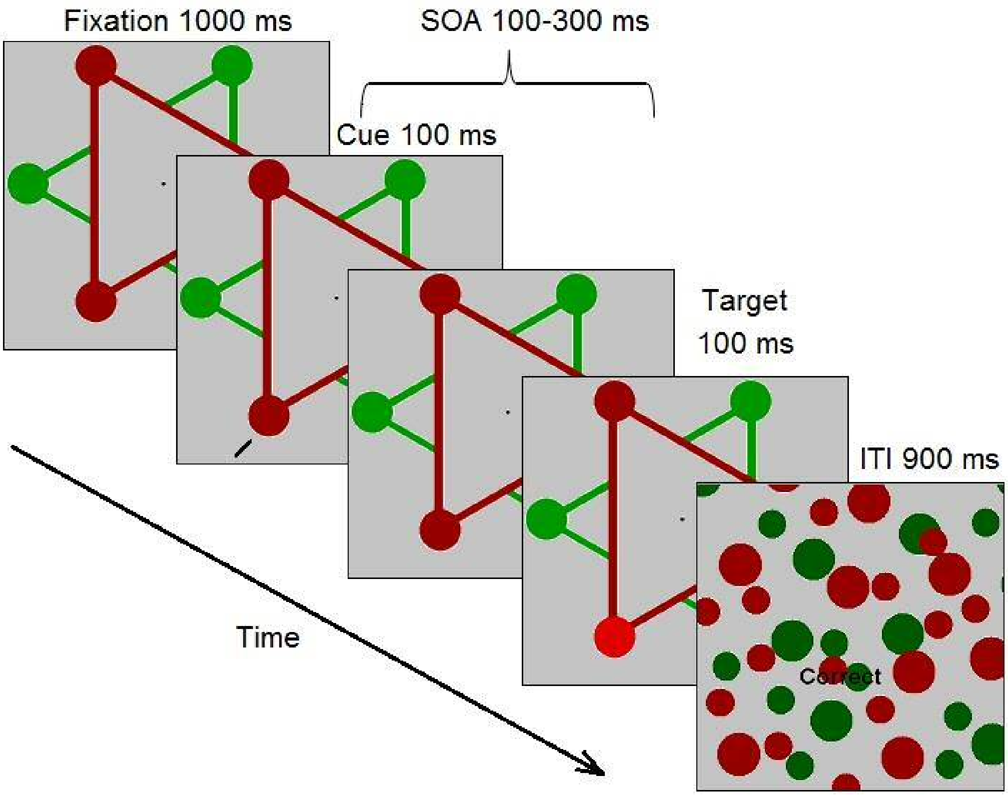

Figure 14: Experiment 2.1 procedure illustration (*two objects* practice trial, red object on top). An example of a cued target with a positive luminance change. ITI = inter trial interval.

The experimental procedure is illustrated on Figure 14. Each trial began by presentation of the relevant object(s) for 1000 ms, followed by the appearance of the cue for 100 ms. In the case of the 100 ms SOA, the target event directly followed the cue offset. For SOAs of 200 ms and 300 ms the target occurred 100 ms and 200 ms after the cue offset, respectively. The duration of the target event was 100 ms, after which the relevant feature returned to its original state until response. Participants were instructed to maintain fixation throughout the duration of the trial, and to balance speed and accuracy when responding. Responses were made with the index fingers placed on the buttons ‘L’ and ‘D’ (highlighted using adhesive labels) on the keyboard, corresponding to ‘lighter’ and ‘darker’, respectively. The position of the response buttons was counterbalanced between participants by switching the adhesive labels on the keys. In order to explicitly remind subjects which key corresponded to which judgement, 0.74° by 1.06° upper case ‘L’ and ‘D’ letters were displayed on the horizontal axis 8.45° lateral of fixation on either side of the stimulus, corresponding to the laterality of the response to which they were mapped. Reaction time (milliseconds) and accuracy (proportion correct) data were collected.

The session was divided into three blocks of 144 trials with self-timed breaks in-between. The SOA and target location combinations varied randomly from trial to trial for all three blocks. However, the proportion of cued and uncued targets was the same for each block in order to maintain equal level of top-down bias throughout the experiment. Participants were aware of the cue-target contingencies and were fully debriefed after the study. The whole procedure lasted approximately 35 minutes.

Results and Discussion

Only reaction times for correct responses were analysed, and responses shorter than 200 ms or longer than 1200 ms were excluded. This trimming procedure resulted in the overall removal of 1.28% of trials for the *one object*, and 2.86% *two objects* condition. Accuracy was not used in the analysis, as performance was overall high and with not enough variation to be informative of any effects.

Reaction time data for target distances 2 and 3 were calculated by averaging performance between the target positions situated at 60° and 120° on either side of the cued location, respectively. Separate 4 (target distance) x 3 (SOA: 100 ms, 200 ms, 300 ms) repeated measures Analyses of Variance (ANOVA) were conducted on each object condition. Whenever the assumption of sphericity was violated, the Greenhouse-Geisser corrected values are reported. Bonferroni correction was used for the significance values of all follow-up analyses of main effects. Statistical interactions were followed up with planned four-way repeated measures ANOVAs at each level of SOA. An alpha level of .05 (adjusted where necessary) was adopted for all tests.

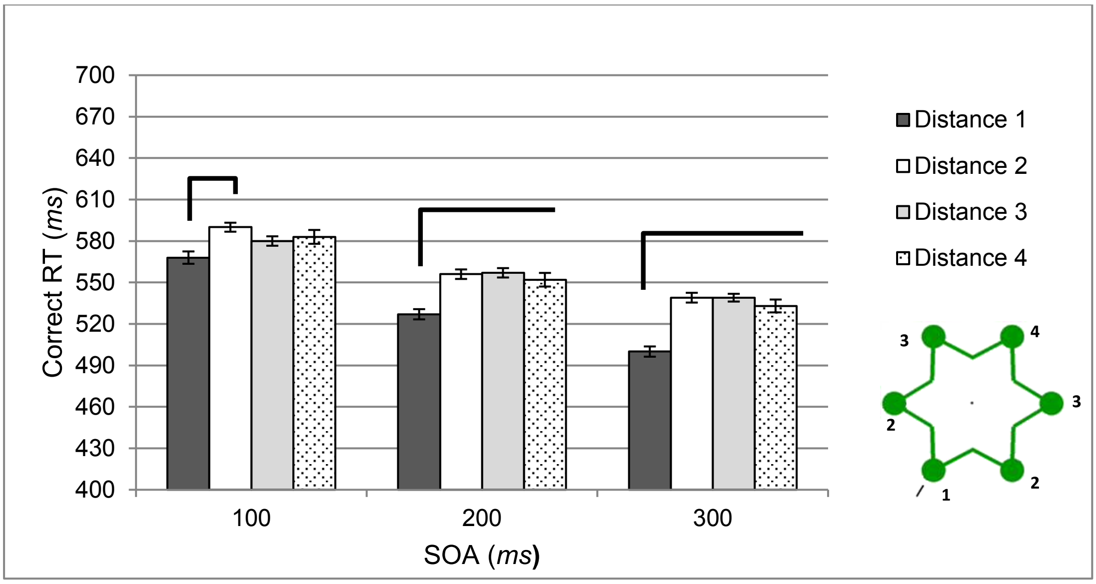

Figure 15: Experiment 2.1 mean correct reaction time (milliseconds) as a function of target distance and SOA for the *one object* condition. Error bars represent standard error for the mean, corrected for between-subject variability (Cousineau, 2005), hereafter referred to as 'corrected SEM', Brackets illustrate the statistical differences between the different distances at $p < .05$ (refer to the text for exact values).

As Figure 15 illustrates, reaction times in the *one object* condition decreased as SOA increased from 100 ms to 300 ms, $F(2, 64) = 169.61$, $p < .001$, $\eta_p^2 = .84$. There was a clear overall trend for fastest reaction times at the cued feature compared to all others, but no effect of cued feature-target feature distance for any of the uncued target locations. Specifically, there was an effect of distance, $F(2.25, 71.85) = 30.75$, $p < .001$, $\eta_p^2 = .49$, which interacted marginally with SOA, $F(6, 192) = 2.23$, $p = .042$, $\eta_p^2 = .07$. This interaction was due to there being a statistical difference only between distance 1 and distance 2 ($p = .002$) for SOA of 100 ms, while for SOA of 200 ms reaction times for distance 1 were the faster than all other positions ($p < .001$ for the comparison with distances 2 and 3, $p = .009$ for distance 4). For SOA of 300 ms responses for targets at distance 1 were also the fastest (all $p_s < .001$), while no other statistical differences were observed (Figure 15). Overall, reaction times for the *one object* condition indicate facilitation for the cued feature that became more pronounced as SOA increased, while reaction times to targets at uncued features did not increase with increasing distance from the cued feature.

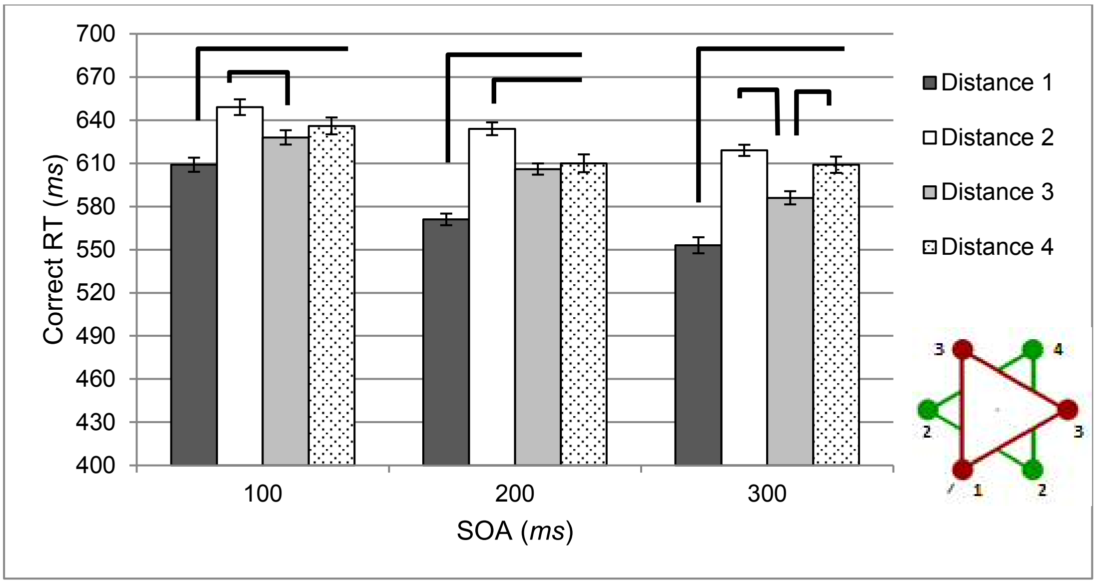

Figure 16: Experiment 2.1 mean correct reaction time (milliseconds) as a function of target distance and SOA for the *two objects* condition. Error bars represent corrected SEM. Brackets illustrate the statistical differences between the four distances at $p < .05$.

In the *two objects* condition (Figure 16), there was once again an overall decrease in reaction times as SOA increased, $F(2, 68) = 47.731$, $p < .001$, $\eta_p^2 = .58$, and the effect of distance, $F(2.306, 78.420) = 47.913$, $p < .001$, $\eta_p^2 = .59$, was again qualified by SOA, $F(6, 204) = 2.654$, $p = .017$, $\eta_p^2 = .07$. As Figure 16 reveals, reaction times for the cued feature (distance 1) were consistently faster than all others, and reaction times at distance 2 were consistently slower than reaction times for targets at the larger cue-target distance 3. While for SOAs of 100 ms and 200 ms there was no difference between reaction times for distance 3 and distance 4, this pattern changed for SOA of 300 ms, such that reaction times for the distances 2 and 4 did not differ statistically (mean difference of 9.66 ms), while reaction times for distance 3 on the cued object were faster than reaction times for any of the uncued object locations (difference of 33.43 ms between distance 2 and 3, $p < .001$; difference of 23.77 ms between distance 3 and 4, $p = .025$). As in the *one object* condition, apart from the advantage at the cued location, there was no evidence of a cost associated with spatial separation of the cued and target features. Rather, the only evident cost of cue-target separation was a function of whether the cued and target features belong to the same object. This was reflected in a reaction times increase for targets rotationally adjacent to the cue and decrease again for targets rotationally more distant from, but belonging to the same object as, the cued feature.

In opposition to an object-oriented account of this effect, it might be argued that object features at distance 3 are connected with the cued location via a continuous straight line, a visual feature that might help expedite the movement of the selection mechanism through space (Crundall et al., 2007). However, a paired samples t-test comparing reaction times at SOA of 300 ms for distance 3 with the line occluded (cued object underneath) (*M* = 591.95, *SE* = 15.09) against distance 3 with the line fully visible (cued object above) (*M* = 591.43, *SE* = 15.41) confirmed no statistical difference, *t*(34) = .06, *p* = .953. Therefore, this effect appears to be truly object-oriented, even when the boundary of the object is itself physically discontinuous within the scene.

As mentioned in the General Introduction, there is some evidence that the spatial distribution of attention has a surround inhibition zone, resulting in a Mexican hat function where attention gradually decreases with distance from the current focus, and then increases again (Müller, Mollenhauer, Rösler, & Kleinschmidt, 2005). Depending on the location of the ‘dip’ in the hypothesised inhibition function, a performance pattern similar to the one observed in the *two objects* condition may be expected. It can therefore be argued that the currently observed object-based effect is nothing more than a result of the spatial distribution of selection, i.e. a space-based effect with an inhibitory surround. However, if this was true, the same pattern should be equally observed in the *one object* condition, which was not the case. Therefore, it is more likely that performance was influenced by the perceived status of the target regarding the cued object, i.e. a genuine object-based effect, rather than its spatial separation from the cue.

Focusing on the results from the *two objects* condition, the fact that performance for the most distant feature (distance 4) was not consistently conforming to the object-based pattern evident at SOA of 300 ms deserves some attention. There appears to be a level of counterintuitive facilitation for the processing of target events at this feature, since both space-based and object-based accounts predict poorest performance there. This unexpected effect is not consistent, since it is not as strongly pronounced for the *one object* condition, but it is nevertheless observable as a weak trend. One possibility is that it is due to the directionality of the cue, which may be perceived as pointing towards the feature at distance 4, since it is situated on the axis passing through its centre. Indeed, in-vivo studies of the striate cortex of small mammals demonstrate intricate long-range horizontal connections between distant neurons with common orientation preference (Bosking, Zhang, Schofield, & Fitzpatrick, 1997). Although this is in the context of stimulus orientation selectivity, which is not specifically relevant here, it suggest a possibility that a given neuron or network can be responsive to a certain stimulus because of input from another distant neuron/ network along an axis.

The symmetrical structure of the stimuli may be, at least to some part, at the root of this counterintuitive effect. It is possible that information is processed with a certain level of independence between the left and right hemifields (Hickey & Theeuwes, 2011; Sereno & Kosslyn, 1991). Assuming that a top-down bias towards the cued feature (since it signifies the most likely target location) can influence the axis of symmetry on each trial, and given the layout of the current stimuli, data for distance 4 is always sampled from a single location in the opposite hemifield directly across the cued feature, while targets at distances 2 and 3 appear in the contralateral hemifield only half of the time. In other words, the top-down bias may not be as influential in the opposite hemifield, leading to a tendency for prioritisation of targets at distance 4 relative to other uncued features on that object (i.e. those located at distance 2). In addition, data for distance 4 is based on averaging from fewer trials than all other distances, which may potentially lead to a noisier measurement, although error bars do not suggest more variation there. It is unclear what this effect may be due to, the most plausible/ parsimonious explanation being cue directionality. In any case, although reaction times for distance 4 are not to be dismissed, the comparisons between distances 1, 2 and 3 provide a cleaner measure for the object versus space-oriented hypothesis.

Finally, the fact that performance in both conditions was consistently superior for the cued location can potentially be interpreted as evidence for space-based selection (Chen, 2012). However, visual cognition is adaptive and can be affected by both bottom-up and top-down factors (Fecteau & Munoz, 2006; Shomstein & Johnson, 2013; Theeuwes, 2010; Yantis & Jonides, 1990). Therefore, strategic control as a result of a known regularity (cue predictability in this case) can lead to prioritised processing for specific stimuli. For example, within the context of the two rectangle paradigm (Egly et al., 1994), low cue validity (i.e. when the cue is not informative about the future target location) can result in equal reaction times for cued and uncued-same object targets, which are nevertheless faster than responses to uncued-different object targets (He, Fan, Zhou, & Chen, 2004). In other words, the benefit of cueing can apply equally to the whole perceptual object when there is no obvious advantage attached to a specific location within this object. Therefore, although the cued feature advantage may be interpreted as a spatial effect, it can also be a strategic orientation effect. Given that there is no evidence of other spatial separation effects in the data, and in fact there is evidence for non-monotonic, object-oriented pattern, then it is much more likely that this is a strategy-driven phenomenon. This pattern of results does not necessarily conflict with a spatially invariant object-based selection, since all uncued features were associated with an equal probability of changing. The only remaining influence on selection for these features was their status regarding the cued feature, i.e. the object belongingness probability.

Since the current conclusions for the lack of spatial separation effects in the *one object* condition are based on null results, additional analyses were conducted to calculate the Bayes Factor (BF) for the relevant order restrictions between distances 1 to 3 for the *one object* condition with SOA collapsed. Data from distance 4 was not included in this analysis due to the issues mentioned above. The order restrictions relevant for the current research questions are: 1) reaction times for distance 1 are faster than reaction times for distance 2 and distance 3, but the latter two are equal (i.e. a performance model supporting equality for responses to uncued targets within the same object), and 2) reaction times for distance 1 are faster than reaction times for distance 2, which in turn are faster than reaction times for distance 3 (a model supporting graded spatial separation effects). BFs for the object-based (equality) model and the monotonic gradient model versus the null hypothesis were calculated, and then these values were directly compared against each other in order to generate BFs that reflect which model is favoured given the data.

The critical comparison is between distance 2 and distance 3, since whether they produce equal reaction times, or distance 2 is faster, determines which model fits the data. Consequently, the object-based model is tested by the restriction that distance 2 and distance 3 elicit the same reaction times, while the monotonic gradient model is reflected in the restriction that distance 2 elicits faster reaction times than distance 3. The latter is calculated by using Markov Chain Monte Carlo algorithm to sample from the posterior distribution of an unrestricted differences model (i.e. a model suggesting that reaction times for distances 2 and 3 differ from each other in either direction). It is then estimated how often the gradient model holds true after a default number of 10000 samples, and this value is divided by the prior odds of obtaining this restriction, which in this case is set to be 0.5. The reason is that there are only two possible directions that the difference can reflect (faster or slower), and an equal weighting is given to both (so adopting an unbiased prior)[2] (refer to Morey (2015) for more detail on the calculation procedures). The analysis suggested that the object-oriented equality pattern is 11.65 times more likely than the graded spatial separation model, whereas the BF for the latter against equality was 0.09. As a reminder for a reference point, BF > 3 is considered as positive evidence, while BF < 0.33 is regarded as evidence against the tested model (Jeffreys, 1961). Therefore, the current data gives strong support that targets within a selected object are equally prioritised for processing, regardless of their distance from the cued object part.

In sum, Experiment 2.1 suggested that there is no evidence of a space-oriented mechanism for selecting information for preferential processing in the visual scene. There is, however, clear evidence of selection operating in an object-oriented fashion, and not one that merely modulates the operation of a primarily space-oriented process. Rather, the evidence points to the primacy of the object structure over space, where spatial separation does not affect responses to

---

[2] Since it was always the case that reaction times for distance 1 were the fastest, no directionality for the comparison between distance 1 and the remaining distances was specified. Only the likelihood that it differed was assessed, since it is equal to the likelihood of it being the fastest. Indeed, it was confirmed that this order was true for all samples from the posterior (i.e. 100% of the time).

targets which are integral parts of the objects in the visual scene. It is important to note that these data do not refute that selection can be spatially graded, instead, it is proposed that such a gradient does not signify space-based selection, but it reflects the probability that the cued and target feature pairs are part of the same object. Since here this probability was very high, the pattern of performance as a function of space reflected the structure of the objects – i.e. non-monotonic function in the *two objects* condition, and a flat function for *one object*. As the locations and distance between features were identical in both conditions, the current results clearly demonstrate that selection was affected exclusively by the objects on the scene and more specifically, by the object relationship of the target event with the cued object feature.

Experiment 2.1 provided strong evidence in favour of pure object-based effects, rather than coexistence of space and object-level factors. However, both the spatial separation and the perceptual organisation of the stimuli were held constant for each participant, as the number of objects was varied between subjects. If visual selection is indeed solely affected by the objects on the scene, then it can be expected that as long as the object-level perceptual organisation remains stable, increasing or decreasing the spatial separation between object features, and thus between cue and target, should not influence performance. That is, manipulating spatial factors (distance and location) while keeping object factors (high probability of target belonging to the object, and the structure of the objects) unchanged should have little or no effect on performance, which should be consistently object-based for all levels of spatial separation. Testing this possibility was the principle aim of the following Experiment 2.2.

Experiment 2.2: Scale-Invariant Object-Based Selection

The key findings from Experiment 2.1 suggest that visual selection may be space-invariant and dependent solely on object structure, as opposed to operating on the basis of spatial separation. This is in line with the argument that in order to be optimal, visual selection needs to be fully object-oriented, since the natural environment is composed of a variety of objects subtending various sizes and often partially occluded by other objects. Prioritising visual information on the basis of space is thus not optimal, since parts of the same object may be discontinuously distributed (due to occlusion). That is, spatial selection may result in prioritising information that is not behaviourally relevant for the current goal, while selecting information with reference to objects is more likely to lead to appropriate execution of the required action.

As pointed out earlier, interacting with the environment leads to shaping the perceptual system in a way that is optimised for the characteristics of the statistical properties of the natural visual scene (Field, 1987; Ruderman, 1994), and in turn these statistical properties, e.g. power spectrum and spatial frequency, are known to be scale-invariant, i.e. do not change with a change in observation scale (Baddeley, 1997; Ruderman, 1997). As the analyses of natural scene statistics indicate, the difference function between two points in an image depends on whether they originate from the same object, rather than directly on the distance between them. The current empirical work so far suggests this is also true for the mechanism of visual selection.

Experiment 2.2 aimed to test the possibility of scale invariance for visual selection, which should be evident if indeed the selection mechanism has evolved in accordance with the statistical properties of the natural scene. In addition, it provides an opportunity to replicate the findings from Experiment 2.1. Therefore, the same types of stimuli were utilised, but instead of SOA, the scale of the display was varied across three levels. Figure 17 illustrates the three scales for each object condition, which were derived by centring the apices of the objects (and the target circles, respectively) at different eccentricities from fixation: 2°, 5° (replicating the visual conditions in Experiment 2.1), and 7°.

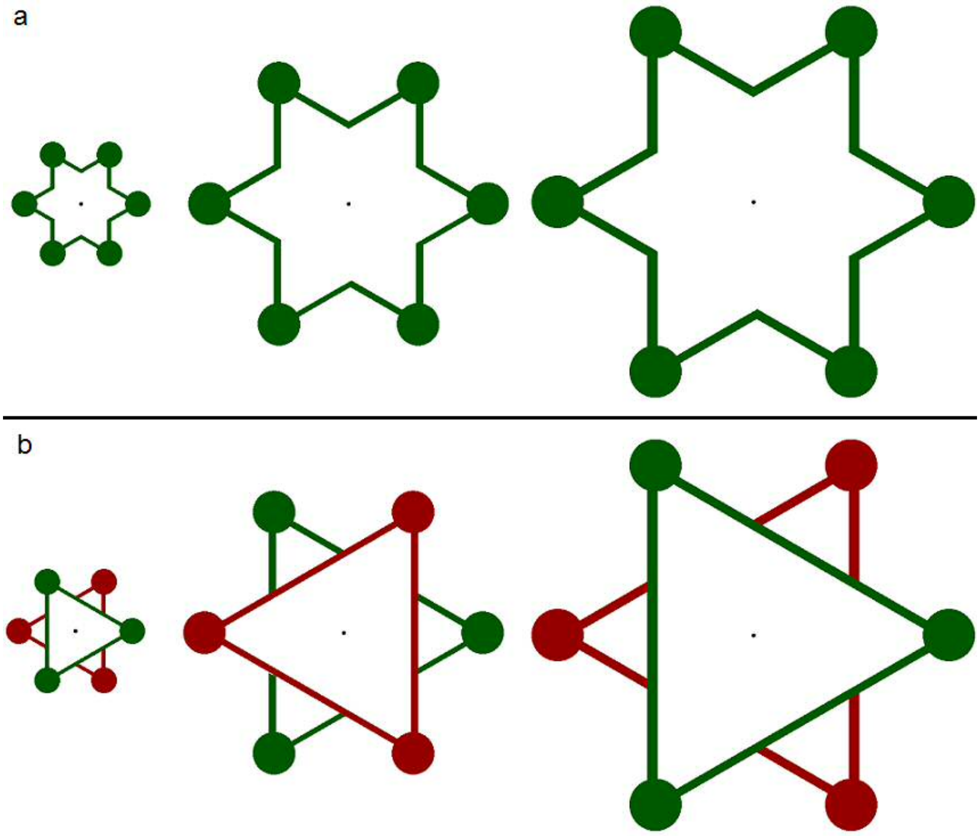

Figure 17: The three object scales of Experiment 2.2. Left to right: 2°, 5°, and 7°; a: *one object* condition, b: *two objects* condition.

Critically, although the spatial separation between object features changes with scale, the probability of object belongingness for the target event remains close to certainty for all scales. For example, a target at distance 2 always represents a change on the uncued object (for the *two objects* condition), but its distance from the cued feature becomes progressively larger with increase in scale. Therefore, while there is a monotonic gradient with reference to space, which is correlated with scale, the gradient with reference to the objects, which in this case is binary – either within the cued object or not, does not change with scale. This setup provides a novel way of examining the possibility that a ‘spatial’ gradient is not truly spatial in nature, but a product of the object-level perceptual organisation of the scene.

Given the structure of these stimuli, if visual selection is adapted to the regularities of the natural environment and the scale invariance of natural scene statistics, it is expected that the object-based effect seen in Experiment 2.1 would be replicated across the three object scales, such that scale would not interact with the non-monotonic (object-based) effect of distance. This is because the perceptual organisation of the stimuli at each scale is the same as in Experiment 2.1. However, if spatial selection plays a critical role in this process, an interaction can be expected between cue-target distance and scale, since there is a larger spatial separation in the largest scale condition compared to the smallest. However, unlike the spatial proximity manipulation of Vecera (1994) discussed earlier, whereby the two rectangles were placed closer together in order to decrease cue-target distance, here the distance manipulation (scaling) is not confounded with changes in the perceptual organisation of the stimuli (refer to Figure 10 in General Introduction). Changing the distance between the rectangles by Vecera (1994) may have also resulted in increasing the probability that they are perceived as a single object, and in addition the target in that study (a transient grey square) was not well integrated within the body of the rectangles. With the current stimuli, the object belongingness probability gradient is not correlated with spatial distance, thus allowing for clearer assessment of the underlying selection mechanism.

Method

Participants

Twenty participants (6 male) took part in the *one object* condition (mean age = 23.95, *SD* = 5.73), and 20 (6 male) participated in the *two objects* condition (mean age = 23.1, *SD* = 6.16). The sample included students and staff from Cardiff University, recruited via the University's online notice board. Participants were paid £4 for participation. The sample size was based on power calculations computed with G*Power 3 software (Faul, Erdfelder, Lang, & Buchner, 2007), using the effect sizes obtained for the interaction effect from Experiment 2.1 for *one* and *two objects*, respectively. The data of one participant was excluded from the *two objects* condition due to low accuracy (under 50% correct), resulting in a sample of 19 participants (mean age = 23.36, *SD* = 6.21) for that condition.

Stimuli and Apparatus

The sizes of the target features and the cue were adjusted in order to correct for variations in visual acuity with changing eccentricity. The quality of visual information deteriorates as the projected locus on the retina moves away from the fovea (Rovamo & Virsu, 1979). Colour discrimination, specifically red-green contrast sensitivity, decreases with increasing eccentricity, but this can be corrected by adjusting the stimuli size (Noorlander, Koenderink, Olden, & Edens, 1983). Given the size of the stimuli in the current study and the fact that 50% change in luminance is a large enough magnitude to detect without difficulties, changing the eccentricity is unlikely to have a big impact on performance that can interfere with measuring the hypothesised effects. Nevertheless, in order to ensure the stimuli across the three eccentricities stimulate approximately equal cortical space and pose similar perceptual demand, they were scaled to correct for potential changes in discriminability. This was achieved through a combination of objective and subjective methods. For the former, the sizes from Experiment 2.1 were adjusted based on a liner cortical magnification factor, as described in Rousselet, Husk, Bennett, & Sekuler (2005) (see also Dougherty et al. (2003)). Following these conversions, finer adjustments (up to ±10 pixels) were made based on the subjective report of two observers, who did not take part in the experiment.

The basic properties of the stimuli and the display layout were the same as in Experiment 2.1, apart from the following changes. The size of the object(s) varied depending on the scaling condition, the scales are labelled based on the eccentricity at which the target features were centred. For the 2° scale, the critical circle features were 0.97° in diameter, and the cue subtended 0.56°. The Euclidian distances between the cued and uncued features (measured centre-to-centre) were 2° (distance 2), 3.4° (distance 3), and 4° (distance 4). The 5° scale was identical to Experiment 2.1, and the 7° scale had features of 1.75° diameter, cue subtending 0.99°, and Euclidian distances corresponding to 7° (distance 2), 12° (distance 3), and 14° (distance 4). The response reminder letters were moved to 9.8° laterally on each side of the stimulus, and the size of the inter-trial-interval mask was increased to 20.25° x 17° in order to cover the area subtended by the largest scale.

Design and Procedure

Each object condition conformed to a 3 (scale: 2°, 5°, 7°) x 4 (target distance) repeated measures design. As before, there were a total of 432 trials, 144 for each scale where the cue predicted the target location on 44 of these. Experiment 2.2 used the timing conditions which elicited the most pronounced object-oriented effect in Experiment 2.1, therefore SOA varied randomly between 300 ms and 350 ms, in order to prevent anticipatory responses. Scale and target distance varied randomly from trial to trial. The procedure was identical to Experiment 2.1.

Results and Discussion

As before, reaction times faster than 200 ms and slower than 1200 ms were excluded from the analyses, resulting in a loss of 1.44% of trials in the *one object* condition, and 2.17% from the *two objects* condition. For the *one object* condition, the effect of target distance, $F(3, 57) = 10.68$, $p < .001$, $\eta_p^2 = .36$, replicated the cueing advantage with a lack of spatial gradient (Figure 18). Namely, reaction times for targets at the cued feature were faster than reactions times for targets at distance 2 and distance 3 (both $p_s < .001$), while performance did not vary between any of the uncued features. In support of this, BF for an object-oriented pattern (equality between distances 2 and 3) was 2.03, while BF for a spatially graded pattern was 0.49. There was an overall influence of scaling, $F(2, 38) = 13.12$, $p < .001$, $\eta_p^2 = .41$, which was due to generally slower reaction times at scale of 7° compared to 2° and 5° (both $p_s < .001$). However, there was no difference between scales of 2° and 5°. Importantly, both factors had independent effects, since there was no interaction between them ($F < 1$), suggesting that the observed cueing and lack of further influence of spatial separation were equally pronounced for all scales, i.e. these effects were scale-invariant.

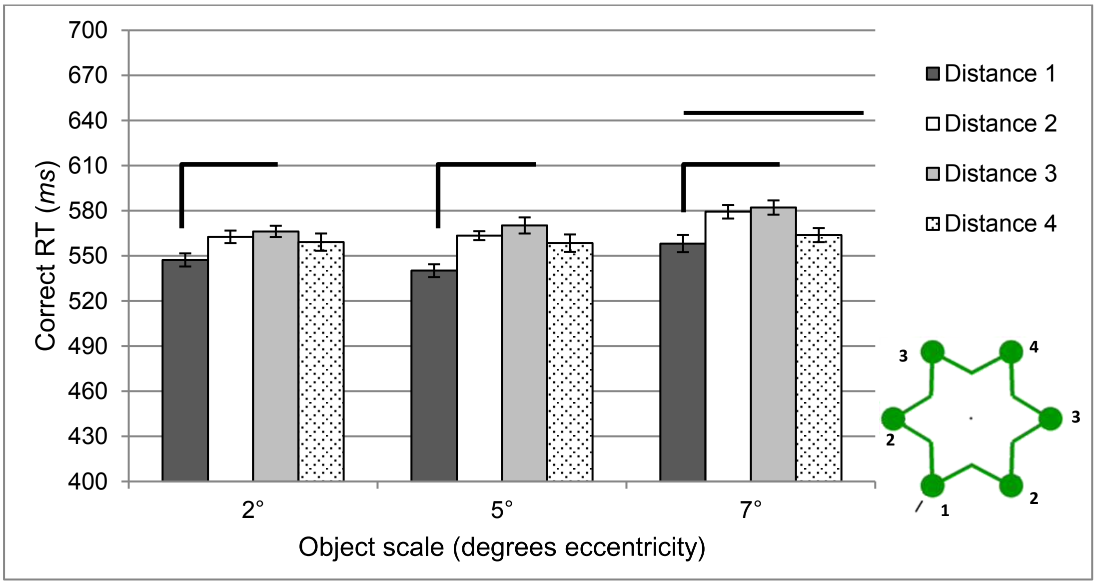

Figure 18: Experiment 2.2 mean correct reaction time for the *one object* condition as a function of object scale and cue-target distance. Error bars represent corrected SEM. Brackets illustrate statistical differences at $p < .05$.

It is also worth noting that targets at distance 4 were again processed faster than other uncued targets, and also to an equal extent compared to targets at the cued feature (as there was no statistical difference between distance 1 and 4). Therefore, the counterintuitive trend observed in Experiment 2.1 was replicated, suggesting that at least one of the discussed possibilities (cue directionality, stimuli symmetry, and hemifield effects) may indeed be taking place.

For the *two objects* condition reaction times were also affected by distance, $F(2.22, 40) = 63.28$, $p < .001$, $\eta_p^2 = .78$, but variations in the scale of the objects had no effect on performance, $F(2, 36) = 2.16$, $p = .130$, $\eta_p^2 = .18$, and there was no interaction, $F(3.68, 66.21) = 1.44$, $\eta_p^2 = .07$, suggesting again that the effect was scale-invariant (Figure 19). As before, the effect of distance reflected an object-based pattern, such that reaction times were consistently fastest for the cued feature (all $p_s < .001$) and followed a non-monotonic variation for uncued features. Responses to targets at distance 2 (uncued object) were the slowest ($p < .001$ for the comparison with distances 1 and 3, $p = .003$ for the comparison with distance 4). Targets at distance 3 (cued object) elicited faster reaction times than targets further away at distance 4 (p = .007). Therefore, in this case the counterintuitive facilitation for distance 4 was still evident, though it was not of the same magnitude as for distance 1.

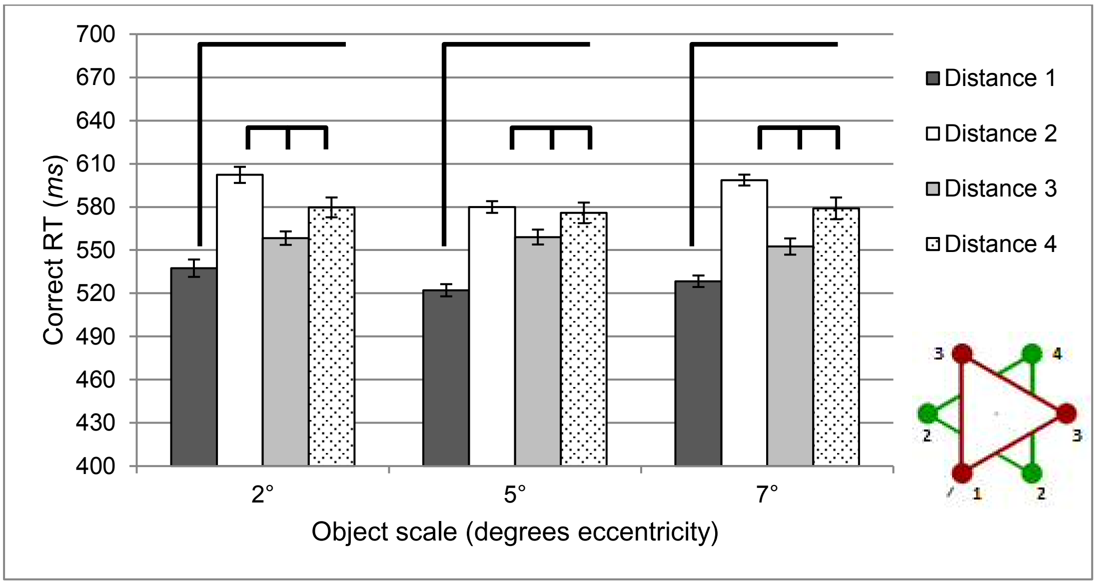

Figure 19: Experiment 2.2 mean correct reaction time for the *two objects* condition as a function of object scale and cue-target distance. Error bars represent corrected SEM. Brackets illustrate statistical differences at $p < .05$.

The data clearly suggest there was an advantage for processing targets integrated within the same object as the cued feature, even if they were spatially more distant than targets forming parts of the uncued object. Although reaction times for distance 4 were faster than reaction times for distance 2, given the potential confounding effects associated with this location, it can be concluded that the object-oriented selection pattern from Experiment 2.1 was successfully replicated. Also, the data clearly supports the influence of object structure, as there were no trends towards spatial separation effects. Most importantly, since there was no interaction between scale and target distance, the results support the hypothesis that visual selection is scale invariant. When a gradient of distance was introduced while keeping constant the probability of shared objecthood between cued and target feature pairs (i.e. same perceptual organisation for each distance level, but different spatial separation depending on the scale condition), performance varied only as a function of objecthood.

It should be noted, however, that the Bayesian evidence in favour of an object-based effect in the *one object* condition was not as compelling as in Experiment 2.1, as the preference for one model over the other was not as pronounced (no BF exceeded the value of 3). That the evidence in support of a monotonic gradient pattern is higher here than in the previous experiment is reflected in the general non-significant tendency of gradually increasing reaction times from distance 2 to distance 3, which can be observed in Figure 18. However, fluctuations are expected, and the evidence is nevertheless in favour of an object-oriented performance.

The lack of interaction between scaling and distance means that there was no difference in the pattern of reaction times even with up to 3.5 times larger Euclidian spatial separation (the scaling factor between 2° and 7°). In fact, in the *two objects* condition scaling had no statistical effect on performance whatsoever. Given the substantial variation in spatial separation, it may be expected that if spatial selection had taken place, it would have had a different effect on the pattern of responses within each scale, or at least a difference should have been evident between the largest and smallest scales. However, since the response patterns remained consistent, the effect is most likely due to the preserved relationship between the object features, which is the critical factor that was held constant.

Overall, the results so far provide strong support for a purely object-oriented and space-independent account of visual selection. However, it may be argued that this object-oriented effect was in fact due to selection being guided by the common colour of features within the same object, rather than object-centred selection per se. For example, in the *two objects* condition facilitation for distance 3 may be caused by the fact that the object feature at that location always shared the same colour as the cued feature. This is typically referred to as feature-based selection, where ‘features’ can be local object characteristics such as colour, orientation, size, etc., and the perceptual system selects items on the basis of such common characteristics (e.g. all items on the visual scene that have a red colour) (Freiwald, 2007). It should be noted, however, that colour is a characteristic of objects, and parts of the same object are likely to have similar colouring, so selection on the basis of common colour does not undermine the current results or object-based selection in general. The purpose of the subsequent experiment was to remove common colour as a cue to shared objecthood, in order to identify if it was a potential ‘confounding’ factor for the object-based results observed so far. It is also an opportunity to provide an indication of the extent to which such colour cues contribute to the object-based effect.

Experiment 2.3: Colour-based versus Object-based Selection

Visual selection is frequently described as operating based on a combination between space-based, object-based, and feature-based mechanisms (Freiwald, 2007). Feature-based selection suggests that attending to a specific feature, e.g. colour, orientation, or motion direction, can lead to the automatic enhancement and increased sensitivity towards this feature across the visual field, as indicated by neural activity indices (Maunsell & Treue, 2006; Serences & Boynton, 2007; Treue & Martinez-Trujillo, 2007). As such, feature-based selection also has a space-invariant property. It diverges from object-based selection in the sense that selecting a specific feature, e.g. colour, can be done across different objects, i.e. the processing benefit of selecting this feature is valid for both task-relevant and task-irrelevant objects (Wegener et al., 2014). In behavioural terms, this translates, for example, into longer time to find a target in a field of distractors when the target and distractors share a common feature critical for the task (e.g. the orientation of a target line needs to be reported, while both target and distractors are tilted to the left), compared to conditions where the searched target is unique (Mounts & Melara, 1999). Within the current context, the implication is that the object-based effects in the previous two studies may have been due to selecting the colour of the cued feature, not the object itself.

The present work supports the idea that there can be one parsimonious account for the mechanism of visual selection, suggesting that it is fully object-based and evidence for space-based or other dimensions can be accommodated within this object-oriented perspective. In terms of feature-based selection, it can also be the case that common features (be it colour, orientation or another dimension) between two or more objects can influence the probability gradient of these objects being in fact one and the same. In other words, there is no need to complicate matters by separating feature-based from object-based effects, as common features between stimuli simply contribute to increasing the probability that they should be equally prioritised by the perceptual system due to the increased likelihood that they are part of the same object. In this sense, feature commonality may contribute to this object-belongingness gradient to a different extent, depending on what other object belongingness cues there are on the scene (e.g. physical or perceived connection between stimuli), and also how the specificities of the task map onto these cues.

In the current experiments, colour commonality between object features was deliberately used as a cue to shared objecthood, so it is expected that it contributed to the obtained object-based pattern of performance. Nevertheless, Experiment 2.3 was specifically designed to test if colour-guided (‘feature-based’) selection alone was responsible for the object-based effects in the *two objects* condition. With the stimuli used so far, the features within each object always had the same colouring, whether it was the *one object*, or the *two objects* condition. Consequently, selection simply on the basis of colour, without any regard for objects per se, may be responsible for the pattern of results.

This issue was addressed by using the *two objects* layout of the stimuli, but both triangles were of the same colour – either red or green. To avoid perceptual regrouping into a single object, i.e. in order to try to maintain the flat within-object belongingness gradient as intended by the design of the original stimuli, the two triangles were made distinct by change in the frame thickness (Figure 20). It was predicted that the non-monotonic, object-based pattern of performance would be replicated since the object-level perceptual organisation and thus the object belongingness of the features is unchanged. The only difference is that now the principal cue used to distinguish one object from the other is the manner in which the features are connected, i.e. the thickness of the connecting contour. However, it may be the case that the object-effect would be of a different magnitude, most likely weaker, since colour is no longer available as an additional cue to object differentiation.

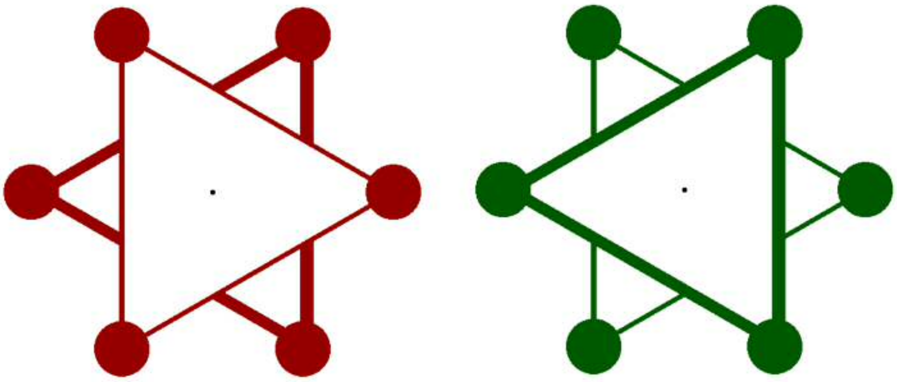

Figure 20: Illustration of the single colour stimuli used in Experiment 2.3.

Method

Participants

Twenty-two participants (1 male, mean age of 21.45, *SD* = 4.87) took part in return of £4 payment. They were recruited via Cardiff University Experimental Management System (EMS).

Stimuli and Apparatus

The characteristics of the stimuli were similar to the ones used in Experiment 2.1 *two objects* condition, with the exception that both objects had the same colour, either red or green, and the frame thickness of either the top or the bottom object was decreased to 4 pixels width. All other aspects and equipment remained unchanged.

Design and Procedure

Since target distance was the only independent variable, this was a simple four-way repeated measured design. As a result, the number of trials per target location was increased. However, the ratio of cued and uncued targets was preserved, so that the predictability of the cue was identical to the previous two experiments. Therefore, there were 432 trials in total, 132 of which were cued (30.5%) and the remaining 300 trials were equally spread between the 5 uncued features (60 trials, or 13.9% each). These trials were structured in 4 blocks of 108 with self-timed break in-between. The proportion of cued and uncued trials was identical in each block, but the order was random. SOA varied randomly within the 300-350 ms range as in Experiment 2.2.

The procedure was identical to Experiment 2.1. On half the trials the position of the thin-framed object was on top, and the colouring of the objects was also evenly distributed between the total number of trials.

Results and Discussion

The usual trimming procedure for reaction times outside the 200-1200 ms range resulted in the removal of less than 1% of trials. The familiar object-based effect was replicated, such that there was a main effect of target distance, $F(2.21,46.4) = 13.45$, $p < .001$, $\eta_p^2 = .39$, reflected in the fact that target events at distance 1 were discriminated faster than target events as distance 2 ($p < .001$), distance 3 ($p = .007$) and distance 4 ($p = .012$). Also, responses to targets integrated within the uncued object at distance 2 were slower than for those at distance 3 on the cued object ($p = .007$) (Figure 21). Distance 4 responses did not differ from responses to any other uncued targets. Therefore, the pattern of performance followed a non-monotonic function reflecting the object structure. There was again some counterintuitive facilitation for targets at distance 4, which were nevertheless not responded to faster than luminance changes at the cued feature, but were as privileged as targets associated with uncued features within the cued object (i.e. distance 3).

Figure 21: Mean correct reaction time for the single-colour stimuli in Experiment 2.3. Brackets illustrate the statistical differences at *p* < .05. Error bars represent corrected SEM.

Given the replication of object-oriented selection, the previously observed object-based effect was not simply due to colour-based grouping, although this certainly had an influence given the smaller (but still considerable) effect size. In fact, the magnitude of the difference between distance 2 and distance 3 in the single colour version of the *two objects* condition is similar to that obtained in Experiment 2.2 where each object had unique colouring (mean difference of 11.71 ms, *SE* = 3.1, and 10.2 ms, *SE* = 4.9, respectively). This is not surprising, since both feature (in this case, colour) similarity and physical connectedness are characteristics of within-object stimuli, and such characteristics can contribute towards perceptual object formation in additive ways, albeit with different weighting as connectedness is considered more powerful (De Winter & Wagemans, 2006).

It is worth noting it is often the case that studies examining the impact of feature-based versus object-based selection use an array of identical stimuli as objects, e.g. a field of equally spaced bars or dots (e.g. Mounts & Melara, 1999). In this case, a common feature such as colour is very probable to increase the likelihood of shared objecthood of the otherwise identical ‘objects’. When an irrelevant distractor object shares features with a target object, this feature similarity is likely to increase the shared objecthood probability, thus making the object-based gradient less steep and leading to the graded selection of the distractor in conjunction with the target. However, when these objects are made clearly distinguishable as different by some other property that can contribute to keeping the probability of shared objecthood low, e.g. connectedness, then colour similarity should be less likely to affect selection, as the current experiment demonstrated. That is, depending on the perceptual organisation of the display, feature commonalities between items may cause different levels of ambiguity regarding the relationship between these items, which in turn affects how they are selected. This is not to say that the perception of objects and its consequences are simply the result of a congregation of features with different weighting regarding object formation. It is the case that these cues to objecthood, such as Gestalt grouping principles, have their origin in the characteristics of natural objects (Strother & Kubovy, 2006). Therefore, perceptual objects and object-based selection are not secondary or emergent phenomena, but a principal influence on range of action-perception mechanisms.

The current results suggest that feature-based and object-based selection are likely to reflect one and the same process, which is ultimately the result of object-level regularities found in the natural environment. This also corroborates the results of the previous two experiments, providing strong support for a purely object-oriented and space-independent account of visual selection. It may be that the inconsistencies with previous research, suggesting that selection is the result of a combination of space, feature, and object-level factors, are due to methodological issues concerning the nature of the stimuli. The currently proposed parsimonious account suggests that there are no other effects than those resulting from object-level factors, but their expression may be misinterpreted as due to space depending on how the nature of the stimuli and their integration with each other relates to the probability that they are parts of the perceived objects on the scene. Low probability leads to weaker selection, and if probability drops with distance, then it may appear as if effects of spatial separation are observed. However, if the probability remains constant, as is the case with the current stimuli, so does the level of selection performance, and this is observed across a variety of scales and sizes. To explore this hypothesis it is necessary to test if the monotonic within-object performance pattern breaks down when the probability of object belongingness of the targets relative to the object is decreased. This was the aim of Experiment 2.4.

Experiment 2.4: The Emergence of Space

The key argument supported by the empirical work so far is that visual selection is exclusively object-based, and effects of spatial separation, such as gradually decreasing response facilitation with increasing cue-target distance, are in fact emerging as a result of object-level factors. Specifically, selection is guided by factors that affect the probability of the target being part of the behaviourally relevant stimuli, which in experimental paradigms are typically a cued object or location, or a set of (presumably) independent objects. As already demonstrated in natural scene statistics, the environment consists of a variety of objects, and the closer two points are together, the more likely they are to belong to the same object (Ruderman, 1994; 1997). Therefore, if there is a high level of uncertainty regarding what a transent targets belong to, targets proximal to the currently relevant visual entity (e.g. the cued object) are more likely to be selected for privileged processing than those further away. As a consequence, performance may appear to follow a spatial gradient. Importantly, this gradient is the result of probabilistic object-level selection, rather than factors relating to space and stimuli distance per se.

So far it was demonstrated that a well integrated target with low ambiguity regarding its object belongingness leads to space-invariant and scale-invariant selection. The current hypothesis states that reducing the probability of same-object belongingness between cued and target location pairs can lead to the manifestation of what looks like a spatial mechanism, i.e. spatially graded performance. Therefore, it is necessary to use the same spatial locations and methodological parameters of the experiments so far, while decreasing the integration of the target event with the cued object. If a spatial gradient is observed under such circumstances, it can be proposed that the so called space-based effects are nothing more than an emergent property of the structure of objects on the visual scene. In other words, whether spatial gradients or flat performance is observed, it is always a function of object-level factors. These factors relate to the probability that the stimuli are parts of whatever objects are perceptually available on the scene.

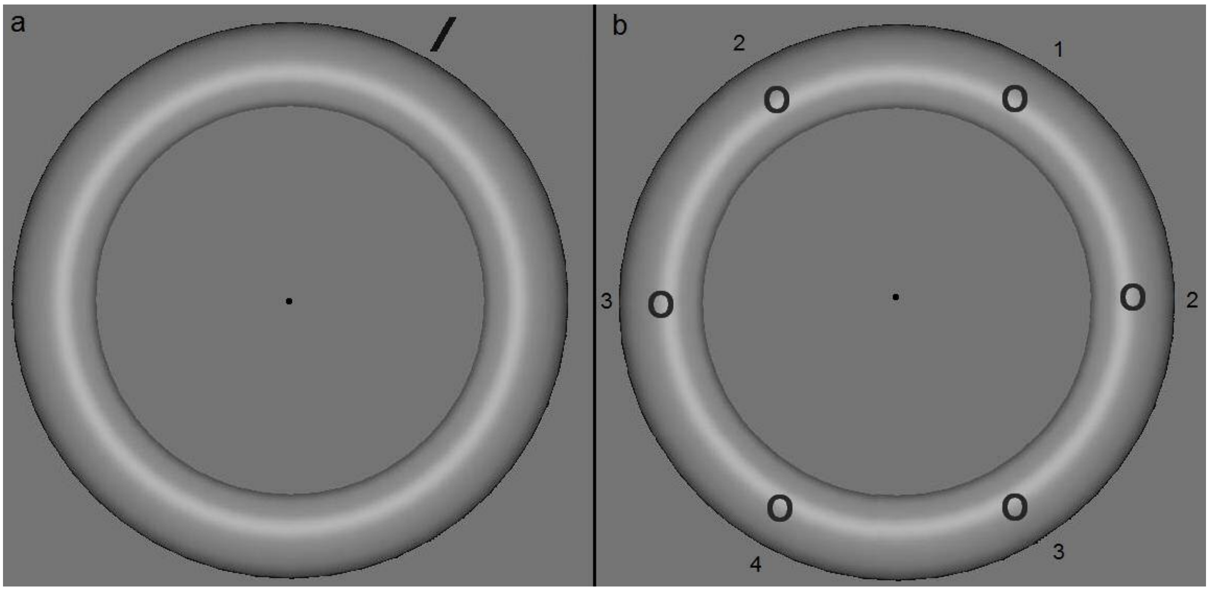

Figure 22: Illustration of the stimuli used in Experiment 2.4; a: ring object with cue (black line); b: potential target locations indicated with target ‘O’ (during the trial only one target is presented at a time). Numbers illustrate the cue-target distances relative to the cued location.

Experiment 2.4 was designed to establish if the purported space-based effects for visual selection are in fact an emergent property of the perceptual organisation of the visual scene. If this is indeed the case, whether a spatial gradient is observed should be controlled by manipulating object-level characteristics (e.g. target-object integration and perceptual organisation of the object features), while keeping the spatial characteristics of the stimuli unchanged. The key principle here was to test whether the spatial and temporal properties of the cue and target used in the three previous experiments can induce a spatial gradient pattern of performance when the targets do not appear to be well integrated with the cued object, i.e. the probability of shared objecthood for cued and target feature pair is reduced. For the purpose, an object stimulus and targets very similar to those used in Hollingworth et al. (2012) were adopted, as they generated evidence that selection is spatially graded within the same object. Specifically, the current stimuli consisted of an apparent three-dimensional ring centred at fixation (Figure 22a), with targets constituting transient appearance of letters X or O, superimposed on the object surface (Figure 22b).

Conducting a study where the experimental parameters from the previous three experiments are preserved, while using a level of target-object integration uncertainty similar to Hollingworth et al., would provide reliable evidence whether there are any confounding factors in the non-physical aspects of the stimuli used here that may be responsible for the previously observed space-invariant selection. That is, the methodological aspects such as target spacing, timing, cue occurring outside the object, cue predictability, and the number of target locations are the same as in the previous experiments from this chapter, while the only key difference is the perceptual organisation of the stimuli, which is the same as Experiment 2 in Hollingworth et al. (2012). Therefore, if a spatial gradient in performance emerges only by changing the object-level characteristics of the current stimuli, then it can be concluded that the main factor for this emergence is the target-object belongingness probability. If, on the other hand, no gradient is obtained under these modified conditions, this may be an indication that there is something about the non-physical characteristics of the stimuli that was critical for observing the previous object-based effect, suggesting a possibility that it was not purely object-based, or at least not directly dependent on the perceived target-object integration.

The reason for using these types of stimuli is because they provide perceptual conditions under which effects of spatial separation have already been demonstrated, albeit with slightly different methodological details[3]. Specifically,

---

[3] Refer to Appendix 2 for details on an additional study which used the same object stimuli as Experiments 2.1-2.3, combined with X/O targets. There was only a trend towards a spatial gradient. It was reasoned that the perceptual target-object segregation was not potent enough, resulting in the decision to use similar stimuli to Hollingworth et al. (2012), which were already successful in establishing spatial separation effects within the same object.

the targets used by Hollingworth et al. (2012) were presented for 70 ms followed by a mask, with five possible cue-target distances. As a comparison, in the present experiments the target duration was 100 ms and unmasked, occupying one of four potential cue-target separations. Also, the cue lasted 100 ms and did not occur on the surface of the object, while in Hollingworth et al.'s study the cue was an apparent three-dimensional bulge on the body of the ring presented for 50 ms. In addition, the SOA for Hollingworth et al. was 120 ms, while the main SOA in the current experiments was approximately 300 ms. Finally, the spatial gradient for reaction times was measured with a simple onset detection task, while here a target discrimination procedure is adopted.

Given the current methodological modifications, it was hypothesised that if spatial separation effects originate from object-level factors, then Experiment 2.4 should result in spatially graded performance. On the other hand, if reaction times here do not vary with distance, it may be the case that there is a potentially confounding factor other than object-level influence contributing to the space-invariant pattern observed in Experiments 2.1 to 2.3. In short, comparing the results from the current experiment with the previous three experiments should give an indication if it is object-based, rather than space-based effects, that are responsible for effects of spatial separation observed in previous studies.

Method

Participants

Twenty-nine participants (6 male) took part in the study in return of partial course credit (mean age = 19.69, *SD* = 2.17). All participants had normal or corrected-to-normal vision, and had not taken part in any of the previous experiments.

Stimuli and Apparatus

The equipment was identical to Experiment 2.1. The three-dimensional ring shape was generated using Blender 2.72 Open Source software. The object had a radius of 5.73° and 1.64° thickness. The targets were 0.5° x 0.5° letters X and O with a 4 pixel stroke and black grey monochromatic colouring (RGB = 40), centred on the body of the object. The properties of the cue and the spatial separation of the targets were the same as in Experiment 2.1. The peripherally displayed reminder letters mapped to the response buttons (now corresponding to X and O) had the same size properties as in the previous experiments, but were moved to the bottom of the screen (i.e. their *y* coordinate was increased relative to the centre of the screen) to minimise any ‘flanker’ interference when identifying the targets.

Design and Procedure

SOA was not manipulated, but it varied randomly between 300 ms and 350 ms. Due to the nature of the experiment, only a *one object* condition was used, and the only independent variable of interest was target distance. Therefore, this was a four-way repeated measures design. The proportion of trials per target distance was the same as in Experiment 2.3. There was an equal amount of X and O targets, and the identity of the target varied at random for each trial.

The procedure was identical to the previous experiments with the following exceptions. Participants had to identify whether the target was a letter X or O by pressing the correspondingly labelled buttons on a standard keyboard. The labels were placed on the L and D keys and were counterbalanced between participants. The duration of the target was decreased to 80 ms (as opposed to 100 ms in the previous experiments) because this timing was judged to be more comparable in difficulty with Experiments 2.1-2.3, i.e. a longer exposure may lead to ceiling effects.

Results and Discussion

Responses faster than 200 ms and slower than 1200 ms were excluded from the analysis (< 1% of the data). Performance was affected by target distance, $F(3, 84) = 9.23$, $p < .001$, $\eta_p^2 = .25$, such that responses to targets at the cued location were faster than responses to targets at distance 2 (p = .048) and distance 3 (p < .001), and also responses to targets at distance 2 were faster than responses for distance 3 (p = .047) (Figure 23). Responses to targets at distance 4 did not differ statistically from any others. This pattern of performance indicates the presence of a spatial gradient of facilitation centred at the cued location and spreading up to distance 3.

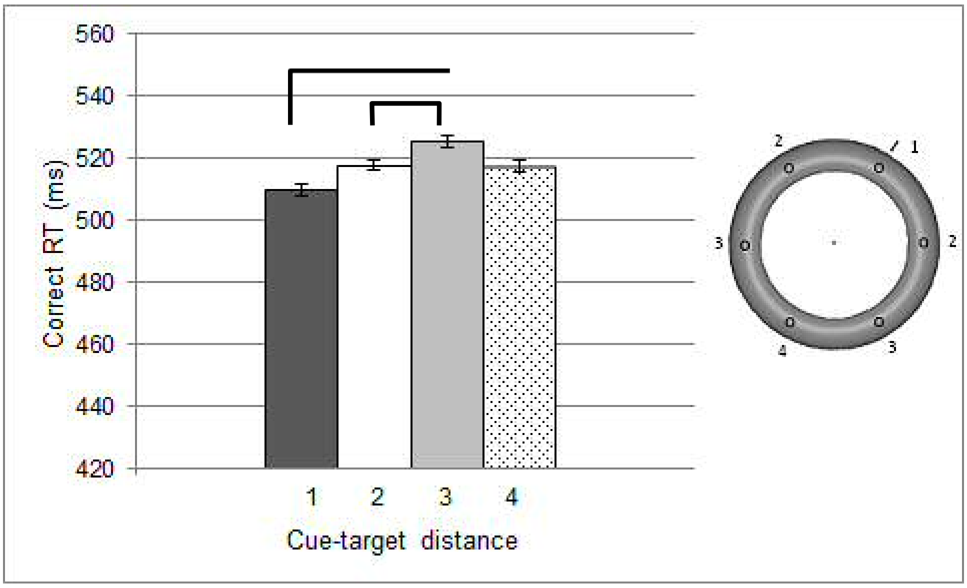

Figure 23: Mean correct reaction time as a function of cue-target distance for Experiment 2.4. Brackets illustrate statistical differences at p < .05. Error bars represent corrected SEM.

Considering the potentially problematic nature of targets at distance 4 discussed previously, the focus is on the significance of the performance difference between distances 2 and 3, which is now in the opposite direction to the *two objects* conditions in the previous three experiments, i.e. distance 2 < distance 3, reflecting a spatially graded pattern. The additional Bayesian analysis corroborated this evidence, suggesting that a gradient pattern of performance was 9.06 times more likely than a flat pattern without distance variation. In fact, the latter had a BF value of 0.11 against the possibility of a gradient, suggesting substantial support that speed of target discrimination was indeed a subject of a monotonic increase in cue-target distance. In terms of the counterintuitive facilitation for the most distant location, it should be noted that a similar pattern can also be observed in Hollingworth et al.'s (2012) data, albeit not as pronounced. In any case, comparing the data from Experiment 2.4 with the data from the three previous experiments clearly indicates that changing the relationship between targets and object had a measurable effect on performance, and this effect favours the possibility that space is not the medium of visual selection.

Given that the spatial separation and locations of the target stimuli, as well as the nature of the cue were similar to the ones in the previous three experiments, the outcome of Experiment 2.4 strongly suggests that altering the integration of the target with the object can lead to a graded performance affected by cue-target spatial separation. Therefore, this result grants further support for the hypothesis that what is normally interpreted as genuine effects of spatial separation can in fact be due to variations in probability that the target is part of the cued object. As this probability was decreased in Experiment 2.4 compared to Experiments 2.1-2.3, the reaction times for responses towards targets within the same object appeared to vary as a function of their proximity to the cued location, which was also the most likely target location. However, it is critical to note that considering all results so far, it is not truly the proximity to the cued location that is the key factor, but it is the *probability* of object belongingness of the target. In other words, all points within a given scene are probabilistically part of some object(s) on that scene. Therefore, the perceptual mechanism needs to solve the problem of what points belong to which object. When the probability of given points belonging together is positively correlated with their proximity, a spatial gradient can emerge. In the previous experiments reported here such correlation was controlled for by the perceptual organisation of the stimuli, i.e. while the spatial separation increased monotonically, the object belongingness likelihood was varied non-monotonically.

Experiments 2.1 to 2.3: Composite Analyses

The results so far provide support for the hypothesis that visual selection is exclusively object oriented, and the expression of effects of cue-target spatial separation is in fact dependent on object-level factors, thus accommodating the notion of space-based selection within a parsimonious object-based account. Nevertheless, there are certain aspects of the results which may pose a challenge to this interpretation. One such aspect is the non-significant trend in the *one object* condition of Experiment 2.2, where performance may be seen as conforming to a spatial gradient of increasing reaction time as a result of increasing cue-target distance (refer to Figure 18). In addition, the BF value corresponding to a spatially graded performance for that condition was noticeably larger than in the other experiments, although it was still less favoured than the flat function model. Another potentially problematic factor is that for the *two objects* condition, data from distance 4 rarely conforms to what would be expected given a pure object-based selection, since it is often not statistically different from distance 3, or is associated with faster reaction times than distance 2.

Given that the most compelling results in favour of the object-based hypothesis were evident in Experiment 2.1 where the sample size was considerably larger than any of the remaining experiments, the issues mentioned above may in fact be result of insufficient power. However, this assumption can be rather problematic, since simply increasing the sample size and testing for the hypothesised effect would eventually result in confirmation of the hypothesis due to chance, i.e. as a consequence of Type I error (Armitage, McPherson, & Rowe, 1969; Wagenmakers, 2007). It should be emphasised that although the sample size for Experiment 2.1 was to some extent arbitrary, the remaining experiments used formal power calculations.

In order to strengthen the conclusions drawn so far, a composite analysis was performed where data from comparable conditions across experiments were pooled together and re-analysed. In this way, if there were any unsystematic variations in performance, they should be cancelled out, allowing for a cleaner measurement of the underlying effects. Also, these were data already collected for this purpose, so there is no bias relating to a 'stopping rule'. However, such analysis is still vulnerable to Type I error as mentioned above, but this issue is addressed by conducting both a standard repeated measures ANOVA, and BF calculations. This is important because Bayesian analysis is not affected by sample size in the same way as orthodox frequentist statistics are. If the alternative hypothesis is true, increasing the sample size has the effect of driving the BF value closer to 0, rather than increasing the chance of a false positive (Dienes, 2011). When Bayesian statistics are adopted, there is generally no need for a stopping rule or any corrections relating to multiple testing. Therefore, for the current purposes both frequentist and Bayesian statistics were conducted on the pooled data. In this case Bayesian analysis was applied to both the *one object* condition, with the same order restrictions as described earlier (distance 1 < distance 2 = distance 3), and also on the composite data from the *two objects* condition. The latter compared a gradient order restriction (distance 1 < distance 2 < distance 3) versus a non-monotonic object-based gradient order restriction (i.e. distance 1 < distance 3 < distance 2).

For the *one object* condition, data were collated from Experiment 2.1, SOA of 300 ms, and Experiment 2.2 with the values collapsed across scale due to the lack of interaction with distance. This resulted in a total sample of 51 participants. For the *two objects* condition the data were combined in the same way from Experiment 2.1 and Experiment 2.2, and also from Experiment 2.3, since it was only based on the *two objects* version of the task. Therefore, the total sample was composed of 76 participants. Separate four-way repeated measures ANOVAs were conducted for each condition. At a first glance it may appear inappropriate to perform a within-subjects analysis on data gathered from different experiments, but what is essentially tested is the effect of cue-target distance, which was a within-subject factor manipulated in the same way for all of these experiments. Furthermore, the individual analyses reported for each study revealed a similar pattern of results (supporting pure object-based effects).

In accordance with the results so far, the composite *one object* condition revealed a distance effect, $F(2.33, 116.58) = 30.98$, $p < .001$, $\eta_p^2 = .38$, which was due to a superior benefit for the cued feature (distance 1), as it resulted in faster reaction times compared to all other object features (all $p_s < .001$), and no further reaction time variations with increasing distance (Figure 24, left panel). In addition, the BF in favour of equality between distance 2 and distance 3 was 6.56, compared to 0.15 for a gradient pattern. For the composite *two objects* condition, spatial separation also had an effect, $F(2.55, 191.41) = 74.20$, $p < .001$, $\eta_p^2 = .50$, which reflected a consistent cueing benefit where targets at distance 1 produced the fastest reaction times (all $p_s < .001$). Most importantly, responses for the uncued object features followed an object-based pattern, such that distance 3 (associated with the cued object) elicited faster performance than distance 2 ($p < .001$) and distance 4 ($p = .001$), which correspond to features on the uncued object (Figure 24, right panel). Performance for targets at distance 4, however, was faster than for distance 2 ($p = .019$), demonstrating again the counterintuitive facilitation for the most distant feature, although in this case it was not more privileged than targets on the cued object (i.e. distance 3).

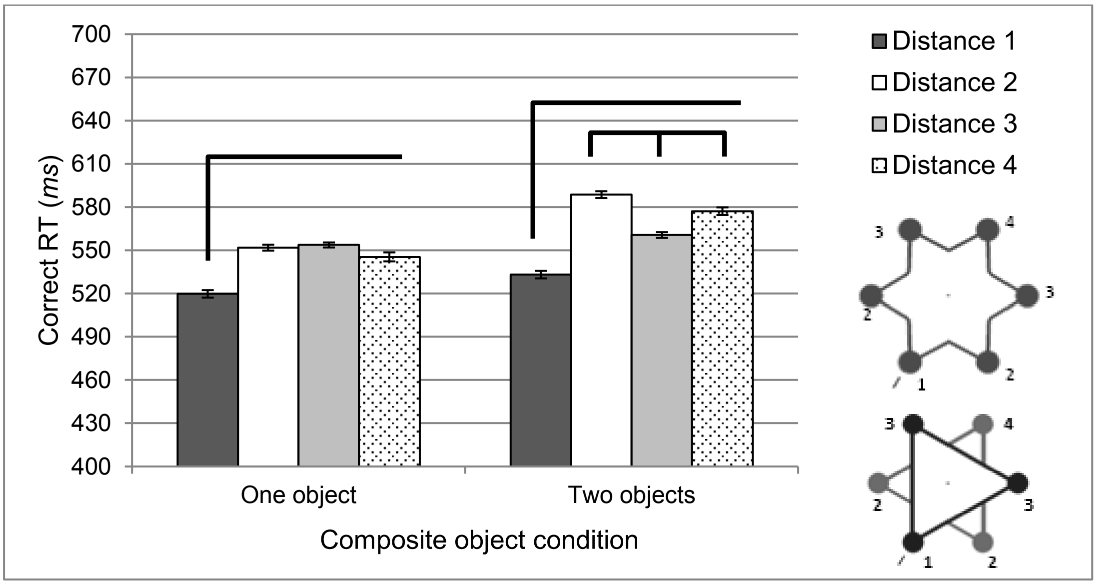

Figure 24: Mean correct reaction times as a function of object condition and target distance, combined across Experiments 2.1-2.2 (*one object*) and Experiments 2.1-2.3 (*two objects*). Error bars represent corrected SEM, and brackets illustrate the statistical differences at $p < .05$.

In terms of Bayesian evidence for the *two objects* composite condition, since all of the 10000 samples of the posterior distribution were consistent with the non-monotonic object-based restriction, and none followed a gradient pattern, the final BF value for the gradient model was calculated to be 0, while the non-monotonic model was supported to infinity. Since such level of certainty is rarely to be adopted, especially in statistical tests, it can be assumed that a gradient distribution is not completely impossible, but it is highly unlikely (what would be portrayed as "$p < .001$" in frequentist terms). Overall, the pattern of responses does not suggest any graded influence due to monotonic spatial separation in either the *one object* or *two objects* condition. The evidence again points to the primacy of objects over space, supporting that selection is purely object-oriented.

In summary, the composite analysis for *one object* and *two objects* conditions corroborates the main findings that visual selection is object-oriented and effects of spatial separation are not evident when the perceptual organisation of the stimuli affords good target-object integration and controls for the confound between probability of shared object belongingness and spatial proximity.

General Discussion

Together, the four experiments in the current chapter demonstrate that visual selection is object-oriented. The manipulation of spatial separation between a cue and a subsequent target, which usually results in a gradient pattern of gradually decreasing facilitation, did not affect speed of response for the discrimination of a visual event. Instead, performance was influenced by whether this event was perceived as part of the cued object, and also whether it occurred on the most expected location as indicated by the cue. This was evidenced by a distribution of reaction times which favoured targets within the cued object, regardless of their distance from the cued object feature, and also a consistent cueing effect (Experiments 2.1-2.3). Importantly, when the probability of the target belonging to the cued location was decreased, i.e. the perceptual integration between the object and the target was more ambiguous, selection was no longer space-invariant and performance displayed a gradual decline as cue-target distance increased (Experiment 2.4).

Therefore, the results provide evidence that spatially graded performance is not a sign of space-based selection per se, but it is the result of object-level perceptual organisation factors, and one such factor is the level of target-object integration. The importance of the perceptual integration of the target for obtaining a standard object-based effect, e.g. for the two rectangle paradigm or flanker tasks, has already been emphasised in previous research (e.g. Richard et al., 2008; Zhao et al., 2013). However, an important and novel finding on the basis of the current experiments is that it is such perceptual organisation factors which can lead to the presence or absence of spatial gradient effects. In other words, effects interpreted as relating to space are in fact object-related. In turn, this supports the hypothesis that space-based effects are not caused by spatial factors per se, such as the spatial separation between cue and target or distractor and target. The reason it appears that spatial proximity plays a key role for selection is because of its correlation with probability of object belongingness between the stimuli (Oyama, 1961; Ruderman, 1997). The work in the current chapter demonstrated that when visual information is organised in a way that does not support a positive correlation between object belongingness probability and distance between stimuli, performance no longer reflects the typical space-based effect.

Based on the present results, it can be suggested that what happens directly with the initial stimulus presentation is the perceptual segmentation into objects, so within the current context either a single star or two overlapping triangles. Next, following the cue presentation, the cued feature is selected and prioritised since it is expected that it is the one most likely to change. Importantly, there is concurrent selection of all other features which are perceived to be part of the same object as the cued feature, while features belonging to a different object are not prioritised. Consequently, the readiness with which the target event is responded to is directly affected by whether it is part of the cued object. A direct consequence of this process is that since it is the objects that get selected, spatial separation makes no difference to performance. This can be observed in the lack of within-object gradients and the lack of distance effects when the scale of the objects is varied. This is also in accordance with the scale-invariance of statistical properties of objects in the natural scene (Ruderman, 1997).

How can the appearance of a spatial gradient in performance in Experiment 2.4 be accommodated within this account? The poorly integrated targets do not enjoy the same status as the luminance change targets of the objects in Experiments 2.1-2.3. Since visual selection is object-oriented, they may be coded in relation to the level of probability that they belong to the same object as the cued feature, and this probability decreases with distance, leading to the observed gradient effect. Importantly, when the perceptual system is faced with low level of ambiguity regarding the status of the target (so either within the cued object or not), spatial separation does not matter.

The possibility for equal prioritisation of information within a cued object is supported by neuroimaging evidence suggesting that with the use of the two rectangle paradigm (Egly et al., 1994), cueing one end of the rectangle results in automatic neural enhancement to an equal extent throughout its surface prior to target presentation (Müller & Kleinschmidt, 2003). In other words, the whole object is selected and prioritised relative to the uncued object, which fits well with the current results in terms of the neural mechanism of object-based selection. Regarding the emergent gradient effect demonstrated in Experiment 2.4, there are no studies to date which can be used to explain it from a neural perspective. Most paradigms test performance for two equidistant locations from the cued location/feature, which belong to separate objects (i.e. standard two rectangle layout), and this type of setup does not afford the test of within-object spatial separation effects. In any case, the current explanation for the emergent spatial gradient is well accommodated within the study of natural scene statistics, since the probability of two points belonging to the same object within an image drops with increasing distance (Ruderman, 1997). Therefore, both spatial separation effects and a flat performance function without distance variation can be accommodated within a gradient mechanism – but it is an *object-based* gradient mechanism linked to the probability of object belongingness between the stimuli on the visual scene.

In terms of the factors that can influence this object belongingness probability gradient, it is possible that all of the information on the visual scene interacts to jointly determine the object structures in a probabilistic fashion. This results in biasing the pattern of prioritisation towards or away from what appears to be space-based or feature-based selection. That is, since within-object pixel correlations (for natural image analyses) are close to 1 (Baddeley, 1997), then there may be a tendency to select two (or more) stimuli of identical colour or another similar property, and treat them as the same perceptual unit (object). This would then result in the so called feature-based attention, and also in the interaction or additive effects of feature-based and object-based attention. For example, in the two rectangle paradigm, if the objects are of the same shape or colour, the reaction time difference between uncued-same object and uncued-different object targets is less pronounced than when the two objects are made more dissimilar (Kravitz & Behrmann, 2011). In other words, the magnitude of the within-object benefit is not as strong when the ‘different’ objects share identical characteristics. Although Kravitz & Behrmann (2011) argue that this outcome is due to an interaction between object-based and feature-based attention, it can be explained by object-based selection alone.

A relevant point is the ability of the perceptual system to readily estimate average statistics relating to the properties of similar stimuli (e.g. average size or colour of a set of circles), known as ensemble statistics (Alvarez, 2011; Ariely, 2001; Chong & Treisman, 2005). This phenomenon is also evident in visual short-term memory, such that when a property of a stimulus needs to be remembered, e.g. colour or size, the memory at recall is biased towards the average colour or size of all similar stimuli that were present during the study phase (Brady & Alvarez, 2011). For example, when the specific colour hue of a memory probe has to be reported at test, this could result in reporting a hue which is very close to the average of all similarly coloured stimuli that were present during the memorising stage of the task. Again, this potentially indicates the stimuli sharing a common characteristic, such as all circles with a shade of blue, were selected and remembered as an ensemble.

Within the context of the current studies, Experiment 2.3 demonstrated that when the probability that given stimuli are parts of the same object is kept high by, in this case, connectedness cues, then removing another commonality between these stimuli (colour) does not affect the strength of the object belongingness probability. However, it may be the case that different cues to objecthood differentially affect the likelihood to select two or more points as a single object, but the key premise is that whatever those cues are (e.g. symmetry, perceived continuity, feature similarity, etc.) and the strength of their additive effect, they can be traced in the structure of objects in the natural environment.

In sum, the results so far make a strong case for object-oriented, scale-invariant visual selection. The strength of selection is modulated by a likelihood gradient that the task-related stimuli, and stimuli in the visual field in general, are integrated as one or more objects. Therefore, the perceptual organisation of the stimuli and the strength of target-object integration in experimental tasks can lead to different performance patterns ranging from a flat function, non-monotonic variations, or a linear gradient, depending on what object-level information is available. Importantly, this evidence is in line with data from natural scene statistics which exhibit the same characteristics. In turn, this corroborates the fact that selection has evolved to provide for the needs of the organism in accordance with its natural environment, and these needs pertain to making decisions and performing actions towards objects (Simoncelli & Olshausen, 2001).

A global point of importance is that experimental paradigms, hypothesis testing, and result interpretations are often limited by reasoning which is not formally based on statistical regularities of the environment, when in fact perceptual mechanisms are directly influenced by these regularities (Geisler, 2008). As a result, research can often lead to interpretations which do not provide a realistic account of the studied constructs. The studies here demonstrate that visual selection can exhibit the same characteristics as natural scene statistics, i.e. it is scale-invariant and object-based. Previous interpretations of spatial gradients relying on the role of space do not take into account these aspects, and therefore cannot readily explain the effects observed here, namely that space may simply be an emergent property of the perception of objects.

Chapter 3

Object-Based Perception and Visuo-Spatial Short Term Memory

Introduction

The aim of this chapter is to investigate the effect of object-level factors within the context of visuo-spatial short term memory, as it is an aspect of visual cognition closely linked to visual perception. Visuo-spatial short term memory (VSTM) is the system supporting key cognitive processes, such as mental rotation (Prime & Jolicoeur, 2010), visual imagery (Borst, Ganis, Thompson, & Kosslyn, 2012), and orientation and navigation in the environment (Baumann, Skilleter, & Mattingley, 2011). Generally, VSTM is associated with maintaining and processing online visual and spatial information when it is no longer available in the immediate environment (Baddeley, 2010). However, it is not only related to higher order processes but can also be linked to lower level perceptual mechanisms, such as trans-saccadic integration (Prime, Vesia, & Crawford, 2011).

There is considerable evidence that short term memory in general functions by recruiting the same neural mechanisms as those involved in action and perception, so within parietal and frontal brain regions there are qualitative and quantitative commonalities in neural activation patterns between active experience and interaction with the environment, and retention of the same type of information in memory (Gao, Li, Yin, & Shen, 2010; Ikkai & Curtis, 2011; Jonides, Lacey, & Nee, 2005; Postle, 2006). Therefore, there is a strong link between visual selection mechanisms (online perception) and VSTM (often regarded as post-perception), such that they can be affected and constrained by similar types of phenomena, or even considered to be one and the same process (Awh & Jonides, 2001; Theeuwes, Belopolsky, & Olivers, 2009).

An example of a common phenomenon is the cueing effect, which can also be observed in the context of VSTM with the use of a change detection paradigm. This is a standard assessment tool for the properties of VSTM, involving a brief presentation of a to-be-remembered display with various items (study phase), followed by a retention period and a probe display (test phase), which requires a same/ different judgement (Rouder, Morey, Morey, & Cowan, 2011). The test phase may represent an altered version of the study display, thus requiring a ‘different’ response, or an identical version for a ‘same’ response. In addition, the probe display may or may not contain an indication of which item participants are required to make the decision about (in the case of multiple items). However, whether the test phase limits the decision to a single item or not does not make a substantial difference to performance (Luck & Vogel, 1997). When cueing is introduced, the study phase is either preceded or immediately followed by a cue indicating a possible location for the probe stimulus at test.

Similarly to the visual selection experiments addressed in the earlier chapters, cueing leads to improved VSTM for items associated with the cued location (Berryhill, Richmond, Shay, & Olson, 2012; Matsukura, Cosman, Roper, Vatterott, & Vecera, 2014; Schmidt, Vogel, Woodman, & Luck, 2002). That is, participants have a higher rate of correct same/ different identification for the cued item than for uncued items (i.e. percentage correct responses). Therefore, visual selection and VSTM can be affected by similar factors (in this case a cueing manipulation), resulting in a similar outcome - prioritisation of a subset of information from the visual display. The question arises then, since the work so far indicated that visual selection is object-oriented, what are the functional units of VSTM, and to what extent is VSTM object-oriented too.

As a starting point it can be emphasised that there is a pronounced tendency to remember information from the display at an object level. For example, it is argued that the capacity of VSTM is around four features, such as colour or orientation, but this number can increase dramatically (up to sixteen) if these features are grouped into objects (Vogel, Woodman, & Luck, 2001; Xu, 2006). Specifically, it is harder to remember two types of features belonging to two separate sets of objects on the display, e.g. the colour of circles and the orientation of bent lines, relative to remembering an equal amount of information when it is contained within the same object – i.e. the lines and circles connected together (Xu, 2006). This effect is reminiscent of the object-based selection in divided attention tasks discussed earlier, where making judgements about two attributes of the same object is more efficient than judging two attributes of two separate objects (e.g. Lavie & Driver, 1996).

In addition, remembering a certain feature (designated as a target) of an object can lead to the automatic encoding of task-irrelevant features within the same object (Gao et al., 2010; Jiang, Chun, & Olson, 2004; Shen, Tang, Wu, Shui, & Gao, 2013). This effect is inferred from poorer VSTM (lower accuracy and/ or longer reaction time) for the target dimension of the object when the irrelevant aspect is also changed at the test display. For example, if a task requires detecting a change in the position of a gap on a ring object (or in other words, a change in the two-dimensional orientation of a semi-circle), performance is impaired if at test the colour of this object (i.e. a task-irrelevant object dimension) is also changed relative to conditions where this task-irrelevant dimension is constant (Gao et al., 2010). Also, change detection for the colour of an object is impaired when at test its shape it also changed (Shen et al., 2013).

It can be argued that this impairment following an alteration in a task-irrelevant object dimension is simply due to an additional change on the display, rather than an object-based effect. However, the evidence suggests that there is no difference between conditions where only the task-relevant dimension is changed at test (colour), compared to changing the irrelevant dimension only (shape) or both dimensions together (Gao et al., 2010, Experiment 4). Therefore, there was no additional reduction in accuracy due to changing two, as opposed to one dimension of the object. Moreover, a conjunction change of the irrelevant dimension, i.e. swapping the colours of the two midline planes of a triangle when a change in the orientation of the triangle is the relevant dimension, elicits similar performance to trials where the triangle colour remains constant. In comparison, changing the colours (as opposed to exchanging their position within the triangle) impairs accuracy (Gao et al.,2010, Experiment 3). It can be suggested that while in both cases there is an abrupt change on the display, the conjunction change preserved a higher probability of the triangle being the same object at test, while introducing new colouring makes it more likely that it is a different object, affecting change detection judgements for the task-relevant dimension (orientation). Such effects of automatically encoding the information within an object suggest that VSTM has an object-oriented propensity, in a similar manner to visual selection.

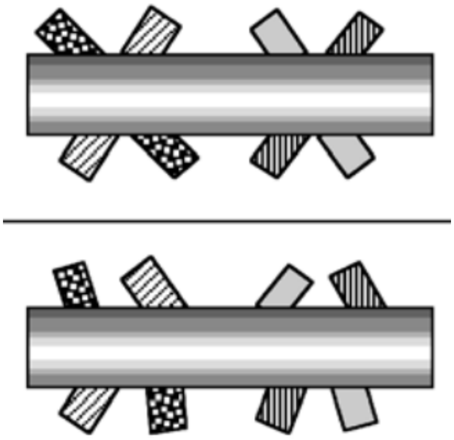

Figure 25: Illustration of the study display stimuli used in Walker & Davies (2003). Top panel: eight rectangles perceptually organised into four occluded objects. Bottom panel: perception of eight independent rectangles. Patterns indicate different colours. The top and bottom rows each contain the same set of four colours, but arranged in a different sequence, which had to be remembered by participants.

Similarly, detecting a change in the order of a set of colours may be substantially improved when the to-be-remembered colours are perceptually organised to appear contained within four as opposed to eight objects (Walker & Davies, 2003). This is demonstrated in a task where participants are shown a set of eight coloured rectangles arranged in two rows of four with a cylindrical occluder placed between the two rows, so there is no visible gap between them. Importantly, the rectangles are either oriented to look as four crossed objects (long rectangles) occluded in the middle (Figure 25 top), or eight independent, unconnected rectangles (Figure 25 bottom). At test the top or bottom row is fully occluded, and participants have to decide if the order of the visible parts of the coloured rectangles has changed. Change detection accuracy was superior when the study display consisted of four perceptually completed objects, as opposed to eight independent objects. This result corroborates the proposal that VSTM is encoded in terms of object structures, since performance was facilitated when fewer objects had to be remembered. It also emphasises the critical role of perceptual organisation.

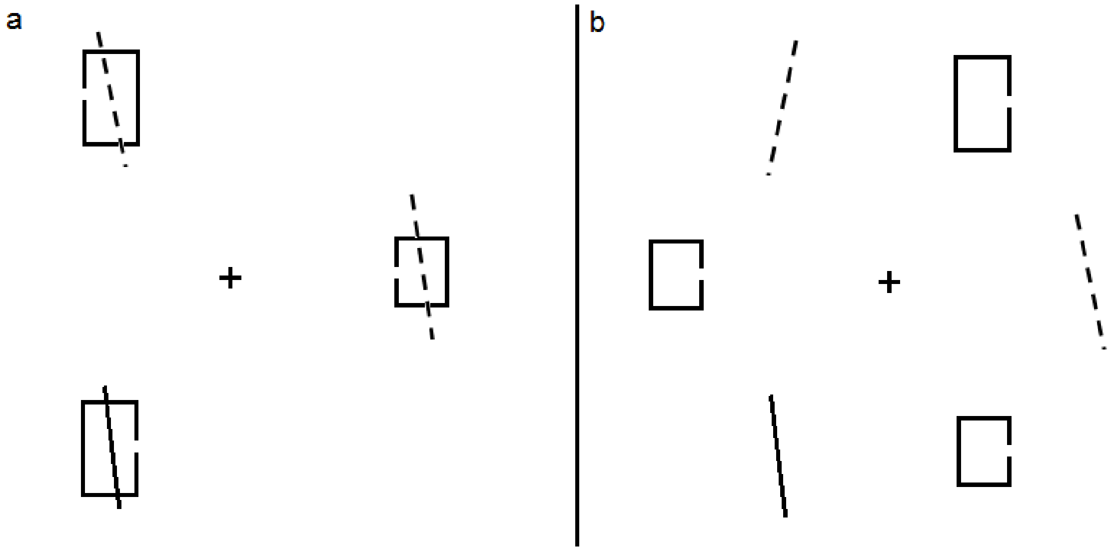

Figure 26: Reproduction of the stimuli used in Lee & Chun (2001); a: superimposed condition; b: separate condition. Both conditions contain six objects (three lines and three boxes), which also vary in colour (red or green) to enhance perceptual segmentation.

The object-based nature of VSTM, and its commonality with visual selection mechanisms, can also be observed within the context of the location-based versus object-based issue. Specifically, varying the number of locations occupied by objects while keeping the number of objects constant has no effect on VSTM (Lee & Chun, 2001). This is demonstrated with a variation of the divided attention paradigm (Duncan, 1984), where a box and line can be presented superimposed or spatially separated. In this case, six objects (three lines and three boxes) were presented either superimposed (occupying three locations, Figure 26a) or separated (occupying six locations, Figure 26b), always located at the same eccentricity along an imaginary circle. As in the original task, the boxes could vary in size and location of a gap on the contour, and the lines could vary in texture and orientation. At test, participants were presented with one of the six objects at its original location, and asked to detect a possible change in either of these dimensions. Response accuracy did not vary as a function of the spatial arrangement of the study display (i.e. superimposed or separated), suggesting that information was encoded in an object-based manner, rather than depending on the number of spatial locations, i.e. space-based encoding.

An important point stemming from the object-oriented nature of VSTM is that information from the display is often encoded configurally, i.e. the relational properties of the elements are also memorised, leading to an overall encoding of the global configuration of all items (Boduroglu & Shah, 2014; Timothy F. Brady & Tenenbaum, 2013). One implication of this is impaired VSTM when a single item is presented for recognition at test, relative to when it is presented in its original configural context with all other items on the display (Boduroglu & Shah, 2009, 2014; Mutluturk & Boduroglu, 2014; Patterson, Bly, Porcelli, & Rypma, 2007). Another consequence of configural encoding is that displacing non-target items at test leads to impaired VSTM for the target's original position, as the memory for its location is also displaced in accordance with the overall spatial configuration (Katshu & d'Avossa, 2014). In other words, the study display is often remembered as a holistic superordinate representation, which can have problematic implications for change detection methodologies that use only a single item as a probe at test, due to associated costs in parsing the holistic representation encoded in VSTM.

The key implication here is that the individual items on the display are not necessarily remembered in terms of separate, independent objects, but also as a global configuration where the relationship between the constituent items is also encoded. Also, this process is likely to be automatic, since a holistic representation has no strategic benefit in a task where memory for individual items on the display is probed. This is problematic if trying to assess capacity limits, since a single unit of information in VSTM is not readily identifiable and higher order aspects relating to perceptual organisation need to be taken into account (Brady & Tenenbaum, 2013). Moreover, these units may in fact have a dynamic nature, dependent on the incidental perceptual organisation, and the global task setting. Specifically, a display of four similar squares may be remembered as a single higher-order square, or as two rectangles, or some alternative configuration, depending on the symmetrical properties of the scene.

Thus, the same amount of visual information can lead to different perceptual outcomes and consequences for VSTM.

Given the powerful influence of perceptual organisation, i.e. the automatic tendency to group features into coherent objects (Katz, 1950; Wagemans, Feldman, et al., 2012), and the holistic global processing of information (Boduroglu & Shah, 2014), it can be suggested that any manipulation of the study or test displays can potentially lead to unanticipated effects on VSTM. Specifically, attempting to vary, for instance, the number of to-be-remembered items or their complexity can also result in altering the global properties of the display, i.e. how the items are perceived in relation to one another, and whether this affords the formation of higher order perceptual objects. Consequently, the measured VSTM property, e.g. capacity or precision, may be unintentionally affected, leading to invalid conclusions and misinterpretation (Orhan & Jacobs, 2014a, 2014b).

As an illustration of this point, the emergent properties of the display may act as confounding variables if not controlled for. For example, this may lead to concluding that under certain conditions VSTM can hold *N* number of items. However, if the display encourages perceptual restructuring of the presented, for example, six objects into fewer composite objects, then it may be challenging to quantify *N* (i.e. capacity), since the intended number of to-be-remembered objects does not equal the perceived number of objects. A relevant example is the colour-sharing effect, whereby VSTM capacity, as measured by colour change detection accuracy with varied set size (1-7 to-be-remembered items), is higher when the study display contains two items of the same colour relative to a study display where all items have unique colouring (Quinlan & Cohen, 2012). Importantly, the benefit of shared colour is not only evident when one of the duplicated items is probed at test, but it also translates to the uniquely coloured items from the same display (Morey, Cong, Zheng, Price, & Morey, 2015). This illustrates the powerful influence of perceptual organisation factors on aspects of VSTM such as apparent capacity limits.

The issues addressed so far relate to the tendency towards object-oriented coding of information, and how this tendency can influence VSTM. Accordingly, this chapter is focused on how VSTM is affected by the object-level perceptual organisation of the display. More specifically, investigating the possibility that the same quantity and spatial arrangement of to-be-remembered items can result in different VSTM performance depending on how these items are grouped into higher-order perceptual objects. This was of interest within the context of a cueing paradigm, since this is among the standard methods of studying visual selection, and for the purpose of consistency with Chapter 2. Moreover, as emphasised earlier, a cueing effect has already been established for change detection tasks (e.g. Berryhill et al., 2012; Schmidt et al., 2002), but there has not been a lot of focus on object-based effects for the uncued items, which is the key theme addressed here.

One exception is Woodman, Vecera, & Luck (2003), who used a cued change detection paradigm, where one of four locations on the screen was cued, followed by the study phase consisting of either four squares positioned at the corners of an imaginary square centred at fixation (set size 4, left panel of Figure 27), or a similar arrangement where an additional square was placed between either the two horizontal or the two vertical pairs of squares (set size 6, right panel of Figure 27). The latter condition results in the perception of two horizontal or vertical objects by virtue of proximity cues. Following a 900 ms retention interval, participants were shown the display again with one of the squares (the probe) surrounded by a black frame, indicating that a decision needs to be made whether this square has changed colour relative to the study display. Change detection was found to be superior when the probe matched the cued location, i.e. demonstrating a standard cueing effect. There was also no difference in change detection for probes equidistant from the cue in the set size 4 condition. Critically, in the set size 6 condition, performance was better for the probe perceptually grouped with the cued location, either vertically or horizontally. Therefore, the perceptual organisation of the stimuli led to an object-based benefit analogous to that observed in visual selection experiments (Chen, 2012).

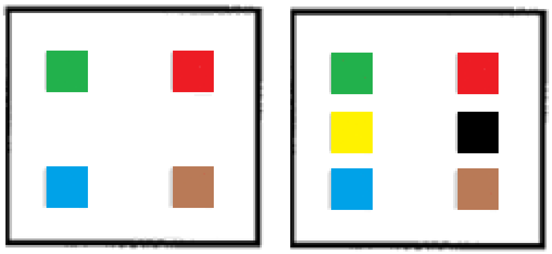

Figure 27: Reproduction of the study display stimuli used in Woodman et al. (2003). Left panel: set size 4 condition; right panel: set size 6 condition with vertical grouping. This is for illustration purposes only, the colours do not match closely the originals.

Although these results provide evidence for object-based coding in VSTM, the alterations used to manipulate the perceptual organisation of the display involve direct changes in the number and arrangement of the memory items (i.e. manipulations of set size). This may lead to confounding changes in the overall display complexity and perceptual load, which means that the differences in performance between the manipulated conditions (i.e. set size 6 and set size 4) may be at least partially due to other factors than the intended independent variables (i.e. changes in the perceptual organisation of the display). Therefore, it is more appropriate to assess the role of perceptual object formation under conditions that vary the perceptual organisation of the display without affecting the number or spatial arrangement of the target memory items, i.e. keeping the critical task-relevant information and the overall visual complexity constant between conditions.

In addition to the role of perceptual organisation, another issue addressed here relates to the measurement of VSTM. Change detection paradigms typically measure accuracy of same/ different judgements (e.g. percentage correct responses), while visual selection studies adopt in addition, or alternatively, a more continuous indication of processing – reaction time. Categorical yes/no or same/different responses to supra-threshold changes (e.g. colour, shape, or location) may not be sensitive enough to detect more subtle processes relating to VSTM representations and retrieval (Bays & Husain, 2008). Therefore, measuring the time that it takes to make a response regarding the memory probe may be indicative of aspects in the underlying structure which are subtle and not readily identifiable with a categorical measurement. Moreover, using reaction time as an additional dependent measure can be informative of differences in performance when accuracy is not likely to show much variation, for example due to ceiling effects (Olson & Marshuetz, 2005). Therefore, the current VSTM studies adopted a combination of complementary performance measurements based on signal detection theory (*d’*) and processing speed (reaction time).

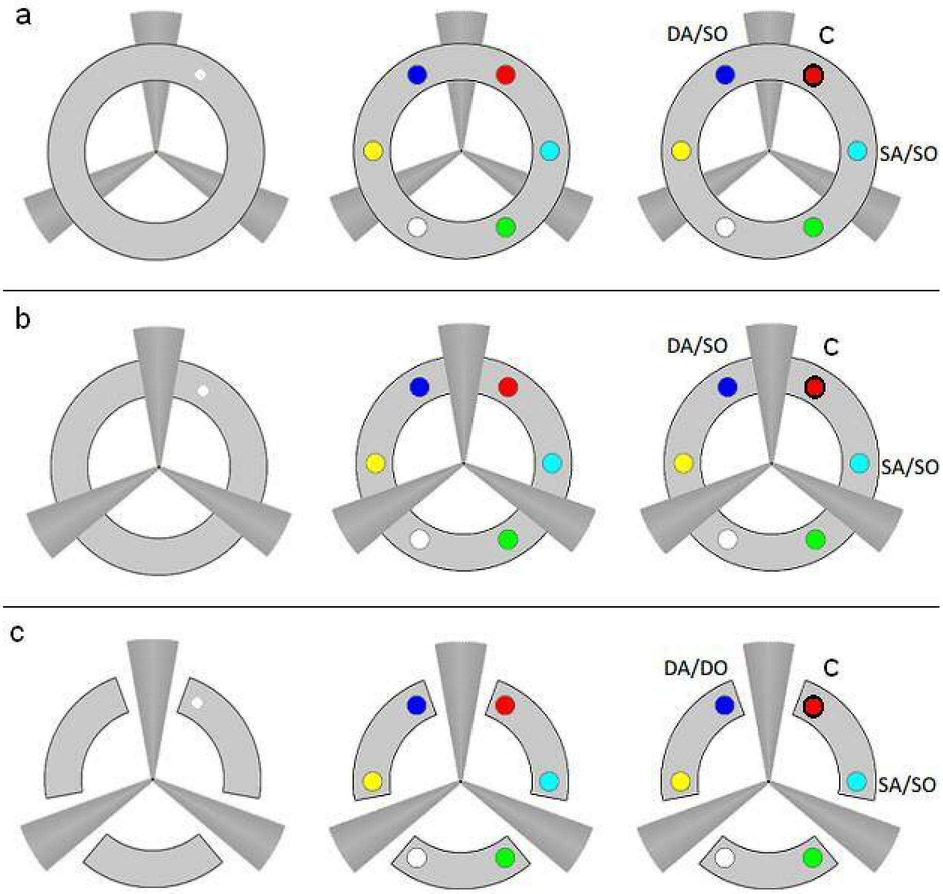

Figure 28: Object formation conditions in Experiment 3.1; a: *intact* object; b: *completed* object; c: *segmented* object. The left panel illustrates the cue (white dot) on the surface of the object(s), the middle panel depicts the object(s) together with the six memory items (coloured circles), and the right panel illustrates a test display with a cued probe (outlined in black). Letters proximal to the items code the critical locations of interest relative to the cue: C = cued; DA/SO = different arc/ same object; SA/SO = same arc/ same object; DA/DO = different arc/ different object.

This chapter consists of two colour change detection experiments aiming to study the influence of perceptual object formation on VSTM, and test for a same-object advantage analogous to that found in visual attention experiments. The principle behind Experiment 3.1 is to keep the spatial properties and number of target memory items constant, while varying the perceptual organisation of the display with minimal impact on visual complexity or other potentially confounding factors. Specifically, the task always entails memorising the colours of six equally spaced circles arranged around fixation, but the perceptual context within which these items appear encourages different object-level grouping (Figure 28).

The perceptual object formation is varied such that all six items appear within the same circular object, which can be either physically intact (Figure 28a) or perceptually completed via an occluder breaking the integrity of the object into three arcs (Figure 28b), or the items appear grouped in pairs within the three separate arcs of a segmented object (Figure 28c). Critically, although the *intact* and *completed* conditions differ visually, the perceptual system should treat them as functionally equivalent in terms of integrity because in both cases the resulting perception is one of an integral circular object. In contrast, the spatial characteristics of the *segmented* condition display were very similar to *intact* and *completed*, but in this case the items should no longer be perceived as being on the same object because the ring should not be percetually completed. Together, these three conditions test for a perceptual same-object advantage when performance is compared between uncued items on either adjacent side of the cue.

The items on the study display were preceded by a brief transient cue (left panels of Figure 28), and at test one of the six items was probed (indicated by a black outline) for a same/different judgement (right panels of Figure 28). The critical change detection comparisons were made between three probe types: *cued*, *same arc*, and *different arc*. Importantly, the latter two are equidistant from, but differ in terms of their perceptual status with the cued item. Probes corresponding to *same arc-same object*, and *different arc-same object* (SA/SO and DA/SO, respectively on Figure 28a/b) are expected to result in similar performance due to being on the same perceptual object, regardless of the physical discontinuity. On the other hand, probes corresponding to *different arc-different object* are expected to result in poorer performance compared to *same arc-same object* probes (DA/DO and SA/SO, respectively on Figure 28c). This would translate into an interaction between object formation condition (*complete, intact, segmented*) and probe type (*cued, same arc, different arc*). If this interaction reflects an object-based advantage for VSTM, it should be translated into no difference between *same arc* and *different arc* probes for *intact* and *completed* objects (hence, their functional equality), but superiority for *same arc* probes compared to *different arc* probes for the *segmented* condition. Within all conditions, performance was expected to be best for the *cued* probe, replicating the standard cueing effect. Both accuracy and reaction time were used as dependent measures.

Since the key proposal put forward here is that VSTM is object-oriented, Experiment 3.2 aimed to test VSTM under conditions where the perceptual organisation (and thus the object-level relationship between the memory items) is kept constant, but their spatial location and scale is varied from study to test. Specifically, the same stimuli as in Experiment 3.1 *segmented* condition were adopted (Figure 28c), but the overall scale was either changed (increased or decreased) from study to test, or it remained the same (replicating the conditions from Experiment 3.1). These scaling changes also resulted in displacement of the spatial coordinates of the memory items, thus they no longer occupied the same locations when the probe was presented at test. However, the perceptual organisation of the same three segmented arcs was preserved. It is therefore hypothesised that the object-based advantage (better performance for *same arc* probes compared to *different arc* probes) would also be preserved. The superior performance for *cued* probes should also be replicated, even though the absolute spatial location of the probes would be displaced. In other words, an identical same-object advantage is expected for the *changed* and *unchanged* scale conditions, marking the importance of perceptual organisation over space.

Experiment 3.1: Perceptual Completion and Visuo-Spatial Short Term Memory[4]

The aim of this experiment was to test the role of perceptual object-formation in VSTM while controlling for the amount and complexity of visual information across the different perceptual organisation conditions. Specifically, the experiment tested for a same-object advantage in VSTM with a cued colour change detection paradigm. This was accomplished by varying object-level properties of the display, in a way that also allows examining the role of perceptual completion. That is, assess if the object-based effect (i.e. the same-object advantage) is equally evident for perceptually intact and phenomenally completed objects (i.e. objects whose integrity or presence is inferred). Indeed, there is evidence in the visual attention literature, as discussed earlier, that object-based effects are equally exhibited for physically intact, modally (via the induced perception of illusory contours) or amodally (via occlusion) completed objects (Moore et al., 1998; Pratt & Sekuler, 2001). However, the effects of perceptual completion have not been studied extensively in relation to VSTM and change detection.

As described in the Introduction section above, two key variables were manipulated: the perceptual object formation afforded by the display (*intact*, *completed*, and *segmented*), and the location, with reference to the cue, of the memory probe at test (*cued*, *same arc*, and *different arc)* (refer to Figure 28). The procedure involved initial exposure to the background display (the ring and occluder), followed by cueing one of the six possible target locations. The study display was subsequently presented, containing the six to-be-remembered coloured circles in their fixed spatial locations within the background object(s). The items were then removed during the retention interval, and finally the test display included the six memory items with one of them indicated as a probe by a black frame. Participants had to decide if the probe had changed colour relative

---

[4] This experiment features in a published study: Nikolova & Macken, 2015.

to the study display. Any of the six items could be probed with equal probability, but only the critical three probes indicated above were of interest.

To summarise the aims of Experiment 3.1, it was hypothesised that an object-based effect will occur, such that there would be no difference in change detection performance for *same arc* and *different arc* probes in the *intact* and *completed* conditions, since both of these probes are perceptually within the same object. However, it was expected that for the *segmented* condition *same arc* probes would elicit better performance than *different arc* probes, because they are situated in different objects relative to the cue. It should also be noted that only the six to-be-remembered circles were emphasised as task-relevant, while the context within which they appeared was not of critical importance for completing the task. Therefore, any effects resulting from perceptual organisation are the result of implicit and automatic processes.

Method

Participants

Twenty-eight undergraduate and postgraduate students (4 males), mean age 22.31 years (*SD* = 2.94) were recruited using the Cardiff University, School of Psychology online recruitement system (EMS). The sample had normal or corrected-to-normal vision, and normal colour vision. Participants were paid £5 for participation.

Stimuli and Apparatus

Unless stated otherwise, the size of the stimuli is reported in degrees of visual angle calculated on the basis of 70 centimetres viewing distance. Each target circle was 1.0° in diameter, centred at 4.7° from fixation. The six to-be-remembered items were equally spaced, with an angular deviation of 60° relative to the central fixation point. The cue was a small filled circle, 0.52° in diameter, centred on the same axis as the target circles.

All stimuli were presented on a grey background (RGB: 212, 201, 200). The colours of the to-be-remembered items were chosen randomly without replacement from the following set, with the corresponding RGB coordinates in parentheses: brown (205, 133, 63), red (255, 0, 0), yellow (255, 255, 0), green (0, 255, 0), blue (0, 0, 255), cyan (0, 255, 255) and white (255, 255, 255). The cue was coloured in ‘blanched almond’ white (255, 235, 205).

Targets were centred within a ring of 1.95° width, and blue-grey colouring (RGB: 200, 200, 200). For the *segmented* condition, the ring was intersected by three gaps between each pair of memory items. Cone objects of 7.37° length passed through the middle of each gap, with the thin-end points linked at fixation (Figure 28c). To give the cones the illusion of solid 3-dimensionality, their colouring was graded from RGB: 160, 160, 160 on the edges, increasing in steps of 2 units to RGB: 180, 180, 180 at the centre. For the *intact* condition, the ring appeared superposed on the cones. For the *completed* condition, the cones occluded the ring. A bold black outline of 0.21° thickness surrounded the probe circle. The whole display (ring and cone shapes) subtended a total of 14.69° x 14.69° centred at fixation.

The experiment was conducted using a Windows XP operating system on a 17-inch monitor with 1280 x 1024 pixel resolution and 32-bit colour quality with a 60 Hertz refresh rate. A standard keyboard was used to record input. Visual Basic 6.0 was used to program and run the task.

Design and Procedure[5]

A 3 (object formation: *intact*, *completed*, *segmented*) x 3 (probe type: *cued*, *same arc*, *different arc*) repeated measures design was used. The *cued* probe coincided with the location of the cue (Figure 28, item labelled C). For the purpose of comparison between conditions, the location of the memory items relative to the interpolated cones was used as a landmark to label the two types of uncued probes. *Same arc* probes were located within the same uninterrupted arc as the cue (Figure 28, item labelled SA), while the *different arc* probes

---

[5] Refer to Appendix 3 for details on a preceding study used as a basis for the current stimuli and design characteristics.

appeared on the other side of the separating cone (Figure 28, item labelled DA). Thus, in the *segmented* condition *different arc* probes were not located on the same intact object as the cue, while in the *intact* and *completed* conditions they were on the same perceptual object as the cued location.

Participants were tested one at a time in semi-closed booths. Each participant underwent a brief practice session with 15 randomised trials, 5 from each object formation condition. The stimulus without target circles was initially presented for 1500 ms, after which the cue was presented for 50 ms. Participants were instructed that the cue was not informative of the probe's location and should be ignored. Fifty ms after cue offset, the six coloured circles were presented for 100 ms. Following a 900 ms retention interval, the six circles were displayed again with one of them (the probe) surrounded by a bold black outline, indicating that a decision needed to be made about whether it had changed colour from its initial presentation (Figure 29). Participants responded on a standard keyboard by pressing the '<' key for '*same*' and the '>' key for '*different*' judgements. These keys were labelled 'S' and 'D', respectively.

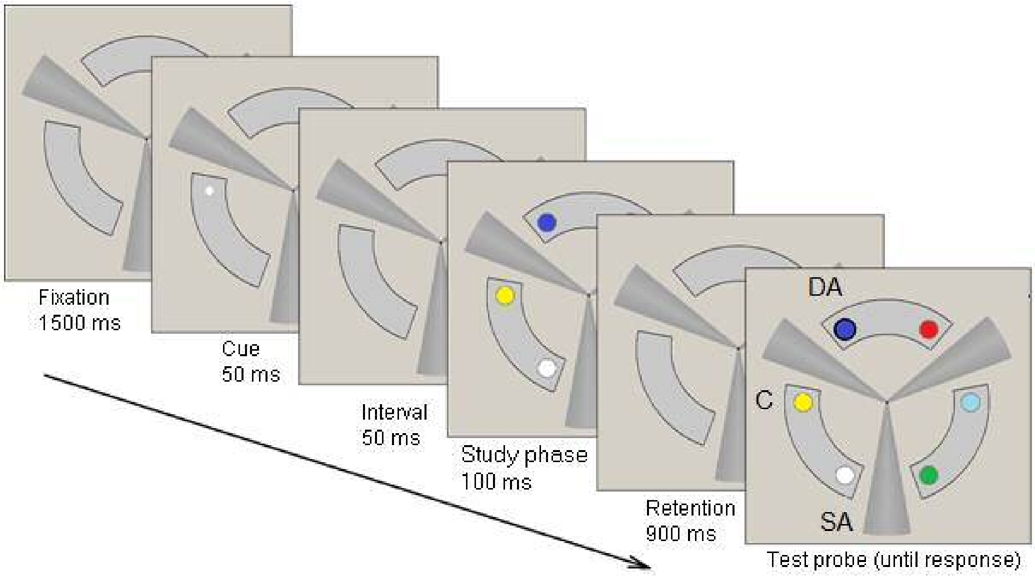

Figure 29: Experiment 3.1 procedure (*segmented* condition). The test phase illustrates a *different arc* probe requiring a "same" response. C = *cued* probe; DA = *different arc* probe; SA = *same arc* probe.

There were 144 trials for each of the three object formation conditions (432 trials in total). Within each of these, the cue and test probe appeared at random, but with equal probability on each of the six possible locations. Therefore, by the end of the 144 trials, 24 responses were made for each cue-probe relationship. Half of the trials involved a change in probe colour from initial to subsequent presentation, while half were no-change trials. Also, on half of the 144 trials, the location of the dividing cones (and gaps between segments) was randomly rotated by 40° to make sure all possible pairings of targets were used.

The 1480 ms inter-trial interval was filled with a dynamic visual noise mask. It consisted of rapidly alternating images of randomly generated black and white pixels and a negative image of the same stimulus configuration as the one in the immediately preceding trial in order to minimise afterimage effects. There were three blocks, the order of which was counterbalanced between participants, with self-timed breaks in-between. Each block contained a single type of object formation condition. Accuracy (*d'*) and reaction times (ms) were recorded for the three critical locations. The procedure lasted about 45 minutes.

Results and Discussion

The data from two participants were excluded from the analysis. One had a consistently low performance around 50% correct, and for the other a programming error occurred and more than half of the data was not recorded. As a result, the analysis included the data from twenty-six participants. A separate 3 (object formation: *intact, completed, segmented*) x 3 (probe type: *cued, same arc, different arc*) repeated measures Analysis of Variance (ANOVA) was conducted on accuracy and reaction time for correct responses from the three locations of interest. No responses were trimmed due to prolonged reaction times (the adopted threshold for discarding a trial was 3000 ms). Whenever the assumption of sphericity was violated, Greenhouse–Geisser correction is reported. Bonferroni corrections were applied to all follow-up pairwise comparisons of main effects. Bayes Factors (BFs) were also calculated to investigate in more detail the null effects, as it is critical for the current hypotheses to establish if a lack of statistical difference between *same arc* and *different arc* probes is due to a genuine equality, or lack of evidence in the data. Change detection accuracy was measured by transforming the proportion of hits (i.e. when a changed probe was correctly identified as *different*) and false alarms (when the probe colour was unchanged, but the response was *different*) into z scores to calculate *d'* (Macmillan & Creelman, 2004).

The patterns in performance differed for the different dependent variables. The contrast can be clearly observed by comparing Figure 30, illustrating reaction time, and Figure 31, illustrating accuracy. Overall reaction time was not affected by changes in object formation, $F(2, 50) = 1.224$, $p = .303$, $\eta_p^2 = .05$, but it did vary as a function of probe type, $F(1.51, 38.8) = 30$, $p < .001$, $\eta_p^2 = .55$ (Figure 30). Most importantly, there was an interaction between probe type and object formation condition, $F(4, 100) = 2.71$, $p = .034$, $\eta_p^2 = .10$. Planned comparisons at each level of object formation revealed that for the *intact object* condition, responses for *cued* probes were faster than for *same* and *different arc* probes ($p < .001$) while there was no difference between *same* and *different arc* probe types ($p > .99$). This equality pattern for uncued probes was supported with a BF of 3.92 against a possibility that *same arc* and *different arc* probes differed in any direction (supported only by BF = 0.25).

Reaction times for the *completed object* condition followed the same pattern, with faster responses for *cued* probes compared to *same* ($p = .001$) and *different arc* probes ($p < .001$), and no difference between the latter two ($p > .99$). Again, this equality pattern for uncued probes was supported by BF = 3.79 versus an unrestricted differences model (BF = 0.26). Within the *segmented* condition, however, responding to *cued* probes was faster than responding to *same* ($p = .001$) and *different* ($p < .001$), but responses for *same* were also faster than responses for *different arc* probes ($p = .018$). The BF for this same-object benefit as compared against an equality pattern was 25.38, thus providing substantial evidence in favour of object-based VSTM. Critically, therefore, even though they were physically segregated from the cued location, *different arc* probes were processed as readily as *same arc* probes if object-formation processes led to them being on the same perceptual object as the cued location. On the other hand, if the physical segregation supported the perception that items were contained within separate objects, then those *different arc* probes were disadvantaged relative to equidistant locations lying on the same intact object as the cued location.

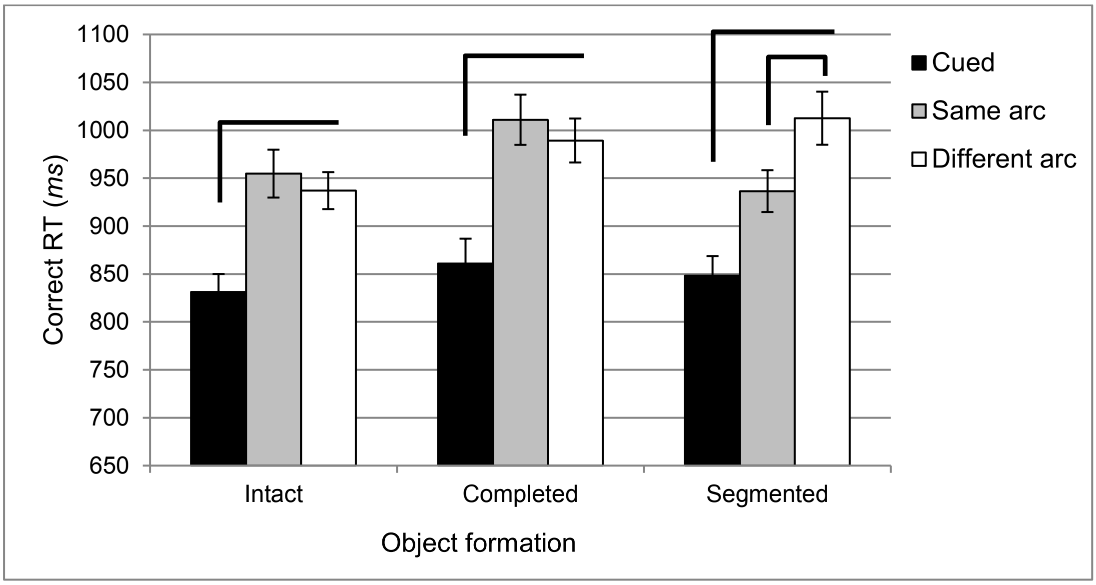

Figure 30: Mean reaction time (milliseconds) for correct responses as a function of probe type and object formation for Experiment 3.1. Brackets illustrate statistical differences within object formation at p < .05 (see text for details). Error bars represent corrected SEM.

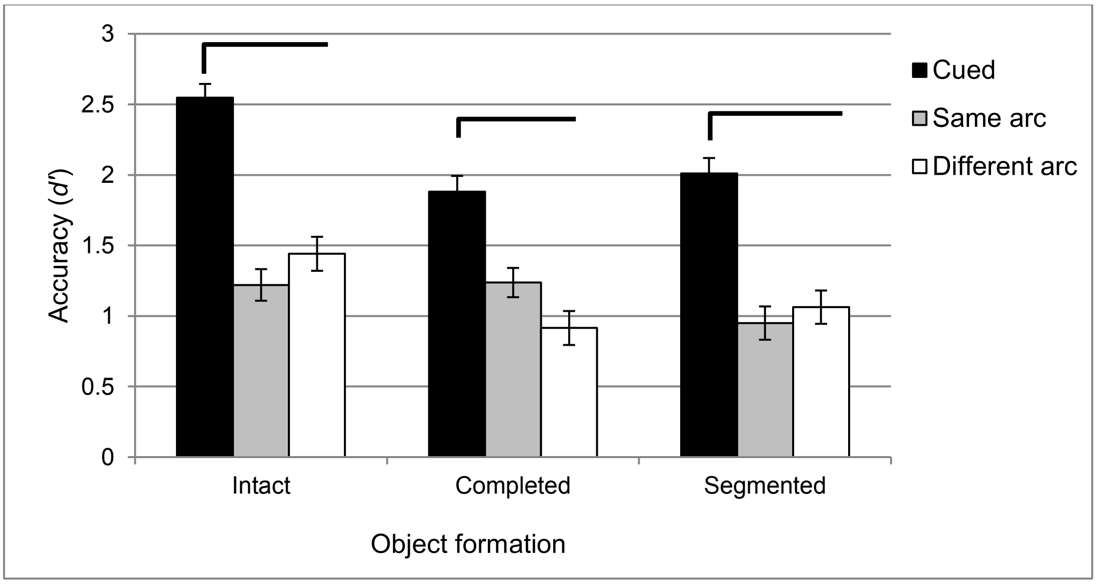

Figure 31: Accuracy (*d'*) as a function of probe type and object formation for Experiment 3.1. Brackets illustrate statistical differences within object formation at p < .05 (see text for details). Error bars represent corrected SEM.

With regards to accuracy (Figure 31), main effects of both object formation, *F*(2, 50) = 55.38, $p < .001$, $\eta_p^2 = .69$, and probe type, *F*(2, 50) = 8.74, $p = .001$, $\eta_p^2 = .26$, were statistically significant, as was the interaction between them, *F*(4, 100) = 3.3, $p = .014$, $\eta_p^2 = .18$. However, the critical interaction took a different form to that observed with the reaction time results. The pattern across probe types was identical for the three object formation conditions, such that accuracy was superior for the *cued* probe (all $p_s < .001$, except for *cued* compared to *same arc* probes within the *completed* object condition, where $p = .005$), and there were no differences between *same* and *different arc* probes in any object formation condition (all $p_s > .096$). BF for the equality model versus differences model for uncued probes (the cued probe always fixed as eliciting highest values) for the *intact* condition was 3.26, while for the *completed* condition it was only 0.94. However, the difference model for the latter was only supported with a BF of 1.06 against an equality model, suggesting that the null effect may reflect insufficient evidence in the data. Finally, for the *segmented* condition equality was preferred with a BF of 5.07 compared to an unrestricted differences model, and with a BF of 3.59 compared to a restricted differences model specifying a same-object benefit (*same arc* probes > *different arc* probes).

Therefore, unlike reaction time, accuracy showed neither a same-object advantage nor an effect of perceptual organisation of the display. Rather, the interaction was due to higher accuracy for *cued* probes in the *intact* object condition compared to those in the *completed* ($p < .001$) and *segmented* ($p = .003$) object conditions, and higher accuracy for *different arc* probes in the *intact* compared to *completed* ($p = .03$) condition. This pattern is unlikely to be due to a floor effect in performance for uncued locations given that *d'* for those locations is consistently at or above 1. It may be the case that the overall superior performance in the *intact* condition, which appears to be the source of this interaction, is due to the circular object being on top of the fan-shaped object, and hence closer (in terms of apparent depth) to the viewer, potentially granting it higher behavioural priority. In the *completed* condition the apparent depth ordering is reversed, while for the segmented object it can be ambiguous. In any case, the observed effect is not critical for assessing the role of object-based mechanisms which is based on performance variation between probe type within the different object formation contexts.

Given the current pattern of performance, an issue worth addressing is whether there is any evidence of speed-accuracy trade-offs, since *different arc* reaction time was slower than *same arc* reaction time in the *segmented* condition, but also *d'* for *different arc* was slightly higher than the *same arc d'* value. Such a pattern may suggest a strategy that compensates for better accuracy by taking longer to make a decision. However, this is unlikely for a number of reasons. Firstly, speed-accuracy trade-offs are typically dependent on the decision criterion starting point, i.e. whether speed or accuracy is emphasised during task instructions, and also the probability of a target occurrence (Wagenmakers, Ratcliff, Gomez, & McKoon, 2008). In the current experiment, however, participants were instructed to balance speed and accuracy, and one was not stressed more than the other. In addition, the probe type was not predicted by the cue, and each of the six possible memory targets was probed an equal amount of times. Participants were explicitly informed about this at the start of the experiment, so there was unlikely to be any bias in expectation. In any case, the Bayesian statistics confirmed that the accuracy response pattern for the *segmented* condition was supported by an equality model more than a model indicating any difference between *same* and *different arc* probes, suggesting that an element of a trade-off is highly unlikely.

In sum, change detection decisions were affected by the perceptual organisation of the stimulus, such that information was more readily retrieved for target information located perceptually on the same object as the cued location. Critically, items within the same perceptual object as the cued feature were retrieved with the same speed, regardless of the presence of a physical discontinuity in the form an occluding object. Therefore, the physical boundary did not in itself lead to a cost associated with the cued location being on a different object to the target. Rather, the perceptual organisation of the scene determined the efficiency of processing the probed information in VSTM. Importantly, this effect was only observable for reaction time and not evident for accuracy, as there was no statistical difference in performance between *same* and *different arc* probes within any of the object formation conditions.

Given that the study demonstrated object-based effects for VSTM by varying perceptual organisation, and given that perceptual organisation itself can vary greatly depending on factors such as proximity between items, similarity, co-linearity and other Gestalt cues (e.g. Wagemans et al., 2012), it may be suggested that experimental manipulations of the number of items and features can affect the manner in which elements in the display are subject to object-formation processes. In turn, this raises potential issues regarding the assumption that the number of objects or features intended by the experimenter actually equals the number of objects encoded in VSTM (Luck & Vogel, 1997; Zhang & Luck, 2008). For instance, presenting an array of six squares may indeed result in the encoding of six items, as intended, but it can alternatively lead to encoding two or three perceptual objects by virtue of grouping on the basis of whatever cues are available in the display (e.g. proximity or similarity). Therefore, there may be a mismatch in the inferred number of objects in VSTM, and the actual capacity (Brady & Tenenbaum, 2013).

Object-formation may also be critical to the interpretation of change detection results depending on the type of methodology used. For example, change detection is often assessed by presenting a stimulus array with multiple items during the study phase, followed by a single probe in isolation indicating the target for which a decision needs to be made at test, which can also be accompanied by empty placeholders occupying the locations of the remaining non-target items (Hardman & Cowan, 2015; Oberauer & Eichenberger, 2013; Rerko, Oberauer, & Lin, 2014). However, given the tendency of VSTM, and the perceptual system in general, to encode information in terms of holistic, object-level representations (e.g. Banks & Prinzmetal, 1976; Boduroglu & Shah, 2014), using a test display which is a segment of what may have originally been encoded as a holistic representation may substantially change performance compared to a test condition which preserves the original perceptual organisation and thus affords object-level matching (Macken et al., 2015). Therefore, other than the intentionally manipulated independent variables, performance may depend on whether feature-level or object-level matching is afforded at test, and the way the items of the memory display map onto a higher-level object representation.

Experiment 3.1 suggested that VSTM is object-oriented, and supported the possibility that information may be coded in terms of composite object structures afforded by the visual display, rather than only on individual item basis. That is, the perceptual objects formed by the memory items exhibit an influence on VSTM, and this may be evident to a different level depending on the type of measurement (reaction time being more sensitive). If the perceptual organisation of the items is indeed of key importance, then manipulating the spatial coordinates of the target items should not affect performance as long as the perceptual integrity is preserved at the object level. Namely, it is the object-level factors that exhibit a primary influence on VSTM, rather than space itself (i.e. the specific location of the stimuli). Testing this possibility was the key aim of Experiment 3.2.

Experiment 3.2: The Scaling of Memory

The purpose of Experiment 3.2 was to test the robustness of the established object-based effect under conditions where the spatial characteristics of the display (e.g. size and position) are changed from study to test within the same trial, without affecting the overall perceptual organisation. Exploring this possibility is important because it would provide an indication of the role of space versus the role of perceptual objects for VSTM.

There already is evidence to suggest that the preservation of the configural properties of the display (i.e. the relationship between individual to-be-remembered items) is important for colour and shape change detection accuracy. This can be illustrated by presenting participants with a set of, for example, coloured squares at study, and then at test one of the squares may have changed colour (thus colour change detection needs to be performed), but the items on the test display may have also changed their locations (Jiang, Olson, & Chun, 2000). Preserving the spatial configuration of the display at test while changing the absolute location of the memory items (by means of expansion) does not impair VSTM for colour or shape relative to a standard test display where the original locations of the items are preserved. However, if the location change involves also a change in the overall configuration of the items, then performance suffers. Shrinking of the display at test has also been found to not impair colour change detection performance (Ma, Husain, & Bays, 2014). Therefore, it is not the spatial location of the items per se that gets encoded, but rather the holistic representation of the display. Also, when an item does change colour at test, it is more often reported as unchanged when the global configuration is preserved than when it is disrupted (Boduroglu & Shah, 2009). This bias suggests that the holistic-level matching of information at retrieval may supersede individual level processing when making a change detection decision. Moreover, these effects occur automatically, since the configuration of the items is task-irrelevant.

Although the studies above provide evidence for the importance of perceptual organisation, they do not employ a cueing technique prior to the test display. Therefore, it is even more likely that the information is encoded in a holistic manner, since there is no exogenous event that can potentially lead to prioritisation of a specific item prior to change detection. Given the propensity of the visual system to encode information in configural terms, the preserved VSTM performance for displaced relative to unchanged set of items is not unexpected. However, adopting a cueing methodology provides a specific reference point (the cued location or item), in order to assess the potential contribution of both the object-level perceptual organisation of the items and their spatial positions relative to the cue.

As Experiment 3.1 demonstrated, cueing a specific location prior to the study display leads to faster change detection for items within the same perceptual object, relative to items of equal distance but in another object. A question which arises then is whether this cueing benefit would also be immune to scaling changes from study to test, or would performance reveal a specific preference for the absolute cued location. Such preference can be expected to result in less pronounced object-based and cueing effects when the absolute spatial coordinates of the items are changed, even though the overall object-level perceptual organisation is preserved. In other words, if information is specifically remembered with reference to space, then probing VSTM for an item which no longer occupies its original location should lead to poorer performance compared to probing an item at the same location. On the other hand, if the object-level structure of the items is the key factor, and it remains unchanged after spatial rescaling, then performance should be comparable to an unchanged scale condition.

There is some evidence that may provide an answer to this question, albeit not without limitations. Makovski, Sussman, & Jiang (2008) used a colour and shape change detection paradigm where a set of six coloured circles or four novel shapes (objects) were equally spaced around an imaginary circle at study, followed by a blank retention interval. During this interval, on some trials a central arrow cue was presented, indicating the future location of the probe item at test. Other trials contained no cue during retention, so participants did not know which item may have changed at test. The subsequent test display could either consist of the items in their original spatial locations, or the items could be spatially displaced away from the centre of the display (i.e. via expansion). Participants had to decide if one of the items had changed. If the trial had contained a cue during the retention interval, the decision was limited to the cued item (i.e. it was the only item that may have changed). Change detection performance for both colour and shape (indicated by accuracy and reaction time) was better when the target location was cued, compared to when no cue was available. Importantly, the cueing benefit was of equal magnitude for both expanded and unchanged test displays, suggesting it was not the spatial coordinate of the cued item that mattered. However, this cue was located at the centre of the screen, and it represented an arrow pointing towards a location previously occupied by one of the items of the study display. That is, the cue was indicating a general direction (e.g. up, down, left or right), as opposed to the specific location occupied by the target. As a result, whether the subsequent test display consisted of the items at their original or displaced locations, their general position with reference to the indicated direction by the arrow was always the same. Therefore, the methodology is not fully informative about the role of space, since both the holistic configuration and the spatial location of the items with reference to the cue were preserved.

The current experiment aimed to test if the object-based effect established in Experiment 3.1 under conditions of exogenous spatial pre-cue (i.e. introduced prior to the study display) would hold true when the spatial locations of the items are displaced at test, but nevertheless preserve their object-level perceptual organisation. Given that VSTM codes information in terms of object-level representation, rather than location per se (e.g. Lee & Chun, 2001; Luck & Vogel, 1997), it is likely that even if a pre-cue indicates a specific location in space, the cueing benefit will relate to the perceptual object associated with this location, rather than its specific Euclidian coordinates. In other words, if the object is moved or spatially displaced in any way, the cueing and object-based benefits will be expected to move with it as long as the perceptual integrity is preserved. Although there is some research, as illustrated above, which explores the effect of spatial displacement in the face of preserved or altered perceptual organisation, this has not yet been investigated within the context of same-object advantage for VSTM. For example, research has been focused on effects regarding the cued item only, or the memory display as a whole.

To test this possibility, the current study adopted the stimuli and procedure from the *segmented* condition in Experiment 3.1, but an additional condition was included where the scale of the display was varied from study to test. By changing the scale of the stimuli, either by expansion or contraction, all items are displaced in absolute terms, but remained the same in terms of relative, object-level representation (refer to Figure 32). The effect of changed scale from study to test was assessed relative to a condition where the scale remained unchanged within the same trial (as in Experiment 3.1). This type of stimulus was chosen based on the previous experiment, as it successfully demonstrated a same-object advantage for VSTM speed of retrieval. In this sense, it also served as an opportunity to replicate the previous results.

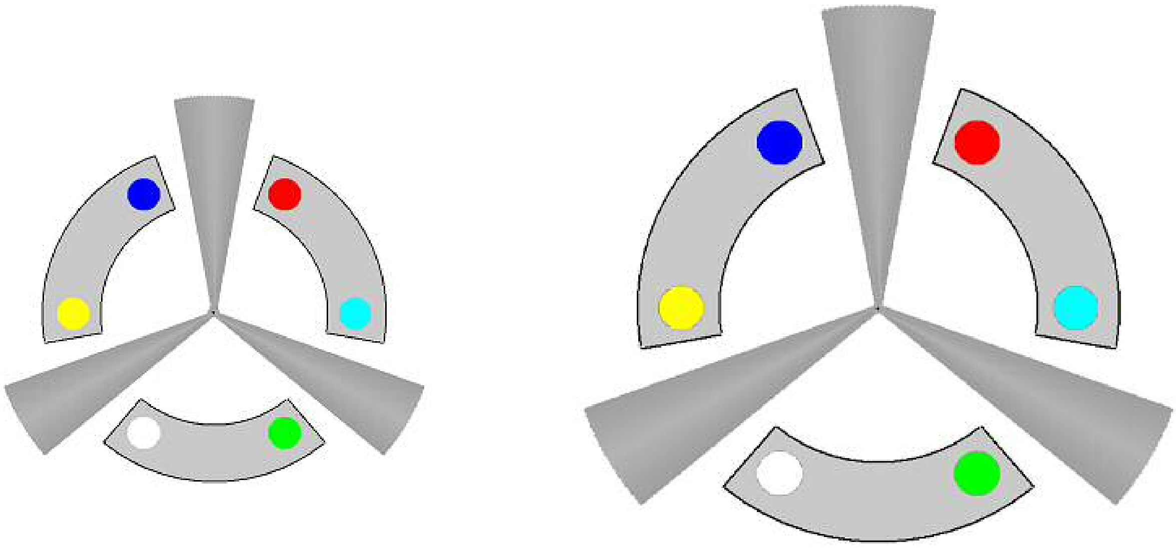

Figure 32: Stimuli scale illustration for Experiment 3.2 (study display sample): standard scale matching the sizes from Experiment 3.1(left), and large scale following magnification of the display by a factor of 1.5 (right).

Experiment 3.2 manipulated two key variables: object scale (*changed* or *unchanged*) between study and test, and probe type (*cued*, *same arc*, or *different arc*), aiming to establish if VSTM and the same-object advantage demonstrated earlier is scale-invariant and based on relational, object-level coding. If this is the case, no interaction between scaling and probe type is hypothesised, and the object-based and cueing effects should be replicated to an equal extent for the *changed* and *unchanged* scale conditions. However, if there is a level of absolute spatial location coding, it is expected that while there may be an object-based effect for the *unchanged* scale (that is, superior performance for *same arc* compared to *different arc* probes), this would not be observed (or it would be of a lesser magnitude) for the *changed* condition. As before, change detection performance was measured by accuracy (*d'*) and reaction time.

Method

Participants

Twenty-six (4 male) undergraduate and postgraduate students at Cardiff University took part in the study for a payment of £5.00. The sample size was aimed to match the one from Experiment 3.1. However, one (female) participant had to be excluded due to being identified as a consistent outlier on Q-Q plots for all trial types. This participant also had a high proportion of excluded trials (12.9%) due to reaction times being outside the specified acceptable range. The final sample size for the analyses consisted of 25 participants (average age of 23.36, *SD* = 5.1).

Stimuli and Apparatus

The perceptual structure of the stimuli was identical to Experiment 3.1, *segmented* condition. The experiment used two scales - the original scale of the display from Experiment 3.1, and an altered scale of the original. The alteration was accomplished by magnifying the original display by a factor of 1.5. Therefore, the proportions and the perceptual organisation of the segmented object were preserved relative to the standard scale, but the size and distances between to-be-remembered items were increased (Figure 32, right panel). For the large scale, the memory items were 1.5° in diameter, and the Euclidian (centre-to-centre) distance between two adjacent items was 7°, equal to the eccentricity at which the items were centred. All other aspects of the stimuli (e.g. colour and structure) and testing apparatus were identical to Experiment 3.1.

Design and Procedure

A 2 (stimuli scale: *unchanged*, *changed*) x 3 probe type (*cued, same arc, different arc*) repeated measured design was employed. The probe type was defined in the same way as in Experiment 3.1, i.e. with reference to the (relative) cued location. For the *unchanged* scale condition, the size of the stimuli remained the same from study to test, while for the *changed* condition the scale could vary either from standard (Figure 32, left panel) to large (Figure 32, right panel), or vice versa. In order to equate the two scaling conditions in terms of perceptual demand, for half of the trials for the *unchanged* condition the scale was standard, and for the other half it was large. Respectively, for the *changed* condition half of the time the scale was large at study and standard at test, while the reverse was true for the remainder of the trials.

The directionality of the scale change (i.e. whether it changed from large to small or vice versa within the trial) was not of interest for the present study and it was averaged out. Previous studies with scaling manipulations found equivalent effects of contracting and expanding the test display (Jiang et al., 2000; Ma et al., 2014), and the key aspect of interest in the current context was the effect of changing the absolute location of the memory items while preserving the perceptual organisation. As before, there were a total of 432 trials. Half of these had an *unchanged* scale and the other half had a *changed* scale. Within each scale condition, each memory item was probed 36 times at test.

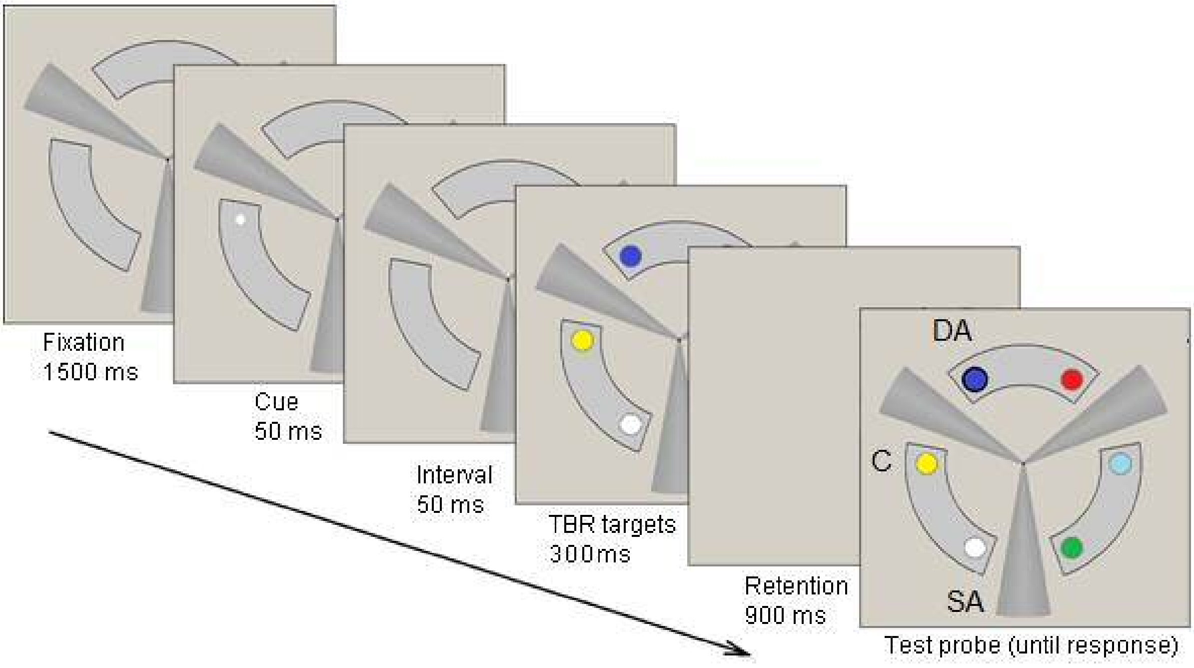

Figure 33: Experiment 3.2 procedure (*unchanged* trial with a *different arc* probe). The labels on the far right panel illustrate the three critical probe types used for the analysis. C: *cued*; SA: *same arc*; DA: *different arc*. TBR = to-be-remembered.

The procedure was similar to Experiment 3.1, except for the following changes. The study display duration was extended from 100 ms to 300 ms, and the offset of the memory items was accompanied by the offset of all visible stimuli, resulting in a blank retention interval of 900 ms (Figure 33). This was necessary in order to introduce a changed scale at test, depending on the trial type. The study time was increased since the ensemble offset and onset of all stimuli made the task more difficult, so performance with a 100 ms study time resulted in very poor accuracy levels (under 50% correct, based on 3 participants which were not included in the current sample). All other aspects of the procedure were identical to Experiment 3.1. The process took approximately 45 minutes.

Results and Discussion

Separate 2 x 3 ANOVAs were conducted on accuracy (*d')* and correct reaction time (*ms*) data. Responses longer than 3000 ms were not included in the analysis (< 1% of the total amount of trials). There were no violations of the assumption of sphericity. As before, additional Bayesian analyses were also conducted. Reaction times were only affected by probe type, $F(2, 48) = 14.16$, $p < .001$, $\eta_p^2 = .37$, such that fastest responses were generated for *cued* probes compared to *same arc* ($p = .005$) and *different arc* probes ($p < .001$) (Figure 34). The overall difference between *same* and *different* probes (29 ms, compared to 76 ms in Experiment 3.1) did not reach statistical significance ($p = .198$), resulting in a lack of a pronounced object-based effect. However, the trend was in the hypothesised direction, as reaction times for *different arc* probes were the slowest, and this was consistent across scale. Whether the scale varied from study to test or remained unchanged did not affect responses, $F(1, 24) = 1.69$, $p = .206$, $\eta_p^2 = .07$, and there was no interaction between scale and probe type ($F < 1$).

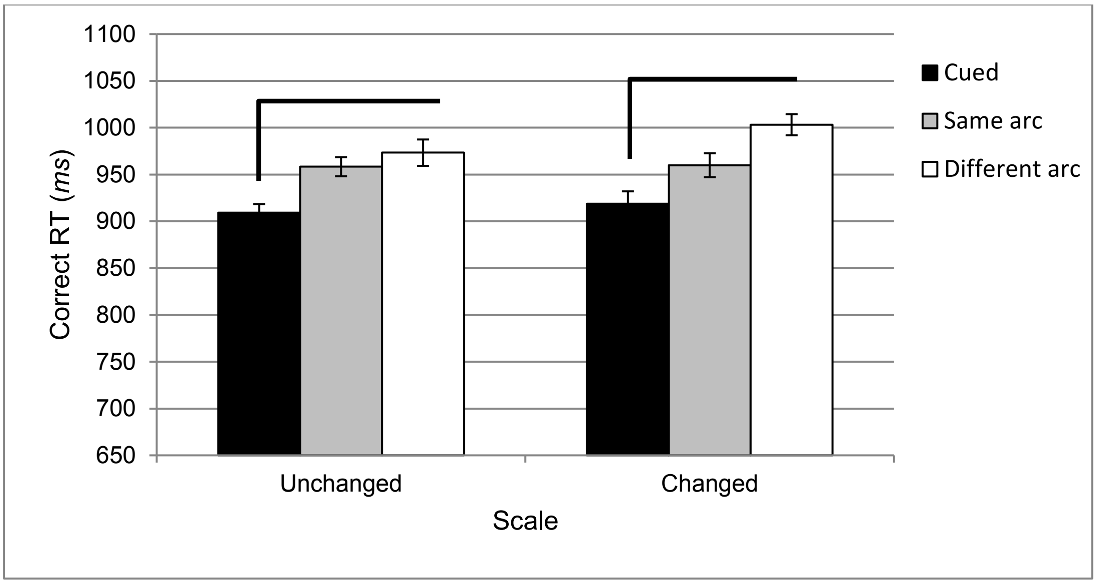

Figure 34: Experiment 3.2 mean correct reaction time (*ms*) as a function of scale and probe type. Error bars represent corrected SEM. Brackets illustrate statistical differences at *p* < .05 (see text for details).

Since the original prediction referred to a main effect of probe type in favour of object-based VSTM, such that *same arc* probes elicit faster correct responses than *different arc* probes across scale, and there was indeed a non-significant trend in this direction, this possibility was followed up with additional Bayesian analyses. The order restrictions of interest reflect an object-based effect resulting in same-object benefit (*cued* < *same* < *different* probes for reaction time), and an equality model between uncued probe types (*cued* < *same* = *different* probes). The lack of difference between *same arc* and *different arc* probes may be expected in the case of space-based coding because these targets are equally distant from the cued location, and also in absolute terms, during the test phase they occupy different locations on the display compared to the study display. The assumption that the *cued* probe elicits fastest responses is kept constant for both models, as the null (indicating no difference in performance between any probe types) was not favoured against any order restriction. This is because, as the pronounced cueing effect suggests, performance for *cued* probes is always superior to all others, so a condition of complete equality is highly unlikely (i.e. where *cued* = *same arc* = *different arc* probes). The critical comparison is thus between *same arc* and *different arc* probes.

In sum, the BF comparisons of interest relate to the evidence in favour of the object-based model compared to the evidence for equality between *same* and *different arc* probes. Based on the current reaction time data (scaling collapsed), the BF ratio for the object-based model versus the equality model was 10.69, compared to a BF of 0.09 in favour of the equality model. Therefore, the possibility that *same-arc* probes were responded to faster than *different arc* probes was approximately 10 times more likely than a pattern of equal performance between the two, reflecting substantial evidence in favour of a same-object advantage. Although the object-based trend for reaction time did not reach statistical significance with the standard *p* value, the Bayesian analysis strongly supports this hypothesis.

It should be noted that even though there was no statistical interaction between scale and probe type, there was a stronger trend towards object-based effects in the *changed* scale condition. Specifically, there is a difference of 15 ms between *different arc* and *same arc* probes for the *unchanged* condition (BF = 1.88 in favour of an object-based versus equality model), compared to 44 ms difference in the *changed* condition (BF = 11.4 in favour of an object-based versus equality model). This pattern goes against the prediction associated with a role of space. If the items are remembered with reference to their spatial location in combination with perceptual organisation cues (i.e. some form of co-existence of space-based and object-based coding of information), the object-based effect should be more pronounced in, or even exclusive to, the *unchanged* condition because the spatial coordinates are preserved. Alternatively, a pure space-based account would predict the complete lack of a same-object benefit, which is clearly not the case for the current data.

The reason that there is a tendency for more pronounced object-based VSTM (albeit not statistically significant based on *p* values) in the *changed* context may be because varying the scale of the stimuli from study to test makes the perceptual organisation of the display more salient, due to the abrupt change in a short time interval (900 ms retention). However, it should be noted that a BF of 1.88 (same-object benefit for *unchanged* scale) reflects insufficient evidence in the data, rather than evidence for or against any account. Therefore, no conclusion can be established with certainty, but a strong trend towards object-based VSTM is highly likely.

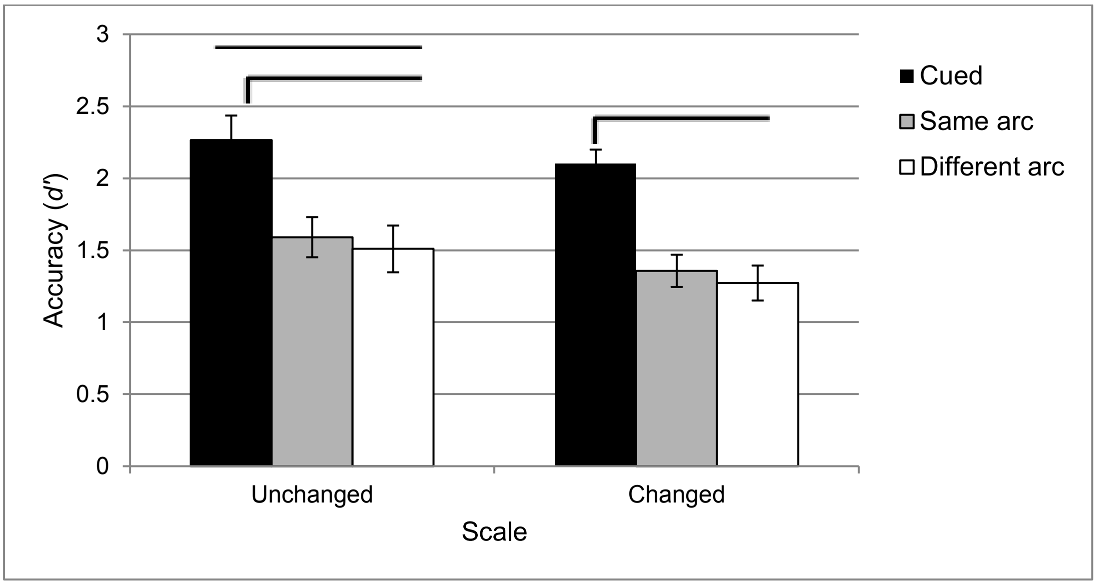

Figure 35: Experiment 3.2 accuracy (*d'*) as a function of scale and probe type. Error bars represent corrected SEM. Brackets illustrate statistical differences at *p* < .05.

In terms of accuracy, change detection performance was independently affected by probe type, *F*(2, 48) = 13.56, *p* < .001, η_p^2 = .36, and scale, *F*(1, 24) = 4.88, *p* = .037, η_p^2 = .17, while the two variables did not interact (*F* < 1) (Figure 35). These results suggest that performance was overall better in the *unchanged* condition, and for *cued* probes compared to *same arc* (*p* < .001) and *different arc* probes (*p* = .001), thus replicating the pronounced cueing effect and lack of same-object benefit for accuracy observed in Experiment 3.1. The BF in support of a main object-based effect for accuracy (*cued* > *same* > *different arc* probes) was 0.88, while the BF for the respective equality model (*cued* > *same* = *different arc* probes) was 1.13. Since a BF between 0.3 and 3 is typically considered as insufficient to distinguish between model suitability (e.g. Dienes, 2014), the current analysis suggests there was not enough evidence in the data to establish if indeed equality is more likely compared to a same-object benefit. Critically, however, scale and probe type did not interact, suggesting the cueing effect was of equal magnitude for both conditions, which emphasises the importance of perceptual organisation. That is, the benefit associated with cueing the location of a to-be-remembered item was not less pronounced when this item occupied a displaced location at test, reltive to when it was presented in its original (study phase) location.

In summary of Experiment 3.2, there are two important points to be noted. First of all, manipulating the absolute spatial locations of the stimuli from study to test does not result in a strong cost for VSTM. In terms of accuracy, there was a small overall advantage when the study and test displays were identical, but for reaction time there was no influence of scaling whatsoever. In both cases, performance was superior when VSTM was probed for a *cued* item, so the cueing effect was of equal magnitude even if the probe did not match the absolute location of the cue (in the case of a *changed* context). This provides support for the hypothesis that VSTM codes information in terms of object-level representation, dependent on the holistic perceptual organisation of the items. If it was the case that information is coded in space-based terms, then it may be expected that the probe type and scale would interact, such that the cueing effect would only be pronounced in the *unchanged* condition.

The second point of importance is that the perceptual organisation of the stimuli did not lead to a pronounced same-object advantage, i.e. there was no statistically evident benefit for reaction time or accuracy when making change detection decisions for *same arc* compared to *different arc* probes. Nevertheless, reaction time exhibited a strong trend in this direction, which was supported by Bayesian statistics. The observation for a lack of same-object benefit was based on $p < .05$, while the Bayesian evidence suggested that the object-based pattern for reaction times was 10 times more likely compared to a pattern of an equal level of performance for these critical probes. Therefore, the results do provide evidence that VSTM is object-oriented, and space is not a critical factor for information coding.

General Discussion

The general aim of this chapter was to explore the influence of perceptual organisation on VSTM without changing the number and arrangement of task-relevant properties. Specifically, the interest was on investigating the way object formation affects information representation in VSTM, with the use of a cued colour change detection paradigm. It was predicted that performance would be influenced by the way the memory targets are organised in terms of higher order perceptual objects, resulting in a same-object advantage for targets situated within the same perceptual object as the cue, relative to targets of equal distance from the cue but in a different object.

Experiment 3.1 supported this hypothesis, demonstrating that when the local spatial properties of the memory targets are kept constant (i.e. location coordinates and size), VSTM is influenced by the manner in which these targets are organised into higher order objects afforded by the display. Moreover, this object-based effect was equally evident when there was a physical discontinuity between the cued and probed locations, as long as the perceived integrity of the underlying object was preserved. Therefore, the results also demonstrated functional equality between intact and perceptually completed objects. Another important point is that the object-based effect was only observed when the reaction time to make a same/different colour judgement was measured, and it was not reflected in response accuracy. Therefore, these types of measurements may be differentially sensitive to the effects of object formation on VSTM. In any case, change detection accuracy and reaction time were both superior for *cued* memory probes, thus replicating the cueing effect evident in both the visual selection and VSTM literature (e.g. Posner, 1980; Schmidt et al., 2002).

While Experiment 3.1 varied the perceptual organisation whilst maintaining the same spatial coordinates of the items, Experiment 3.2 involved keeping the object-level perceptual organisation constant while varying the spatial locations of the target stimuli within the same trial. The key aim was to compare the evidence for a same-object advantage and cueing effect between conditions where the spatial properties of the display remained unchanged, compared to conditions where the display was scaled up or down at test. This allowed exploring the role and contribution of individual item level space-based information versus higher order object-based perceptual organisation representations in VSTM.

The results provided partial support for a scale-invariant, object-based VSTM, such that varying the scale did not affect performance, and the cueing effect was replicated for both accuracy and reaction time. An object-based effect was observed in the form of a pronounced trend towards faster responses to *same arc* compared to *different arc* probes, which was fully supported by Bayesian analyses. In terms of accuracy, there was not enough evidence to conclude whether performance followed an object-based pattern. Overall, the current results suggest that VSTM is affected by object-formation processes. Also, since the higher order perceptual organisation manipulations were not explicitly task-relevant, it can be proposed that they affected performance in an obligatory and automatic fashion. In sum, visual information is coded with respect to the perceptual organisation of the items into coherent objects.

The most consistent result from the current set of experiments was the cueing effect for both accuracy and reaction time, while the evidence for object-based effects, with reference to the differences between *same arc* and *different arc* probes, was only evident for reaction time. Benefits associated with the cued location are robustly established in the memory and attention literature (e.g. Schmidt et al., 2002), and as expected, both types of measurement here were complementary in replicating this effect. Exogenous cueing is known to produce pronounced benefits for VSTM, evidenced in superior change detection for the cued memory item, regardless of whether the cue predicts the location of the memory probe or not (as in the case of the current experiments) (Schmidt et al., 2002; Woodman et al., 2003). However, a continuous measure, such as reaction time, may be more sensitive to detecting aspects of processing of *uncued* probes, compared to accuracy for a categorical same/different judgement. In addition, participants were not time-limited when responding to the memory probe, which was visible until a response was made. Therefore, reaction times may be more suited to reveal the underlying structure of information in VSTM, reflected in the requirement for extended retrieval time in order to support a judgement when the probed item lies on an object other than the cued one.

Given the results from both experiments, an issue that stands out is the fact that the object-based effect for reaction time obtained in Experiment 3.1 was not fully replicated in Experiment 3.2, i.e. it was not evident when data was analysed with traditional frequentist methods. In terms of accuracy, the effects were identical for both experiments, such that there was pronounced superiority for the *cued* memory probe, and no advantage for *same arc* compared to *different arc* probes (i.e. no same-object advantage). In fact, Experiment 3.2 provided a more stable pattern of mean accuracy responses across conditions, while Experiment 3.1 produced larger variability. This may be traced to the fact that the conditions in Experiment 3.1 involved variations in object formation affordances (i.e. *intact, completed* and *segmented*), while Experiment 3.2 had a stable object-level perceptual organisation. However, even though the same type of stimulus was used (*segmented*), reaction time data in Experiment 3.2 did not conform strongly to the object-based prediction. Nevertheless, the Bayesian analysis provided clear support that when the evidence for equal performance for equidistant probes from the cue is compared to the evidence that *same arc* probes elicit faster responses than *different arc* probes, the data favours the latter, same-object advantage pattern.

One reason for this potential discrepancy between the two experiments may be rooted in the structure of the experiment itself. Specifically, in Experiment 3.1 participants were exposed to three different types of perceptual organisation, so there was overall more, albeit task-irrelevant, variability (as participants were only required to remember the colours of the six circles). This variability may have made the perceptual organisation more obvious, thus contributing to a stronger object-based effect. On the other hand, Experiment 3.2 had only one type of perceptual organisation, the *segmented* condition, and the fact that there was no other variation in object level context may have made it likely to habituate to this background stimulus, as it was also not relevant for performing the actual task. Indeed, the BF in support of a same-object effect for the *segmented* condition in

Experiment 3.1 was approximately 25, compared to approximately 10 for Experiment 3.2. This possibility is also consistent with the observation that the object-based trend seems more emphasised for the *changed* scale, where the context varies from study to test and thus possibly becomes more salient. Indeed, there is a possibility that performance can be influenced by statistical regularities in the overall trial structure, thus affecting the tendency of participants to habituate to task-irrelevant factors (Shomstein, 2012). In other words, participants anticipate and adapt to what is relevant for performing the task optimally, so factors which have no value to performance may be ignored.

It should be noted that the possibility that participants may experience habituation to the visual context does not mean that the role of perceptual organisation and object-level factors is ignored. One of the arguments put forward here is that in fact perceptual organisation has an obligatory influence on VSTM and perception in general. Repetitive laboratory-based tasks and artificial stimuli settings may suffer a lot of limitations when assessing the ecological role of objects for perception (Orhan & Jacobs, 2014b), but in any case both experiments presented here provided supporting evidence for object-based effects, albeit of different magnitude. Also, it should be noted that in Experiment 3.1 the perceptual organisation was blocked, so although for the duration of the experiment participants were exposed to varied perceptual object formation layouts, this was not on a trial-to-trial basis per se. Therefore, there was also a repetitive exposure to the same object formation context, but the same-object advantage was clearly visible.

It should also be noted that the effect of cueing in Experiment 3.2 was equally pronounced for *changed* and *unchanged* trials, suggesting there was indeed object-level coding of information, since the *changed* trials 'carried' the cueing benefit together with the displaced memory item. As a potential criticism, however, it can be suggested that the displacement of the items was not very large, as the magnification factor for producing the large scale was 1.5, and so not enough to induce a measurable change in performance due to space-based coding of information (if such process is taking place). In any case, this level of scale change was adopted on the basis of previous research (Makovski et al., 2008), so it is within the standard parameters for this type of manipulation. It is important to consider the results of the two experiments in conjunction, rather than separate. Therefore, within the context of Experiment 3.1, where the *segmented* perceptual organisation produced an object-based effect for VSTM retrieval, the results from Experiment 3.2, which demonstrate a strong bias in the same direction, can be considered as positive support for the object-based hypothesis. Moreover, this was strongly supported by Bayesian statistics. In any case, the current results, albeit not without limitations, do demonstrate that abstract-level rather than absolute space-based coding of information takes place in VSTM.

The fact that the object-based advantage was not replicated to the same magnitude can potentially be attributed to methodological factors, but to conclude this with certainty follow-up work needs to be done. For example, an additional condition may be introduced, where the perceptual integrity of the object is changed from study to test while the spatial coordinates are kept constant. Such a manipulation, for example displaying an *intact* object at study and a *segmented* object at test, would provide additional insight for the importance of preserved perceptual organisation. For instance, if object level factors are critical, then memorising the display with a *completed* context should not lead to impaired performance when the context is changed to *intact* at test, relative to unchanged context (since the implied perceptual organisation is identical – the same integral object). However, VSTM may be expected to suffer if the study phase contains a *segmented* context, while the test phase is *completed* or *intact* (or vice versa).

Varying the object formation context from study to test also allows testing the relative importance of the integrity of the study display compared to the test display, i.e. if it is the encoding (study) or retrieval (test) context that affects performance more. If there is a difference between the two, it may be expected, for example, that memorising the items in a *segmented* context and testing in an *intact* context may actually result in a same-object advantage, not evident in the reversed condition. Such a result may imply more importance for the context at test, i.e. at the time of information retrieval. This manipulation may also be performed without cueing and probing elements, i.e. looking at the overall level of performance for changed versus unchanged object formation contexts within a trial, rather than a same-object advantage (as in this case there would be no reference point, i.e. cued location).

Another methodological alteration may involve object rotation. Rotating the circular object(s) from study to test can be used as an alternative way of preserving perceptual organisation while exploring the effect of changed spatial location and the implied role of space. In this way, the absolute cued location of the study display would be occupied by an uncued (with reference to the object structure) item at the test phase. Under these conditions, the cueing and same-object benefits may ‘move’ together with the rotated object and thus be associated with probes at the original locations during study (object-based effect), or these benefits may remain anchored to the absolute spatial locations, which would no longer correspond to the original, object-based probes (space-based effect).

It should be considered, however, that the rotation design may need additional alterations in object structure in order to avoid confounding problems due to the rotation element. For example, with the current object appearance, rotation in any direction would visually result in the same object, leading to potential confusion about directionality and thus probe type. Alternatively, the rotation may be smoothly performed with the object(s) visible during the retention interval (with the memory item locations as empty place holders). In any case there are a number of issues to consider, e.g. distraction effects due to the visible turning during retention, task difficulty, cognitive load associated with a mental rotation element at test, etc. Therefore, additional control conditions and pilot testing may be necessary.

In terms of limitations of the current design, a further point worth considering is that presenting a cue prior to the study array may be exerting an effect on the quality of perceptual input into VSTM, rather than internal VSTM processes per se. However, this possibility has been examined in detail by studies comparing pre- and post-cues (i.e. cues presented during the retention interval between study and test), which found no difference between the two types - both in terms of accuracy benefits for items at the cued location (Griffin & Nobre, 2003; Schmidt et al., 2002), and object-based benefits for items perceptually grouped with the cued location (Woodman et al., 2003). Considering also the overlap between mechanisms responsible for perception, attention, and VSTM (Awh & Jonides, 2001; Theeuwes, Kramer, & Irwin, 2011), it is unlikely that the object-based effects here are due to this particular aspect of the cueing methodology. However, this would remain to be confirmed with a follow-up study.

Finally, it should be noted that a common limitation for both experiments, which may help explain the lack of object-based effects for accuracy, is related to the manner of presenting the memory items. In the current paradigm, the visual context responsible for the perceptual organisation of the items was almost permanently present throughout the trial, with the exception of the retention interval in Experiment 3.2. The only dynamically changing aspects were the appearance and disappearance of the cue and the memory items, while the perceptual organisation was constantly visible. As already mentioned, abrupt onset events have the power to attract attention even if irrelevant (e.g. Yantis & Jonides, 1984). Therefore, the manner of stimulus presentation may actually encourage grouping all of six coloured items by virtue of ‘common fate’ due to the salience of the dynamic event (Flombaum & Scholl, 2006; Scholl, 2001). It may be the case that there is a form of competition between the two forms of perceptual organisation: the object formation displays grouping the memory items into one (*intact* and *completed*) or three (*segmented*) objects, and the temporal dynamics grouping the six items into a single object (or event). Given the possibility that reaction time and accuracy measurements are not equally sensitive to the object-level VTSM processes, perhaps reaction time was influenced more by the object formation displays (hence, the same-object advantage). Thus, grouping by common fate and temporal coincidence may be decreasing the potential influence of the visual context, and this may be differently manifested depending on the performance measurement used.

This issue can be linked back to the importance of target-object integration discussed in Chapter 2[6]. That is, if the to-be-remembered items are not truly perceived as parts of the higher order objects afforded by the perceptual organisation, then no difference in performance can be expected between *same arc* and *different arc* probes. In other words, classifying the memory items as being within the same object or not would be meaningless, since they would not be associated with the manipulated visual context at all. One way to address this potential limitation would be to have the memory items appear and disappear together with the perceptual objects (i.e. the ring and or the ring segments). However, this may lead to further complications, e.g. decreasing preview time of the object and increased difficulty of the task. Also, it would involve cueing an empty spatial location prior to displaying the object(s) and the memory items, which may also lead to weaker object-based effects, as the object itself would not be cued per se. On the other hand, it may be a good method to compare for potential differences in the strength of the cueing and object-based effects when the surface of the object is cued (as in the current experiments), versus cueing a location which is later occupied by the object (and the memory item contained in it). Alternatively, the methodology can be kept in its present form, but placeholders or monochrome features introduced at the locations of the memory items, which change into coloured circles at the time of the study and test display phases. This modification may contribute to perceiving the memory items as integral parts of the objects by decreasing the occurrence of abrupt onsets and novel events.

In terms of the wider implications from this set of studies, it can be suggested that a continuous measure of VSTM constructs should be used in a combination with a categorical one. Utilising reaction time to evaluate the time it takes to make a correct response can reveal additional information about underlying processes which can be harder to detect only with accuracy measurement. Another implication relates to the importance of accounting for the perceptual organisation of the stimuli. As the perceptual system forms superordinate objects depending on the available cues on the display, this can consequently lead to better memory

---

[6] Target-object integration issues are also highlighted in the pilot work of Appendix 3.

for a set of items which may appear as part of the same perceptual object (e.g. as the colour-sharing effect discussed earlier, cf. Morey et al., 2015).

In relation to this, it is worth noting there is a possibility that object-based effects were not strongly pronounced in the current paradigm since the *segmented* object has the potential to be perceived as perceptually completed by virtue of collinearly and symmetry cues for the arcs. That is, the intended perception is of three separate objects, but the actual perception may be of a perceptually completed circular object, or a conflict/ uncertainty between these two possibilities. Again, the different measurement-dependent outcome (i.e. *d'* or reaction time) may be linked to a dissociation in sensitivity to this potential conflict of perceptions. In any case, given the obligatory effects of perceptual object formation throughout visual cognition (Wagemans, Feldman, et al., 2012), the role of perceptual organisation is a critical issue which needs to be controlled for when other factors of interest, e.g. set size or item locations, are manipulated.

Given the current results and discussion above, there are a number of important implications relating to the design of appropriate experimental conditions. This design should ensure the control, where possible, of perceptual organisation factors in a way that causes minimal or no interference with the measured VSTM constructs. Indeed, studying the role of perceptual organisation effects provides an important route towards situating the role of VSTM within a broader adaptive functionality, and making links with how it has been shaped by the continuous interaction with the environment (Orhan, Sims, Jacobs, & Knill, 2014). As the environment is composed of objects, cognitive and perceptual mechanisms need to be able to respond accordingly to these objects and this necessitates object-based functioning. Consequently, organisms have evolved to be sensitive to perceptual organisation cues and to form perceived objects, which subsequently affects various aspects of behavioural responses. It is therefore important to quantify these processes in order to understand how VSTM and all aspects of visual cognition operate.

Chapter 4

Object-Based Inhibition

Introduction

In order to respond adaptively to the environment, parts of the visual information which are relevant for the current behavioural goals need to be prioritised over others. Accordingly, the focus of the current project so far has been on the object-based nature of processing facilitation, with regards to visual selection and VSTM. However, in some instances it is beneficial to ignore or inhibit the processing of certain information in order to be optimal, e.g. to avoid redundant actions. Given that the empirical work so far suggests that facilitation mechanisms function on the basis of object-level factors, then it is likely that this mechanism also applies to inhibitory processes. In order to explore this possibility, the current chapter is dedicated to probing the object-based nature of the phenomenon known as inhibition of return, which is closely linked to visual selection (Klein, 2000; Posner & Cohen, 1984).

Inhibition of return (IoR) can be described as the process of preventing the selection of a previously attended stimulus or location (Posner, Rafal, Choate, & Vaughan, 1985). This is typically observed in cueing paradigms using a non-informative peripheral cue (i.e. not predicting a target’s location). The usual facilitation effect for subsequent visual targets at the cued location (e.g. faster reaction times and improved accuracy relative to uncued locations) turns into inhibition as SOA increases over approximately 300 ms (Klein, 2000). That is, responses to targets at the cued location slow down compared to responses to targets at uncued locations. In other words, as the name of the phenomenon suggests, re-selecting, or returning to the previously selected (cued) location/target is inhibited.

There is evidence to suggest that a condition for this effect to occur is that attention needs to be disengaged from the cued location before target stimulus display (i.e. during the SOA interval), usually by a flash at fixation or another location (Terry, Valdes, & Neill, 1994). However, disengagement from the cued location is not always necessary, as the inhibitory effect has been replicated without a post-cue attractor (Martín-Arévalo, Kingstone, & Lupiáñez, 2013). Another way in which the IoR is exhibited is by biasing visual search away from previously searched items, both in terms of covert selection (Bao et al., 2013; Taylor, Chan, Bennett, & Pratt, 2015) and execution of eye movements (Klein & MacInnes, 1999; Torbaghan, Yazdi, Mirpour, & Bisley, 2012), including free-viewing of naturalistic scenes (Bays & Husain, 2012). The latter is manifested in longer fixation latencies before revisiting a previously fixated item (i.e. maintaining gaze on the current target for a longer period if the next fixation target is an item which has been fixated before, relative to a novel item), and lower frequency of saccades towards already explored items in the environment. Although the exact mechanisms behind this effect are still under investigation, its functional purpose is most likely rooted in facilitating foraging by encouraging exploration of novel items (Klein & MacInnes, 1999). Therefore, it has an adaptive functionality related to information prioritisation.

In light of the issues discussed in this work, namely the necessity of object-based processing of information for an adaptive functioning, it can be hypothesised that the phenomenon of IoR is also constrained by object-level factors. However, as with the facilitatory side of visual selection, the evidence suggests that IoR displays both space-based and object-based characteristics (Hu, Samuel, & Chan, 2011; List & Robertson, 2007). For example, using the two rectangle paradigm with increased SOA to 1000 ms leads to slowest reaction times for target detection at the cued location, followed by the uncued-same object location, and finally the uncued different-object location (Jordan & Tipper, 1999). As with facilitation, this suggests strongest inhibition effect for the target sharing both the location and object of the cue, since the slowest reaction times are associated with the cued location. However, there is some evidence for object-based inhibition only, without an additive effect of shared location with the cue. That is, equal level of inhibition for the cued and the uncued-same object location relative to the uncued-different object location within the context of a two-rectangle paradigm with a central arrow endogenous cueing (Weger, Abrams, Law, & Pratt, 2008). Therefore, the cued location does not necessarily have a privileged status compared to other locations within the selected (inhibited) object.

It has also been suggested that there is an additive location and object inhibition effect, evidenced by cases where IoR is more pronounced with static compared to moving objects (Tipper, Driver, & Weaver, 1991; Tipper, Weaver, Jerreat, & Burak, 1994). This is demonstrated by cueing one of two square objects on either side of fixation with brief flickering of the object's outline, after which the objects can move to new locations or remain stationary. A smaller white square (the target) then appears in one of these objects, and participants have to respond to its occurrence by a button press. Consequently, there is less inhibition for cued object targets (so faster target detection) when the squares moved their location relative to when they remained in their original positions, suggesting an additive effect of cued location and cued object inhibition (Tipper et al., 1991). Nevertheless, when the locations of the square objects are rotated by 180°, i.e. at the time of target appearance the locations are reversed and the cued object ends up occupying the previous position of the uncued object, IoR is still evident for the cued object, not the cued location. Therefore, the inhibition effect is confined to the cued object itself.

More evidence implicating the important role of objects comes from demonstrating IoR with phenomenally experienced objects. Specifically, inhibition in a stationary display is stronger when the cued location is perceived as the surface of an illusory object (a Kanizsa square, Figure 36, location A) compared to an 'empty' space (when the perceptual organisation is changed so the display features no longer induce object contours, Figure 36, location B) (Jordan & Tipper, 1998). The authors interpret this as co-existence of space-based and object-based effects, i.e. an additive result from both. However, since in this case the inhibition effect is more pronounced when the cued location represents a defined object compared to when it cannot be clearly linked to a coherent object structure on the display, the result can also be interpreted as a probabilistic object-belongingness effect. Specifically, cueing the perceptual object leads to its selection, and subsequent targets within that object have a strong probability to be part of it, thus leading to inhibition of their processing. On the other hand, cueing the empty space and then presenting a target at that location may lead to less inhibition because this target has a higher probability of being a novel object on the scene compared to a target within a cued (i.e. previously selected) object. Based on this probability, the potentially novel target should be prioritised more than a target which is part of an already explored object. Therefore, as with facilitation effects, results interpreted as the co-existence of space and object-related mechanisms can be potentially accommodated within a pure object-based account.

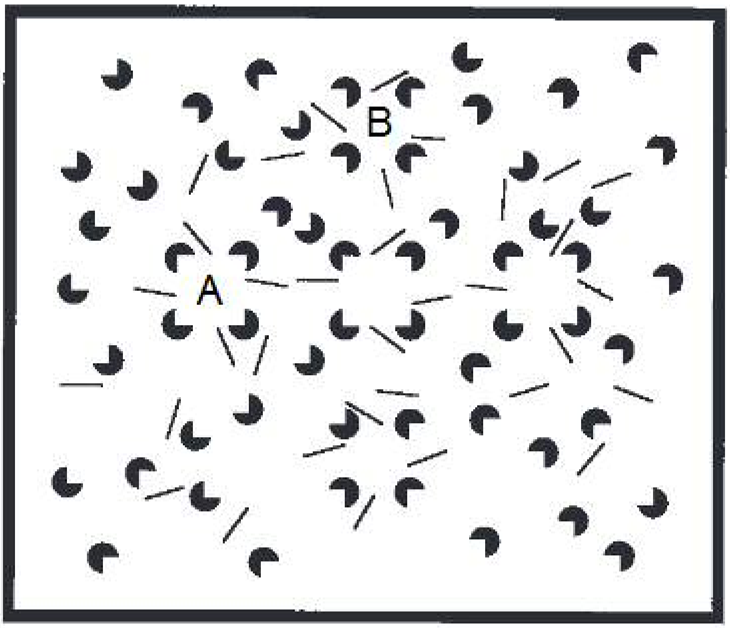

Figure 36: Illustration of the stimuli used in Jordan & Tipper (1998). A: Location inside an illusory object; B: location not belonging to an illusory object. Letters are only for the purpose of indicating possible target locations and were not present during the task, the target was a solid black square centred inside the illusory object (if applicable).

A question arises then as to whether the patterns observed in IoR are constrained by perceptual organisation in a similar fashion as with the facilitatory effect of visual selection, specifically relating to the link between proximity and object belongingness (Oyama, 1961). Indeed, the strength of inhibition had been found to gradually drop with increasing distance of the target from the cued location, displaying a spatial gradient pattern (Bennett & Pratt, 2001; Jay Pratt, Adam, & McAuliffe, 1998). This graded performance is especially pronounced between 300 and 1500 ms SOA, after which inhibition can continue to exert influence up to 3000 ms, but in a uniform fashion, i.e. without spatial gradient effects (Samuel & Kat, 2003). Importantly, however, these studies utilise onset detection tasks where cue and target represent transient stimuli (e.g. dots) on a blank display, so the only perceptual organisation cues to objecthood are linked to cue-target proximity. Therefore, a spatial gradient of inhibition may emerge, since visual stimuli closer together are more likely to be part of the same object, and thus more likely to be inhibited.

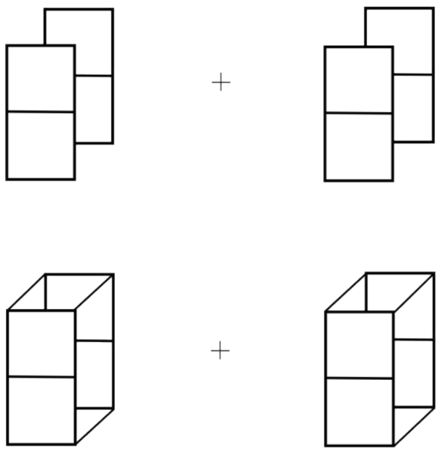

Figure 37: Reproduction of the object stimuli used in Bourke et al. (2006). Top row illustrates the four objects condition and bottom row depicts the two objects condition.

A study which illustrates the importance of perceptual organisation for inhibition and potential (mis)interpretation of object-based as space-based effects was conducted by Bourke, Partridge, & Pollux (2006). The methodology involved cueing one of four identical rectangular figure-of-eight objects (by highlighting its outline), which were situated in two pairs at the left and right of fixation. The pairs on either side differed in depth position, i.e. one object was partially occluded by the other (Figure 37, top panel). After an SOA of approximately 900 ms the contours of one of the objects were partially deleted, turning either into an “S” or an “H” shape to be identified by the participant. Therefore, this was a target discrimination task and the targets were integral featured of the objects, i.e. they represented structural changes.

Critically, IoR was strongest when the cue and target were within the same object, and there was also some inhibition for the uncued object on the same side of fixation, but different depth plane relative to uncued objects on the other side of fixation. In fact, there was no difference in reaction time for objects on the different side of fixation from the cued object, regardless of depth. The authors interpreted this as a combination of space-based and object-based inhibition, where inhibition is evident for the uncued object on the cued side because there is a shared retinotopic location. However, in a different condition where the overlapping object pairs were joined to form the perception of a single cuboid (Figure 37, bottom panel), the inhibition was equal for targets within the cued object, regardless of the depth of the target. That is, in this case there was no gradation of inhibition for targets on the same side of fixation, as they were perceptually organised within the same object. Again, responses were slower relative to the uncued cuboid object to an equal level for both depths. Bourke et al. (2006) conclude that space-based IoR is two-dimensional, while object-based IoR is three-dimensional. However, these results can be interpreted in an alternative way, without resolving to space-based mechanisms.

The alternative explanation for Bourke et al.'s (2006) data relies solely on an object-level representation. In the case of four figure-of-eight objects, the positioning of each overlapping pair predisposes a perception that the pair is in fact a single object, but with a level of uncertainty. This is due primarily to cues of proximity and similarity. Therefore, the inhibition that spreads to the different depth object in this case may be because of the possibility that it is a part of the cued object, but since the probability that the two stimuli are the same object is not high, the inhibition is less than for the actual cued surface. On the other hand, when the two objects are joined into a single cuboid, the probability of object belongingness is increased, resulting in equal IoR for targets at both depths, because they are parts of the same object.

This pure object-based explanation is corroborated by the fact that there is no difference in reaction time for uncued targets of either depth on the uncued side of fixation, because in both the single cuboid and separate objects conditions, targets on the uncued side are always equally likely to be part of a different object to the cued one. Therefore, the observed two-dimensional space-based effect may be a result of an object-belongingness gradient, i.e. the transfer of IoR may depend on the level of perceived objecthood for the displayed items. This also raises the issue regarding the importance of incidental perceptual organisation and its impact on the measured variables, as the authors of the paper intended to present participants with four versus two objects, but it may be the case that this distinction was not clearly made by the perceptual system.

The importance of taking into consideration the global visual scene is also demonstrated for saccadic IoR. For example, saccade latency is increased when a previously fixated item is revisited within a complex visual scene, but this inhibition is abolished when the visual scene is removed and a saccade is consequently executed towards a blank location (Klein & MacInnes, 1999). Specifically, the experiment required participants to search for a character (a wizard) in a visually cluttered scene. A small disk was superimposed on the image on some trials, which was a signal for participants to stop searching and fixate this disk target. The position of the fixation target was manipulated to either coincide with the previously fixated location (based on the search behaviour of the participant) or gradually increase in distance from it. As expected, IoR was observed and participants were slower to execute a saccade to the fixation target when it coincided with the region of the previous fixation on the scene, relative to when the target appeared in another (previously unexplored) location. Critically, on some trials when the fixation target was presented, the cluttered search display was simultaneously erased, leaving the target on a blank grey screen. Under these circumstances, IoR was no longer observed. Therefore, there is a possibility that the IoR effect was anchored to the fixated objects within that context, and not to their location in space. Although there was a lack of a control condition to confirm whether the abolished IoR was due to the abrupt change on the display (as a result of erasing the visual context), this result implicates the role of object-based factors for visual selection inhibition.

Given the evidence for spatial gradients of IoR and the purported coexistence of space-based and object-based effects, the experiment in this chapter (Experiment 4.1) was focused on examining the possibility that these patterns are due to object-based factors in a similar way to the facilitation effects discussed earlier. A similar cueing paradigm was adopted as in Chapter 2, using the same type of objects and stimuli in order to pit object-based versus space-based effects with well-defined objects and an integrated target. The key purpose here was to test if IoR can display a pure object-based pattern without regards for cue-target distance as long as the cue and target are unambiguously located within the same object.

Experiment 4.1

The principal change in the methodology relative to the visual selection experiments in Chapter 2 relates to the SOA duration, which was increased in order to induce inhibition, as opposed to facilitation effects. Experiment 4.1 also adopted a central attractor stimulus embedded in the middle of the SOA interval. As before, the task involved target discrimination as opposed to onset detection. It is important to note that IoR was originally believed to be observed only with detection tasks (Terry et al., 1994), but it was later established that it is evident for discrimination tasks too, albeit with a longer SOA over 400 ms (Gabay, Chica, Charras, Funes, & Henik, 2012; Lupiáñez, Milán, Tornay, Madrid, & Tudela, 1997; Jay Pratt & Abrams, 1999). Therefore, for the purpose of consistency with Chapter 2, the current experiment preserved the task in its original form of luminance change discrimination.

There were three main modifications to the design. First of all, a visual event was introduced during the SOA interval, whose purpose was to lead to involuntary disengagement from the recently cued feature. In this way, the subsequent target appearance leads to re-selecting the feature (if the target is at distance 1) and/ or the object stimulus (if the target is within the cued object, i.e. at distance 3 for the *two objects* condition, or any uncued feature for the *one object* condition). There is evidence to suggest that this disengagement from the cued item during the SOA interval is a necessary condition to induce IoR (Terry et al., 1994), and it is also a standard practice for most IoR studies (e.g. Jordan & Tipper, 1998; Tipper et al., 1997). The second key modification was extending the SOA duration. In this case, SOA was manipulated across two different intervals, 900 ms and 1300 ms. The SOA of 1300 ms was chosen as this is in the range of the upper boundary for IoR (Reppa & Leek, 2003; Samuel & Kat, 2003)[7]. The third modification was making the cue uncorrelated with the target location, so each object feature had equal probability of changing its luminance. Using a non-informative cue is important for detecting IoR because if there is any level of top-down incentive to focus on the cued location, facilitation is likely to be robust at long SOAs (Klein, 2000).

The choice of the methodological parameters in the current study was additionally motivated by evidence that object-based IoR with very similar stimuli and task was only observed with a central attractor and a long SOA. Specifically, Chou & Yeh (2005) used six circles connected into two overlapping triangles forming a Star-of-David shape, and target event discrimination (luminance change of one of the circles). Cueing was performed by highlighting the frame of one of the objects (triangles), thus simultaneously cueing the entire object surface. Subsequently, the luminance of one of the six features (cued or uncued object) changed and participants had to discriminate its polarity. Results indicated object-based IoR evidenced by slower reaction times to changes in any of the three cued object features relative to the three uncued object features. However, this was only observed with long SOA of 1360 ms, but not with an SOA of 884 ms, and only in the presence of a central attractor stimulus during the cue-target interval. Therefore, the study suggested that the interim attractor is a necessary condition for inducing IoR. It should be noted that although the object stimuli appearance was very similar to the ones adopted in the current study, the cueing

---

[7] The use of a central attractor stimulus and the specified range of SOA were also motivated by a pilot study reported in Appendix 4. This small-scale pilot used SOAs of 300, 600, and 900 ms and no interim attractor, resulting in null effects with a pronounced trend towards object-based facilitation.

methodology was very different. As opposed to cueing the whole object, the current study used a cue which was specifically oriented towards a single object feature, thus probing for possible gradient effects with reference to this feature.

The attractor stimulus in the current experiment was designed to be very similar to the one used by Chou & Yeh (2005), as it was proven to be successful in inducing IoR. Namely, this was a dashed line centred at fixation, which spanned the length of the object stimulus, but it varied in orientation, such that it passed in-between any two adjacent object features and never overlaped the features (Figure 38). Its orientation was randomly varied on a trial-to-trial basis. This stimulus was task-irrelevant and presented in the middle of the SOA interval.

In sum, the manipulated variables were SOA and target distance from the cued feature. As before, this was done separately for *one object* and *two objects* conditions. Given that the cue in this study was not correlated with target location, and participants were made aware of this, it was anticipated that a pure object-based effect would be manifested in no difference for the cued feature (distance 1) relative to uncued features on the same object. In other words, the effect of cueing would be equally pronounced for all features on the cued object when compared to features on the uncued object. Although an abrupt visual onset (e.g. an exogenous cue) typically produces an involuntary capture, this capture effect can be considerably modulated by top-down (volitional) factors (Egeth & Yantis, 1997; Theeuwes, 2010). Taking in consideration the evidence that within the two rectangle paradigm an uninformative cue can lead to equal benefit for cued and uncued-same object locations (He et al., 2004), a within-object equality model should be supported with the current stimuli.

It was hypothesised that object-based inhibition for the *two objects* condition would be manifested by slower reaction times for targets at distance 1 and distance 3 (cued object), compared to targets at distance 2 and distance 4 (uncued object). No variation of performance was predicted for the *one object* condition, as all targets corresponded to changes within the same object. However, if there is manifestation of space-based IoR, then it may be the case that responses would be slowest for distance 1, and then gradually decrease with increasing cue-target distance in both object conditions. Both of the SOA levels were expected to produce IoR, as they are well above the proposed threshold of 300 ms (Klein, 2000), but it could be the case that the pattern of the effect may vary. For example, if space-based and object-based IoR coexist, this may depend on SOA.

Method

Participants

Twenty-one participants (4 males, mean age = 22.65, *SD* = 3.21) took part in the *one object* condition, and twenty-one different participants (4 males, mean age = 23, *SD* = 1.48) took part in the *two objects* condition in return of £4 payment. The sample size was pre-determined with power calculations using G*Power software and anticipated effect size of $\eta_p^2 = .22$ based on the main effect of distance in a pilot experiment (Appendix 4).

Stimuli & Apparatus

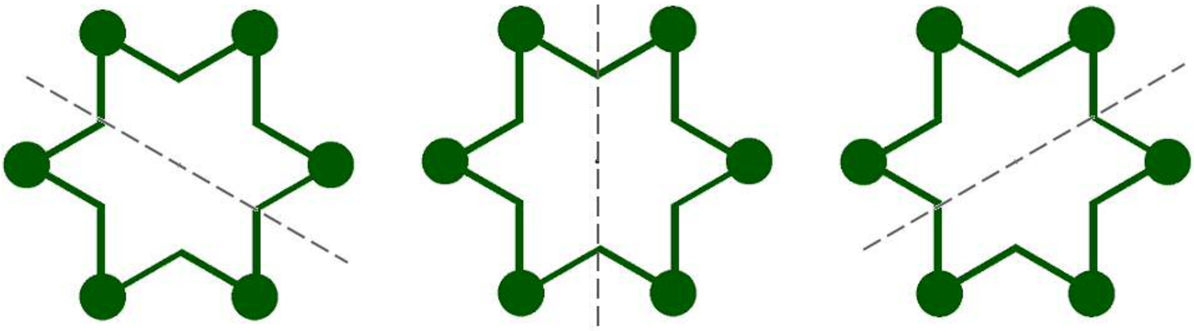

Figure 38: Illustration of possible attractor orientations for Experiment 4.1 (*one object* condition). From left to right: - 60° angular offset from vertical, vertical, + 60° angular offset from vertical.

The equipment and physical characteristics of the stimuli were identical to the stimuli in Chapter 2, Experiment 2.1, with the addition of the line stimulus presented during the SOA interval (the attractor). This was a dashed line of light grey monochromatic colouring (RGB = 150), measuring 10° visual angle and 3-pixel stroke. It was always centred at fixation, but its orientation could either be vertical or oblique at ± 60° angular deviation from the vertical meridian (Figure 38). Therefore, it always passed in-between two adjacent object feature pairs.

Design and Procedure

The manipulated variables were SOA (900 ms and 1300 ms), and target distance (four levels), combined in a 2 x 4 repeated measures design for each object condition. The cue did not predict the target location, so each object feature was subjected to luminance change an equal number of times, resulting in 36 trials for each of the six features per SOA. Target location and SOA varied randomly from trial to trial, and the experiment was structured into 4 blocks of 108 trials with self-paced breaks between each block.

The procedure was similar to the one employed in Experiment 2.1 (Chapter 2), however, the timings were slightly different due to the different SOAs utilised here, and also the additional attractor presentation in the middle time section of the SOA. The procedure is graphically represented in Figure 39. The attractor was presented for 50 ms and its onset varied at random within a 200 ms time window during the SOA of the respective trial. This time window represented the middle portion of the time interval between cue offset and target onset, such that if the trial was with an SOA of 900 ms, the attractor could appear between 300 and 500 ms after cue offset, while if the trial was with an SOA of 1300 ms, the attractor was presented in the 500-700 ms time interval after cue offset. This was done in order to introduce an element of uncertainly that would prevent complete habituation towards this stimulus, in order to ensure it captured attention. Similarly, its orientation was varied in a random fashion, adopting one of the three possible positions described above (Figure 38). The procedure took approximately 35 minutes.

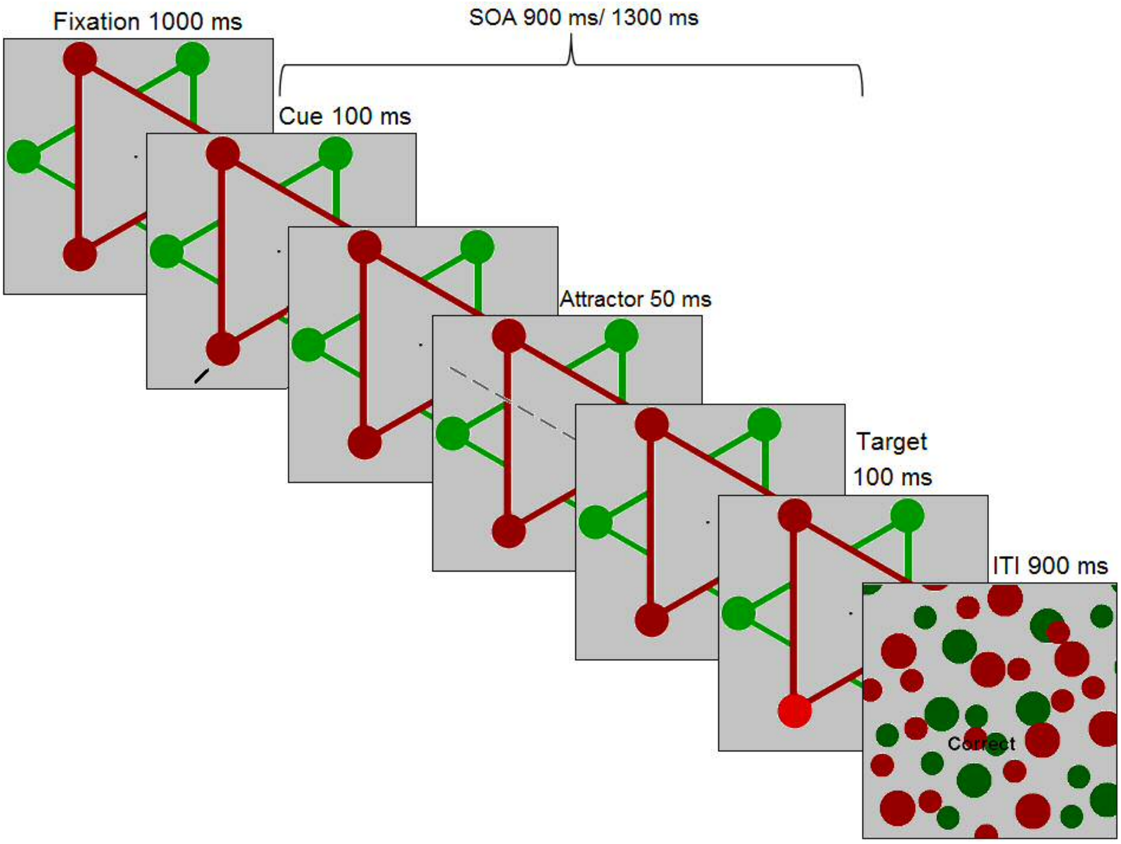

Figure 39: Experiment 4.1 procedure illustration (*two objects* condition). This example represents a positive luminance change at distance 1. ITI = inter-trial-interval.

Results and Discussion

The standard trimming of correct reactions times resulted in a loss of < 1% of the data for the *one object* condition, and 1.5% for the *two objects* condition. Correct reaction times for the *one object* condition were not affected by target distance, $F(3, 60) = 2.61$, $p = .06$, $\eta_p^2 = .12$, but varying SOA did produce an effect, $F(1, 20) = 11.51$, $p = .003$, $\eta_p^2 = .37$, such that SOA of 900 ms resulted in overall slower responses than SOA of 1300 ms, with a mean difference of 10.2 ms ($SE = 3$) (Figure 40). There was no interaction between the variables ($F < 1$).

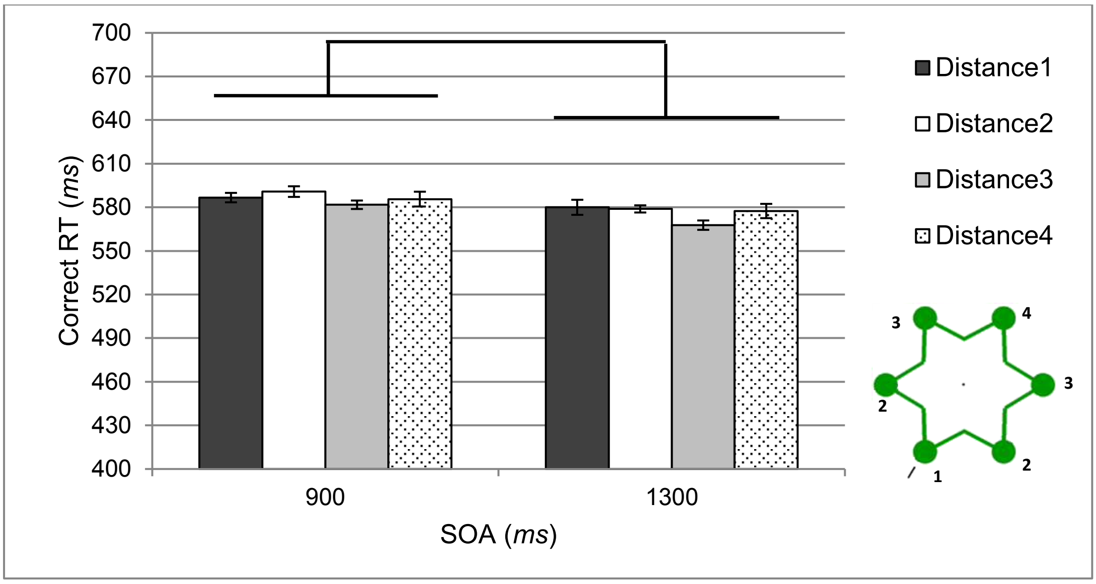

Figure 40: Mean correct reaction time (*ms*) as a function of SOA and target distance for Experiment 4.1, *one object* condition. Brackets illustrate statistical differences at $p < .05$. Error bars represent corrected SEM.

The lack of distance effects is in support of object-level selection, but similarly to the empirical work in Chapter 2, these conclusions are based on null effects. In addition, it should be noted that there is a trend for distance 2 to elicit slower responses compared to distance 3, which may be interpreted as a trend towards space-based IoR. In other words, it may be the case that inhibition decreases with increasing cue-target distance, although there is no other trend in the data that suggests spatial effects. In any case, the non-significant effect for target distance was marginal at $p = .06$, so Bayesian analysis was conducted in a similar fashion to Chapter 2 in order to investigate this effect further.

Since in this case a cueing effect was not expected, an object-based model is the same as the null, as it predicts no differences between distances. On the other hand, a gradient model is defined as an order restriction where reaction time decreases with increasing distance due to less inhibition further away from the cued location (i.e. distance 1> distance 2 > distance 3). In order to be consistent, this analysis excluded distance 4 for the potential confounding reasons stated earlier (refer to Chapter 2, pages 61-62). Also, the data were collapsed across SOA, as the main variable of interest was cue-target distance, and there was no interaction. As before, the BF values represent the preference for one model over the other, and in this case the analysis revealed a BF of 3.58 in favour of a gradient model, compared to BF of 0.28 in support of an equality model.

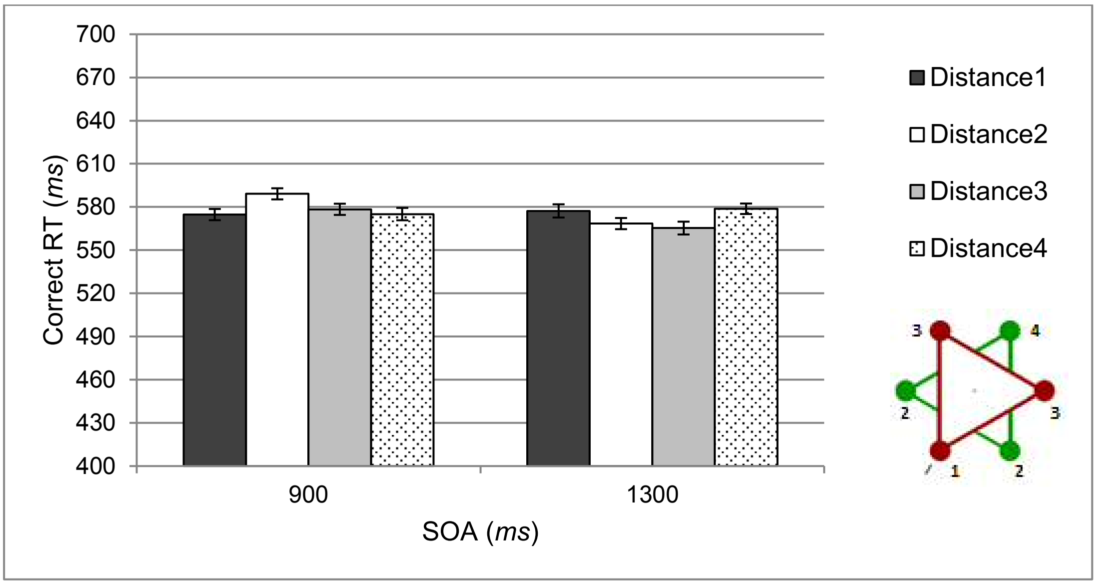

Figure 41: Mean correct reaction time (*ms*) as a function of SOA and target distance for Experiment 4.1, *two objects* condition. Error bars represent corrected SEM.

For the *two objects* condition there was also no effect of target distance $F(3, 60) = 1.22$, $p = .309$, $\eta_p^2 = .06$, and in this case SOA did not influence performance, $F(1, 20) = 2.55$, $p = .126$, $\eta_p^2 = .11$. However, there was an interaction between SOA and target distance, $F(3, 60) = 3.96$, $p = .012$, $\eta_p^2 = .17$ (Figure 41). Therefore, the effect of target distance was investigated separately for each SOA level. Reaction time was affected by distance for 900 ms SOA, $F(3, 60) = 3.02$, $p = .037$, $\eta_p^2 = .13$, but no statistical pair-wise differences were observed between the four levels of target distance (all $p_s > .05$), suggesting that that any potential differences did not survive multiple comparison adjustments. For SOA of 1300 ms target distance had a marginal effect on performance, but not reaching statistical significance, $F(3, 60) = 2.54$, $p = .065$, $\eta_p^2 = .12$.

The original hypothesis for the *two objects* condition stated that performance would follow a non-monotonic, object-oriented pattern, where targets at distances 1 and 3 would elicit slower reaction times than targets at distances 2 and 4 (although performance for distance 4 is subject to speculation). The data, however, did not conform to this prediction, as no pronounced trend was revealed.

Given that there was an interaction between distance and SOA, the Bayesian analysis was performed separately at each SOA level. Also, given the inconsistent trends in the data, the comparison models of interest in this case refer to an inhibition gradient, as was the case for the *one object* condition (denoted as D1>D2>D3), but also an object-based facilitation model. The latter reflects a possibility that distance 2 (uncued object) is related to slower responses than distances 1 and 3 (D1<D2>D3), as this was the trend observed for 900 ms SOA (Figure 41). It is of interest to see how much the data support a model of object-based inhibition (D1>D2<D3), since this was the original hypothesis, albeit not reflected in the mean trends. For the purpose of simplicity of comparisons, the three critical models illustrated in Figure 42 represent the respective BFs as assessed against the null hypothesis (i.e. equality of means: D1=D2=D3).

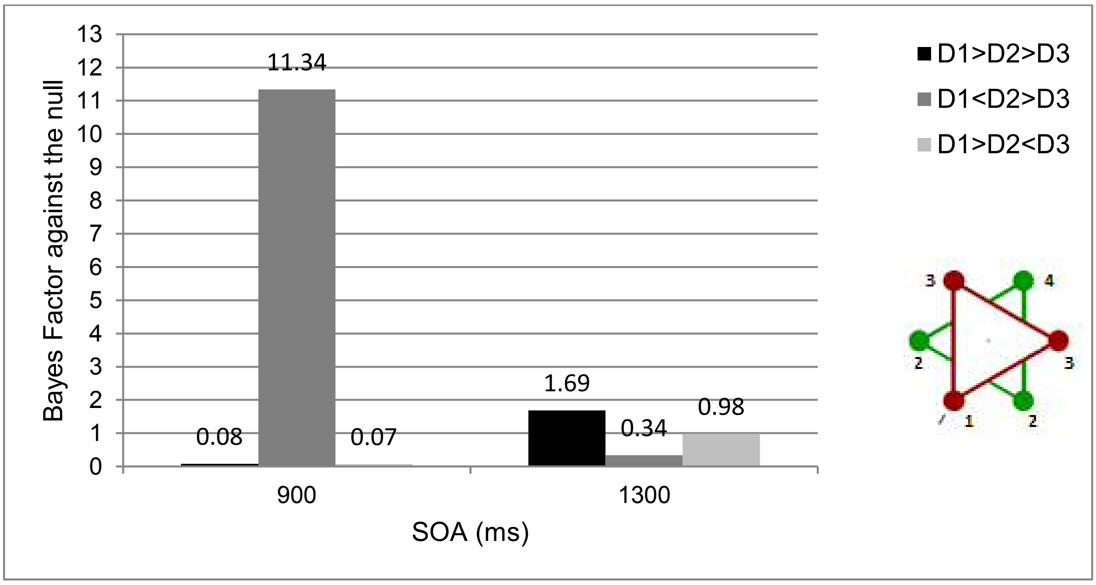

Figure 42: BF ratio comparisons for a gradient inhibition model (D1>D2>D3), object-based facilitation model (D1<D2>D3), and object-based inhibition model (D1>D2<D3) assessed against the null hypothesis for Experiment 4.1, *two objects* condition. D1 = distance 1, D2 = distance 2, D3 = distance 3.

The BF evidence for the *two objects* condition suggests that object-based facilitation is well supported for SOA of 900 ms (BF = 11.34), and all other difference patterns are less likely than the null (BFs < 0.33). On the other hand, none of the models provide compelling evidence as assessed against a no-difference model for 1300 ms SOA. Overall, the data for the *two objects* condition indicate a strong trend towards object-based facilitation for 900 ms SOA, and not enough evidence to meaningfully assess the variations in performance for 1300 ms SOA.

In sum Experiment 4.1 did not indicate spatially non-monotonic, object-based IoR. Instead, there was some tendency for spatially graded decrease in reaction time for the *one object* condition. For the *two objects* condition there was residual object-based facilitation still having an effect on performance at 900 ms SOA, and no performance variations as the cue-target interval was increased to 1300 ms. There are propositions that facilitation and inhibition can take place simultaneously in a net outcome (Reppa & Leek, 2003; Tipper et al., 1997). If this is the case for the current experiment, then it may explain the lack of reaction time variation (i.e. cancelling out between the two effects). Nevertheless, it is hard to distinguish this proposal from an actual null effect. In any case, the BF values for *two objects*, 1300 ms SOA suggest that the null was more likely than any other account.

Regarding the gradient tendency for the *one object* condition, it may be that increasing the time delay between cue and target leads to a level of perceptual segregation of the target event from the underlying object feature. In turn, this may contribute to performance being influenced by spatial proximity factors, as opposed to object structure. There is some evidence that perceptual object formation is not linearly related to stimulus exposure, i.e. the object perception is initially strengthened, and after prolonged exposure it can start to break down (Feldman, 2007), although this is specifically relevant to extracting object perceptions from low-level ambiguous configurations, e.g. a set of lines. Long exposure to a novel or complex object may also lead to forming alternative configurations and switching between representations (Long & Toppino, 2004). However, all of these possibilities are rather unlikely as the luminance change is a salient event linked to a modification of a property of the object, so there should be little room for ambiguity regarding what it represents. Also, based on the empirical work in Chapter 2, these types of stimuli represent perceptually salient objects. Any representation breakdown is thus unlikely given the well-defined structure and the consistent object-based effects observed earlier.

Alternatively, there is evidence to suggest that SOA duration can in fact shift the prioritisation strategy from configural (object-oriented) for short SOAs to contextual (probability-oriented) for longer SOAs (Shomstein & Yantis, 2004). Specifically, the study involved a version of the two rectangle paradigm (Egly et al., 1994) varying SOA (200 ms, 400 ms, 600 ms) and target location probability relative to the cue. The cued location was most likely to contain the target (rotated letter ‘T’ or ‘L’), but the equidistant uncued locations had different probability for target appearance (41.7% against 8.3% of trials, counterbalanced between the two locations). Short SOAs resulted in additive object-based and probability-based effects, but the longest SOA led to abolishment of the object-based effect. Therefore, reaction time for target identification was influenced only by target location probability as cue-target interval increased.

Although the stimuli and SOA range in Shomstein & Yantis (2004) were different from the current experiment, there is some possibility that the same mechanism of prioritisation shifting may be taking place. The overall null effects in Experiment 4.1 may be due to stronger influence of target appearance probability, which is equal for all object features. Accordingly, the trend for gradient effects may also be linked to this shift in probability orientation, such that the object-belongingness probability of the target may shift from being close to certainty within the cued object, to being correlated with cued feature proximity. This, however, is a very tentative proposition, given that there were no pronounced effects for the cued feature, or indeed no pronounced gradient pattern other than the BF evidence of 3.58 for *one object,* which can still be considered as marginal in Bayesian terms (Dienes, 2014).

Given that the IoR effect is linked to increased SOA, it is worth focusing on the effect of SOA manipulation in the current experiment. The standard threshold for switching from facilitation to inhibition is typically accepted to be 300 ms (Klein, 2000), although there can be a substantial variation. The SOA intervals adopted here (900 ms and 1300 ms) were meant to be well above this threshold in order to maximise the chances of obtaining IoR, given that IoR is considered to require more time for target discrimination tasks (Lupiáñez et al., 1997). For the *one object* condition there was a reduction in overall reaction time as SOA increased. This may be linked to IoR for 900 ms, which starts to disappear for 1300 ms. However, to establish this it would be necessary to have a baseline against which these overall reaction time changes are assessed.

Performance at SOA of 300 ms may be a suitable reference point, given it is the accepted threshold for the facilitation-inhibition boundary. The composite analysis for the *one object* condition performed in Chapter 2 was based on pooled data from Experiments 2.1 and 2.2 using the same type of stimuli where SOA was approximately 300 ms, and cueing resulted in facilitation effects. These pooled reaction times add up to a rounded average of 542 ms (i.e. the average reaction time for the four distances). In comparison, the mean reaction time in Experiment 4.1 varied between 586 ms (SOA = 900 ms) and 576 ms (SOA = 1300 ms). Although not based on formal statistical analyses, this observation suggests that mean reaction time for 300 ms SOA was considerably faster than SOA of 900 ms (a difference of 44 ms), and less so than SOA of 1300 ms (a difference of 34 ms). This may provide some support for object-based IoR in the current data, but it cannot be established with certainty based on the current experimental manipulations alone.

Analogous comparison between the *two objects* condition in Experiment 4.1 and the composite analysis in Chapter 2 is not as informative, as there are additional variations due to object belongingness (cued or uncued object) making the comparison more complicated. For the *two objects* condition it is the relative difference between cued and uncued object features that is of interest. Faster reaction times for cued object features relative to uncued object features is taken to signify object-based facilitation, whereas the opposite pattern (slower reaction times for cued object features) is to be associated with object-based inhibition. No evidence was found for the latter. In addition, the overall reaction time difference between the various SOAs is not as large as for the *one object* condition comparison. Specifically, in Experiment 4.1 reaction times for *two* *objects* varied between an average of 579 ms (SOA = 900 ms) and 572 ms (SOA = 1300 ms), while for the composite data for SOA of 300 ms the average was 564 ms (i.e. 15 ms and 8 ms differences, respectively). Therefore, it can be suggested that for *two objects* there was not a lot of increase in overall reactions times as a result of extended SOAs.

Given the inconclusive results, it is worth considering what variables may have hindered the expression of IoR, regardless of whether it follows a non-monotonic or a gradient pattern. It may be that the study was underpowered. Although the sample consisted of 21 participants for each object condition (based on actual power calculations) and this sample size is within the range used in IoR experiments, in an IoR meta-analysis study Samuel & Kat (2003) note that it is only with a large pooled sample that they found a stable IoR at longer SOAs (up to 1600 ms). This effect was otherwise inconsistent when examining individual studies. Also, in their follow-up study, Samuel & Kat used a sample of 40 participants and found that IoR can last up to 2000 ms for the cued location. The task, however, required target detection as opposed to discrimination. In relation to this, it should be emphasised again that IoR for discrimination tasks is more difficult to obtain. When a target needs to be discriminated, IoR occurs later (around 700 ms), and disappears earlier than for detection tasks (Lupiáñez et al., 1997; Lupiáñez, Milliken, Solano, Weaver, & Tipper, 2001). Given that it is challenging to establish a cut-off point for IoR duration, it may be the case that the peak of the effect with respect to optimal SOA duration was missed out in Experiment 4.1, i.e. it was not at any of the tested time intervals and may have occurred earlier.

In addition, the SOA range within a block of trials has been found to have an influence on the onset of IoR. For example, in the case of randomly alternating between three SOA intervals, when they span a longer range (e.g. 100-400-700 ms) IoR appears earlier than with a shorter range (e.g. 100-300-500 ms) (Cheal & Chastain, 2002). Therefore, Experiment 4.1 may have been more successful in producing IoR if a larger SOA range was used, e.g. 600-900-1300. In addition, Cheal & Chastain (2002) note that another factor which affects the crossover from facilitation to inhibition is the number of potential target locations (placeholders), with earlier IoR onset related to fewer locations (they tested 2 versus 4 locations). It is complicated to see how this may apply to the current study, since there were six potential target locations in any condition, but they were organised within a single object or within two objects. From an object-based perspective, however, the number of target locations per se should not make a difference, but rather the number of objects is what counts. By this logic, inhibition may be expected to appear earlier on average for the *one object* than for the *two objects* condition. At present there is not enough evidence in the data to explore this possibility. Nevertheless, the trend for inhibition in the *one object* condition (based on informal comparisons with Chapter 2) contrasted with the tendency for object-based facilitation in the *two objects* condition for the same SOA (900 ms). Therefore, the effect of perceived number of objects versus locations may be an interesting question to follow up.

Overall, Experiment 4.1 was not fully successful in demonstrating IoR, and consequently it cannot be concluded whether IoR can exhibit a pure object-based pattern without effects of cue-target spatial separation. There was some evidence of uniform inhibition in the *one object* condition, inferred from slower reaction times compared to experiments with 300 ms SOA from Chapter 2. Also, the results were useful in supporting the possibility that features belonging to the same object can be equally selected (in this case, potentially inhibited) when there is no top-down preference towards the cued feature. This can be suggested based on the lack of a cueing effect for distance 1. However, these propositions remain to be confirmed with a replication and a suitable control condition. Also, since no clear pattern of object or space-oriented IoR was found, the study raises questions about the methodological aspects that influence IoR onset and expression.

As a future direction it may be useful to establish the optimal conditions for IoR with the current stimuli (e.g. SOA duration and range), and then investigate whether IoR is affected by the number of perceived objects or the absolute number of possible target locations. Another useful aspect to study would be the level of target-object integration and its effect on IoR expression. In general, a good practice would be to establish the methodological parameters that can induce a stable spatial gradient pattern of IoR (e.g. sample size, timings, visual stimuli specifications), and then incrementally introduce manipulations towards integrating the target into perceptual objects. It should be emphasised again that a graded pattern of performance, whether it is related to facilitation or inhibition, does not necessarily suggest that selection is space-based, as the gradient may in reality be due to a probability of shared objecthood. Therefore, the incremental manipulations towards target-object integration would help to establish a profile of the conditions that influence object belongingness probabilities. If the functionality of IoR is to be a foraging facilitator (Klein & MacInnes, 1999), then it is valid to hypothesise that IoR should be object, rather than space-oriented. As a result, with appropriate perceptual organisation manipulations a gradient pattern should be eventually transformed into distance-independent, object-based inhibition. Therefore, additional empirical work is needed to establish if an IoR pattern can corroborate the findings from object-based facilitation in the previous chapters.

Chapter 5

General Discussion

Summary of Empirical Aims & Hypotheses

The present empirical work aimed to develop the understanding of visual selection mechanisms, specifically relating to the role of objects in determining which part of the visual scene gets prioritised for processing. This research was focused specifically on evaluating the role of objects as counterposed to the role of space, since previous research proposes that spatial factors, such as stimulus location, have a key influence on visual perception (e.g. see Carrasco, 2011; Chen, 2012 for a review). Although selection is recognised to be object-based, i.e. there is processing prioritisation of information contained within a selected object relative to information outside this object, this effect is typically explained with reference to space, namely, suggesting that it is a congregation or an array of spatial locations grouped by the object boundary that gets selected, rather than the object itself (e.g. O'Grady & Müller, 2000; Vecera, 1997). The current work proposes an alternative and more parsimonious possibility, which attempts to explain visual selection from an object-oriented perspective.

The main hypothesis of the current project predicted the primacy of objects over space, and supported the notion that effects appearing to be a result of spatial factors, such as the spatial separation between stimuli, are in fact originating from the structure and probabilistic relationship between and within objects in the environment. Therefore, what is considered as evidence for selection on the basis of space is actually selection on the basis of how the available information is perceived to be integrated into objects. In turn, spatial selection as such does not exist as a genuine phenomenon, but is created, i.e. emerges, as a consequence of object-level factors. In other words, there is only one 'type' of selection, and that is object-based selection. One of the most challenging aspects of this account, and incidentally a point vulnerable to criticism, is anticipating what the perceptual system would class as an object. It is therefore challenging to define what a perceptual object is, as it can vary depending on the perceptual organisation of the information on the visual scene and also the subjective state of the perceiver (e.g. Macken, Taylor, & Jones, 2015).

One of the key points argued here is that what information gets prioritised (i.e. selected) depends on the extent to which it is perceived to be a part of a behaviourally-relevant object. Consequently, potential perceptual uncertainty regarding what belongs to which object can lead to a modulation of the consequences of selection, be it facilitation or inhibition of processing. That is, two stimuli are selected together and equally prioritised when there is a high probability that they are part of the same object. Similarly, when an object is selected, all of its features are likely to be selected to the same extent, i.e. without a level of gradation with reference to distance or any other aspect. The critical factor then is the probability associated with what information is part of which object. Related to this, however, is the underlying problem of defining an object. These issues are resolved most likely by Bayesian-type inference based on the prior experience with regularities and structures of the visual world, and probability learning during interaction with the environment (Geisler & Diehl, 2003; Kersten, Mamassian, & Yuille, 2004; Quinn, Bhatt, & Hayden, 2008). Therefore, deciding what information is to be integrated into a single perceptual object can be affected by the presence and ratio of certain ‘cues to objecthood’, which originate from the properties of objects in the natural environment. These are properties such as similarity, good continuation, symmetry, proximity, and other factors relating to non-accidental regularities (Strother & Kubovy, 2006).

Considering this possibility, the current empirical work aimed to explore the prospect that selection is purely object-based, and that perceptual objects determine what information is prioritised, both in terms of immediate behavioural response, and delayed responding involving short-term memory processes. Using a cueing paradigm, the experiments in Chapters 2 and 4 were focused on testing the hypothesis that when the probability of the target belonging to the cued object is high, there is no influence of spatial factors on target selection (assessed in terms of facilitation and IoR). What guides selection is whether this target is part of the cued object or not (that is, of course, controlling for potential confounding factors, such as eccentricity, luminance, top-down expectations, etc.).

From this hypothesis it follows that target selection is directly dependent on the probability of it being part of the cued object. This possibility was tested with the use of object stimuli that address common limitations in previous studies, such as a confounding correlation between spatial proximity and object belongingness, as well as issues regarding the perceptual integration of the target. Chapter 3 was concerned with the effects of perceived object belongingness of memory targets in a cued colour change detection task. It was hypothesised that VSTM is superior for targets within the same perceived object as an uninformative pre-cue, compared to targets situated on a different (uncued) object. This possibility was assessed under conditions where the cue-target distance for same- and different-object memory probes was held constant, and also for cases when this distance was varied (while object belongingness probability was held constant) between the time of memorising (study phase) and the time of remembering the information (test phase).

Summary and Implications of Findings

In summary, Chapter 2 provided evidence that varying cue-target distance makes no difference for the speed of identification of a target event. Instead, what affects performance is whether this target is part of the cued object, resulting in a performance function that reflects the probability of object-belongingness of the target (Experiment 2.1). That is, if there is a high probability that the target is an integral part of the cued object, and this probability of object belongingness is constant across space within the same object, varying target distance from the cued object feature/ location makes no difference to reaction times for identifying uncued targets (*one object* condition). Accordingly, when cue-target distance is non-monotonically related to the probability of object belongingness of the target, i.e. there is an alternation between being in the same object as the cued feature or in a different object as distance increases, then performance alternates in a similar fashion, favouring features on the cued object (*two objects* condition). Other than object belongingness probability, another factor influencing performance is related to whether the target constitutes a change to the cued feature itself, or to any of the remaining five object features. Priority is given to cued feature targets, reflecting the well established cueing effect (e.g. Posner, 1980).

This object-based pattern of selection holds true for various object scales. When the scale of the objects is increased or decreased, the distance between object features, and thus cue-target distance, is accordingly changed too. However, performance remains unaffected by these scaling variations, and instead is independently guided by the perceptual organisation of the stimuli into either one or two objects (Experiment 2.2). Moreover, these effects are not simply due to some feature-based grouping of information, e.g. on the basis of common colour (Experiment 2.3). Although it is ecologically valid that parts or features of the same object are also likely to share common colouring, and thus common colour can be a cue to shared objecthood, it is all of the information on the scene which is evaluated to reach a perceptual decision. Therefore, under more visually-impoverished circumstances where colour commonality may be the only available cue, this could indeed be a key factor in determining which parts of the scene get prioritised together. However, with multiple information available, such as perceived physical connections and symmetry, common colour does not have as much weighting. Importantly, this is not to say that the resulting behavioural outcome is a combination of feature-based and object-based selection (Kravitz & Behrmann, 2011), because it is in fact the same type of process. The perceptual system may use different cues to estimate which aspects are most likely to be perceived as an object unit, and the importance if these cues can vary depending on the combined information and task demands.

A critical demonstration in favour of the case of pure object-based selection was evident in Experiment 2.4, which preserved the spatial coordinates of the cue and targets and cueing probabilities from the earlier experiments, but it decreased the likelihood that the targets were part of the cued object. This manipulation resulted in a spatially graded performance within the same object, as opposed to the flat, distance-invariant function observed in Experiments 2.1 - 2.3 where targets were assigned a high probability of being within the cued object. Reaction times in

Experiment 2.4 gradually increased with increasing cue-target separation, even though these targets were always situated on top of the cued object. It was proposed that this gradient is a consequence from the dynamics of target presentation – transient, superimposed stimuli – encouraging segmentation from the object. In turn, this segmentation leads to alterations in the probability of object belongingness of the targets, such that cue-target pairs are more likely to be proximal if they originate from the same object and thus higher prioritiation is given to targets closer to the cue. Importantly, this graded effect is normally regarded as evidence for combined object-based and space-based selection (e.g. Hollingworth et al., 2012), but the set of studies in Chapter 2 revealed that it is not truly due to spatial factors. It is much more likely the result of object-oriented processes involved in calculating probabilities of object formation.

Given these results, it can be suggested that performance is indeed always a function of a gradient, but this is not a *spatial* gradient, it is an *object* gradient. Under some circumstances, such as in Experiment 2.4, this object gradient may be correlated with the spatial separation of the stimuli, but space is nevertheless not the key factor. Effects of spatial separation emerge since two points close together are more likely to belong to the same object than two points further apart (Ruderman, 1997), which also explains the fact that proximity has been established as a very powerful cue to perceptual object formation (Claessens & Wagemans, 2005; Compton & Logan, 1993; Pomerantz & Schwaitzberg, 1975). In other words, the gradient arises from the distribution of selection based on probability of object belongingness. Therefore, ‘spatial’ effects emerge from object-level factors, and not vice versa, suggesting the primacy of objects over space.

The stimuli utilised in Chapter 2 aimed to create conditions where spatial separation and target belongingness were not correlated in a linear fashion, thus addressing challenges and limitations in previous research that may have led to the erroneous conclusion that space is a primordial factor for selection. For example, the critical comparison between same-object and different-object targets is typically done under conditions of equal cue-target separation. In this case the probability of object belongingness with reference to spatial proximity is kept constant, and only the perceptual organisation cues relating to object structure are manipulated (e.g. the two rectangle paradigm of Egly et al., 1994). In addition, attempts to manipulate spatial separation and perceptual organisation independently may also suffer from limitations, such that bringing two similar and symmetrical objects closer together may increase the probability of them being treated as a single perceptual unit (e.g. Vecera, 1994). However, taken together, the results from Chapter 2 revealed that it is not space that matters for information selection, but it is rather the object-level perceptual organisation of the stimuli.

The critical role of objects is also evident in the domain of VSTM, which was the focus of Chapter 3. It was demonstrated that for the same distance, colour change detection is executed faster when the memory probe is situated within the same perceived object as the cue (introduced prior to the study display), compared to a probe in a different object (Experiment 3.1). Importantly, this same-object advantage is equally observable regardless of whether the object is visibly intact or perceptually completed via occlusion cues. The occlusion creates a physical discontinuity between the cued and probed object fragments, but the object structure is nevertheless experienced as complete. Since the spatial aspects of the targets were identical in all three object formation conditions (*intact*, *completed*, and *segmented*), this effect was clearly due to the perceived status of the targets relative to the cued part of the object(s) – namely, whether the probed target was part of the cued object or not.

These results provide evidence that object structures are automatically extracted from the visual scene (since the perceptual organisation of the memory items was not relevant for completing the task), and consequently they determine how information is selected and prioritised for further processing. Objects in the natural world are most often occluded by one another, so the perceptual system extracts all the information to arrive at a coherent interpretation and render the most plausible object formation (Geisler & Perry, 2009). Therefore, just as the results from Experiment 3.1 demonstrated, perceptually completed objects are functionally equivalent to physically intact ones.

Further emphasis for the role of objects in VSTM was provided in Experiment 3.2. Changing the scale, and thus displacing the spatial location, of the stimuli from the time of cueing and memorising the targets to the time of recalling the colour of the test probe made overall no difference to the pattern of performance. However, the same-object benefit (i.e. faster reaction times for uncued probes within the *same arc* object as the cue, compared to uncued probes on the *different arc* object) was not very robust in this experiment. This lack of pronounced same-object advantage can be potentially argued to result as a consequence of space-based coding of information, assuming such coding exists, since *same arc* and *different arc* probes are at equal distance from the cue. However, even though the standard ANOVA failed to detect a significant difference between change detection for *same arc* and *different arc* probes, the Bayesian analysis confirmed that there is substantial support for object-based effects.

Additionally, the Bayesian evidence for a same-object advantage was stronger in the *changed* scale than in the *unchanged* scale condition, which goes against the possibility for a combined space-based and object-based effect. A combination effect is more likely to lead to the opposite pattern, whereby there is a stronger same-object advantage for the *unchanged* scale. This is so because in this condition all spatial coordinates are preserved from study to test, while in the *changed* condition the stimuli appear at different absolute locations. Therefore, it may be the case that the change of object scale within the same trial emphasised the perceptual organisation of the stimuli, thus strengthening the object-based effect. In any case, a possibility for space-based VSTM was not supported.

Taken together, the results from Chapter 3 support the notion that information processing for VSTM is influenced by object-based, as opposed to space-based factors. When cue-target distance is held constant, the speed of change detection is clearly affected by the perceived structure of the objects within which the stimuli are situated. Moreover, this appears to be enhanced when the objects change their scale at recall. Also, the object-based advantage was evident only in the pattern of reactions times, as accuracy was only affected by whether the memory probe was directly cued or not, i.e. accuracy was insensitive to whether the probed item was elsewhere within the cued object arc. This emphasises the importance of using measurements that are sufficiently sensitive to identify underlying effects. In conjunction with the results from Chapter 2, it is clear that object-level perceptual organisation has a profound and prime influence on how information is processed, whether this is for the purpose of an immediate perceptual decision (luminance change identification), or remembering visual information for a delayed change detection.

An important point to be emphasised here is the robust cueing effect observed in all experiments in Chapter 2 and 3. This is often considered as evidence for space-based selection. However, in Experiments 2.1-2.4, the cue predicted the most likely target location, and thus it had valuable information for performing the task. Therefore, superior performance when responding to changes of the cued feature is not directly related to space, but it is a top-down, i.e. voluntary and intentional, strategic orientation. Nevertheless, this effect was clearly observed in the VSTM experiments, where the cue was *not* correlated with the location of the target. Even when participants had no incentive to place more importance on the cued location, change detection performance was superior for *cued* probes compared to all others. It may be suggested that if the experiments in Chapter 3 were to provide evidence for a pure object-based effect, it should be the case that there is no difference between *cued* and *same arc* probes, and they are both associated with faster responses than *different arc* probes. In turn, the fact that there is a level of gradation, i.e. *cued* probes are superior, followed by *same arc* probes, followed by *different arc* probes, may be taken to suggest the potential interplay of space-based and object-based effects. However, since in Experiment 3.2 the benefit for the *cued* probe was of an equal magnitude for *changed* and *unchanged* scale, while the spatial location in the *changed* condition did not match the cued location, it can be concluded that this effect was not necessarily due to spatial coding of information.

There are of course various methodological differences between the experiments in these two chapters, making a direct comparison rather problematic. For example, in Chapter 2, the cue did not appear inside the object feature that it indicated, while in Chapter 3 the cue was internal to the object arc, and it was also visually similar to the targets (i.e. a filled circle appearing on the object surface). In any case, when all the evidence is taken into account, i.e. equal level of prioritisation for information within the same perceptual object as the cue (as indicated from the empirical work so far), it is clear that information is selected and prioritised with reference to the object structures in the environment, even if these objects appear spatially discontinuous.

The object-based effect, however, was not observed for the inhibition of responses. Chapter 4 aimed to test if IoR can follow the same gradient of object belongingness probability as the one observed for facilitation effects in Chapter 2. Specifically, a flat cue-target distance function of luminance change identification for targets within the same object as the cued feature, and a non-monotonic pattern of performance in the *two objects* condition. SOA was increased in Experiment 4.1, so the prediction was of slower reaction times for targets within the cued object compared to targets representing luminance changes in the uncued object (i.e. an object-based inhibition pattern). Although the same types of stimuli were used as in Chapter 2, the results failed to indicate strong support for IoR. When all targets were within the same object (*one object* condition), there was an indication of overall slower reaction times at 900 ms SOA compared to 1300 ms. There was also some tendency towards a spatial gradient of gradually decreasing reaction times for the *one object* condition. However, when the same information was organised into two objects, there was a strong trend towards facilitation, following the object-based, non-monotonic pattern for 900 ms SOA. For the longer SOA of 1300 ms there was no clear pattern of performance, as reactions times were overall unaffected by target distance.

Given that the experiment in Chapter 4 indicated no cueing effect in either direction (towards either relative facilitation or inhibition), while there was some evidence towards object-based facilitation, it may indeed be the case that when the cue does not encourage strategic orientation towards a specific object feature, all features of the cued object are equally selected. Although the experiments in Chapter 3 also employed a non-informative cue and there was a strong cueing effect despite that, the tasks in the two chapters were quite different, and so was the time interval between cue and target (it being much longer for Chapter 4 experiments). In any case, at this stage no definitive conclusions can be reached on the basis of the experiment in Chapter 4 alone. It may be the case that IoR does not follow the same pattern as facilitation does, or the adopted range of SOAs did not afford inhibition with this type of stimuli and task.

Considering the broader context within which the current work is situated, the results contribute to the understanding of visual selection mechanisms by proposing a novel perspective. This perspective aims to emphasise the critical role of objects, suggesting that information from the environment is processed on the basis of how it is organised into objects. To navigate in the environment and complete any task, the organism needs to engage and interact with objects. The behaviourally relevant object needs to be selected, leading to concurrent processing of all of its parts, in order to calibrate the necessary action (Allport, 1989; Neumann, 1987). To perform this optimally, the incoming visual information needs to be assessed in terms of how it is organised into objects - categorise which bits of information belong to which object. This process is accomplished by relying on certain cues or heuristics learnt through continuous interaction with the environment and experiencing objects, leading to a mechanism that can be described as a probability gradient of object formation, or object belongingness. In turn, this object-based probabilistic mechanism influences which parts of the visual scene become prioritised for processing. The end result is that any behavioural effect should be explainable by analysing the object-level perceptual organisation the environment, and the actions it affords. More specifically, the current work demonstrated how this object-oriented mechanism can accommodate effects which were previously attributed to selection on the basis of space.

At this point it is worth discussing how the current object-oriented perspective differs from the grouped array account (e.g. Vecera & Farah, 1994), which also recognises the role of objects in influencing selection, such that the locations corresponding to the object get activated. An extension of the grouped array account attempts to explain visual selection in terms of *object-directed location selection* (Kim & Cave, 2001). The argument is that object-level factors (e.g. any

Gestalt cues to objecthood) can guide the allocation of spatial selection. More specifically, after a target is selected, other perceptually grouped locations can be automatically prioritised over locations associated with other objects (or perceptual groups). This account can in theory explain the non-monotonic pattern of performance in the *two objects* conditions from a space-based point of view. That is, the locations within the cued object were prioritised over the locations in the uncued object, thus eliminating the potential ‘benefit’ of cue-target spatial proximity for distance 2. Similarly, for the *one object* condition the locations within the object may be equally prioritised, rather than the object itself. Such location-based account, however, would have difficulties explaining the results of the emerging spatial gradient in Experiment 2.4, and the scale-invariant effects in Experiment 2.2, as well as previous evidence in the literature implicating the importance of the strength of object representation and target-object integration (e.g. Zhao et al., 2013). If selection is genuinely location-based, then it should not be dependent on probability of object belongingness of the target. Therefore, it is important to develop this notion further, and explore in more detail (i.e. with additional empirical work) the variables that lead to the emergence of spatial gradients. At this point the debate becomes also philosophical, rather than purely empirical, leading back to the question of what defines an ‘object’ and what is the origin of object-based effects.

Limitations and future directions

The experiments featured in the current work are of course not without limitations. Perhaps the most prominent point to be noted here, which is relevant to all experiments in Chapter 2 and Chapter 4, is the pattern of performance for targets at distance 4, directly opposite the cued object feature. Reaction times for these targets exhibited a consistent counterintuitive trend towards facilitation, which is not in line with any of the tested hypotheses. That is, it goes against both object-oriented non-monotonic and gradient performance patterns. As mentioned earlier, the most likely and simple explanation is the directionality of the cue, which is always oriented along the axis passing directly through the cued feature and the feature at distance 4. Therefore, the cue may be perceived as pointing towards the feature at distance 4. Although this possibility was identified, for the sake of consistency the appearance of the cue was preserved as it is throughout the studies, but the interpretations of the results were often not with consideration for targets at distance 4 (e.g. in the case of Bayesian analyses). However, this phenomenon needs to be investigated further, and the most appropriate way would be to preserve all methodological aspects as they are, while testing performance with various types of cue (in terms of physical appearance). Perhaps the most appropriate modification would be changing the cue to a dot, still centred just outside the cued feature on the same axis. In this way, the cued feature should be unambiguously indicated, without providing any implicit directionality. If the facilitation for distance 4 persists, the symmetry of the object stimulus may also be manipulated.

In addition to the point above, if selection is indeed space-invariant, the circular object features need not be equally spaced to obtain the same effects. Also, the connections between the circles may be repositioned, forming shapes other than overlapping triangles, and thus formulating a variety of perceptual organisations with different object-belongingness gradients versus the same cue-target spatial separations (distances 1-4). Alternatively, or in addition, more potential targets can be added, allowing for more complex shapes and a wider range of distances. Finally, the task can also be adapted into a divided attention paradigm, where two object features, either belonging to the same or different objects, can simultaneously change their property (e.g. shape or colour) to be compared as *same* or *different*. The main principle of all these manipulations would be to test different conditions where object belongingness is not correlated with spatial proximity, and compare results with conditions where this correlation is preserved. If the results consistently support the pure object-oriented hypothesis, this would provide additional evidence it is a genuine effect and space-based patterns of selection are emerging from object structures.

In relation to the emergence of space point, the stimuli in Experiment 2.4 may be developed further, in order to test if the observed spatial gradient of performance can revert back to a flat function when the targets are well integrated within the ring stimulus. For this purpose, the star object from the *one object* condition in Experiments 2.1-2.3 can be modelled with 3D software to appear as three-dimensional, combined with two-dimensional, superimposed target letters ‘X’ or ‘O’ (as the studies in Appendix 2 suggested that superimposing these letter targets on the standard two-dimensional star stimulus did not change the space-invariant pattern). If the object-based hypothesis is supported, this manipulation should result in the emergence of a spatially graded performance. Alternatively, the apparent three-dimensional ring from Experiment 2.4 can be used with a different type of targets, which appear as integrated features. For example, the task may be to discriminate whether the target was a round dent or bulge into the surface of the object (Figure 43). Under these conditions, it can be expected that performance would not vary with distance. However, it should also be ensured that the task is comparable in difficulty with Experiment 2.4.

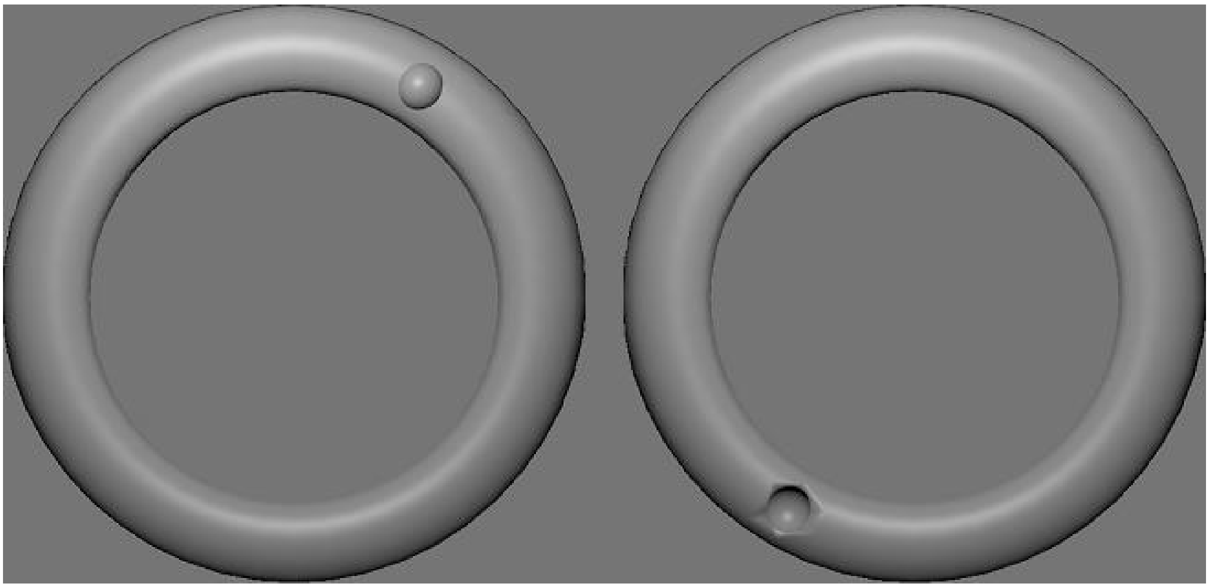

Figure 43: Sample illustration of an apparent 3D ring stimulus with integrated targets: left - a bulge target; right - a dent target. This is an example of the potential final stage of the target. In order to support an integrated perception, the bulge or dent would be gradually formed across a couple of frames, as opposed to abruptly appearing on the surface.

An additional manipulation which can be informative regarding the object belongingness hypotheses can involve introducing the same luminance changes as targets, with identical spatial and temporal properties as in Experiment 2.1 (300 ms SOA), but without any object to contain them. In other words, the cue and target would appear on a blank, uniform background without placeholders or any constantly visible stimuli that can encourage some form of perceptual organisation. Since in this case there are no obvious cues to objecthood other than the proximity of the stimuli, it would be hypothesised that performance will follow a spatial gradient of increasing reaction time with increasing cue-target distance.

The change detection experiments in Chapter 3 can also benefit from further developments. For example, comparing the results from Experiment 3.1 with a similar experiment where a post-cue is utilised, i.e. introducing the cue after the offset of the study display. Replicating the object-based effect in this case would provide further support that object-oriented mechanisms are at play also with abstract representations within VSTM. Specifically, all items may be remembered with an equal weight at the time of presentation, but cueing after their encoding in VSTM may still lead to enhancement or prioritisation of the cued object. It is worth noting that even tasks that are not considered as studying abstract VSTM aspects, but only immediate responses to directly observable stimuli, cannot fully separate one from the other. For example, any cueing paradigm involves a VSTM element, since what was indicated by the cue needs to be ‘held’ in memory in order to complete the task.

An additional investigation of the functional equivalence between *intact* and *completed* objects can be conducted by manipulating the perceptual integrity of the object from study to test. Specifically, it may be expected that remembering the items within a *completed* object, and then performing change detection with an *intact* object (or vice versa) should not impair performance relative to when the object remains unchanged within a trial. This is because in either case the perceptual organisation of the display is the same - the occluded object is the same as the physically intact object (Kellman, 2003; Kellman & Shipley, 1991). On the other hand, switching between *segmented* and *intact/completed* conditions may disturb VSTM because the global context and object-level relationships between the items would be altered, even though their spatial locations would remain unchanged. In relation to this point, Experiment 3.2 may be followed up by a similar methodology, but a larger displacement of the memory items from study to test, since the current manipulation may be criticised for being too conservative in terms of spatial displacement. As already mentioned, additional manipulations may involve rotation during the retention interval, such that the absolute cued location would no longer correspond to the relative (object-based) cued location at test. Therefore, this would allow an alternative method of pitting object versus space-based selection.

In summary, the present work offered an integral and parsimonious perspective on visual selection, which was successfully backed up by empirical evidence, albeit with the need for replication and further investigation. This perspective, namely that visual selection is fully object-oriented, is in line with an adaptive functionality for selection and the idea that visual perception is shaped by the properties and regularities of the visual environment (Ruderman, 1994; Simoncelli & Olshausen, 2001). It is no doubt counterintuitive to suggest that space-based selection does not exist, and that it is just an emergent phenomenon due to object-oriented processes, but the data presented above support the possibility that this may indeed be the case. Space has been considered as a primary domain for vision not only because it seems intuitively logical that all objects are situated in 'space', i.e. space precedes them, but also because there is a lot of focus on the fact that information in the visual cortex is initially coded in retinotopic format (Cavanagh, 2011; Hubel & Wiesel, 1962). However, there is an increasing focus towards the idea that ensemble neural networks can explain much more emergent functionalities that the brain has evolved for, and studying the properties of single cells for the purpose of explaining any cognitive process is very limiting, simplistic and incomplete (Duncan, 2013; Hannus et al., 2005; Yuste, 2015). Objects are functional units for perception and for action (Kellman, 2003; Neumann, 1996), and thus the most obvious and meaningful consequence is that objects are also the unit for visual selection.

References

Albrecht, A. R., List, A., & Robertson, L. C. (2008). Attentional selection and the representation of holes and objects. *Journal of Vision*, *8*(13), 8–8. http://doi.org/10.1167/8.13.8

Allport, A. (1987). Neurophysiological considerations of attention and action. In H. Heuer & A. Sanders F. (Eds.), *Perspectives on Perception and Action* (pp. 395–419). Hillsdale, N.J.: Erlbaum Associates.

Allport, A. (1989). Visual Attnetion. In M. I. Posner (Ed.), *Foundations of Cognitive Science* (pp. 631–682). Cambridge: The MIT Press.

Allport, A. (1992). Attention and control: Have we been asking the wrong questions? A critial review of twenty-five tears. In D. E. Meyer & S. Kornblum, *Attention and Performance XIV* (pp. 183–218). Cambridge: The MIT Press.

Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends in Cognitive Sciences*, *15*(3), 122–131. http://doi.org/10.1016/j.tics.2011.01.003

Alvarez, G. A., & Franconeri, S. L. (2007). How many objects can you track?: Evidence for a resource-limited attentive tracking mechanism. *Journal of Vision*, *7*(13), 14–14. http://doi.org/10.1167/7.13.14

Anderson, D. E., Vogel, E. K., & Awh, E. (2013). Selection and storage of perceptual groups is constrained by a discrete resource in working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(3), 824–835. http://doi.org/10.1037/a0030094

Ariely, D. (2001). Seeing Sets: Representation by Statistical Properties. *Psychological Science*, *12*(2), 157–162. http://doi.org/10.1111/1467-9280.00327

Armitage, P., McPherson, C. K., & Rowe, B. C. (1969). Repeated Significance Tests on Accumulating Data. *Journal of the Royal Statistical Society. Series A (General)*, *132*(2), 235. http://doi.org/10.2307/2343787

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

Awh, E., Belopolsky, A. V., & Theeuwes, J. (2012). Top-down versus bottom-up attentional control: a failed theoretical dichotomy. *Trends in Cognitive Sciences*, *16*(8), 437–443. http://doi.org/10.1016/j.tics.2012.06.010

Awh, E., Dhaliwal, H., Christensen, S., & Matsukura, M. (2001). Evidence for Two Components of Object-Based Selection. *Psychological Science*, *12*(4), 329–334. http://doi.org/10.1111/1467-9280.00360

Awh, E., & Jonides, J. (2001). Overlapping mechanisms of attention and spatial working memory. *Trends in Cognitive Sciences*, *5*(3), 119–126.

Baddeley, A. D. (2010). Working memory. *Current Biology*, *20*(4), R136–R140.

Baddeley, R. (1997). The correlational structure of natural images and the calibration of spatial representations. *Cognitive Science*, *21*(3), 351–371.

Baldauf, D., & Desimone, R. (2014). Neural Mechanisms of Object-Based Attention. *Science*, *344*(6182), 424–427. http://doi.org/10.1126/science.1247003

Banks, W., & Prinzmetal, W. (1976). Configurational effects in visual information processing. *Perception & Psychophysics*, *19*(4), 361–367.

Bao, Y., Lei, Q., Fang, Y., Tong, Y., Schill, K., Pöppel, E., & Strasburger, H. (2013). Inhibition of Return in the Visual Field: The Eccentricity Effect is Independent of Cortical Magnification *The first two authors contributed equally to this study. *Experimental Psychology*, *60*(6), 425–431. http://doi.org/10.1027/1618-3169/a000215

Baumann, O., Skilleter, A. J., & Mattingley, J. B. (2011). Short-Term Memory Maintenance of Object Locations during Active Navigation: Which Working Memory Subsystem Is Essential? *PLoS ONE*, *6*(5), e19707. http://doi.org/10.1371/journal.pone.0019707

Baylis, G. C., & Driver, J. (1992). Visual parsing and response competition: The effect of grouping factors. *Perception & Psychophysics*, *51*(2), 145–162.

Bays, P. M., & Husain, M. (2008). Dynamic Shifts of Limited Working Memory Resources in Human Vision. *Science*, *321*(5890), 851–854. http://doi.org/10.1126/science.1158023

Bays, P. M., & Husain, M. (2012). Active inhibition and memory promote exploration and search of natural scenes. *Journal of Vision*, *12*(8), 8–8. http://doi.org/10.1167/12.8.8

Behrmann, M., Zemel, R. S., & Mozer, M. C. (1998). Object-based attention and occlusion: evidence from normal participants and a computational model. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(4), 1011–1036.

Behrmann, M., Zemel, R. S., & Mozer, M. C. (2000). Occlusion, symmetry, and object-based attention: Reply to Saiki (2000). *Journal of Experimental Psychology: Human Perception and Performance*, *26*(4), 1497–1505. http://doi.org/10.1037//0096-1523.26.4.1497

Bennett, P. J., & Pratt, J. (2001). The Spatial Distribution of Inhibition of Return. *Psychological Science*, *12*(1), 76–80. http://doi.org/10.1111/1467-9280.00313

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. http://doi.org/10.1016/j.visres.2006.10.019

Berryhill, M. E., Richmond, L. L., Shay, C. S., & Olson, I. R. (2012). Shifting attention among working memory representations: Testing cue type, awareness, and strategic control. *The Quarterly Journal of Experimental Psychology*, *65*(3), 426–438. http://doi.org/10.1080/17470218.2011.604786

Boduroglu, A., & Shah, P. (2009). Effects of spatial configurations on visual change detection: An account of bias changes. *Memory & Cognition*, *37*(8), 1120–1131. http://doi.org/10.3758/MC.37.8.1120

Boduroglu, A., & Shah, P. (2014). Configural representations in spatial working memory. *Visual Cognition*, *22*(1), 102–124. http://doi.org/10.1080/13506285.2013.875499

Borst, G., Ganis, G., Thompson, W. L., & Kosslyn, S. M. (2012). Representations in mental imagery and working memory: Evidence from different types of visual masks. *Memory & Cognition*, *40*(2), 204–217. http://doi.org/10.3758/s13421-011-0143-7

Bosking, W. H., Zhang, Y., Schofield, B., & Fitzpatrick, D. (1997). Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. *The Journal of Neuroscience*, *17*(6), 2112–2127.

Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature*, *226*, 177–178.

Bourke, P. A., Partridge, H., & Pollux, P. M. J. (2006). Additive effects of inhibiting attention to objects and locations in three-dimensional displays. *Visual Cognition*, *13*(5), 643–654. http://doi.org/10.1080/13506280544000309

Brady, T. F., & Alvarez, G. A. (2011). Hierarchical Encoding in Visual Working Memory: Ensemble Statistics Bias Memory for Individual Items. *Psychological Science*, *22*(3), 384–392. http://doi.org/10.1177/0956797610397956

Brady, T. F., & Tenenbaum, J. B. (2013). A probabilistic model of visual working memory: Incorporating higher order regularities into working memory capacity estimates. *Psychological Review*, *120*(1), 85–109. http://doi.org/10.1037/a0030779

Brawn, P. T., & Snowden, R. J. (2000). Attention to overlapping objects: Detection and discrimination of luminance changes. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(1), 342–358. http://doi.org/10.1037//0096-1523.26.1.342

Brown, H., & Friston, K. J. (2012). Free-Energy and Illusions: The Cornsweet Effect. *Frontiers in Psychology*, *3*. http://doi.org/10.3389/fpsyg.2012.00043

Caparos, S., & Linnell, K. J. (2010). The spatial focus of attention is controlled at perceptual and cognitive levels. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(5), 1080–1107. http://doi.org/10.1037/a0020367

Carrasco, M. (2011). Visual attention: The past 25 years. *Vision Research*, *51*(13), 1484–1525. http://doi.org/10.1016/j.visres.2011.04.012

Castiello, U., & Umiltà, C. (1990). Size of the attentional focus and efficiency of processing. *Acta Psychologica*, *73*(3), 195–209. http://doi.org/10.1016/0001-6918(90)90022-8

Cavanagh, P. (2011). Visual cognition. *Vision Research*, *51*(13), 1538–1551. http://doi.org/10.1016/j.visres.2011.01.015

Cheal, M., & Chastain, G. (2002). Timing of facilitatory and inhibitory effects of visual attention. *Visual Cognition*, *9*(8), 969–1002. http://doi.org/10.1080/13506280143000467

Chen, Z. (2012). Object-based attention: A tutorial review. *Attention, Perception, & Psychophysics*, *74*(5), 784–802. http://doi.org/10.3758/s13414-012-0322-z

Chen, Z., & Cave, K. R. (2006). Reinstating object-based attention under positional certainty: The importance of subjective parsing. *Perception & Psychophysics*, *68*(6), 992–1003.

Chen, Z., & Cave, K. R. (2008). Object-based attention with endogenous cuing and positional certainty. *Perception & Psychophysics*, *70*(8), 1435–1443. http://doi.org/10.3758/PP.70.8.1435

Chong, S. C., & Treisman, A. (2005). Statistical processing: computing the average size in perceptual groups. *Vision Research*, *45*(7), 891–900. http://doi.org/10.1016/j.visres.2004.10.004

Chou, W.-L., & Yeh, S.-L. (2005). Object-based inhibition of return: Evidence from overlapping objects. *Chinese Journal of Psychology*, *47*(1). Retrieved from http://ntur.lib.ntu.edu.tw/handle/246246/136957

Claessens, P. M., & Wagemans, J. (2005). Perceptual grouping in Gabor lattices: Proximity and alignment. *Perception & Psychophysics*, *67*(8), 1446–1459.

Cohen, E. H., & Tong, F. (2013). Neural Mechanisms of Object-Based Attention. *Cerebral Cortex*. http://doi.org/10.1093/cercor/bht303

Compton, B. J., & Logan, G. D. (1993). Evaluating a computational model of perceptual grouping by proximity. *Perception & Psychophysics*, *53*(4), 403–421.

Cosman, J. D., & Vecera, S. P. (2012). Object-based attention overrides perceptual load to modulate visual distraction. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(3), 576–579. http://doi.org/10.1037/a0027406

Cousineau, D. (2005). Confidence intervals in within-subjects designs: A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology*, *1*(1), 42–45.

Cowan, N., Elliott, E. M., Scott Saults, J., Morey, C. C., Mattox, S., Hismjatullina, A., & Conway, A. R. A. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology*, *51*(1), 42–100. http://doi.org/10.1016/j.cogpsych.2004.12.001

Crundall, D., Cole, G. G., & Galpin, A. (2007). Object-based attention is mediated by collinearity of targets. *The Quarterly Journal of Experimental Psychology*, *60*(1), 137–153. http://doi.org/10.1080/17470210600654792

Davis, G., & Driver, J. (1997). SPREADING OF VISUAL ATTENTION TO MODALLY VERSUS AMODALLY COMPLETED REGIONS. *Psychological Science*, *8*(4), 275–281. http://doi.org/10.1111/j.1467-9280.1997.tb00438.x

Davis, G., Welch, V. L., Holmes, A., & Shepherd, A. (2001). Can attention select only a fixed number of objects at a time? *Perception*, *30*(10), 1227–1248. http://doi.org/10.1068/p3133

De Freitas, J., Liverence, B. M., & Scholl, B. J. (2014). Attentional rhythm: A temporal analogue of object-based attention. *Journal of Experimental Psychology: General*, *143*(1), 71–76. http://doi.org/10.1037/a0032296

De Winter, J., & Wagemans, J. (2006). Segmentation of object outlines into parts: A large-scale integrative study. *Cognition*, *99*(3), 275–325. http://doi.org/10.1016/j.cognition.2005.03.004

Dienes, Z. (2011). Bayesian Versus Orthodox Statistics: Which Side Are You On? *Perspectives on Psychological Science*, *6*(3), 274–290. http://doi.org/10.1177/1745691611406920

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology*, *5*. http://doi.org/10.3389/fpsyg.2014.00781

Dodd, M., & Pratt, J. (2005). Allocating visual attention to grouped objects. *European Journal of Cognitive Psychology*, *17*(4), 481–497. http://doi.org/10.1080/09541440440000122

Dougherty, R. F., Koch, V. M., Brewer, A. A., Fischer, B., Modersitzki, J., & Wandell, B. A. (2003). Visual field representations and locations of visual areas V1/2/3 in human visual cortex. *Journal of Vision*, *3*(10), 1–1. http://doi.org/10.1167/3.10.1

Downing, C. J. (1988). Expectancy and visual-spatial attention: effects on perceptual quality. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(2), 188.

Downing, C. J., & Pinker, S. (1985). The spatial structure of visual attention. In M. I. Posner & O. Marin (Eds.), *Attention and Performance XI* (pp. 171–187). Hillsdale, N.J.: Erlbaum.

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.

Duncan, J. (1984). Selective Attention and the Organisation of Visual Information. *Journal of Experimental Psychology: General*, *113*(4), 501–517.

Duncan, J. (1993). Similarity between concurrent visual discriminations: Dimensions and objects. *Perception & Psychophysics*, *54*(4), 425–430.

Duncan, J. (2006). EPS Mid-Career Award 2004: Brain mechanisms of attention. *The Quarterly Journal of Experimental Psychology*, *59*(1), 2–27. http://doi.org/10.1080/17470210500260674

Duncan, J. (2013). The Structure of Cognition: Attentional Episodes in Mind and Brain. *Neuron*, *80*(1), 35–50. http://doi.org/10.1016/j.neuron.2013.09.015

Duncan, J., & Nimmo-Smith, I. (1996). Objects and attributes in divided attention: Surface and boundary systems. *Perception & Psychophysics*, *58*(7), 1076–1084.

Egeth, H. E., & Yantis, S. (1997). Visual attention: Control, representation, and time course. *Annual Review of Psychology*, *48*(1), 269–297.

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.

Elder, J. H., & Goldberg, R. M. (2002). Ecological statistics of Gestalt laws for the perceptual organization of contours. *Journal of Vision*, *2*(4), 324–353. http://doi.org/doi: 10.1167/2.4.5

Emberson, L. L., & Amso, D. (2012). Learning to sample: Eye tracking and fMRI indices of changes in object perception. *Journal of Cognitive Neuroscience*, *24*(10), 2030–2042.

Emmanouil, T. A., & Magen, H. (2014). Neural evidence for sequential selection of object features. *Trends in Cognitive Sciences*. http://doi.org/10.1016/j.tics.2014.04.007

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

Eriksen, C. W., & James, J. D. S. (1986). Visual attention within and around the field of focal attention: A zoom lens model. *Perception & Psychophysics*, *40*(4), 225–240.

Eriksen, C. W., Pan, K., & Botella, J. (1993). Attentional distribution in visual space. *Psychological Research*, *56*(1), 5–13.

Eriksen, C. W., & Yeh, Y. (1985). Allocation of attention in the visual field. *Journal of Experimental Psychology: Human Perception and Performance*, *11*(5), 583.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191.

Fecteau, J., & Munoz, D. (2006). Salience, relevance, and firing: a priority map for target selection. *Trends in Cognitive Sciences*, *10*(8), 382–390. http://doi.org/10.1016/j.tics.2006.06.011

Feldman, J. (1999). The role of objects in perceptual grouping. *Acta Psychologica*, *102*(2), 137–163. http://doi.org/10.1016/S0001-6918(98)00054-7

Feldman, J. (2003). What is a visual object? *Trends in Cognitive Sciences*, *7*(6), 252–256. http://doi.org/10.1016/S1364-6613(03)00111-6

Feldman, J. (2007). Formation of visual ‘objects’ in the early computation of spatial relations. *Perception & Psychophysics*, *69*(5), 816–827.

Fernandez-Duque, D., & Johnson, M. L. (1999). Attention metaphors: How metaphors guide the cognitive psychology of attention. *Cognitive Science*, *23*(1), 83–116.

Festman, Y., & Braun, J. (2010). Does feature similarity facilitate attentional selection? *Attention, Perception & Psychophysics*, *72*(8), 2128–2143. http://doi.org/10.3758/APP.72.8.2128

Fiebelkorn, I. C., Saalmann, Y. B., & Kastner, S. (2013). Rhythmic Sampling within and between Objects despite Sustained Attention at a Cued Location. *Current Biology*, *23*(24), 2553–2558. http://doi.org/10.1016/j.cub.2013.10.063

Field, D. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Sociaty of America*, *4*(12), 2379–2394.

Field, D. (1989). What the statistics of natural images tell us about visual codin. *Human Vision, Visual Processing, and Digital Display*, *1077*, 269–276.

Flombaum, J. I., & Scholl, B. J. (2006). A temporal same-object advantage in the tunnel effect: Facilitated change detection for persisting objects. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(4), 840–853. http://doi.org/10.1037/0096-1523.32.4.840

Forster, S., & Lavie, N. (2008). Failures to ignore entirely irrelevant distractors: The role of load. *Journal of Experimental Psychology: Applied*, *14*(1), 73–83. http://doi.org/10.1037/1076-898X.14.1.73

Fox, E. (1998). Perceptual grouping and visual selective attention. *Perception & Psychophysics*, *60*(6), 1004–1021.

Franconeri, S. L., Alvarez, G. A., & Cavanagh, P. (2013). Flexible cognitive resources: competitive content maps for attention and memory. *Trends in Cognitive Sciences*, *17*(3), 134–141. http://doi.org/10.1016/j.tics.2013.01.010

Freiwald, S. (2007). Attention to objects made of features. *Trends in Cognitive Sciences*, *11*(11), 453–454. http://doi.org/10.1016/j.tics.2007.08.012

Fuller, S., & Carrasco, M. (2006). Exogenous attention and color perception: Performance and appearance of saturation and hue. *Vision Research*, *46*(23), 4032–4047. http://doi.org/10.1016/j.visres.2006.07.014

Gabay, S., Chica, A. B., Charras, P., Funes, M. J., & Henik, A. (2012). Cue and target processing modulate the onset of inhibition of return. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(1), 42–52. http://doi.org/10.1037/a0023675

Gao, Z., Li, J., Yin, J., & Shen, M. (2010). Dissociated Mechanisms of Extracting Perceptual Information into Visual Working Memory. *PLoS ONE*, *5*(12), e14273. http://doi.org/10.1371/journal.pone.0014273

Geisler, W., & Diehl, R. L. (2003). A Bayesian approach to the evolution of perceptual and cognitive systems. *Cognitive Science*, *27*(3), 379–402. http://doi.org/10.1016/S0364-0213(03)00009-0

Geisler, W. S. (2008). Visual Perception and the Statistical Properties of Natural Scenes. *Annual Review of Psychology*, *59*(1), 167–192. http://doi.org/10.1146/annurev.psych.58.110405.085632

Geisler, W. S., & Perry, J. S. (2009). Contour statistics in natural images: Grouping across occlusions. *Visual Neuroscience*, *26*(1), 109. http://doi.org/10.1017/S0952523808080875

Gibson, J. (1966). *The Senses Considered as Perceptual Systems*. New York: Praeger.

Gillmeister, H., Cantarella, S., Gheorghiu, A. I., & Adler, J. (2013). Object-Guided Spatial Selection in Touch Without Concurrent Changes in the Perceived

Location of the Hands. *Experimental Psychology (Formerly Zeitschrift Für Experimentelle Psychologie)*, *60*(1), 64–70. http://doi.org/10.1027/1618-3169/a000180

Griffin, I. C., & Nobre, A. C. (2003). Orienting attention to locations in internal representations. *Journal of Cognitive Neuroscience*, *15*(8), 1176–1194.

Han, S., Humphreys, G. W., & Chen, L. (1999). Uniform connectedness and classical Gestalt principles of perceptual grouping. *Perception & Psychophysics*, *61*(4), 661–674.

Hannus, A., Cornelissen, F. W., Lindemann, O., & Bekkering, H. (2005). Selection-for-action in visual search. *Acta Psychologica*, *118*(1–2), 171–191. http://doi.org/10.1016/j.actpsy.2004.10.010

Hardman, K. O., & Cowan, N. (2015). Remembering Complex Objects in Visual Working Memory: Do Capacity Limits Restrict Objects or Features? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(2), 325–347. http://doi.org/10.1037/xlm0000031

He, X., Fan, S., Zhou, K., & Chen, L. (2004). Cue Validity and Object-Based Attention. *Journal of Cognitive Neuroscience*, *16*(6), 1085–1097. http://doi.org/10.1162/0898929041502689

Henderson, J. M. (1991). Stimulus discrimination following covert attentional orienting to an exogenous cue. *Journal of Experimental Psychology: Human Perception and Performance*, *17*(1), 91.

Henderson, J. M., & Macquistan, A. D. (1993). The spatial distribution of attention following an exogenous cue. *Perception & Psychophysics*, *53*(2), 221–230.

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

Hickey, C., & Theeuwes, J. (2011). Context and competition in the capture of visual attention. *Attention, Perception, & Psychophysics*, *73*(7), 2053–2064. http://doi.org/10.3758/s13414-011-0168-9

Ho, M.-C. (2011). Object-based attention: Sensory enhancement or scanning prioritization. *Acta Psychologica*, *138*(1), 45–51. http://doi.org/10.1016/j.actpsy.2011.05.004

Hoffman, J. E., & Nelson, B. (1981). Spatial selectivity in visual search. *Perception & Psychophysics*, *30*(3), 283–290.

Holcombe, A. O., & Chen, W.-Y. (2013). Splitting attention reduces temporal resolution from 7 Hz for tracking one object to <3 Hz when tracking three. *Journal of Vision*, *13*(1), 12–12. http://doi.org/10.1167/13.1.12

Hollingworth, A., Maxcey-Richard, A. M., & Vecera, S. P. (2012). The spatial distribution of attention within and across objects. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(1), 135–151. http://doi.org/10.1037/a0024463

Hu, F. K., Samuel, A. G., & Chan, A. S. (2011). Eliminating inhibition of return by changing salient nonspatial attributes in a complex environment. *Journal of*

Experimental Psychology: General, *140*(1), 35–50. http://doi.org/10.1037/a0021091

Huang, L., & Pashler, H. (2005). Attention capacity and task difficulty in visual search. *Cognition*, *94*(3), B101–B111. http://doi.org/10.1016/j.cognition.2004.06.006

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology*, *160*(1), 106.

Humphreys, G. W. (1999). Neural representation of objects in space: a dual coding account. In G. W. Humphreys, J. Duncan, & A. Treisman (Eds.), *Attention, Space and Action* (pp. 165–182). Oxford: Oxford University Press.

Ikkai, A., & Curtis, C. E. (2011). Common neural mechanisms supporting spatial working memory, attention and motor intention. *Neuropsychologia*, *49*(6), 1428–1434. http://doi.org/10.1016/j.neuropsychologia.2010.12.020

Jans, B., Peters, J. C., & De Weerd, P. (2010). Visual spatial attention to multiple locations at once: The jury is still out. *Psychological Review*, *117*(2), 637–682. http://doi.org/10.1037/a0019082

Jeffreys, H. (1961). *The Theory of Probability* (3rd ed.). Oxford: Oxford University Press.

Jiang, Y., Chun, M. M., & Olson, I. R. (2004). Perceptual grouping in change detection. *Perception & Psychophysics*, *66*(3), 446–453. http://doi.org/10.3758/BF03194892

Jiang, Y., Olson, I. R., & Chun, M. M. (2000). Organization of visual short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(3), 683.

Jonides, J., Lacey, S. C., & Nee, D. E. (2005). Processes of working memory in mind and brain. *Current Directions in Psychological Science*, *14*(1), 2–5.

Jordan, H., & Tipper, S. P. (1998). Object-based inhibition of return in static displays. *Psychonomic Bulletin & Review*, *5*(3), 504–509.

Jordan, H., & Tipper, S. P. (1999). Spread of inhibition across an object's surface. *British Journal of Psychology*, *90*(4), 495–507.

Joseph, J. S., & Nakayama, K. (1999). Amodal representation depends on the object seen before partial occlusion. *Vision Research*, *39*(2), 283–292.

Kahneman, D., & Henik, A. (1981). Perceptual organization and attention. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual Organization* (pp. 181–211). Hillsdale, N.J.: Erlbaum.

Kanizsa, G. (1976). Subjective contours. *Scientific American*, *234*(4), 48–52.

Katshu, M. Z. U. H., & d'Avossa, G. (2014). Fine-Grained, Local Maps and Coarse, Global Representations Support Human Spatial Working Memory. *PLoS ONE*, *9*(9), e107969. http://doi.org/10.1371/journal.pone.0107969

Katz, D. (1950). *Gestalt Psychology: Its Nature and Significance*. New York: Ronals Press.

Kellman, P. . (2003). Interpolation processes in the visual perception of objects. *Neural Networks*, *16*(5–6), 915–923. http://doi.org/10.1016/S0893-6080(03)00101-1

Kellman, P., & Shipley, T. (1991). A theory of visual interpolation in object perception. *Cognitive Psychology*, *23*(2), 141–221.

Kersten, D., Mamassian, P., & Yuille, A. (2004). Object Perception as Bayesian Inference. *Annual Review of Psychology*, *55*(1), 271–304. http://doi.org/10.1146/annurev.psych.55.090902.142005

Kersten, D., & Yuille, A. (2003). Bayesian models of object perception. *Current Opinion in Neurobiology*, *13*(2), 150–158. http://doi.org/10.1016/S0959-4388(03)00042-4

Kiefer, R., & Siple, P. (1987). Spatial constraints on the voluntary control of attention across visual space. *Canadian Journal of Psychology*, *41*(4), 474–489. http://doi.org/10.1037/h0084168

Kim, M.-S., & Cave, K. R. (2001). Perceptual grouping via spatial selection in a focused-attention task. *Vision Research*, *41*(5), 611–624.

Klapp, S. T., & Jagacinski, R. J. (2011). Gestalt principles in the control of motor action. *Psychological Bulletin*, *137*(3), 443–462. http://doi.org/10.1037/a0022361

Klapp, S. T., Nelson, J. M., & Jagacinski, R. J. (1998). Can People Tap Concurrent Bimanual Rhythms Independently? *Journal of Motor Behavior*, *30*(2), 301–322.

Klein, R. M. (2000). Inhibition of return. *Trends in Cognitive Sciences*, *4*(4), 138–147.

Klein, R. M., & MacInnes, W. J. (1999). Inhibition of return is a foraging facilitator in visual search. *Psychological Science*, *10*(4), 346–352.

Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in Psychtoolbox-3. *Perception*, *36*(14), 1–1.

Koffka, K. (1922). Perception: an introduction to the Gestalt-Theorie. *Psychological Bulletin*, *19*(10), 531–585. http://doi.org/10.1037/h0072422

Korjoukov, I., Jeurissen, D., Kloosterman, N. A., Verhoeven, J. E., Scholte, H. S., & Roelfsema, P. R. (2012). The Time Course of Perceptual Grouping in Natural Scenes. *Psychological Science*, *23*(12), 1482–1489. http://doi.org/10.1177/0956797612443832

Kramer, A. F., & Jacobson, A. (1991). Perceptual organization and focused attention: The role of objects and proximity in visual processing. *Perception & Psychophysics*, *50*(3), 267–284.

Kramer, A. F., Weber, T. A., & Watson, S. E. (1997). Object-based attentional selection—Grouped arrays or spatially invariant representations?: Comment on Vecera and Farah (1994). *Journal of Experimental Psychology: General*, *123*(1), 3–13.

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. http://doi.org/10.1037/0096-1523.34.2.298

Kravitz, D. J., & Behrmann, M. (2011). Space-, object-, and feature-based attention interact to organize visual scenes. *Attention, Perception, & Psychophysics*, *73*(8), 2434–2447. http://doi.org/10.3758/s13414-011-0201-z

Kubovy, M. (1981). Concurrent-pitch segregation and the theory of indispensable attributes. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual Organisation* (pp. 55–98). Hillsdale, N.J.: Lawrence Erlbaum.

Kubovy, M., Holcombe, A. O., & Wagemans, J. (1998). On the lawfulness of grouping by proximity. *Cognitive Psychology*, *35*(1), 71–98.

Kubovy, M., & van den Berg, M. (2008). The whole is equal to the sum of its parts: A probabilistic model of grouping by proximity and similarity in regular patterns. *Psychological Review*, *115*(1), 131–154. http://doi.org/10.1037/0033-295X.115.1.131

Kubovy, M., & Van Valkenburg, D. (2001). Auditory and visual objects. *Cognition*, *80*(1), 97–126.

Kubovy, M., & Wagemans, J. (1995). GROUPING BY PROXIMITY AND MULTISTABILITY IN DOT LATTICES:. A Quantitative Gestalt Theory. *Psychological Science*, *6*(4), 225–234. http://doi.org/10.1111/j.1467-9280.1995.tb00597.x

Laberge, D., & Brown, V. (1986). Variations in size of the visual field in which targets are presented: An attentional range effect. *Perception & Psychophysics*, *40*(3), 188–200.

Lamy, D. (2000). Object-based selection under focused attention: A failure to replicate. *Perception & Psychophysics*, *62*(6), 1272–1279.

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

Law, M. B., & Abrams, R. A. (2002). Object-based selection within and beyond the focus of spatial attention. *Perception & Psychophysics*, *64*(7), 1017–1027.

Lee, D., & Chun, M. M. (2001). What are the units of visual short-term memory, objects or spatial locations? *Perception & Psychophysics*, *63*(2), 253–257. http://doi.org/10.3758/BF03194466

Lee, D., Koch, C., & Braun, J. (1999). Attentional capacity is undifferenciated: Concurrent discrimination of form, color, and motion. *Perception & Psychophysics*, *61*(7), 1241–1255.

Lee, H., Mozer, M. C., Kramer, A. F., & Vecera, S. P. (2012). Object-based control of attention is sensitive to recent experience. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(2), 314–325. http://doi.org/10.1037/a0025597

Levi, D. M., & Carney, T. (2009). Crowding in Peripheral Vision: Why Bigger Is Better. *Current Biology*, *19*(23), 1988–1993. http://doi.org/10.1016/j.cub.2009.09.056

Lin, S.-Y., & Yeh, S.-L. (2012). Every moment counts: smooth transitions of object boundaries reflect constant updating in object-based attention. *Attention, Perception, & Psychophysics*, *74*(3), 533–539. http://doi.org/10.3758/s13414-011-0245-0

Linnell, K. J., & Caparos, S. (2011). Perceptual and cognitive load interact to control the spatial focus of attention. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(5), 1643–1648. http://doi.org/10.1037/a0024669

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. http://doi.org/10.1037/0096-1523.33.6.1322

Livne, T., & Sagi, D. (2007). Configuration influence on crowding. *Journal of Vision*, *7*(2), 4.

Long, G. M., & Toppino, T. C. (2004). Enduring Interest in Perceptual Ambiguity: Alternating Views of Reversible Figures. *Psychological Bulletin*, *130*(5), 748–768. http://doi.org/10.1037/0033-2909.130.5.748

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*(6657), 279–281. http://doi.org/10.1038/36846

Lupiáñez, J., Milán, E. G., Tornay, F. J., Madrid, E., & Tudela, P. (1997). Does IOR occur in discrimination tasks? Yes, it does, but later. *Perception & Psychophysics*, *59*(8), 1241–1254.

Lupiáñez, J., Milliken, B., Solano, C., Weaver, B., & Tipper, S. P. (2001). On the strategic modulation of the time course of facilitation and inhibition of return. *The Quarterly Journal of Experimental Psychology Section A*, *54*(3), 753–773. http://doi.org/10.1080/713755990

Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nature Neuroscience*, *17*(3), 347–356. http://doi.org/10.1038/nn.3655

Macken, B., Taylor, J., & Jones, D. (2015). Limitless capacity: a dynamic object-oriented approach to short-term memory. *Frontiers in Psychology*, *6*. http://doi.org/10.3389/fpsyg.2015.00293

Macmillan, N. A., & Creelman, C. D. (2004). *Detection Theory: A User's Guide* (2nd ed.). New York: Cambridge University Press.

Makovski, T., Sussman, R., & Jiang, Y. V. (2008). Orienting attention in visual working memory reduces interference from memory probes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(2), 369–380. http://doi.org/10.1037/0278-7393.34.2.369

Mangun, G. R., & Hillyard, S. A. (1987). The spatial allocation of visual attention as indexed by event-related brain potentials. *Human Factors:*, *29*, 195–211. http://doi.org/10.1177/001872088702900207

Mangun, G. R., & Hillyard, S. A. (1988). Spatial gradients of visual attention: behavioral and electrophysiological evidence. *Electroencephalography* *and Clinical Neurophysiology*, *70*(5), 417–428. http://doi.org/10.1016/0013-4694(88)90019-3

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.

Marois, R., & Ivanoff, J. (2005). Capacity limits of information processing in the brain. *Trends in Cognitive Sciences*, *9*(6), 296–305. http://doi.org/10.1016/j.tics.2005.04.010

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.

Martín-Arévalo, E., Kingstone, A., & Lupiáñez, J. (2013). Is 'Inhibition of Return' due to the inhibition of the return of attention? *The Quarterly Journal of Experimental Psychology*, *66*(2), 347–359. http://doi.org/10.1080/17470218.2012.711844

Martínez, A., Teder-Sälejärvi, W., Vazquez, M., Molholm, S., Foxe, J. J., Javitt, D. C., … Hillyard, S. A. (2006). Objects are highlighted by spatial attention. *Journal of Cognitive Neuroscience*, *18*(2), 298–310.

Matsukura, M., Cosman, J. D., Roper, Z. J. J., Vatterott, D. B., & Vecera, S. P. (2014). Location-specific effects of attention during visual short-term memory maintenance. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(3), 1103–1116. http://doi.org/10.1037/a0035685

Matsukura, M., & Vecera, S. P. (2011). Object-based selection from spatially-invariant representations: evidence from a feature-report task. *Attention, Perception, & Psychophysics*, *73*(2), 447–457. http://doi.org/10.3758/s13414-010-0039-9

Maunsell, J. H. R., & Treue, S. (2006). Feature-based attention in visual cortex. *Trends in Neurosciences*, *29*(6), 317–322. http://doi.org/10.1016/j.tins.2006.04.001

McAvinue, L. P., Vangkilde, S., Johnson, K. A., Habekost, T., Kyllingsbæk, S., Robertson, I. H., & Bundesen, C. (2012). The relationship between sustained attention, attentional selectivity, and capacity. *Journal of Cognitive Psychology*, *24*(3), 313–328. http://doi.org/10.1080/20445911.2011.628653

McCormick, P., & Klein, R. M. (1990). The spatial distribution of attention during covert visual orienting. *Acta Psychologica*, *75*(3), 225–242. http://doi.org/10.1016/0001-6918(90)90014-7

Mesulam, M. M. (1985). Attention, confusional states, and neglect. In M. M. Mesulam (Ed.), *Principles of Behavioural Neurology* (pp. 125–168). Philadeplhia: F. A. Davis.

Mitchell, J., F., Stoner, G., R., & Reynolds, J., H. (2004). Object-based attention determines dominance in binocular rivalry. *Nature*, *429*(6990), 410–413. http://doi.org/10.1038/nature02554

Moore, C. M., & Fulton, C. (2005). The spread of attention to hidden portions of occluded surfaces. *Psychonomic Bulletin & Review*, *12*(2), 301–306. http://doi.org/10.3758/BF03196376

Moore, C. M., Yantis, S., & Vaughan, B. (1998). Object-based visual selection: Evidence from perceptual completion. *Psychological Science*, *9*(2), 104–110. http://doi.org/10.1111/1467-9280.00019

Moore, T., & Fallah, M. (2003). Microstimulation of the Frontal Eye Field and Its Effects on Covert Spatial Attention. *Journal of Neurophysiology*, *91*(1), 152–162. http://doi.org/10.1152/jn.00741.2002

Morey, C. C., Cong, Y., Zheng, Y., Price, M., & Morey, R. D. (2015). The color-sharing bonus: Roles of perceptual organization and attentive processes in visual working memory. *Archives of Scientific Psychology*, *3*(1), 18–29. http://doi.org/10.1037/arc0000014

Morey, R. (2015). BayesFactor: Software for Bayesian inference: Multiple Comparisons with BayesFactor, Part 1. Retrieved from http://bayesfactor.blogspot.co.uk/2015/01/multiple-comparisons-with-bayesfactor-1.html

Mounts, J. R., & Melara, R. D. (1999). Attentional selection of objects or features: Evidence from a modified search task. *Perception & Psychophysics*, *61*(2), 322–341.

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

Müller, N. G., Mollenhauer, M., Rösler, A., & Kleinschmidt, A. (2005). The attentional field has a Mexican hat distribution. *Vision Research*, *45*(9), 1129–1137. http://doi.org/10.1016/j.visres.2004.11.003

Mutluturk, A., & Boduroglu, A. (2014). Effects of spatial configurations on the resolution of spatial working memory. *Attention, Perception, & Psychophysics*. http://doi.org/10.3758/s13414-014-0713-4

Naber, M., Carlson, T. A., Verstraten, F. A. J., & Einhauser, W. (2011). Perceptual benefits of objecthood. *Journal of Vision*, *11*(4), 8–8. http://doi.org/10.1167/11.4.8

Nakayama, K., & Shimojo, S. (1992). Experiencing and perceiving visual surfaces. *Science*, *257*(5075), 1357–1363.

Navon, D. (1984). Resources—A theoretical soup stone? *Psychological Review*, *91*(2), 216.

Neumann, O. (1987). Beyond capacity: A functional view of attention. In H. Heuer & A. Sanders F. (Eds.), *Perspectives on Perception and Action* (pp. 361–394). Hillsdale, N.J.: Erlbaum Associates.

Neumann, O. (1996). Theories of attention. In *Handbook of Perception and Action* (Vol. 3, pp. 389–446).

Newston, D., & Engquist, G. (1976). The perceptual organization of ongoing behavior. *Journal of Experimental Social Psychology*, *12*(5), 436–450.

Nikolova, A., & Macken, B. (2015). The objects of visuospatial short-term memory: Perceptual organization and change detection. *The Quarterly Journal of Experimental Psychology*, 1–12. http://doi.org/10.1080/17470218.2015.1083595

Noorlander, C., Koenderink, J., J., Olden, R., J., & Edens, B. W. (1983). Sensitivity to spatiotemporal colour contrast in the peripheral visual field. *Vision Research*, *23*(1), 1–11.

Oberauer, K., & Eichenberger, S. (2013). Visual working memory declines when more features must be remembered for each object. *Memory & Cognition*, *41*(8), 1212–1227. http://doi.org/10.3758/s13421-013-0333-6

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. http://doi.org/10.1038/44134

O'Grady, R., & Müller, H. J. (2000). Object-based selection operates on a grouped array of locations. *Perception & Psychophysics*, *62*(8), 1655–1667.

Olson, I. R., & Marshuetz, C. (2005). Remembering 'what' brings along 'where' in visual working memory. *Perception & Psychophysics*, *67*(2), 185–194.

Orhan, A. E., & Jacobs, R. A. (2014a). Are performance limitations in visual short-term memory tasks due to capacity limitations or model mismatch. *Manuscript under Review*. Retrieved from http://hopf.cns.nyu.edu/~eorhan/papers/OrhanJacobs_plosone2014.pdf

Orhan, A. E., & Jacobs, R. A. (2014b). Toward ecologically realistic theories in visual short-term memory research. *Attention, Perception, & Psychophysics*, *76*(7), 2158–2170. http://doi.org/10.3758/s13414-014-0649-8

Orhan, A. E., Sims, C. R., Jacobs, R. A., & Knill, D. C. (2014). The Adaptive Nature of Visual Working Memory. *Current Directions in Psychological Science*, *23*(3), 164–170. http://doi.org/10.1177/0963721414529144

Oyama, T. (1961). Perceptual grouping as a function of proximity. *Perceptual and Motor Skills*, *13*(3), 305–306.

Palmer, S., & Rock, I. (1994). Rethinking perceptual organization: The role of uniform connectedness. *Psychonomic Bulletin & Review*, *1*(1), 29–55.

Patterson, M. D., Bly, B. M., Porcelli, A. J., & Rypma, B. (2007). Visual working memory for global, object, and part-based information. *Memory & Cognition*, *35*(4), 738–751.

Pelli, D. G., & Tillman, K. A. (2008). The uncrowded window of object recognition. *Nature Neuroscience*, 1129–1135. http://doi.org/10.1038/nn.2187

Pestilli, F., & Carrasco, M. (2005). Attention enhances contrast sensitivity at cued and impairs it at uncued locations. *Vision Research*, *45*(14), 1867–1875. http://doi.org/10.1016/j.visres.2005.01.019

Pomerantz, J. R. (2006). Colour as a Gestalt: Pop out with basic features and with conjunctions. *Visual Cognition*, *14*(4–8), 619–628. http://doi.org/10.1080/13506280500195052

Pomerantz, J. R., & Schwaitzberg, S. D. (1975). Grouping by proximity: Selective attention measures. *Perception & Psychophysics*, *18*(5), 355–361.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*(1), 3–25. http://doi.org/10.1080/00335558008248231

Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. In H. Bouma & D. Bouwhuis G. (Eds.), *Attention and performance X: Control of language processes* (X, Vol. 32, pp. 531–556). Hillside, New Jersey: Erlbaum. Retrieved from http://wknapp.com/class/cog/spring14/posner.pdf

Posner, M. I., Rafal, R. D., Choate, L., & Vaughan, J. (1985). Inhibition of return: Neural basis and function. *Cognitive Neuropsychology*, *2*(3), 211–228.

Postle, B. R. (2006). Working memory as an emergent property of the mind and brain. *Neuroscience*, *139*(1), 23–38. http://doi.org/10.1016/j.neuroscience.2005.06.005

Pratt, J., & Abrams, R. A. (1999). Inhibition of return in discrimination tasks. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(1), 229–242. http://doi.org/10.1037//0096-1523.25.1.229

Pratt, J., Adam, J., J., & McAuliffe, J. (1998). The spatial relationship between cues and targets mediates inhibition of return. *Canadian Journal of Experimental Psychology = Revue Canadienne de Psychologie Experimentale*, *52*(4), 213–216.

Pratt, J., & Fischer, M. (2002). Examining the role of the fixation cue in inhibition of return. *Canadian Journal of Experimental Psychology = Revue Canadienne de Psychologie Experimentale*, *56*(4), 294–301.

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. http://doi.org/10.3758/BF03196209

Prime, D. J., & Jolicoeur, P. (2010). Mental rotation requires visual short-term memory: Evidence from human electric cortical activity. *Journal of Cognitive Neuroscience*, *22*(11), 2437–2446. http://doi.org/10.1162/jocn.2009.21337

Prime, S. L., Vesia, M., & Crawford, J. D. (2011). Cortical mechanisms for trans-saccadic memory and integration of multiple object features. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1564), 540–553. http://doi.org/10.1098/rstb.2010.0184

Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism*. *Spatial Vision*, *3*(3), 179–197.

Quinlan, P. T., & Cohen, D. J. (2012). Grouping and binding in visual short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(5), 1432–1438. http://doi.org/10.1037/a0027866

Quinn, P. C., Bhatt, R. S., & Hayden, A. (2008). Young infants readily use proximity to organize visual pattern information. *Acta Psychologica*, *127*(2), 289–298. http://doi.org/10.1016/j.actpsy.2007.06.002

Reppa, I., & Leek, E. C. (2003). The modulation of inhibition of return by objectinternal structure: Implications for theories of object-based attentional selection. *Psychonomic Bulletin & Review*, *10*(2), 493–502.

Reppa, I., Schmidt, W. C., & Leek, E. C. (2012). Successes and failures in producing attentional object-based cueing effects. *Attention, Perception, & Psychophysics*, *74*(1), 43–69. http://doi.org/10.3758/s13414-011-0211-x

Rerko, L., Oberauer, K., & Lin, H.-Y. (2014). Spatial transposition gradients in visual working memory. *The Quarterly Journal of Experimental Psychology*, *67*(1), 3–15. http://doi.org/10.1080/17470218.2013.789543

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. http://doi.org/10.1037/0096-1523.34.4.842

Rouder, J. N., Morey, R. D., Morey, C. C., & Cowan, N. (2011). How to measure working memory capacity in the change detection paradigm. *Psychonomic Bulletin & Review*, *18*(2), 324–330. http://doi.org/10.3758/s13423-011-0055-3

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology*, *56*(5), 356–374. http://doi.org/10.1016/j.jmp.2012.08.001

Rousselet, G. A., Husk, J. S., Bennett, P. J., & Sekuler, A. B. (2005). Spatial scaling factors explain eccentricity effects on face ERPs. *Journal of Vision*, *5*(10), 1.

Rovamo, J., & Virsu, V. (1979). An estimation and application of the human cortical magnification factor. *Experimental Brain Research*, *37*(3), 495–510.

Ruderman, D. L. (1994). The statistics of natural images. *Network: Computation in Neural Systems*, *5*(4), 517–548.

Ruderman, D., L. (1997). Origins of scaling in natural images. *Vision Research*, *37*(23), 3385–3398.

Ruderman, D. L., & Bialek, W. (1994). Statistics of natural images: Scaling in the woods. *Physical Review Letters*, *73*(6), 814–817.

Saarela, T. P., Sayim, B., Westheimer, G., & Herzog, M. H. (2009). Global stimulus configuration modulates crowding. *Journal of Vision*, *9*(2), 5–5. http://doi.org/10.1167/9.2.5

Samuel, A. G., & Kat, D. (2003). Inhibition of return: A graphical meta-analysis of its time course and an empirical test of its temporal and spatial properties. *Psychonomic Bulletin & Review*, *10*(4), 897–906.

Schmidt, B. K., Vogel, E. K., Woodman, G. F., & Luck, S. J. (2002). Voluntary and automatic attentional control of visual working memory. *Perception & Psychophysics*, *64*(5), 754–763. http://doi.org/10.3758/BF03194742

Scholl, B. (2001). Objects and attention: the state of the art. *Cognition*, *80*, 1–46. http://doi.org/10.1016/S0010-0277(00)00152-9

Scholl, B. J. (2007). Object persistence in philosophy and psychology. *Mind & Language*, *22*(5), 563–591.

Scholl, B., & Pylyshyn, Z. W. (1999). Tracking Multiple Items Through Occlusion: Clues to Visual Objecthood,. *Cognitive Psychology*, *38*(2), 259–290. http://doi.org/10.1006/cogp.1998.0698

Sekuler, A. B., & Bennett, P. J. (2001). Generalized common fate: Grouping by common luminance changes. *Psychological Science*, *12*(6), 437–444.

Serences, J. T., & Boynton, G. M. (2007). Feature-Based Attentional Modulations in the Absence of Direct Visual Stimulation. *Neuron*, *55*(2), 301–312. http://doi.org/10.1016/j.neuron.2007.06.015

Sereno, A. B., & Kosslyn, S. M. (1991). Discrimination within and between hemifields: A new constraint on theories of attention. *Neuropsychologia*, *29*(7), 659–675.

Shariat Torbaghan, S., Yazdi, D., Mirpour, K., & Bisley, J. W. (2012). Inhibition of return in a visual foraging task in non-human subjects. *Vision Research*, *74*, 2–9. http://doi.org/10.1016/j.visres.2012.03.022

Shen, M., Tang, N., Wu, F., Shui, R., & Gao, Z. (2013). Robust object-based encoding in visual working memory. *Journal of Vision*, *13*(2), 1.

Shomstein, S. (2012). Object-based attention: strategy versus automaticity. *Wiley Interdisciplinary Reviews: Cognitive Science*, *3*(2), 163–169. http://doi.org/10.1002/wcs.1162

Shomstein, S., & Behrmann, M. (2008). Object-based attention: Strength of object representation and attentional guidance. *Perception & Psychophysics*, *70*(1), 132–144. http://doi.org/10.3758/PP.70.1.132

Shomstein, S., & Johnson, J. (2013). Shaping Attention With Reward: Effects of Reward on Space- and Object-Based Selection. *Psychological Science*, *24*(12), 2369–2378. http://doi.org/10.1177/0956797613490743

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

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

Shulman, G. L., Wilson, J., & Sheehy, J. B. (1985). Spatial determinants of the distribution of attention. *Perception & Psychophysics*, *37*(1), 59–65.

Shulman, G., Sheehy, J., & Wilson, J. (1986). Gradients of spatial attention. *Acta Psychologica*, *61*(2), 167–181. http://doi.org/10.1016/0001-6918(86)90029-6

Sigman, M., Cecchi, G. A., Gilbert, C. D., & Magnasco, M. O. (2001). On a common circle: natural scenes and Gestalt rules. *Proceedings of the National Academy of Sciences*, *98*(4), 1935–1940.

Simoncelli, E. P., & Olshausen, B. A. (2001). Natural image statistics and neural representation. *Annual Review of Neuroscience*, *24*(1), 1193–1216.

Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception-London*, *28*(9), 1059–1074.

Sperling, G., & Hsu, A. (2014). Deriving the acuity and the capacity of visual spatial attention. *Journal of Vision*, *14*(10), 521–521.

Stigchel, S. V. der, Belopolsky, A. V., Peters, J. C., Wijnen, J. G., Meeter, M., & Theeuwes, J. (2009). The limits of top-down control of visual attention. *Acta Psychologica*, *132*(3), 201–212. http://doi.org/10.1016/j.actpsy.2009.07.001

Strother, L., & Kubovy, M. (2006). On the Surprising Salience of Curvature in Grouping by Proximity. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(2), 226–234. http://doi.org/10.1037/0096-1523.32.2.226

Swallow, K. M., Zacks, J. M., & Abrams, R. A. (2009). Event boundaries in perception affect memory encoding and updating. *Journal of Experimental Psychology: General*, *138*(2), 236–257. http://doi.org/10.1037/a0015631

Taylor, J. E. T., Chan, D., Bennett, P. J., & Pratt, J. (2015). Attentional cartography: mapping the distribution of attention across time and space. *Attention, Perception, & Psychophysics*. http://doi.org/10.3758/s13414-015-0943-0

Terry, K. M., Valdes, L. A., & Neill, W. T. (1994). Does ‘inhibition of return’ occur in discrimination tasks? *Perception & Psychophysics*, *55*(3), 279–286.

Theeuwes, J. (2010). Top–down and bottom–up control of visual selection. *Acta Psychologica*, *135*(2), 77–99. http://doi.org/10.1016/j.actpsy.2010.02.006

Theeuwes, J., Belopolsky, A., & Olivers, C. N. L. (2009). Interactions between working memory, attention and eye movements. *Acta Psychologica*, *132*(2), 106–114. http://doi.org/10.1016/j.actpsy.2009.01.005

Theeuwes, J., Kramer, A. F., & Irwin, D. E. (2011). Attention on our mind: The role of spatial attention in visual working memory. *Acta Psychologica*, *137*(2), 248–251. http://doi.org/10.1016/j.actpsy.2010.06.011

Theeuwes, J., Mathôt, S., & Grainger, J. (2013). Exogenous object-centered attention. *Attention, Perception, & Psychophysics*, *75*(5), 812–818. http://doi.org/10.3758/s13414-013-0459-4

Theeuwes, J., Mathôt, S., & Kingstone, A. (2010). Object-based eye movements: The eyes prefer to stay within the same object. *Attention, Perception, & Psychophysics*, *72*(3), 597–601. http://doi.org/10.3758/APP.72.3.597

Tipper, S. P., Driver, J., & Weaver, B. (1991). Short report: Object-centred inhibition of return of visual attention. *The Quarterly Journal of Experimental Psychology Section A*, *43*(2), 289–298. http://doi.org/10.1080/14640749108400971

Tipper, S. P., Rafal, R., Reuter-Lorenz, P. A., Starrveldt, Y., Ro, T., Egly, R., … Weaver, B. (1997). Object-based facilitation and inhibition from visual orienting in the human split-brain. *Journal of Experimental Psychology: Human Perception and Performance*, *23*(5), 1522.

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

Treue, S., & Martinez-Trujillo, J. C. (2007). Attending to Features inside and outside the Spotlight of Attention. *Neuron*, *55*(2), 174–176. http://doi.org/10.1016/j.neuron.2007.07.005

Tsal, Y., & Lamy, D. (2000). Attending to an object's color entails attending to its location: Support for location-special views of visual attention. *Perception & Psychophysics*, *62*(5), 960–968.

Tsal, Y., & Lavie, N. (1988). Attending to color and shape: The special role of location in selective visual processing. *Perception & Psychophysics*, *44*(1), 15–21.

Turatto, M., Mazza, V., & Umiltà, C. (2005). Crossmodal object-based attention: Auditory objects affect visual processing. *Cognition*, *96*(2), B55–B64. http://doi.org/10.1016/j.cognition.2004.12.001

Vecera, S. P. (1994). Grouped locations and object-based attention: Comment on Egly, Driver, and Rafal (1994). *Journal of Experimental Psychology: General*, *123*(3), 316–320.

Vecera, S. P. (1997). Grouped arrays versus object-based representations: Reply to Kramer et al.(1997). Retrieved from http://psycnet.apa.org/journals/xge/126/1/14/

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

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions, and objects in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(1), 92–114.

Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., & von der Heydt, R. (2012). A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure–ground organization. *Psychological Bulletin*, *138*(6), 1172–1217. http://doi.org/10.1037/a0029333

Wagemans, J., Feldman, J., Gepshtein, S., Kimchi, R., Pomerantz, J. R., van der Helm, P. A., & van Leeuwen, C. (2012). A century of Gestalt psychology in visual perception: II. Conceptual and theoretical foundations. *Psychological Bulletin*, *138*(6), 1218–1252. http://doi.org/10.1037/a0029334

Wagenmakers, E., Wetzels, R., Borsboom, D., & van der Maas, H. L. J. (2011). Why psychologists must change the way they analyze their data: The case of psi: Comment on Bem (2011). *Journal of Personality and Social Psychology*, *100*(3), 426–432. http://doi.org/10.1037/a0022790

Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems of p values. *Psychonomic Bulletin & Review*, *14*(5), 779–804.

Wagenmakers, E.-J., Ratcliff, R., Gomez, P., & McKoon, G. (2008). A diffusion model account of criterion shifts in the lexical decision task. *Journal of Memory and Language*, *58*(1), 140–159. http://doi.org/10.1016/j.jml.2007.04.006

Walker, P., & Davies, S. J. (2003). Perceptual completion and object-based representations in short-term visual memory. *Memory & Cognition*, *31*(5), 746–760.

Wannig, A., Stanisor, L., & Roelfsema, P. R. (2011). Automatic spread of attentional response modulation along Gestalt criteria in primary visual cortex. *Nature Neuroscience*, *14*(10), 1243–1244. http://doi.org/10.1038/nn.2910

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

Wegener, D., Galashan, F. O., Aurich, M. K., & Kreiter, A. K. (2014). Attentional spreading to task-irrelevant object features: experimental support and a 3-step model of attention for object-based selection and feature-based processing modulation. *Frontiers in Human Neuroscience*, *8*. http://doi.org/10.3389/fnhum.2014.00414

Weger, U. W., Abrams, R. A., Law, M. B., & Pratt, J. (2008). Attending to objects: Endogenous cues can produce inhibition of return. *Visual Cognition*, *16*(5), 659–674. http://doi.org/10.1080/13506280701229247

Whitney, D., & Levi, D. M. (2011). Visual crowding: a fundamental limit on conscious perception and object recognition. *Trends in Cognitive Sciences*, *15*(4), 160–168. http://doi.org/10.1016/j.tics.2011.02.005

Woodman, G. F., Vecera, S. P., & Luck, S. J. (2003). Perceptual organization influences visual working memory. *Psychonomic Bulletin & Review*, *10*(1), 80–87. http://doi.org/10.3758/BF03196470

Woolgar, A., Williams, M. A., & Rich, A. N. (2015). Attention enhances multi-voxel representation of novel objects in frontal, parietal and visual cortices. *NeuroImage*. http://doi.org/10.1016/j.neuroimage.2014.12.083

Xu, Y. (2006). Understanding the object benefit in visual short-term memory: The roles of feature proximity and connectedness. *Perception & Psychophysics*, *68*(5), 815–828.

Yantis, S. (1992). Multielement visual tracking: Attention and perceptual organization. *Cognitive Psychology*, *24*, 295–340.

Yantis, S., & Jonides, J. (1984). Abrupt Visual Onsets and Selective Attention: Evidence From Visual Search. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(5), 601–621.

Yantis, S., & Jonides, J. (1990). Abrupt visual onsets and selective attention: voluntary versus automatic allocation. *Journal of Experimental Psychology: Human Perception and Performance*, *16*(1), 121.

Yuste, R. (2015). From the neuron doctrine to neural networks. *Nature Reviews Neuroscience*. http://doi.org/10.1038/nrn3962

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.

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, *453*(7192), 233–235. http://doi.org/10.1038/nature06860

Zhao, J., Kong, F., & Wang, Y. (2013). Attentional spreading in object-based attention: The roles of target–object integration and target presentation time. *Attention, Perception, & Psychophysics*, *75*(5), 876–887. http://doi.org/10.3758/s13414-013-0445-x

Zhao, L., Cosman, J. D., Vatterott, D. B., Gupta, P., & Vecera, S. P. (2014). Visual statistical learning can drive object-based attentional selection. *Attention, Perception, & Psychophysics*, *76*(8), 2240–2248. http://doi.org/10.3758/s13414-014-0708-1

Zhao, L., Gao, Q., Ye, Y., Zhou, J., Shui, R., & Shen, M. (2014). The Role of Spatial Configuration in Multiple Identity Tracking. *PloS One*, *9*(4), e93835.

Appendix 1: Piloting Stimuli Perceptual Organisation

The following experiment represents the original attempt to directly pit object-oriented versus space-oriented selection, and it was performed by using stimuli closely based on the cueing paradigm used by Brawn & Snowden (2000). In their study, the outlines of two overlapping triangles (one red and one green) formed a Star of David-shaped stimulus. Cueing was accomplished by brightening one of the triangles, after which the target event was a change in luminance in one of the circles situated at the corners of the two triangles. Object-based facilitation was observed when the change occurred at any of the three circles belonging to the cued triangle. The study provided evidence for object-based selection in the context of overlapping stimuli. This pattern was less pronounced when the triangle outlines were removed, resulting in a layout consisting of six circles (three green and three red) arranged on an imaginary circle centred at fixation.

For the current purposes the methodology of Brawn & Snowden (2000) was modified to study the spreading of processing facilitation following a non-informative spatial cue, which directly overlapped a potential target location. In addition, the current study aimed to compare the effects of perceptually organising the same set of features (potential target locations) into a single object, and into two overlapping objects where any two neighbouring features belong to one of two different objects. Therefore, there were two conditions of perceptual organisation – *one object*, where a single superordinate object (a circle) was formed by virtue of arranging six circles symmetrically around fixation, and *two objects*, where the same stimuli were grouped by means of connectedness and shading into two separate overlapping triangles (Figure 44).

These stimuli layouts allow testing the effects of perceptual organisation while keeping absolute spatial characteristics of the stimuli constant across conditions. Therefore, performance for target identification can be assessed for the same cue-target distances under different conditions of perceptual organisation. If selection is guided by space, then the same pattern of performance is expected for both *one object* and *two objects*, since the spatial layout of the stimuli is identical. This pattern should reflect a spatial gradient of facilitation centred at the cued feature, i.e. a gradual increase in reaction times and decrease in accuracy as cue-target distance is incremented. If there is a combination between space-based and object-based selection, it is expected that a spatial gradient will be observed for the *one object* condition, and possibly no difference between any uncued features in the *two objects* condition, due to competing mechanisms. Finally, if selection is purely object-based, then no performance variation is expected for the *one object* condition, and a non-monotonic variation is predicted for *two objects*. In other words, the performance function should only be affected by whether the target is within the cued object or not, regardless of cue-target distance. However, under all hypotheses it is expected that the cued feature may be prioritised, as it is directly activated by the cue (i.e. a cueing effect should be evident).

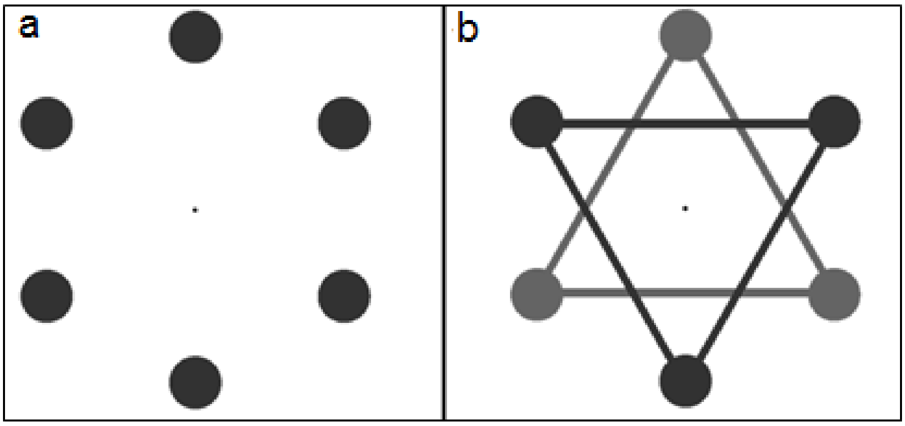

Figure 44: Appendix 1 perceptual organisation of the stimuli into objects; a: *one object* (dark grey); b: *two objects* (dark grey on top).

Method

Participants

Twenty-five undergraduate students (3 male, mean age of 19.32, *SD* = 1.52) from Cardiff University took part in the experiment in return of partial course credit. Participants had normal or corrected-to-normal vision.

Stimuli and Apparatus

The experiment was conducted using a Windows XP operating system on a 23-inch monitor with 1920 x 1080 pixel resolution, 32-bit colour quality, and a 60 Hertz refresh rate. A standard keyboard was used to record input. The task was programmed using Matlab R2012a and run with Psychophysics Toolbox 3 (Kleiner et al., 2007).

The size of the stimuli is reported in degrees of visual angle, unless stated otherwise. These sizes are calculated on the basis of approximately 70 centimetres viewing distance. All stimuli were presented on a monochromatic grey background (RGB: 200, 200, 200).

In both *one* the *two object*s conditions the target stimuli were six equally spaced filled circles, centred at 5° from fixation. Each target circle was with a diameter of 1.5°. The Euclidian distance between the centres of two neighbouring circles was 5°. The exact angular coordinates of the stimuli corresponded to 30°, 90°, 150°, 210°, 270° and 330° with reference to fixation. In the *one object* condition these circles were not connected, forming a circular shape by virtue of groping by similarity and symmetry (Kubovy & Van den Berg, 2008; Wagemans et al., 2012; Dodd & Pratt, 2005) (Figure 44a). For the *two objects* condition, non-adjacent circles were connected with 10-pixel think lines to form two overlapping equilateral triangles (objects), much like the Star of David shape used in Brawn & Snowden (2000) (Figure 44b). The perception of two separate objects was strengthened by colouring each triangle in a different shade of grey (light shade RGB = 100, 100, 100; dark shade RGB = 50, 50, 50). Which object appeared on top varied randomly from trial to trial. In the *one object* condition all circles were of the same colour, which also changed randomly to light or dark grey for each trial. The cue consisted of a circle outline with the same colour as the background, appearing at the centre of one of the grey circles and subtending half of its size (Figure 45). The target event was a ±50% change in the RGB values of one of the grey circles.

Design and Procedure

A 2 (object condition: *one object* or *two objects*) x 4 (cue-target distance) repeated measures design was employed. The two levels of object condition represented variations in the perceptual organisation of the stimuli (Figure 44), while cue-target distance (hereafter referred to by *distance*) had four levels defined with reference to the location of the cue (Figure 45). Distance 1 corresponded to targets appearing at the cued location, i.e. at 0° distance from the cue. Distance 2 corresponded to targets appearing at locations of 60° distance on either side of the cued location. Consequently, distance 3 was associated with targets at 120°, and distance 4 with targets appearing at 180° (directly opposite) to the cued location. For the *two objects* condition, targets at distance 2 and distance 4 appeared on the uncued triangle, and are therefore also referred to as uncued-different object targets. Distance 3 on the other hand is associated with uncued-same object targets, as it is perceptually grouped with the cued location. For the *one object* condition all distances corresponded to targets on the same object as the cue.

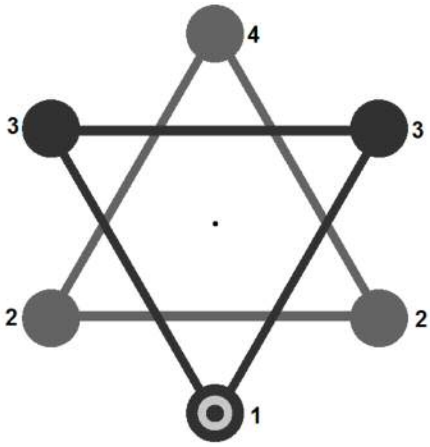

Figure 45: Appendix 1 cue-target distances (*two objects* example). Numbers illustrate the corresponding distances. The cue is depicted at distance 1.

The cue was purely exogenous, i.e. it was not predictive of the target location. For each perceptual organisation, each combination of cue location, target location, stimuli colour (dark or light grey for the *one object* condition), object position (top or bottom for the *two objects* condition), and target change polarity (lighter or darker) appeared an equal number of times. This resulted in 288 trials per object condition. Within these, the target appeared 48 times on each of the six possible locations.

The study was approved by the Cardiff University Ethics Committee. The experiment was organised in two blocks of 288 trials with a self-timed break in-between. Each block contained only one type of object condition, with all other factors varying at random. The order of blocks was counterbalanced between participants.

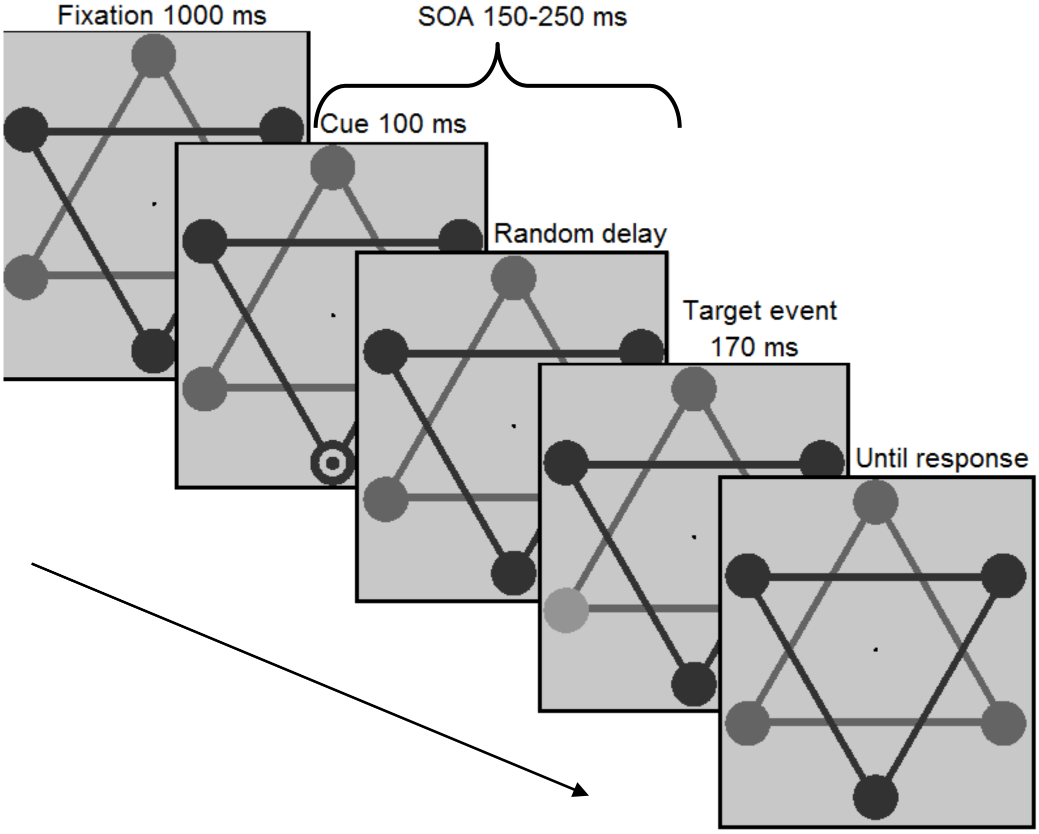

Figure 46: Appendix 1 procedure illustration (*two objects*). In this example the correct response is 'L' for 'lighter', and the target appears at distance 2 (uncued-different object location).

Participants were tested individually in semi-enclosed booths. Each session initiated with 10 randomly selected practice trials, 5 of each object condition. Each trial began with a 1000 ms passive exposure to the relevant stimulus, followed by presentation of the cue for 100 ms (Figure 46). The stimulus onset asynchrony (SOA) was between 150 ms and 250 ms, and varied at random for each trial in order to avoid anticipatory responses. The duration of the target event was 170 ms, after which the relevant feature came back to its original colouring. Participants were instructed to maintain fixation throughout the duration of the trial and to respond as quickly and as accurately as possible what was the polarity of the change. Responses were made with the index fingers placed on the buttons ‘L’ and ‘D’ on the keyboard, corresponding to ‘lighter’ and ‘darker’, respectively. The position of the response buttons was counterbalanced between participants using adhesive labels. In addition, 0.74° by 1.06° upper case ‘L’ and ‘D’ letters were displayed on the horizontal axis 8.45° lateral of fixation on either side of the stimulus. The position of the letters corresponded to the index finger they were mapped to, and they functioned as a reminder to participants.

Accuracy feedback (‘Correct’ or ‘Incorrect’) was displayed at fixation in the 1000 ms inter trial interval of the practice session. The subsequent experimental session did not provide feedback, but the word ‘Ready!’ was displayed instead. These words subtended approximately 1.93°. Participants were aware the cue was not correlated with the location of the target, and were fully debriefed after the study. The whole procedure lasted approximately 30 minutes.

Results and Discussion

For both accuracy and reaction time, the data for distance 2 and distance 3 were obtained by averaging the performance for the two target locations situated at ±60°, and at ±120° angular distance from the cued location, respectively. Prior to averaging, paired samples t-tests were performed to ensure no statistical difference was present. Participants whose performance was not suitable for averaging, and participants scoring under 50% on accuracy were excluded from the analysis (N = 7). As a result, the final analysis was conducted on a sample of 18 participants. Separate 2 (object condition) x 4 (cue-target distance) repeated measures Analyses of Variance (ANOVA) were conducted on accuracy (proportion correct response) and reaction time (milliseconds). Whenever the assumption of sphericity was violated, the Greenhouse-Geisser corrected values are reported. All simple comparisons were subjected to a Bonferroni correction.

Accuracy

The proportion of correct responses was overall higher in the *one object* condition, $F(1,17) = 28.93$, $p < .001$, $\eta_p^2 = .63$, while there was no main effect of distance, $F(1.86, 31.65) = 2.91$, $p = .121$, $\eta_p^2 = .12$. However, there was an interaction between object condition and distance, $F(3, 51) = 6.53$, $p = .001$, $\eta_p^2 = .28$. This effect was followed-up by separate four-way repeated measures ANOVAs for each object condition. As Figure 47 illustrates, for *one object* there was no effect of distance whatsoever ($F < 1$), while the effect of distance in the *two objects* condition, $F(2, 34.1) = 6.76$, $p = .003$, $\eta_p^2 = .28$, indicated higher accuracy for distance 3 compared to distance 1 ($p = .022$) and distance 2 ($p < .001$). There were no other statistical differences. Therefore, contrary to a prediction based on space-oriented selection, accuracy did not appear to follow a spatial gradient centred at the cued location. In fact, in both object conditions there was no advantage for the cued location (distance 1). Interestingly, for the *two objects* condition performance was superior when the target did not match the cued location, but was nevertheless part of the cued object. The results indicate a level of object-based facilitation, but also a counterintuitive inhibition for targets at the cued location.

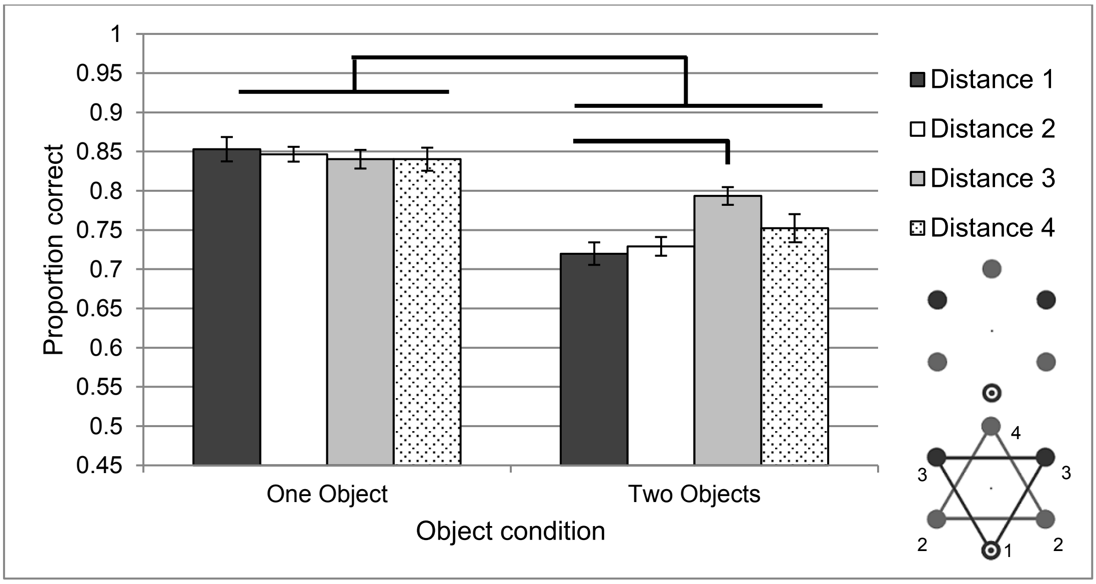

Figure 47: Appendix 1 accuracy (proportion correct) as a function of object condition and cue-target distance. Error bars represent SEM (corrected for between-subject variability). Brackets illustrate statistical difference at $p < .05$. Distances are also indicated by numbers next to the *two objects* stimulus depiction for ease of interpretation.

Reaction time

Only reaction times for correct responses were analysed. In addition, trials which generated responses faster than 200 ms or slower than 1000 ms were excluded from the analysis, resulting in the removal of 15.2% of the total data. As with accuracy, the 2 x 4 ANOVA indicated that performance was overall 37 ms faster for the *one object* condition, $F(1, 17) = 6.37$, $p = .022$, $\eta_p^2 = .27$, and again there was no main effect of distance ($F < 1$). There was an interaction between object condition and distance, $F(3, 51) = 3.58$, $p = .02$, $\eta_p^2 = .17$, but when it was followed up with individual analyses, distance had no effect for *one object* ($F < 1$), and also marginally failed to affect reaction times for *two objects,* $F(3, 51) = 2.77$, $p = .051$, $\eta_p^2 = .14$.

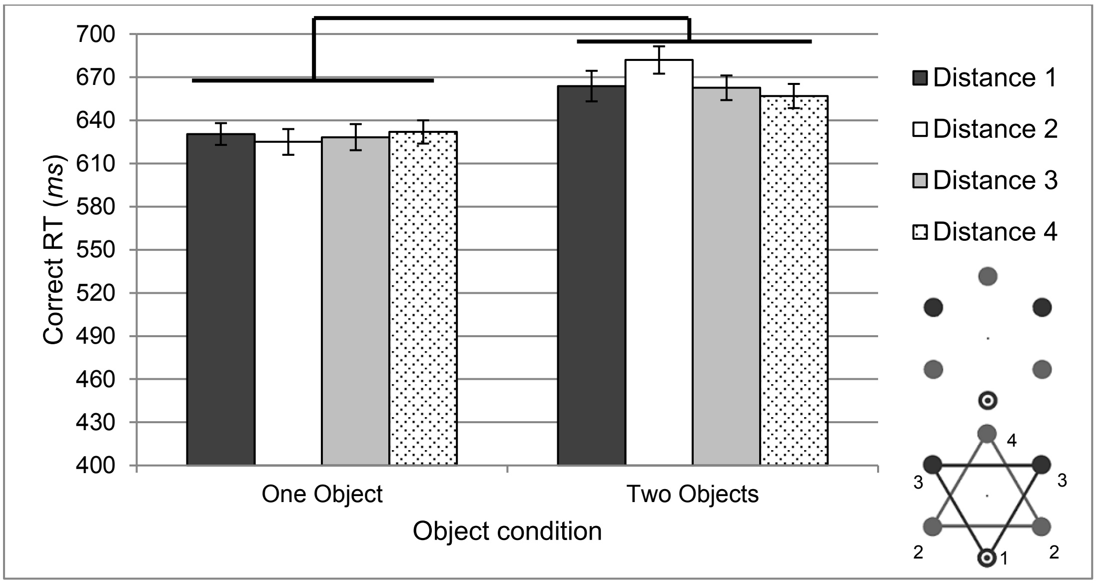

Figure 48: Appendix 1 correct mean reaction time (milliseconds) as a function of object condition and cue-target distance. Error bars represent SEM (corrected for between-subject variability). Brackets illustrate statistical difference at $p < .05$.

However, there was a tendency in the *two objects* condition for distance 2 to be the slowest (Figure 48). Distance 2 being slower than distance 3 suggests a possibility for object-based selection, since this difference goes in the opposite direction to a gradient pattern, and is based on the object structure of the stimuli (features at distance 2 belong to the uncued object). At the same time, no difference between distance 1 and distance 3 also conforms to an object-based pattern of performance, as features at these distances belong to the same (cued) object. However, the fact that reaction times for distance 4 did not display a tendency to differ from any other goes against both space-based and object-based selection models.

In any case, there were no statistical differences based on cue-target distance for either *one object*, or *two objects*. The observed interaction was most likely due to differences for the same distance across object conditions. However, interpreting such patterns is not meaningful in the current context. The reason is that, considering the current purposes, it is more important to investigate the effect *within* each object condition in order to observe how the specific perceptual organisation affects which parts of the scene are prioritised for selection. In addition, given the main effect of object condition for both accuracy and reaction time, it appears that the two conditions differed in terms of difficulty levels, such that the task was easier in the *one object* condition.

It should be noted that the lack of pronounced distance effects may be due to ineffectiveness of the cue. Exogenous cues are known to have a powerful attraction effect, and thus involuntarily lead to a processing facilitation even when not predicting the future target (e.g. Theeuwes, Mathôt, & Grainger, 2013; Yantis & Jonides, 1984). Therefore, the lack of a cueing effect for the current experiment is unusual. At present, the cue is visually similar to the target event, which may have led to some form of perceptual masking of the target when it coincided with its location. This may explain the consistent lack of facilitation for cued targets at distance 1. However, the flat performance function for the *one object* condition may be interpreted as object-based selection where all features were equally selected, so there is no priority for one over another, given that the cue was not informative of the target location (He et al., 2004). On the other hand, the lack of facilitation for distance 1 for *two objects* accuracy combined with improved performance for distance 3 suggests there may be a problem with processing the cued location. In any case, the reaction time data cannot be readily accommodated within this possibility, as there was no statistically pronounced variation for either object condition, although the trend may be interpreted as object-oriented facilitation.

In sum, the results of the experiment are inconclusive about the nature of visual selection, but it appears that the properties of the cue may be masking the expression of any potential effects. In order to investigate this possibility, an additional small-scale study was conducted (N = 8). The stimuli and design were identical to the experiment described above, but the cue was changed to a small black bar of 0.82° length, situated at 0.38° distance from the respective cued circle. This bar was oriented along the axis passing through the centre of the cued feature and the centre of the screen (i.e. fixation). All other factors of the experimental procedure remained the same. The results, albeit not demonstrating any statistical effects, indicated a weak trend towards reaction time facilitation for the cued location without any other variation for uncued features. This suggests the line cue may be more compatible with the current stimuli, due to being perceptually distinct and not spatially overlapping the cued feature. Therefore, this cue was adopted for the remaining experiments based on this paradigm. In addition, a decision was made that the cue should provide a level of predictability for the target location, in order to ensure it is consciously processed and can be used as a reference point for measuring the effect of cue-target distance.

Another important point concerning the current results relates to the perceptual organisation of the stimuli. The differentiation between the two triangles in the *two objects* condition may be difficult to make. The two shades of grey may not be salient enough to override a highly familiar symbol, such as a Star of David, leading to a possibility of perceiving the stimuli as a single object. It was deemed appropriate to use more salient distinction cues for the future versions of the task. The modifications of the perceptual organisation involve rotation of the circle stimuli with 30° in order to discourage the formation of a single Star of David shape for the *two objects* condition. In addition, the use of distinctive colouring for each object may be more appropriate, e.g. red and green as in the original Brawn & Snowden (2000) version, and ensuring that the object conditions are visually more similar. The latter involves introducing connecting lines between the circles in order to attempt equating perceptual load between the two conditions. All of these modifications were adopted for the main follow-up experiments in the empirical chapters, aiming to study the nature of visual selection.

Appendix 2: Issues of Target-object Integration

This set of two experiments represents the original attempt to introduce targets that are poorly integrated with the object(s) on the display, in order to test the hypothesis that a spatial gradient in performance is an emergent property of the perceived structure of the observed stimuli, and it is thus nothing more than an object-based phenomenon. For the purpose, the same object stimuli as those used in Experiments 2.1-2.3 were adopted, but the nature of the target was changed into superimposed letters “X” or “O”. The same type of targets was used by Hollingworth et al. (2012), who demonstrated a spatial gradient within the same object. The aim here was to keep all other aspects of the experimental procedure the same as the experiments in Chapter 2, which demonstrated space-invariant object-based selection, and vary only the level of target-object integration.

It was hypothesised that by decreasing the probability that the transient targets are integral parts of the object, selection will favour targets proximal to the cued feature because they have a higher probability of being perceptually integrated with it. Experiment I tested this possibility by simply changing the type of targets and preserving all other aspects the same as Experiment 2.2 scale of 5º condition. Experiment II was an extension of Experiment I by introducing additional six target locations situated in-between the object features, resulting in a total of seven cue-target distances. This was an attempt to further break down the perception that the targets are integrated within the circular object features.

Experiment I Method

Participants

Twenty-one participants (2 male, mean age = 21.5, *SD* = 0.32) took part in the *one object* condition, and also twenty-one participated in the *two objects* condition (1 male, mean age = 23.3, *SD* = 0.67) in return of partial course credit. All participants had normal or corrected-to-normal vision.

Stimuli and Apparatus

The physical characteristics of the background and the object stimuli were identical to those in Experiment 2.2, scale of 5º eccentricity. The critical change was the nature of the target. Instead of introducing a luminance change in one of the object circular features, the target was a superimposed capital letter “X” or “O”, which measured 0.5° x 0.5° with a 4-pixel stroke and dark grey monochromatic colouring (RGB = 40). The targets were centred inside the object features (Figure 49). The same equipment was used as for all experiments in Chapter 2.

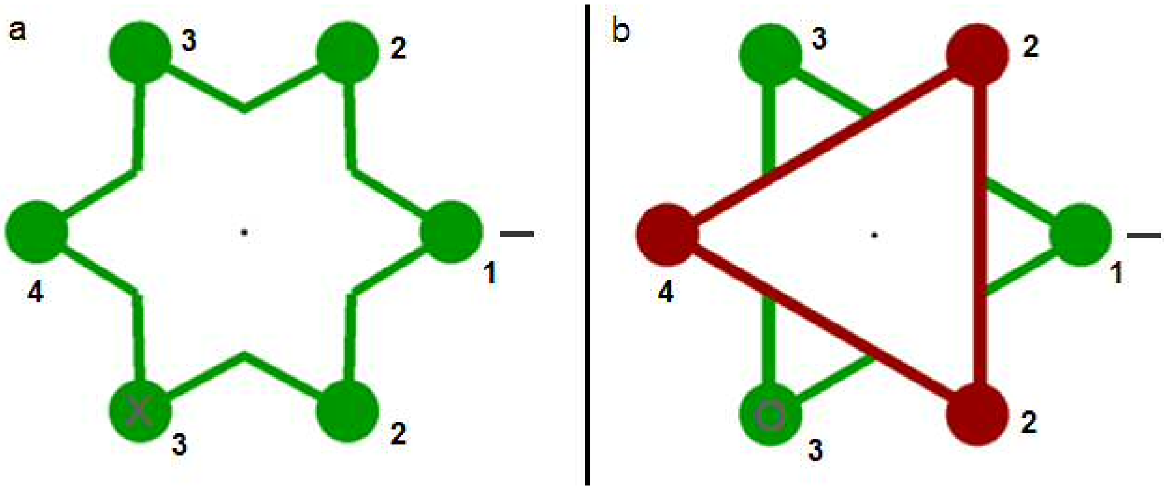

Figure 49: Appendix 2 stimuli illustration for Experiment I. Numbers illustrate the four cue-target distances relative to the location of the cue (black line); a: *one object* with target “X” at distance 3; b: *two objects* with target “O” at distance 3. During the progression of the trial, the cue and target appear in succession and are never visible at the same time.

Design and Procedure

The only variable of interest was cue-target distance with four levels, so this was a four-way repeated measures design. There were a total of 432 trials, 132 of which contained a cued feature, and the remaining trials were equally spread between each of the five uncued features. Therefore, the cue was predicting the target location on 30.5% of the trials. Also, the response button reminders located on either side of the stimuli on the screen appeared at the bottom of the display, instead of being on the same level as fixation. This was a precaution against potential response interference due to incompatible target and reminder identity, since both the targets and reminder labels in this case were letters.

The procedure was similar to Experiment 2.2 and Experiment 2.3. The duration of the stimuli was the same, except that the target was presented for 80 ms instead of 100 ms. Whether the target was “X” or “O” varied randomly from trial to trial, and there was an equal number of each type of target. The procedure was divided into three blocks of 144 trials each. The cue-target distance varied in a semi-random fashion from trial to trial, observing the restriction about cue predictability (i.e. higher target frequency at the cued feature). The procedure took approximately 30 minutes and participants were fully debriefed afterwards.

Experiment I Results and Discussion

Reaction times for correct responses were analysed with a four-way repeated measures Analysis of Variance (ANOVA), separately for each object condition. Responses faster than 200 ms and slower than 1000 ms were excluded from the analysis (< 1% of the data). For the *one object* condition there was only a cuing effect, $F(1.41, 28.27) = 20.36$, $p < .001$, $\eta_p^2 = .50$, where responses to targets at distance 1 were the fastest (all $p_s < .001$) and there were no other statistical differences. The results were identical for *two objects,* $F(3, 60) = 34.2$, $p < .001$, $\eta_p^2 = .63$, where reaction times for distance 1 were also the fastest (all $p_s < .001$) (Figure 50).

Although statistically only a cueing effect was present, the non-significant trend for the one object condition indicates a graded pattern of selection. There was a 13.3 ms difference (*SE* = 5.17) in reaction time between responses for distance 2 and distance 3, which is in contrast with the typical flat performance function observed in Experiments 2.1-2.3. However, the 95% CIs for all pairwise differences (other than comparisons with distance 1) include 0, which is why there is a lack of statistical effect. However, in the context of the studies conducted so far, this trend is meaningful in the sense that it suggests the typical object-based effect observed so far can be reverted into a spatial gradient by decreasing the target-object integration probability.

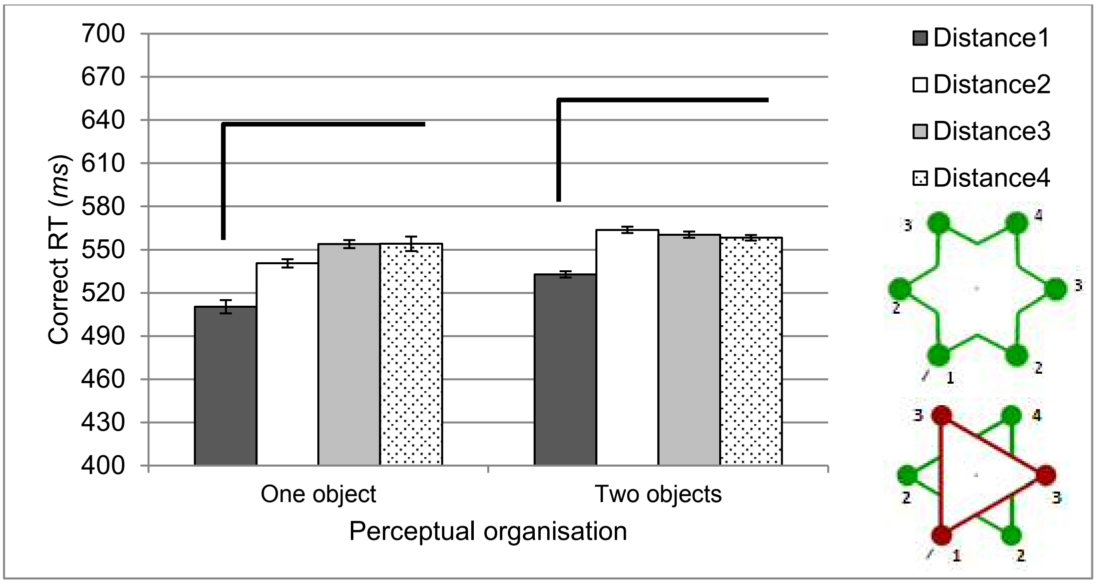

Figure 50: Appendix 2 mean correct reaction time (*ms*) as a function of cue-target distance and perceptual organisation for Experiment I. Brackets illustrate statistical differences at *p* < .05, and error bars represent SEM (corrected for between-subject variability).

Performance in the two objects condition did not show an effect or even a trend towards either spatially graded or non-monotonic pattern of selection. The lack of response variation for uncued targets may in fact be explained by resolving to the perceptual organisation of the stimuli. Since it is expected that a gradient-like selection would result in responses for distance 2 being faster than responses for distance 3, while object-based selection produces the opposite effect, a combination of the two is likely to result in a flat performance, as is currently the case. However, it should be noted that this combination (if it is taking place) is not to be interpreted as the co-existence of space-based and object-based selection, but can be accommodated, as all other findings, within a pure object-based account. That is, the probability of the targets being integrated with the objects is less than in the main experiments of Chapter 2 (Experiments 2.1-2.3), but it is nevertheless not equal to 0, since the targets do overlap the object features. This uncertainly may lead to privileged processing of targets at distance 2 (as two points close together are likely to belong together), but equally so for targets at distance 3, because they appear on the cued object. There is less of a conflict in the one object condition since all targets are superimposed on a single object. Therefore, the uncertainty due to ambiguous target-object integration results in a tendency towards processing on the basis of proximity, but some perceptual integration with the object is still possible as the gradient differences were not large enough to reach statistical significance.

The results from Experiment I suggest a tendency towards performance in the shape of a spatial gradient, i.e. gradual reduction in the quality of processing with increasing cue-target distance. Given the results from Experiments 2.1-2.3, which so far suggest robust object-based selection with the same stimuli, it may be concluded that the observed trend in the current experiment was due to the poor integration between the target and the underlying object. The fact that the gradient performance was not very pronounced, i.e. only evidenced in the form of a non-significant trend, may be because the targets were still reasonably perceived as being part of the objects. This is possible since all stimuli were two-dimensional and the targets were symmetrically centred and well placed within the circular object features. One way of validating this possibility is to make the targets appear less integrated, which should result in a stronger gradient. This was the purpose of the follow-up study, Experiment II, where six additional target locations were introduced in-between the single object features. This manipulation was expected to reduce the perception that the superimposed targets are part of the object, since they can also appear outside its body.

Experiment II Method

Participants

Twenty-one undergraduate psychology students from Cardiff University took part in the experiment for partial course credit. One participant was later excluded due to consistently scoring under 50% on accuracy. The analyses were therefore performed on a sample of 20 participants (all female, mean age = 22.6, *SD* = 0.7) with normal or corrected-to-normal vision.

Stimuli and Apparatus

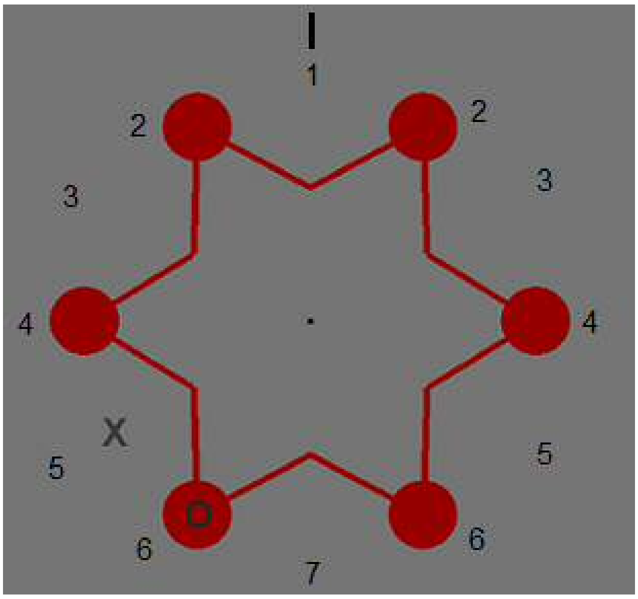

Figure 51: Appendix 2 stimuli used for Experiment II. Numbers illustrate the cue-target distances. The cue (black line) indicates distance 1. Two targets are illustrated: "O" at distance 6, and "X" at distance 5. This is for illustration purposes only, during trial presentation only one target is displayed at a time, following cue offset.

This experiment employed only the object stimulus from the *one object* condition. All aspects of the appearance of the object were the same as in Experiment I. The physical characteristics of the targets were also identical to the previous experiment. The critical difference was that six additional potential locations for the target and cue were added, situated in the space between each neighbouring pair of object features. The additional cue and target coordinates were generated by a 30° angular displacement of the original six coordinates, so all twelve target locations had the same eccentricity (5° from fixation) (Figure 51). Due to the difference in luminance between the background and the colour of the object, the colouring of the target was perceptually adjusted to appear equal when it happened to be on top of the object (RGB = 40), or in-between the features (RGB = 60). As before, for half of the trials the object was coloured red and for the other half it was green, and the colour varied randomly from trial to trial. All other aspects and equipment were the same as in Experiment I.

Design and Procedure

Adding the additional location resulted in a total of 7 cue-target distances (Figure 51). In the cases when an object feature was cued (i.e. distance 1 was on the object feature), this results in 4 distances within the object, and 3 distances outside the object. The opposite was true for when the space between two object features was cued. The total number of trials was the same as before, 432, but the proportion of cued-uncued targets was altered due to doubling the potential target locations. The target appeared at distance 1 on 102 of the trials, and 30 times for each of the remaining 11 locations, which averaged to 60 trials for distances 2-6 and 30 trials for distance 7 (because data for distances 2-6 was obtained by averaging between the two equidistant locations on either side of the cue). Therefore, the cue was predicting the target location on 23.6% of the trials, as opposed to the 30.5% in Experiment I. However, it still had an informative value, as it was the most likely target location and participants were made aware of this contingency.

The procedure was identical to Experiment I, except that the experiment was divided into four blocks of 108 trials in order to provide more opportunity for breaks.

Experiment II Results and Discussion

Overall linear analysis

The hypothesis pertaining to the current study was that a spatial gradient of performance will be evident as a result of the poor target-object integration. However, performing a 7-way repeated measures ANOVA revealed no effect of distance. Therefore, there was no evidence of a spatial gradient, and no evidence of a cueing effect either. However, there was a significant linear contrast trend in the data, $F(1, 19) = 7.65$, $p = .003$, $\eta_p^2 = .29$, suggesting that overall the performance fits best with a monotonic change function.

Exploratory analyses based on target and cue location categorisation

An additional 2 (target location: inside or outside the object) x 7 (cue-target distance) ANOVAs were conducted to test if a gradient is likely to appear only when the targets are outside the body of the object, which would be reflected as an interaction between the two factors. There was a statistical effect only for target location, $F(1, 19) = 59.89$, $p < .001$, $\eta_p^2 = .76$, demonstrating that when the target appeared inside the object features, performance was overall 19.6 ms slower than when targets were situated outside the object ($p < .001$). However, regardless of whether the targets were internal or external to the object, there was no difference in reaction time based on cue-target distance, i.e. no gradient in performance (Figure 52). It may be the case that more effort is needed when targets are within the object. Alternatively, this could be a confounding contrast effect, since targets on the background may be more readily identifiable.

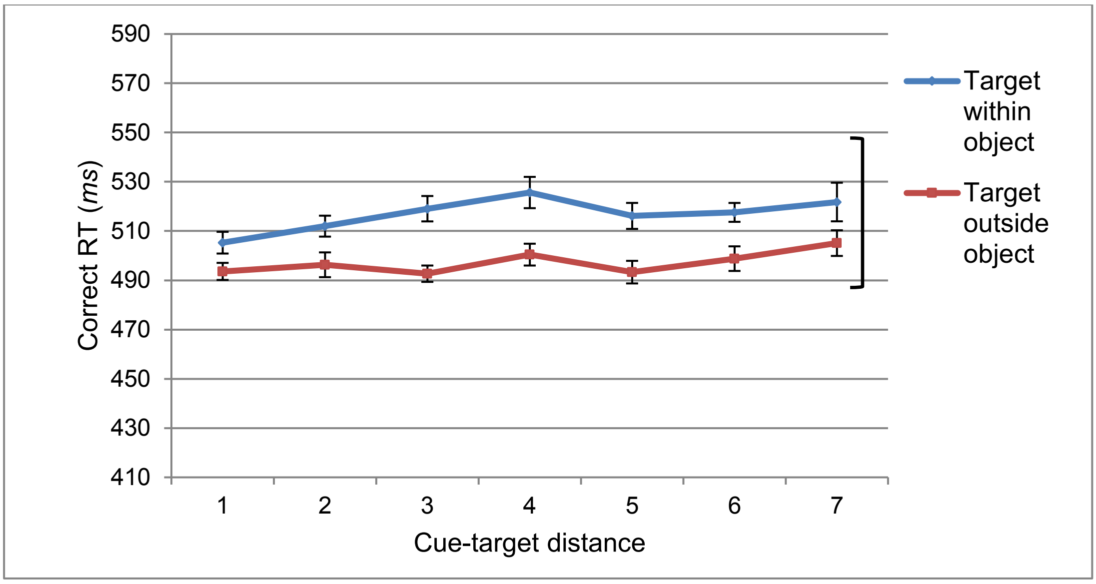

Figure 52: Appendix 2 mean correct reaction time as a function of cue-target distance and target location with reference to the object for Experiment II. Error bars represent corrected SEM. The bracket illustrates statistical difference at *p* < .05.

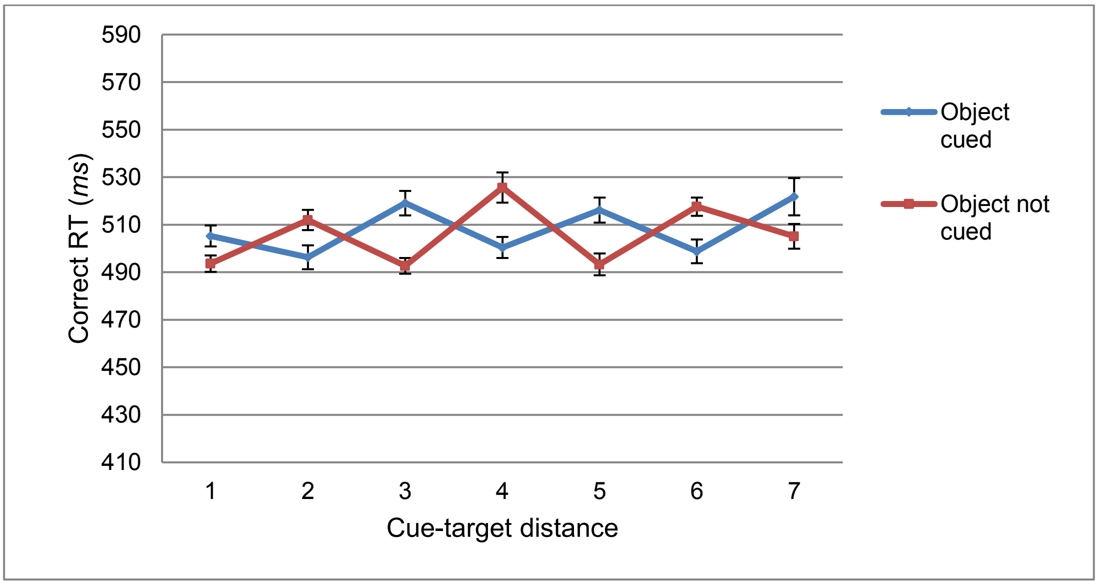

Figure 53: Appendix 2 mean correct reaction time as a function of cue-target distance and cue location with reference to the object for Experiment II. Error bars represent corrected SEM. Differences were observed between odd and even distances only when the object was not cued, i.e. when the space between object features was cued (red line, refer to the text for details).

Finally, the effect of cue-target distance on reaction time was analysed based on whether the cue indicated an object feature, or the space between two object features, using a 2 (cued feature, cued space in-between) x 7 (cue-target distance) ANOVA. There was no effect of distance and no variation based on where the cue was situated, but there was an interaction between these two variables, $F(6, 114) = 7.59$, $p < .001$, $\eta_p^2 = .29$. Figure 53 illustrates the data as a function of what was indicated by the cue, and target distance. Therefore, this is essentially the same data graph as in Figure 52, but regrouped for ease of illustration of the current analysis.

Follow-up tests involved separate 7-way analyses for cases when the object was cued, and when the space between features was cued. There was no effect of distance when the cue indicated an object feature, but there was an effect when the cue was between features, $F(6, 114) = 7.26$, $p < .001$, $\eta_p^2 = .28$. Pairwise comparisons indicated that although reaction times for distance 7 did not differ from any others, the differences for the remaining distances revealed a clear odd versus even distance differentiation. Specifically, distance 1 was faster than distance 2 ($p = .009$), distance 4 ($p = .013$), and distance 6 ($p = .006$). Also, distance 3 was faster than distance 4 ($p = .012$), and distance 6 ($p = .004$); distance 4 was faster than distance 5 ($p = .045$), and finally distance 5 was faster than distance 6 ($p = .038$). In simpler terms, when the space between object features was cued, all targets at locations external to the object generated faster reaction times than targets at locations internal to the object (refer to Figure 51 as a reference).

This result is expected given that the earlier analysis revealed that targets within the object are processed more slowly, and when the space between features is cued all even distances happen to be situated within the object. Therefore, these even-numbered distances are associated with slower reaction times. However, what is informative from this analysis is that this internal-external effect is only pronounced when the object is *not* cued. When the cue predicts one of the object features, then the difference between distances inside and outside the object is not as large, although as Figure 53 illustrates, it is still pronounced as a strong trend in the same direction. It may be possible that when one of the object features is cued, the whole object gets activated, i.e. prioritised for selection. In turn, this results in boosted performance for the targets overlapping the object and therefore less of a difference between internal and external targets. Although this suggests some form of object-based selection, it does not explain the lack of gradient for targets outside the object, and also the lack of cueing effect which until now was a consistent occurrence.

It should be noted that these analyses were largely explorative, as they were performed post-hoc and not planned in advance (other than the initial overall linear trend analysis). Therefore, results should be interpreted with caution and only as a possible indication of the underlying effects. In any case, no strong effect was observed other than faster processing of external targets. There is no clear interpretation at this stage, other than suggesting that there was an object-based facilitation when any of the object features was cued, resulting in enhanced perceptual processing for targets overlapping the object surface. This enhancement may override potential visual contrast issues, which otherwise lead to faster responses for targets appearing on the grey background compared to targets within the object. In other words, the internal-external target difference is modulated by the location of the cue. Finally, the lack of a spatial gradient may also be due to the decrease in the predicting power of the cue and the doubling of possible target locations. If the cue is not treated as informative, it is possible that performance would not be strongly graded in relation to its location.

Experiments I and II: General Discussion

Experiments I and II aimed to test if a spatial gradient in performance would be observed with the same stimuli as in Experiments 2.1-2.3 (which elicited robust object-based selection in the face of timing, scaling, and feature salience manipulations), given that the target represented a superimposed letter stimulus. Given that Experiment I suggested a non-significant trend towards graded performance and a possibility that this is due to the targets not being readily segregated from the objects, Experiment II aimed to increase the perception of poor target-object integration by introducing additional target locations outside the object. Nevertheless, no evidence of spatially graded performance was found.

However, since the total amount of trials in Experiment II was not increased, this resulted in reducing the predictive power of the cue, which in turn may be responsible for the lack of a cueing effect (so no pronounced advantage for targets at distance 1). In turn, it is not certain whether the cue was processed to a level necessary to use it as a reference point for measuring distance effects on performance. It may be the case that participants did not have a top-down strategy of prioritising the cued location, and instead had more or less equal expectation for each possible target location. Reaction time was nevertheless overall slower when targets were inside the object features, which may simply be due to a colour contrast issue. Finally, considering the fact that in Experiment I, *two objects* condition, performance also showed no distance variation trend, within the context of the current findings, it may be the case that the type of target and object need to be made more separable from one another.

Given the current results, an additional investigation of the role of target-object integration should preserve the ratio of cued-uncued targets, and also keep the potential target locations the same as in the experiments which demonstrated object based effects (Experiments 2.1-2.3). However, the perceptual integration between the target and the object needs to be decreased further, for example using an apparent three-dimensional object with a two-dimensional target, as in Hollingworth et al. (2012), who successfully demonstrated a within-object gradient.

Appendix 3: Visuo-Spatial Short Term Memory and Strength of Perceptual Organisation

This study represents the original attempt to test the effects of object-level perceptual organisation on visuo-spatial short term memory (VSTM). The principle aim is identical to that of Experiment 3.1, namely to control for the number and spatial distribution of the memory items, while varying the perceptual organisation of the stimuli in order to test for same-object advantage following a non-informative spatial cue in a colour change detection task.

The manipulation of the perceptual organisation relates to the tendency of the display to encourage the formation of higher-level perceptual objects, which in turn are expected to influence VSTM. The prediction is that when performance for two equidistant memory probes from the cue is compared, change detection would be superior when the probe is situated on the same object as the cue compared to a different perceptual object. Importantly, it is hypothesised that when the equidistant probes appear to be located within the same object, no difference in performance should be observed, and this would hold true even if the integrity of the object is based on an emergent perception, formed by cues of occlusion and illusory contours. Such a pattern of results would demonstrate the important role of perceptual object formation in VSTM, and also the obligatory object-oriented encoding of information.

In order to test the hypotheses outlined above, the condition of object formation was varied across four levels of perceptual organisation for the objects within which the memory items were contained. All other local factors of the memory stimuli remained unchanged between conditions. The four perceptual organisation levels represented an *intact*, *occluded*, *illusory*, and *segmented* object (Figure 54 a-d, respectively). The *intact* condition was formed by an integral circular object appearing on top of three cone shapes. The same circular object with the three cones appearing on top and thus occluding proportions of its surface was used to form the *occluded* object. For the *illusory* object, the circle appeared as modally completed by erasing parts of its contours where the cones meet its outline. Finally, segmenting the intact circle into three arcs with the cone shapes appearing in-between resulted in the *segmented* object.

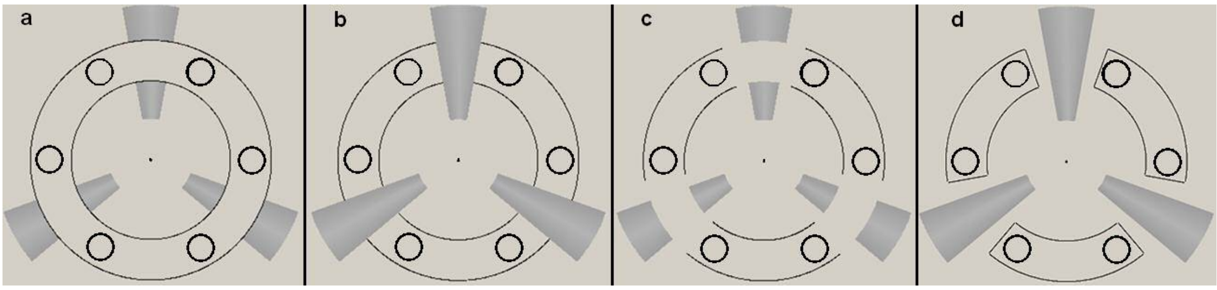

Figure 54: Appendix 3 perceptual organisation illustration; a: *intact* object; b: *occluded* object; c: *illusory* object; d: *segmented* object. The circle outlines mark the locations and size of the memory items. These outlines are for illustration purposes only and were not present during the experiment.

The second independent variable of interest was probe type. After one of six possible locations on the object(s) was cued by a brief flash, six to-be-remembered coloured circles were presented during the study phase with fixed locations within the perceptual objects (illustrated by empty placeholders on Figure 54). Following a retention interval, the six circles were presented again and one of them was indicated as a probe that required a *same/different* colour judgement. Each item was probed with equal probability, but there were three critical probe locations of interest. These probes were defined relative to where the cue was presented prior to the study phase, and relative to the dividing cones. These were the *cued* location, and the two immediately adjacent equidistant locations on either side, located either within the *same arc* as the cued location (relative to the dividing cones), or within a *different arc*. Depending on the perceptual organisation condition, *different arc* probes would either be within the same object (integral or perceptually completed), or within two different objects for the *segmented* condition. Since the memory items always occupied the same absolute locations, variations in performance are unlikely to be due to any spatial factors, but rather as a result of the perceptual organisation of the display.

Based on results from visual selection experiments (Moore et al., 1998; Pratt & Sekuler, 2001), it is expected that the *intact* and perceptually completed objects (*occluded* and *illusory*) would produce functionally equivalent results, i.e. no difference between equidistant (relative to the cue) memory probes. On the other hand, for the *segmented* object it is expected that probes within the *same arc* as the cue would elicit better change detection performance than probes in the *different arc* (i.e. an object-based effect evidenced by same-object advantage). It should be noted that for the purpose of clarity and consistency, the memory probes are labelled on the basis of the arc they belong to relative to the cued location, and an arc is defined as the region between two adjacent cones. Therefore, for the *intact* and perceptually completed conditions *same arc* and *different arc* probes are always within the same object, but for the *segmented* condition they fall within two separate objects. Finally, performance across all conditions is likely to be superior when the *cued* item is probed at test, since the spatial cue is expected to produce an involuntary capture of attention, leading to prioritisation of the subsequently occurring item there (Schmidt et al., 2002). In sum, it is expected that the manipulations of probe type and object formation will interact to reveal an object-based effect which is of equal magnitude for perceptually completed objects.

Method

Participants

Twenty-three undergraduate psychology students (2 males, mean age of 20.22 years, *SD* = 2.31) took part in the experiment for partial course credit. The sample had normal or corrected-to-normal colour vision, and was recruited using Cardiff University Experiment Management System (EMS).

Stimuli and Apparatus

The experiment was conducted using a Windows XP operating system on a 17-inch monitor with 1280 x 1024 pixel resolution and 32-bit colour quality with a 60

Hertz refresh rate. A standard keyboard was used to record input, and Visual Basic 6.0 was used to program and run the task. The size of the stimuli is reported in degrees of visual angle, unless stated otherwise. These sizes are calculated on the basis of 70 centimetres viewing distance, but since no chin rest was used, the measures are approximate.

All stimuli were presented on a grey background (RGB: 212, 201, 200). The colours of the to-be-remembered items were chosen randomly without replacement from the following set, with the corresponding RGB coordinates presented in parentheses: brown (205, 133, 63), red (255, 0, 0), yellow (255, 255, 0), green (0, 255, 0), blue (0, 0, 255), cyan (0, 255, 255) and white (255, 255, 255). The cue was coloured in ‘blanched almond’ white (255, 235, 205).

Each target circle was with a diameter of 1.0°. The distance from the central fixation point and the centre of each circle was 5°. The six to-be-remembered items were equally spaced, with an angular spacing of 60°, relative to the central fixation point. The size of the cue was 0.52°, approximately half of the size of the memory stimulus. The stimuli were centred within a ring of 1.95° width outlined in black. In the case of the *intact* condition, the ring was fully visible and without any discontinuity, while its integrity varied for the perceptually completed and *segmented* conditions (refer to Figure 54). For the *segmented* condition the ring was broken down into three arcs by deleting its outlines at three 40° (angular distance) wide sections between each pair of memory items. These segments were then contained by drawing a line at each end. The three gaps between segments were separated by cone-shaped items of 5.33° length, centred on the same imaginary circle as the to-be-remembered items. Consequently, the cones were placed in the middle of the gaps, and were situated at 20° angular distance from the centre of the memory items on either side. These stimuli were coloured in different shades of grey, ranging from RGB: 160, 160, 160 on the edges to RGB: 180, 180, 180 at the centre. The RGB coordinates increased in units of two from edge to centre on each side. The purpose of this shading technique was to induce a basic perception of depth.

For the *intact* condition the ring appeared on top of the cones, while for the *occluded* condition the cones were drawn on top of the ring. Finally, for the *illusory* condition the outlines of the ring were deleted in a similar manner as for the *segmented* context, but the section of the cones passing on top of the ring was not visible and the ring segments were not closed off with a contour on the side. The colour inside the ring was identical to the background, a condition necessary to encourage modal completion. A black circular frame of 0.21° thickness surrounded the probe stimulus. The whole display (memory targets, ring and cone shapes) subtended a total of 14.69° x 14.69° centred at fixation.

Design and Procedure

The experiment conformed to a 4 (object formation: *intact*, *occluded*, *illusory*, and *segmented* object) x 3 (probe type: *cued*, *same arc*, and *different arc* probe) repeated measures design.

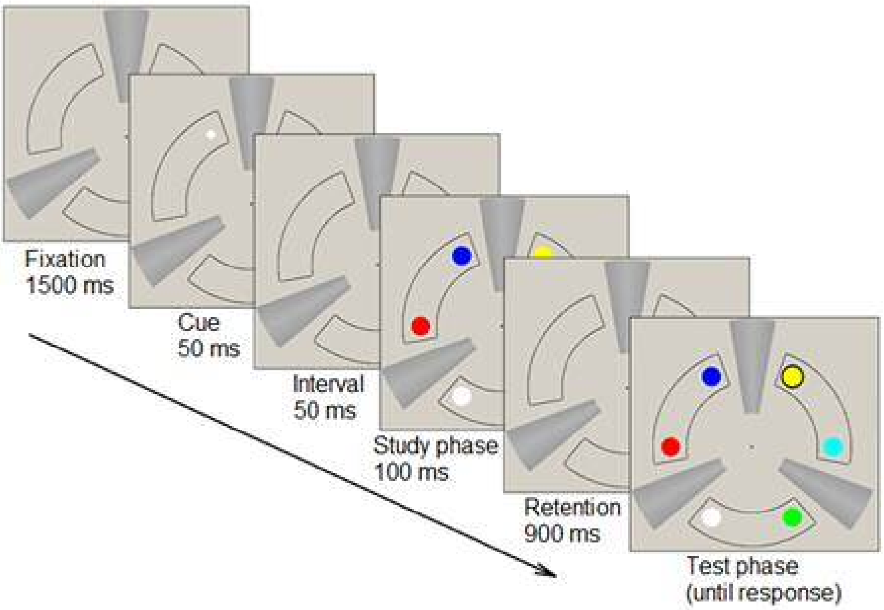

Figure 55: Appendix 3 procedure illustration (*segmented* condition). The last panel depicts a *different arc* probe requiring a ‘same’ response.

Participants were tested one at a time in semi-closed booths. Each participant underwent a brief practice session with 20 randomised trials, 5 from each type of perceptual organisation. The procedure is illustrated in Figure 55. For each trial, the object formation context alone was initially presented for 1500 ms, and then the cue appeared on one of the six memory items locations for 50 ms. Participants were instructed that the cue was not informative of the location of the probe and should be ignored. Following 50 ms after cue offset was the study phase, where the six coloured circles were presented for 100 ms. After a 900 ms retention interval where only the perceptual object(s) were visible, the six memory circles were displayed again. One of them (the probe) was surrounded by a black otuline, prompting participants to make a decision whether its colour is ‘same’ or ‘different’ relative to the study display. Any of the six items could be probed with equal probability, but of crucial importance were the three probe types: *cued*, *same arc* and *different arc* probes (the latter is illustrated on the last panel of Figure 55).

There were 108 trials for each of the four object formation conditions. Within each set of 108, the cue and probe locations followed a quasi-random pattern. Each of the six locations was cued 18 times, and for those 18 times, each memory item was probed 3 times. Therefore, by the end of the 108 trials an average of 18 responses was calculated for each of the 6 possible cue-probe relationships. Half of these required a ‘same’ response. Also, on half of the 108 trials, the location of the dividing cones (and gaps between segments) was rotated by 40° to make sure all possible pairings of targets were used. These variations appeared at random. Participants used the ‘<’ button for ‘same’ and the ‘>’ button for ‘different’ responses. These buttons were labelled ‘S’ and ‘D’, respectively.

Following response, there was a 1480 ms gap before the next trial. This inter-trial interval was filled with a dynamic masking stimulus, in order to minimise afterimage of the recent visual context due to prolonged fixation. The mask consisted of three images, which were repeatedly alternated every 40 ms during this period. Two of the images consisted of randomly generated black and white pixels, creating the perception of dynamic visual noise. The third image consisted of the same visual context as the one in the immediately preceding trial, but with inverted colours resulting in a negative image. It was found from a set of pilot trials that exposure to this type of stimuli in the inter-trial interval was successful in cancelling out the otherwise strong afterimage, which could interfere with the perception of the visual context in the following trial.

Throughout the experiment, all of the 432 trials were presented in a fixed random order. Breaks were introduced after each 108 trials, forming 4 blocks. However, the trials within these blocks were not grouped according to any principle, and participants did not know what type of trial would follow next. Accuracy (*d'*) and reaction time for correct responses (in milliseconds) were recorded. The whole procedure lasted about 45 minutes. Participants were treated in accordance with the British Psychological Society code of ethics, and an ethical approval was obtained from the School of Psychology, Cardiff University prior to testing.

Results and Discussion

A separate 3 x 4 repeated measures ANOVA was conducted for accuracy and reaction times on data from the three locations of interest. Whenever the assumption of sphericity was violated, Greenhouse–Geisser correction is reported. Bonferroni corrections were applied to all follow-up pairwise comparisons of main effects. Change detection accuracy was measured by transforming the proportion of hits (i.e. when a changed probe was correctly identified as *different*) and false alarms (when the probe colour was unchanged, but the response was *different*) into *z* scores to calculate a measure of sensitivity *d'* (Macmillan & Creelman, 2004). For this purpose, the false alarm rate is subtracted from the hit rate. A 3000 ms upper boundary was adopted for trimming reaction time data, whereby responses exceeding this time window were to be discarded as lapses in attention. However, no exclusions were performed on the basis of these criteria.

Accuracy

Manipulating object formation did not have an effect on change detection accuracy, $F(3, 66) = 1.46$, $p = .234$, $\eta_p^2 = .06$, while accuracy did vary as a function of probe type, $F(1.57, 34.62) = 31.98$, $p < .001$, $\eta_p^2 = .59$. Unlike the outlined prediction, there was no interaction between object formation and probe type, *F*(6, 132) = 1.66, *p* = .136, η_p^2 = .07 (Figure 56). Follow up analyses for the main effect of probe type revealed a cueing effect, such that that accuracy for *cued* probes was higher than accuracy for *same arc* and *different arc* probes (all p_s < .001). However, there was no statistical difference between the latter two probe types (*p* = .301), suggesting the lack of a same-object advantage for either type of perceptual organisation.

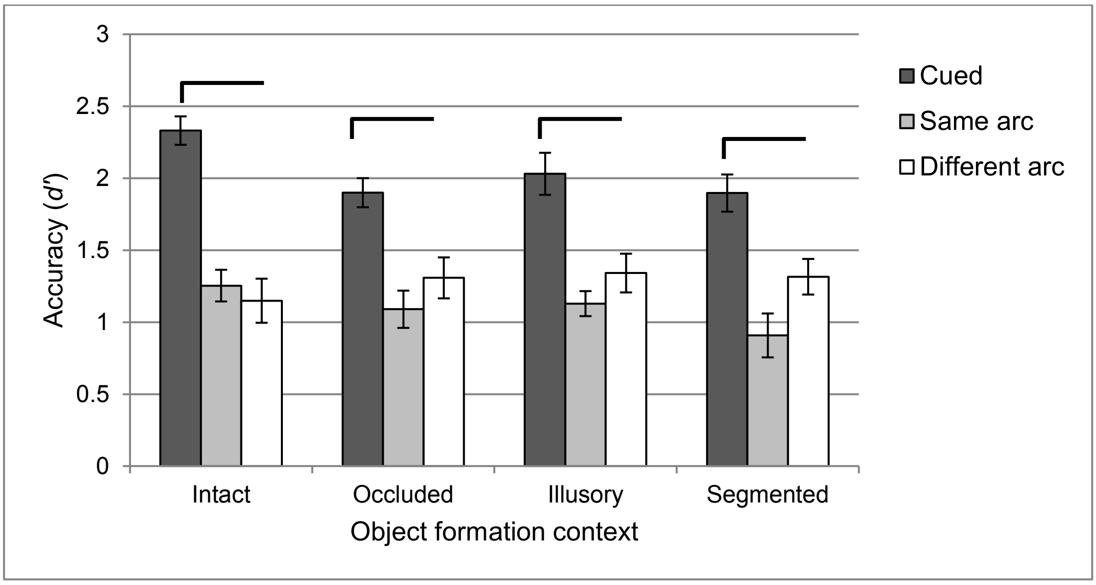

Figure 56: Appendix 3 accuracy (*d'*) as a function of object formation and probe type. Brackets illustrate the statistical differences at *p* < .05 for the main comparisons of probe type. Error bars indicate SEM, corrected for between-subject variability.

Reaction time

As with accuracy, reaction time was not affected by object formation (*F* < 1), but only by probe type, *F*(1.37, 30.18) = 10.54, *p* = .001, η_p^2 = .32 (Figure 57). However, in this case there was a marginal interaction between the two independent variables, *F*(3.53, 77.76) = 2.61, *p* = .048, η_p^2 = .11. The interaction was followed up with separate three-way repeated measures planned comparisons at each of the four levels of object formation. There was a statistical difference between the three critical probe types only for the *occluded* and *illusory* conditions. For the *illusory* condition, pairwise comparisons revealed that responses for *cued* probes were faster than responses for the *same arc* (*p* =

.001) and *different arc* probes (*p* = .008), while the latter two did not differ, thus replicating the established cueing effect. For the *occluded* context, however, responses to *cued* probes were faster than responses to *same arc* probes (*p* = .007), but responses to *same arc* probes were slower than responses to *different arc* probes (*p* = .001). In other words, reaction time did not consistently conform to a cueing effect, and there was no evidence of same-object advantage either. It may be the case that with a *p*-value of .048, following up with simple effects and comparisons is not informative, as there is a high probability that any statistical effects at this level are due to Type I error, i.e. chance fluctuations. In any case, the main effect of probe type indicated a clear cueing effect, where responses to *cued* probes were overall faster than responses to *same arc* (*p* = .005) and *different arc* (*p* = .008) probes, without a difference between the latter two (*p* > .99).

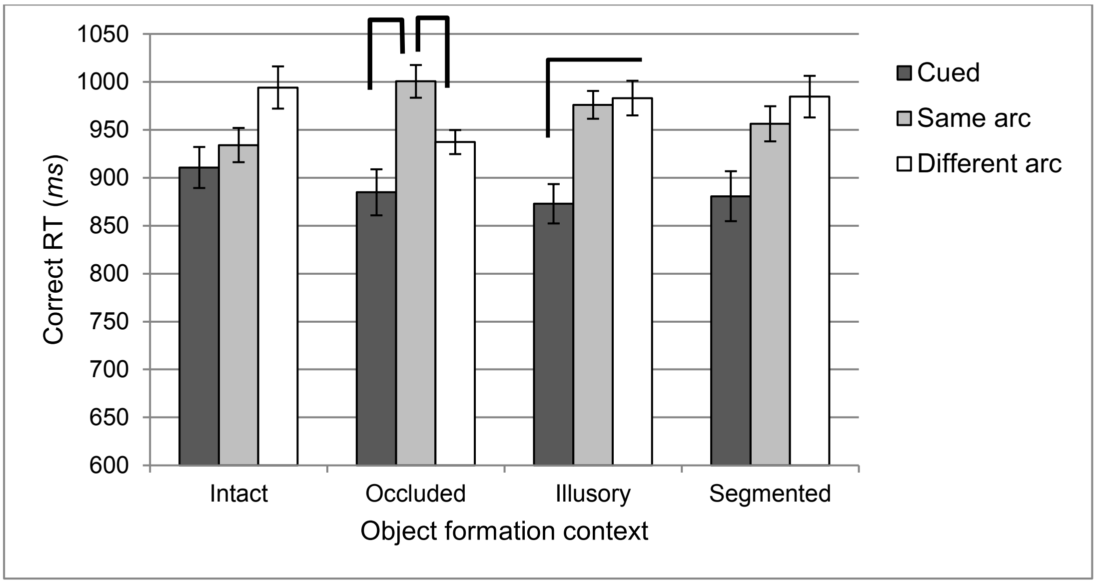

Figure 57: Appendix 3 mean correct reaction time (*ms*) as a function of object formation and probe type. Brackets illustrate the statistical differences between probe types. Error bars indicate SEM, corrected for between-subject variability.

Focusing on accuracy and reaction time together, it may be the case that there was a speed-accuracy tradeoff in some isntances. Inspection of Figure 56 indicates that although there were no statistical effects for the differences between *same arc* and *different arc* probes for any of the object formation conditions, there is a trend towards the opposite direction. In other words, better accuracy for *different arc* compared to *same arc* probes. Looking at reaction time, there is a weak trend towards faster responses for *same arc* probes, hence a possibility of a tradeoff. Exception is the *occluded* context, where performance for *same arc* probes was both poorer and slower compared to *different arc* probes. It should be noted that these trends are to be interpreted with caution, as these are effectivelly null results and may simply be random meaningless fluctuations in the data. In any case, it appears that the task was of high difficulty, particularly when responding to uncued probes, as accuracy is relatively low throughout.

A tentative suggestion at this stage is that the perceptual organisation cues were not salient enough to produce object-based advantages, as the only cues to objecthood used in this study were black outlines on a uniform grey background. In support of this possibility, using a variation of the two rectangle paradigm, Albrecht, List, & Robertson (2008) demonstrated that an object-based advantage for stimulus detection can be completely eliminated if the two rectangles are made to appear as two holes (slits) in an object. When the rectangles are perceived as slits, the stimuli appearing within them are perceived as occupying the same uniform surface underneath the object. Therefore, the object-based effect observed when the slits appear as separate rectangles is eliminated, because in this case the stimuli are always within the same (background) surface, which is partially visible through the slits. A similar process may be occurring with the current stimuli, as the area within the ring is of the same colour as the background, so it may be ambiguous whether they appear on the outlined object, or on the uniform grey surface underneath.

The strength of the object representation can indeed be of crucial importance. Colouring in the outlines of two dimensional objects can contribute to strengthening object-based effects even in the case of very short pre-cue exposure to the visual context (200 ms), which is not typically considered sufficient to induce within-object advantage (Shomstein & Behrmann, 2008). Emphasising cues to objecthood can be critical for the subjective perception of the visual scene and its parsing into separate items (Chen & Cave, 2006). Therefore, introducing additional cues, such as shading the area within the ring in the current experiment, can be beneficial in strengthening the intended perception of the object stimuli.

In addition to these potential limitations, it may also be the case that an alternative perceptual grouping is occurring. The consistently counterintuitive (albeit non-significant) pattern for accuracy is visible for all object formation conditions apart from the *intact* one. This pattern reflects a tendency for poorer performance for *same arc* compared to *different arc* probes. In the *occluded* and *segmented* contexts the three cones were fully visible, and in the case of the *illusory* condition, they were potentially perceptually completed. As a result, there may be a perception of three symmetrical shapes formed by a cone and two circles on each side, which may potentially lead to a grouping of the *cued* item with the item situated on the other side of the cone (i.e. the *different arc* item). Perhaps linking the three cones into a single fan-shaped object can help strengthen its segregation from the principal ring object. Although this alternative perceptual organisation may be unlikely, linking the cones into a single shape may also be beneficial for making the overall display appear less cluttered and more organised (i.e. there would be fewer objects), which may lead to overall improved performance.

There is another element of the current design that may be problematic for detecting a same object advantage, if it exists. The fact that the perceptual organisation changes from trial to trial may disturb the formation of a constant perception. There is conflicting evidence as to whether previous experience affects object-based selection. For example, amodal completion may be affected by whether the object has been seen in its intact form prior to occlusion (Joseph & Nakayama, 1999; Zemel et al., 2002), but this result is not always replicated (Lin & Yeh, 2012; Jay Pratt & Sekuler, 2001). However, it is generally accepted that visual selection can be dynamic and strategy-driven (Shomstein & Johnson, 2013; Theeuwes, 2010), based on statistical regularities extracted from the structure of a set of trials (Lee, Mozer, Kramer, & Vecera, 2012; Sarah Shomstein, 2012). Therefore, it is possible that the unstable visual context resulting from changing the perceptual organisation on trial-to-trial basis may lead to a stronger focus on task-relevant information only. This may at least partially contribute to the lack of object formation effects on the basis of the different perceptual organisation, although it may also be suggested that not changing the context from trial to trial may lead to habituation, so it is not clear how this element may influence object formation.

In summary, the current experiment replicated the well established cueing effect for VSTM (e.g. Schmidt et al., 2002), but failed to demonstrate any specific effects of object formation. Given the limitations of the current stimuli, it may be the case that the lack of effects was due to poor perceptual organisation cues and high difficulty of the task. Therefore, a follow-up experiment should adopt more salient stimuli that provide easier segregation into objects. This can be achieved with simple shading and reduction of the number of total objects visible on the scene, while makg no changes to the memory items per se.

Appendix 4: Inhibition Without Disengagement

The main purpose of this experiment was to test for object-based IoR under conditions very similar to those inducing the object-based facilitation effects observed in Chapter 2, simply by extending the SOA and keeping all configural aspects the same. Specifically, the aim was to manipulate cue-target distance in a way which is not confounded with probability of object belongingness, i.e. create conditions where these two variables are not correlated. In turn, it can be tested if selection of features within the cued object can be inhibited relative to the uncued object, regardless of cue-target distance. That is, slower reaction times are expected for identifying changes at the cued feature and the cued object (i.e. uncued features belonging to the cued object) compared to reaction times for responding to changes of features belonging to the different (uncued) object.

Studies investigating IoR often use a centrally presented transient stimulus during the SOA interval, in order to disengage covert focus from the cued item/ location before target appearance (Klein, 2000). This method aims to cause re-orienting towards the cued location if the subsequent target is to be presented there. The procedure may involve a brief flickering of the fixation point or presentation another item (e.g. a square), and it is always task-irrelevant. There is conflicting evidence regarding whether this interim attractor stimulus is necessary to obtain IoR, as the phenomenon has also been observed without it, i.e. only by extending SOA (Bourke et al., 2006; Gabay et al., 2012; Lupiáñez et al., 1997). The attractor is considered necessary primarily if inhibition is to be induced at early SOAs where facilitation is typically observed, e.g. around 200 ms (J Pratt & Fischer, 2002). Therefore, the current experiment aimed to test if object-based IoR can be achieved without the use of an interim attractor stimulus, especially since the adopted SOAs were longer than 200 ms. Another important factor for detecting the phenomenon is to use a non-informative cue, since knowing that the cue indicates the most likely target location can lead to prolonging the duration of the facilitation effect for cued targets (Klein, 2000). Thus, another key difference from the methodology in Chapter 2 was that the cue in the current experiment was not correlated with target location.

The current experiment involved the manipulation of two independent variables: target distance with four levels reflecting a gradually increasing spatial separation from the cued object feature (as in Chapter 2), and SOA with three levels: 300, 600, and 900 ms. As before, there were two versions of the task with separate participant samples, one of which involved the four cue-target distances to correspond to features of a single star-shaped object (*one object* condition), and another condition where the same spatial coordinates corresponded to two overlapping objects (*two objects* condition). The latter allows the direct comparison between space-based and object-based effects. The physical characteristics of all stimuli were the same as in Chapter 2 (coloured versions of the objects), but in this case the cue was not correlated with the location of the subsequent target. The earliest SOA was chosen to be 300 ms because this corresponds to the main SOA used for inducing facilitation effects in Chapter 2. This allows testing if the object-based effects can be replicated with a non-informative cue, as well as testing if the robust cueing effect observed throughout Experiments 2.1-2.4 would be replicated when there is no top-down incentive to select the cued feature. The longer SOAs of 600 ms and 900 ms were chosen to be of equal intervals within the typical range used in IoR experiments (e.g. 300-1500 ms).

It was hypothesised that there will be no difference in reaction time when responding to targets within the same object as the cued feature, resulting in a flat function for the *one object* condition, and a non-monotonic pattern of response for the *two objects* condition. The prediction was that responses for SOAs longer than 300 ms would be slower for the cued object compared to the uncued object, i.e. distance 1 and distance 3 would elicit slower reaction times than distance 2 and distance 4. In other words, the inverse effect of that observed in Chapter 2 was expected due to object-based IoR. However, another key difference in the prediction here is that a cueing effect was not expected, i.e. no superiority (for 300 ms SOA) or inferiority (for later SOAs) for distance 1, since the cue did not carry any information about the target location. As SOA of 300 ms was still expected to elicit facilitation effects as opposed to inhibition, it was predicted that there should be a trend towards overall faster reaction times for the shortest SOA relative to SOAs of 600 ms and 900 ms.

Method

Participants

Eight participants took part in the *one object* condition (1 male, mean age = 23.23, *SD* = 2.1), and a different sample of eight participants took part in the *two objects* condition (1 male, mean age = 23.75, *SD* = 8.5). Participants received £4 payment, and had normal or corrected-to-normal vision. It should be noted that this was aimed as a pilot experiment in order to identify if there would be a trend towards inhibition with the standard stimuli parameters, or whether additional measures, such as a central attractor stimulus, should be adopted. Therefore, a relatively small sample was used.

Stimuli and Apparatus

The equipment and the physical characteristics of the stimuli were identical to the ones described in Chapter 2, Experiment 2.1.

Design and Procedure

As in Chapter 2, the six possible target locations corresponded to four cue-target distances after averaging data between equidistant features. SOA was also manipulated, so the experiment conformed to a 3 (SOA: 300 ms, 600 ms, 900 ms) x 4 (target distance) repeated measures design for each object condition (*one object* and *two objects*). The procedure was identical to Experiment 2.1 (Chapter 2), so SOA and target distance were randomly intermixed in three blocks with self-timed breaks in-between. However, since in this case the cue was not predicting the location of the target, the target event appeared an equal number of times at each of the six object features. The total amount of trials was 432, and each of the three SOAs was adopted on 144 trials. Within each of these sets of trials, the cue and target appeared 24 times at each of the six object features in random order for each participant. The procedure took approximately 35 minutes.

Results and Discussion

Reaction times shorter than 200 ms and longer than 1200 ms were excluded from the analysis, resulting in excluding < 1% of the data in the *one object* condition, and 1% in the *two objects* condition.

For the *one object* condition, neither SOA, nor target distance made a statistical difference to reaction time performance, and there was no interaction between the two variables (all $F_s < 1$). Inspection of Figure 58 suggests that there were no pronounced trends in the data, including no cueing effect. The lack of cueing was expected, given that the cue had no predictive value for the target feature. Therefore, a pure object-based effect is likely to result in equal performance for all features. At SOA of 300 ms there was a slowing of responses for distance 2, but this is possible to be a random fluctuation due to low power. In any case, the results suggest uniform performance across the whole object, without cue-target distance variations. However, the fact that there was no main effect of SOA suggests the results did not follow the prediction that SOA of 300 ms would lead to faster reaction times than the longer SOAs.

Since the key prediction for the one object condition reflects a null effect of target distance, a Bayesian analysis was conducted in order to investigate if the current null result is related to genuine equality of means, or insufficient evidence in the data. This distinction is to be made based on BF values for the effect of each variable and the interaction as assessed against the null. A BF between 0.33 and 3 is typically considered as not enough evidence to distinguish if the model fits the data, while a BF < 0.33 is regarded as more evidence for the null than the alternative model (e.g. Dienes, 2011; Jeffreys, 1961). Results indicated consistent support for the null. Specifically, the effect of target distance alone was associated with a BF of 0.12 (this is based on testing for mean differences in any direction, i.e. without specifying an order restriction), SOA was reflected by a BF of 0.25, and for the distance-SOA interaction the BF was merely 0.003. Therefore, reaction time was not affected by any of the manipulated variables. Although this goes against the prediction for SOA, which was expected to result in a main effect, the pattern provides some support for object-based effects (i.e. equal performance within the same object). However, the lack of relative differences between SOA levels means it is not possible to conclude if there was any inhibition taking place.

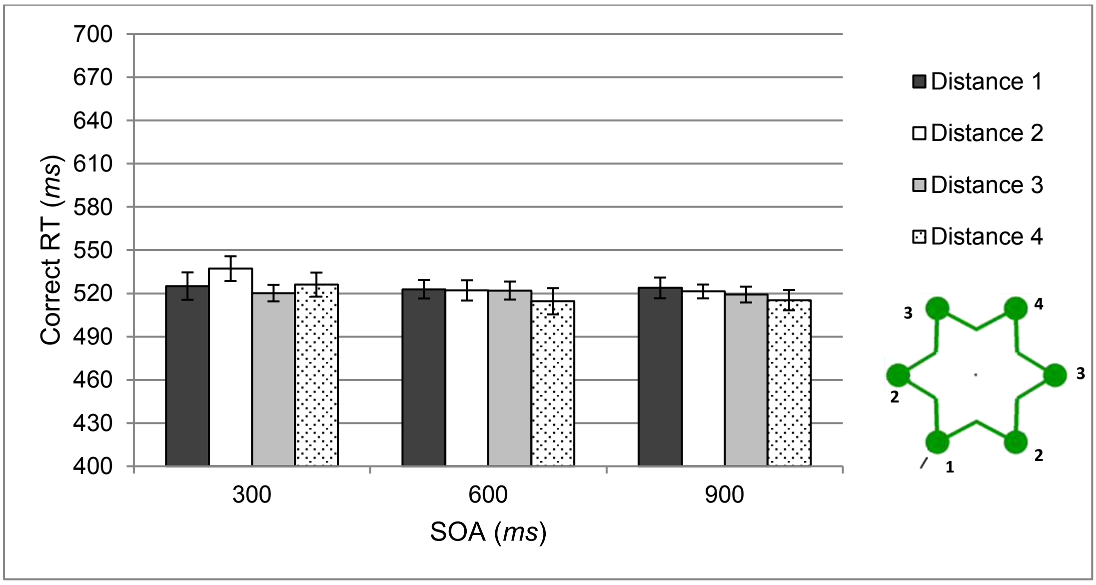

Figure 58: Mean correct reaction time (*ms*) as a function of SOA and target distance for Appendix 4, *one object* condition. Error bars represent corrected SEM.

For the *two objects* condition, correct reaction times were not affected by any of the manipulations. There was no variation in responses based on SOA, $F(2, 14) = 2.88$, $p = .089$, $\eta_p^2 = .29$, or target distance, $F(1.18, 8.28) = 2.01$, $p = .143$, $\eta_p^2 = .22$, and no interaction ($F < 1$). The trends in the data can be observed in Figure 59, suggesting that there was a non-significant tendency towards object-based facilitation, as targets at distances corresponding to the cued object (distance 1 and distance 3) were overall responded to faster than targets belonging to the uncued object (distance 2 and distance 4). Therefore, there was no indication of IoR. Also, as with the *one object* condition, there was no cueing effect. Overall, the trend in the performance pattern is opposite to the one predicted.

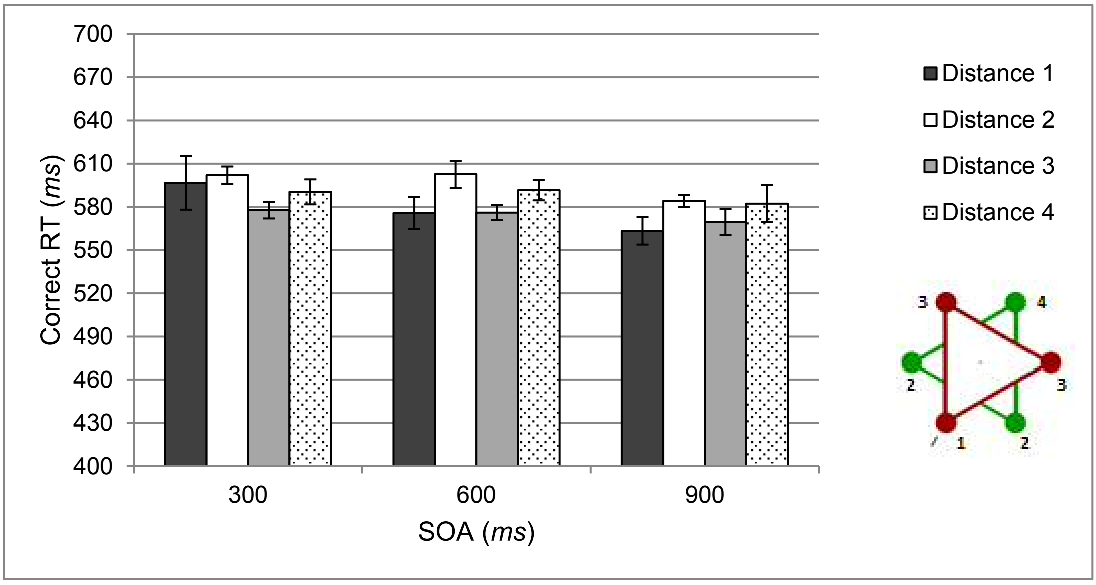

Figure 59: Mean correct reaction time (*ms*) as a function of SOA and target distance for Appendix 4, *two objects* condition. Error bars represent corrected SEM.

Given the null results and what appears to be a strong trend towards object-based facilitation for all SOAs, post-hoc Bayesian analyses were performed to investigate this trend further. It should be noted that this analysis aimed to specifically test the strength of evidence for object-based facilitation, rather than inhibition, as the trend clearly suggests that reaction times were faster within the cued object compared to the uncued object. Therefore, this was an exploratory analysis, not related to the original prediction. The full set of data was used, i.e. including distance 4, with SOA collapsed. The order restriction reflected object-based facilitation, such that reaction times for targets at distance 1 and distance 3 are faster than reaction times for targets at distance 2 and distance 4. This was assessed against the null hypothesis of equality between means, yielding a BF of 6.2 in support of object-based facilitation.

Overall, the results from this experiment indicate that extending the SOA interval alone and making the cue non-predictive of target location may not be enough to obtain IoR. Although the sample is too small to make definite conclusions, the trends in the data suggest that even if the power is increased, the effects likely to emerge would be towards facilitation, rather than inhibition. This was supported by Bayesian statistics for the *two objects* condition. A point worth noting is the lack of cueing effect for distance 1, combined with the object-based trend. This result suggests that selection can indeed be fully object-oriented, i.e. features within the cued object can be equally selected. This poses a challenge to the standard account of the cueing effect, which suggests that it is evidence for space-based selection (e.g. Egly et al., 1994). Instead, the current results propose that it may be due to strategic orientation. Here the cue did not predict the target location, so the cueing benefit was equally pronounced for all features within the cued object. This remains to be investigated further, but it is a realistic possibility given the current data.

Based on the current results, a full-scale experiment may require the use of an attractor stimulus during the SOA interval, as an additional measure that can contribute towards obtaining IoR. In addition, the SOA may need to be increased over 900 ms, given that the currently adopted intervals did not affect performance. The latter modification may also be necessary due to the nature of the task, namely target discrimination as opposed to onset detection, which is associated with a later IoR onset (Gabay et al., 2012).